การวิเคราะห์ข้อมูลผิดปกติโดยวิธีการเหมืองข้อมูลและการควบคุมกระบวนการเชิงสถิติ กรณีศึกษาข้อมูลปริมาณอนุภาคฝุ่นละอองของอุตสาหกรรมผลิตฮาร์ดดิสก์

OUT OF CONTROL ANALYSIS USING DATA MINING AND STATISTICAL PROCESS CONTROL: A CASE STUDY OF PARTICLE DATA IN HARD DISK DRIVE MANUFACTURING

กฤษฎ์ฐมรัฐฎร์ เพ็งอุบล
KRITHAMARATH PHENG-UBON

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศ
คณะเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2556
KMITL-2013-IT-M-001-001

ห้องสมุดคณะเทคโนโลยีสารสนเทศ พระจอมเกล้าลาดกระบัง

การวิเคราะห์ข้อมูลผิดปกติโดยวิธีการเหมืองข้อมูลและการควบคุมกระบวนการเชิงสถิติ กรณีศึกษาข้อมูลปริมาณอนุภาคฝุ่นละอองของอุตสาหกรรมผลิตฮาร์ดดิสก์

OUT OF CONTROL ANALYSIS USING DATA MINING AND STATISTICAL PROCESS CONTROL: A CASE STUDY OF PARTICLE DATA IN HARD DISK DRIVE MANUFACTURING

กฤษฎ์ฐมรัษฎร์ เพ็งอุบล

KRITHAMARATH PHENG-UBON

(copy 2)

เลขหมู่.....
เลขทะเบียน.....7167
วัน,เดือน,ปี.....17 ตค. 2556

.b.11259110
.i.....

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ.2556

KMITL-2013-IT-M-001-001

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

OUT OF CONTROL ANALYSIS USING DATA MINING AND STATISTICAL PROCESS CONTROL: A CASE STUDY OF PARTICLE DATA IN HARD DISK DRIVE MANUFACTURING

KRITHAMARATH PHENG-UBON

A THESIS SUBMITTED IN PATIAL FULFILLMENT

OF THE REQUIREMENT FOR THE DEGREE OF

MASTER OF SCIENCE IN INFORMATION TECHNOLOGY

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2013

KMITL-2013-IT-M-001-001

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COPYRIGHT 2013

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ใบรับรองวิทยานิพนธ์

หัวข้อวิทยานิพนธ์ การวิเคราะห์ข้อมูลผิดปกติโดยวิธีการเหมืองข้อมูลและการควบคุมกระบวนการเชิงสถิติ กรณีศึกษาข้อมูลปริมาณอนุภาคฝุ่นละอองของอุตสาหกรรมผลิตฮาร์ดดิสก์

Out of control analysis using data mining and statistical process control : A case study of particle data in hard disk drive manufacturing

นักศึกษา นางสาวกฤษฎ์ฐมรัษฎร์ เพ็งอุบล

รหัสประจำตัว 51066427

ปริญญา วิทยาศาสตรมหาบัณฑิต

สาขาวิชา เทคโนโลยีสารสนเทศ

อาจารย์ที่ปรึกษาวิทยานิพนธ์ รองศาสตราจารย์ ดร.พรฤดี เนติโสภากุล

คณะกรรมการสอบวิทยานิพนธ์	ลายมือชื่อ
รองศาสตราจารย์ ดร.วรพจน์ กรีสุระเดช	
รองศาสตราจารย์ ดร.นุชรี เปรมชัยสวัสดิ์	
รองศาสตราจารย์ ดร.อาริต ธรรมโน	
รองศาสตราจารย์ ดร.พรฤดี เนติโสภากุล	
ผู้ช่วยศาสตราจารย์ ดร.กันต์พงษ์ วรรัตน์ปัญญา	

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

วัน / เดือน/ ปี ที่สอบ วันอังคารที่ 9 เมษายน 2556 เวลา 13.00 น.

สถานที่สอบ ณ ห้อง 335 (ชั้น 3) คณะเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศรับรองแล้ว

(รองศาสตราจารย์ ดร.จันทร์บูรณ์ สถิตวิริยวงศ์)

คณบดีคณะเทคโนโลยีสารสนเทศ

วันที่ 14 **เดือน** เมษายน **พ.ศ.** 2556

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อ	การวิเคราะห์ข้อมูลผิดปกติโดยวิธีการเหมืองข้อมูลและการควบคุมกระบวนการเชิงสถิติ กรณีศึกษาข้อมูลปริมาณอนุภาคฝุ่นละอองของอุตสาหกรรมผลิตฮาร์ดดิสก์
นักศึกษา	นางสาวกฤษฎ์ฐมรัษฎร์ เพ็งอุบล
รหัสนักศึกษา	51066427
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
แขนงวิชา	วิทยาการสารสนเทศ
พ.ศ.	2556
อาจารย์ที่ปรึกษา	รศ.ดร. พรฤดี เนติโสภากุล

บทคัดย่อ

ในอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟนั้นประกอบด้วยกระบวนการผลิตทั้งหมดภายในห้องควบคุมปริมาณอนุภาคฝุ่นละอองซึ่งมีการควบคุมสภาพแวดล้อมหรือเรียกว่า "คลีนรูม" (Cleanroom) โดยการผลิตภายในห้องคลีนรูมนั้นต้องมีการตรวจสอบและควบคุมปริมาณอนุภาคฝุ่นละออง (Particle) ด้วยอุปกรณ์ตรวจจับหรือเซ็นเซอร์ในทุกๆ กระบวนการของการบรรจุชิ้นส่วนต่างๆ จนถึงการประกอบเป็นฮาร์ดดิสก์ไดร์ฟที่เสร็จสมบูรณ์ก่อนจะนำฮาร์ดดิสก์ไดร์ฟที่ผลิตไปทำการทดสอบคุณภาพด้วยเครื่องตรวจสอบคุณภาพสินค้าในขั้นตอนต่อไปซึ่งอยู่นอกห้องคลีนรูม หากปริมาณอนุภาคฝุ่นละอองที่เกิดขึ้นในกระบวนการผลิตมีปริมาณมากเกินกว่าการควบคุม (out of control) จะทำให้ฮาร์ดดิสก์ดิสก์ไดร์ฟที่ผลิตได้เกิดความเสียหายหรือคุณภาพไม่ดีตามต้องการ ดังนั้นปริมาณฝุ่นละอองที่จัดเก็บได้โดยอุปกรณ์ตรวจจับนั้นจึงมีความสำคัญสำหรับอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟในด้านการควบคุมกระบวนการ

วิทยานิพนธ์ฉบับนี้จึงนำเทคนิคการทำเหมืองข้อมูล (Data Mining) มาประยุกต์ใช้เพื่อวิเคราะห์ข้อมูลผิดปกติร่วมกับการควบคุมกระบวนการทางสถิติ (Statistical Process Control) โดยนำมาทำการเรียนรู้และศึกษาข้อมูลในอดีตเพื่อการค้นหาปัจจัยที่เป็นสาเหตุให้เกิดปริมาณฝุ่นละอองที่เกินควบคุมและตรวจจับค่าผิดปกติซึ่งจะนำไปสู่แนวทางในการปรับปรุงกระบวนการที่เหมาะสมซึ่งจะทำให้กระบวนการผลิตมีประสิทธิภาพมากขึ้นและทำให้มีของเสียออกจากการกระบวนการผลิตน้อยลง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Title	Out of Control Analysis Using Data Mining and Statistical Process Control: A Case Study of Particle Data in Hard Disk Drive Manufacturing.
Student	Krithamarath Pheng-ubon
Student ID.	51066427
Degree	Master of Science
Programme	Information Science
Year	2013
Advisor	Assoc.Prof.Dr. Ponrudee Netisopakul

ABSTRACT

In a hard disk drives manufacturing industry, the environment of all production processes are monitored under the "Cleanroom". The amount of particles is controlled for each process from packing components through completing the hard disk drives assembly using sensors. After the assembly, the hard drive is taken outside the cleanroom for quality control test. If the amount of particle in the production process is more than the controlled limit, it will cause damage or low quality product. Thus the important part of hard disk drive manufacturing industry is the process control, with the sensors to collect amount of particle.

This thesis applies the data mining technique with statistical process control to analyze the irregular data. Historical data is analyzed to find out the factors that cause excessive amounts of particle. Detection of outliers will lead to improve the hard disk drive production process with more efficiency.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความกรุณาจาก รศ.ดร.พรฤดี เนติโสภากุล อาจารย์ที่ปรึกษาวิทยานิพนธ์ที่ได้ให้คำเสนอแนะ แนวคิด ตลอดจนแก้ไขข้อบกพร่องต่างๆ มาโดยตลอด จนวิทยานิพนธ์ฉบับนี้เสร็จสมบูรณ์ ซึ่งข้าพเจ้าขอขอบพระคุณเป็นอย่างสูง

ขอขอบคุณศูนย์วิจัยร่วมเฉพาะทางด้านส่วนประกอบฮาร์ดดิสก์ไดรฟ์ I/UCRC (Industry/University Cooperative Research Center) ภายใต้ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (National Electronics and Computer Technology Center: NECTEC)

ขอขอบคุณวิทยาลัยร่วมด้านเทคโนโลยีการบันทึกข้อมูลและการประยุกต์ใช้งาน (Data Storage Technology and Applications: DSTAR) สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง และบริษัทฮิตาชิ โกลบอล สตอเรจ เทคโนโลยีส์ (ประเทศไทย) จำกัด ที่ให้การสนับสนุนทางด้านทุนวิจัยสำหรับโครงการวิจัย

ขอขอบพระคุณบิดา มารดา และครอบครัว ที่ให้คำปรึกษาและเป็นกำลังใจที่ดีเสมอมา

ขอขอบพระคุณคณาจารย์คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ทุกๆ ท่านที่ได้ประสิทธิ์ประสาทวิชาให้กับข้าพเจ้า

ขอบคุณเพื่อนๆ พี่ๆ น้องๆ ในคณะเทคโนโลยีสารสนเทศ และหน่วยปฏิบัติการการจัดการองค์ความรู้และวิศวกรรมความรู้ (KMAKE LAB)

ขอบคุณเจ้าหน้าที่คณะเทคโนโลยีสารสนเทศทุกๆ คนที่คอยให้ความช่วยเหลือและคำแนะนำต่างๆ เสมอมา รวมทั้งเจ้าหน้าที่ของบริษัทฮิตาชิ โกลบอล สตอเรจ เทคโนโลยีส์ (ประเทศไทย) จำกัด ที่ให้ข้อมูลอันเป็นประโยชน์ต่อการจัดทำวิทยานิพนธ์จนเสร็จสมบูรณ์

และผู้เขียนขอให้ความดีอันเกิดจากการศึกษาค้นคว้าครั้งนี้ มอบแด่บิดา มารดา ครู อาจารย์ และผู้มีพระคุณทุกท่าน ผู้เขียนมีความซาบซึ้งในความกรุณาอันดียิ่งจากทุกท่านที่ได้กล่าวนามมา และขอกราบขอบพระคุณมา ณ โอกาสนี้

กฤษฎ์ฐมรัษฎร์ เพ็งอุบล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย I

บทคัดย่อภาษาอังกฤษ II

กิตติกรรมประกาศ III

สารบัญ IV

สารบัญตาราง VII

สารบัญรูป X

บทที่ 1 บทนำ 1

1.1 ความเป็นมาและความสำคัญของปัญหา 2

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา 2

1.3 ทฤษฎีและแนวคิดที่ใช้ในการวิจัย 2

1.4 ขอบเขตการวิจัย 2

1.5 ขั้นตอนการดำเนินงานวิจัย 3

1.6 ประโยชน์ที่คาดว่าจะได้รับ 3

บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 4

2.1อุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟ 4

2.1.1 กระบวนการผลิตฮาร์ดดิสก์ไดร์ฟ (Hard Disk Drive Manufacturing Process) 4

2.1.2 กระบวนการทำงานของห้องคลีนรูม ในอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟ 6

2.1.3 การควบคุมคุณภาพการผลิต 7

2.2 การควบคุมคุณภาพเชิงสถิติ (Statistical Quality Control) 8

2.2.1 การควบคุมกระบวนการเชิงสถิติ (Statistical Process Control) 8

2.2.2 แผนภูมิควบคุม (Control chart) 10

2.2.3 แผนภูมิควบคุมค่าวัดเชิงเดี่ยว (Individual control chart: X-chart) 13

2.2.4 แผนภูมิค่าพิสัยเคลื่อนที่ (Moving range chart) 14

2.2.5 การตีความแผนภูมิควบคุมคุณภาพ 15

2.2.6 มาตรฐานการควบคุมคุณภาพ 20

2.3 เทคนิคเหมืองข้อมูลชนิดการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจ 21

2.3.1 ปัญหาการไม่สมดุลของข้อมูล (Imbalanced datasets Problem) 23

2.3.2 กระบวนการจำแนกข้อมูล 24

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

2.3.3 การจำแนกข้อมูลด้วยต้นไม้ตัดสินใจ (Decision Tree Classification) 25
2.3.4 เครื่องมือสำหรับการสร้างต้นไม้ตัดสินใจ 27
2.3.5 วิธี Cross-Validation 28
2.3.6 การวัดประสิทธิภาพของแบบจำลอง 30
2.4 งานวิจัยที่เกี่ยวข้อง 31
2.4.1 การใช้เทคนิคเหมืองข้อมูลในการปรับปรุงคุณภาพการผลิต: กรณีศึกษาของโรงงานบรรจุ LCD Driver IC 31
2.4.2 ศึกษาระบบการปรับปรุงคุณภาพอย่างต่อเนื่องของการผลิตแอลอีดีด้วยเหมืองข้อมูล 35
บทที่ 3 วิธีดำเนินการวิจัย 38
3.1 การศึกษาและเก็บรวบรวมข้อมูลปริมาณฝุ่นละอองจากระบบเฝ้าระวังและตรวจจับฝุ่นละออง 39
3.1.1 ลักษณะข้อมูลเบื้องต้นของระบบเฝ้าระวังและตรวจจับฝุ่นละออง 39
3.1.2 การเตรียมข้อมูลสำหรับการวิเคราะห์ข้อมูล 40
3.1.3 การสำรวจข้อมูลเบื้องต้น 45
3.2 การออกแบบมาตรฐานสำหรับการควบคุมกระบวนการเชิงสถิติของกรณีศึกษา 50
3.2.1 ชนิดของแผนภูมิควบคุมที่เหมาะสมสำหรับข้อมูลปริมาณอนุภาคฝุ่นละออง 50
3.2.2 มาตรฐานการควบคุมคุณภาพ 52
3.3 การหาคฺของปัจจัยที่ทำให้ปริมาณอนุภาคฝุ่นละอองเกินควบคุมด้วยเหมืองข้อมูล 53
3.3.1 การแก้ปัญหาความไม่สมดุลของข้อมูล 53
3.3.2 การเลือกใช้เทคนิคเหมืองข้อมูล 53
3.3.3 การแบ่งกลุ่มข้อมูลที่ใช้ทดสอบ 54
3.3.4 การวัดประสิทธิภาพของแบบจำลอง 54
3.3.5 การวิเคราะห์และการประเมินผล 54
บทที่ 4 ผลการทดลองและวิเคราะห์เปรียบเทียบประสิทธิภาพของต้นไม้ตัดสินใจ 56
4.1 ผลการทดลองและการอภิปรายผลการควบคุมกระบวนการเชิงสถิติของข้อมูลปริมาณอนุภาคฝุ่นละออง 56
4.1.1 ผลการสร้างแผนภูมิควบคุมของชุดข้อมูล พื้นที่การทำงาน cleaning 56

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

4.1.2 ผลการสร้างแผนภูมิควบคุมของชุดข้อมูล พื้นที่การทำงาน HDE 63
4.1.3 ผลการสร้างแผนภูมิควบคุมของชุดข้อมูล พื้นที่การทำงาน HSA 69
4.2 ผลการทดลองและการอภิปรายผลส่วนของกฎที่ได้จากการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจ 75
4.2.1 ผลการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจของกลุ่มข้อมูลการควบคุมอนุภาคฝุ่นละอองตามเกณฑ์ของโรงงาน 75
4.2.2 ผลการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจของกลุ่มข้อมูลการควบคุมอนุภาคฝุ่นละอองด้วยการควบคุมกระบวนการเชิงสถิติ 81
4.2.3 ผลการวัดประสิทธิภาพของการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจ 105
4.2.4 การวิเคราะห์และประเมินผล 107
บทที่ 5 บทสรุปและข้อเสนอแนะ 110
5.1 บทสรุป 110
5.2 ข้อจำกัดและข้อเสนอแนะ 113
บรรณานุกรม 114
ภาคผนวก 116
บทความและผลงานวิจัยที่ได้รับการตีพิมพ์ 117
ประวัติผู้เขียน 133

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ หน้า

2.1 มาตรฐานการปนเปื้อนของพื้นที่ควบคุมในอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟ 7

2.2 ระดับความเชื่อมั่นของการยอมรับระดับต่างๆ 11

2.3 วิธีคำนวณค่าพิสัยเคลื่อนที่ เมื่อ m = 2 และ m = 3 15

2.4 ตัวอย่างข้อมูลการซื้อคอมพิวเตอร์ 22

2.5 ปัญหาทางด้านคุณภาพของการบรรจุไดรเวอร์ไอซี 32

2.6 ตัวอย่างแอททริบิวต์และปัญหาคุณภาพ 33

2.7 ตัวอย่างข้อมูลปัญหาคุณภาพ 33

2.8 ผลการปรับปรุงคุณภาพ 34

2.9 ข้อมูลปัญหาคุณภาพของการบรรจุแอลอีดี 35

2.10 เปรียบเทียบแนวคิดเพื่อปรับปรุงงานวิจัยการควบคุมคุณภาพด้วยกระบวนการเชิงสถิติและเหมืองข้อมูล 37

3.1 ตารางอธิบายข้อมูลจากระบบงานเฝ้าระวังฝุ่นละออง 40

3.2 ตารางการแปลงข้อมูลตามมิติพื้นที่การทำงาน 43

3.3 เกณฑ์การควบคุมปริมาณอนุภาคฝุ่นละอองโดยโรงงานอุตสาหกรรม 52

3.4 ปัจจัยนำเข้าของข้อมูลปริมาณอนุภาคฝุ่นละออง 54

4.1 ผลการสร้างแผนภูมิควบคุมจากข้อมูลพื้นที่ cleaning 57

4.2 ผลการสร้างแผนภูมิควบคุมจากข้อมูลพื้นที่ HDE 64

4.3 ผลการสร้างแผนภูมิควบคุมจากข้อมูลพื้นที่ HSA 69

4.4 จำนวนข้อมูลสำหรับการสร้างแบบจำลองต้นไม้ตัดสินใจชุดข้อมูลพื้นที่ cleaning.....76

4.5 ค่าความถูกต้องและค่าเฉลี่ยความคลาดเคลื่อนจากการสร้างแบบจำลองต้นไม้ตัดสินใจชุดข้อมูลพื้นที่ cleaning 76

4.6 กฎจากต้นไม้ตัดสินใจข้อมูลพื้นที่ cleaning 77

4.7 จำนวนข้อมูลสำหรับการสร้างแบบจำลองต้นไม้ตัดสินใจพื้นที่ HDE 77

4.8 ค่าความถูกต้องและค่าเฉลี่ยความคลาดเคลื่อนจากการสร้างแบบจำลองต้นไม้ตัดสินใจพื้นที่ HDE 77

4.9 กฎจากต้นไม้ตัดสินใจข้อมูลพื้นที่ HDE 79

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ หน้า
4.10 จำนวนข้อมูลสำหรับการสร้างแบบจำลองต้นไม้ตัดสินใจพื้นที่ HSA 80
4.11 ค่าความถูกต้องและค่าเฉลี่ยความคลาดเคลื่อนจากการสร้างแบบจำลองต้นไม้ตัดสินใจพื้นที่ HSA 80
4.12 เปรียบเทียบค่าความถูกต้องและค่าเฉลี่ยความคลาดเคลื่อนจากการสร้างแบบจำลองต้นไม้ตัดสินใจของข้อมูลการควบคุมอนุภาคฝุ่นละอองตามเกณฑ์ของโรงงาน 81
4.13 จำนวนข้อมูลสำหรับการสร้างแบบจำลองต้นไม้ตัดสินใจพื้นที่ Cleaning 82
4.14 ค่าความถูกต้องและค่าเฉลี่ยความคลาดเคลื่อนจากการสร้างแบบจำลองต้นไม้ตัดสินใจพื้นที่ cleaning 82
4.15 ค่าประสิทธิภาพของแบบจำลองต้นไม้ตัดสินใจพื้นที่ cleaning 82
4.16 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 1 พื้นที่ cleaning 84
4.17 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 2 พื้นที่ cleaning 85
4.18 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 5 พื้นที่ cleaning 86
4.19 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 6 พื้นที่ cleaning 87
4.20 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 8 พื้นที่ cleaning 88
4.21 จำนวนข้อมูลสำหรับการสร้างแบบจำลองต้นไม้ตัดสินใจพื้นที่ HDE 89
4.22 ค่าความถูกต้องและค่าเฉลี่ยความคลาดเคลื่อนจากการสร้างแบบจำลองต้นไม้ตัดสินใจชุดข้อมูลพื้นที่ HDE 89
4.23 ค่าประสิทธิภาพของแบบจำลองต้นไม้ตัดสินใจพื้นที่ HDE 89
4.24 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 2 พื้นที่ HDE 91
4.25 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 7 พื้นที่ HDE 94
4.26 จำนวนข้อมูลสำหรับการสร้างแบบจำลองต้นไม้ตัดสินใจพื้นที่ HSA 96
4.27 ค่าความถูกต้องและค่าเฉลี่ยความคลาดเคลื่อนจากการสร้างแบบจำลองต้นไม้ตัดสินใจชุดข้อมูลพื้นที่ HSA 96
4.28 ค่าประสิทธิภาพของแบบจำลองต้นไม้ตัดสินใจพื้นที่ HSA 96

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ หน้า

4.29 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 2 พื้นที่ HSA 98

4.30 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 6 พื้นที่ HSA 101

4.31 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 7 พื้นที่ HSA 102

4.32 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 8 พื้นที่ HSA 103

4.33 เปรียบเทียบค่าความถูกต้องในการจำแนก, ค่าความครบถ้วน, ค่าความแม่นยำ และค่าถ่วงดุล ของต้นไม้ตัดสินใจพื้นที่การทำงาน cleaning 106

4.34 เปรียบเทียบค่าความถูกต้องในการจำแนก, ค่าความครบถ้วน, ค่าความแม่นยำ และค่าถ่วงดุล ของต้นไม้ตัดสินใจพื้นที่การทำงาน HDE 106

4.35 เปรียบเทียบค่าความถูกต้องในการจำแนก, ค่าความครบถ้วน, ค่าความแม่นยำ และค่าถ่วงดุล ของต้นไม้ตัดสินใจพื้นที่การทำงาน HSA 107

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่ หน้า

2.1 ส่วนประกอบของฮาร์ดดิสก์ไดร์ฟ 5

2.2 กระบวนการผลิตฮาร์ดดิสก์ไดร์ฟ 6

2.3 โครงสร้างแนวคิดการป้องกันในระหว่างกระบวนการผลิต 9

2.4 แสดงค่า Z และพื้นที่ใต้โค้งปกติ 11

2.5 องค์ประกอบของแผนภูมิควบคุม 12

2.6 การเกิดจุดอยู่นอกขอบเขตการควบคุม 16

2.7 การเกิดการไม่สมดุล (Runs) 17

2.8 การเกิดแนวโน้ม (Trends) 17

2.9 การเกิดจุดอยู่นอกเขตควบคุมเตือน 18

2.10 การเกิดวัฏจักร (Cycle) 18

2.11 การเกิดใกล้เส้นกลาง 19

2.12 การเกิดการผสม 20

2.13 โมเดลต้นไม้ตัดสินใจ 22

2.14 (a) ตัวอย่าง K-nearest neighbors ของข้อมูลที่พิจารณา เมื่อ k=6 23

2.14 (b) ข้อมูลสังเคราะห์ (Synthetic Instance) ที่สร้างขึ้น 23

2.15 กระบวนการจำแนกข้อมูล 24

2.16 โครงสร้างของแผนภูมิต้นไม้ตัดสินใจ 25

2.17 การแบ่งกลุ่มข้อมูลเพื่อทดสอบ Cross-Validation แบบ Hold-out ที่ร้อยละ 50 28

2.18 ขั้นตอนที่ 1 ของการแบ่งกลุ่มข้อมูลเพื่อทดสอบ Cross-Validation แบบ 10-fold 29

2.19 ขั้นตอนที่ 2 ของการแบ่งกลุ่มข้อมูลเพื่อทดสอบ Cross-Validation แบบ 10-fold 30

2.20 โมเดลการควบคุมและปรับปรุงคุณภาพการผลิต 36

3.1 ลักษณะข้อมูลปริมาณอนุภาคฝุ่นละอองจากระบบเฝ้าระวังอนุภาคฝุ่นละออง 39

3.2 ตัวอย่างการจัดระเบียบข้อมูล 41

3.3 ตัวอย่างข้อมูลตรวจจับอนุภาคฝุ่นละอองหลังการเตรียมข้อมูล 45

3.4 กราฟการกระจายของปริมาณอนุภาคฝุ่นละอองเปรียบเทียบวันกับกะการทำงานในพื้นที่การทำงาน cleaning 46

3.5 กราฟการกระจายของปริมาณอนุภาคฝุ่นละอองเปรียบเทียบวันกับกระบวนการทำงานในพื้นที่การทำงาน HDE 47

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่ หน้า

3.6 กราฟการกระจายของปริมาณอนุภาคฝุ่นละอองเปรียบเทียบวันกับกะทำงานในพื้นที่การทำงาน HDE 48

3.7 กราฟการกระจายของปริมาณอนุภาคฝุ่นละอองเปรียบเทียบวันกับกระบวนการทำงานในพื้นที่การทำงาน HSA 49

3.8 กราฟการกระจายของปริมาณอนุภาคฝุ่นละอองเปรียบเทียบวันกับกะทำงานในพื้นที่การทำงาน HSA 50

3.9 ตัวอย่างข้อมูลที่ได้หลังจากการควบคุมคุณภาพด้วยแผนภูมิควบคุม 53

4.1 แผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 1 57

4.2 ภาพขยายของแผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 1 ที่เรคคอร์ด 1989 ถึง 2120 57

4.3 แผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 2 58

4.4 ภาพขยายของแผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 2 ที่เรคคอร์ด 3131 ถึง 3294 58

4.5 แผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 3 59

4.6 ภาพขยายของแผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 3 ที่เรคคอร์ด 3618 ถึง 3749 59

4.7 แผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 4 59

4.8 แผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 5 60

4.9 ภาพขยายของแผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 5 ที่เรคคอร์ด 1557 ถึง 1691 60

4.10 แผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 6 60

4.11 ภาพขยายของแผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 6 ที่เรคคอร์ด 1416 ถึง 1548 61

4.12 แผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 7 61

4.13 แผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 8 61

4.14 ภาพขยายของแผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 8 ที่เรคคอร์ด 1557 ถึง 1798 62

4.15 แผนภูมิควบคุมแผนภูมิค่าพิสัยเคลื่อนที่ของพื้นที่การทำงาน cleaning 62

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่ หน้า

4.16 จำนวนครั้งที่เกินควบคุมในพื้นที่การทำงาน cleaning 63
4.17 แผนภูมิควบคุมของพื้นที่การทำงาน HDE กฎการควบคุมข้อ 1 64
4.18 แผนภูมิควบคุมของพื้นที่การทำงาน HDE กฎการควบคุมข้อ 2 65
4.19 แผนภูมิควบคุมของพื้นที่การทำงาน HDE กฎการควบคุมข้อ 3 65
4.20 แผนภูมิควบคุมของพื้นที่การทำงาน HDE กฎการควบคุมข้อ 4 66
4.21 แผนภูมิควบคุมของพื้นที่การทำงาน HDE กฎการควบคุมข้อ 5 66
4.22 แผนภูมิควบคุมของพื้นที่การทำงาน HDE กฎการควบคุมข้อ 6 67
4.23 แผนภูมิควบคุมของพื้นที่การทำงาน HDE กฎการควบคุมข้อ 7 67
4.24 แผนภูมิควบคุมของพื้นที่การทำงาน HDE กฎการควบคุมข้อ 8 68
4.25 แผนภูมิควบคุมแผนภูมิค่าพิสัยเคลื่อนที่ของพื้นที่การทำงาน HDE 68
4.26 จำนวนครั้งที่เกินควบคุมในพื้นที่การทำงาน HDE 69
4.27 แผนภูมิควบคุมของพื้นที่การทำงาน HSA กฎการควบคุมข้อ 1 70
4.28 แผนภูมิควบคุมของพื้นที่การทำงาน HSA กฎการควบคุมข้อ 2 70
4.29 แผนภูมิควบคุมของพื้นที่การทำงาน HSA กฎการควบคุมข้อ 3 71
4.30 แผนภูมิควบคุมของพื้นที่การทำงาน HSA กฎการควบคุมข้อ 4 71
4.31 แผนภูมิควบคุมของพื้นที่การทำงาน HSA กฎการควบคุมข้อ 5 72
4.32 แผนภูมิควบคุมของพื้นที่การทำงาน HSA กฎการควบคุมข้อ 6 72
4.33 แผนภูมิควบคุมของพื้นที่การทำงาน HSA กฎการควบคุมข้อ 7 73
4.34 แผนภูมิควบคุมของพื้นที่การทำงาน HSA กฎการควบคุมข้อ 8 73
4.35 แผนภูมิควบคุมแผนภูมิค่าพิสัยเคลื่อนที่ของพื้นที่การทำงาน HSA 74
4.36 จำนวนครั้งที่เกินควบคุมในพื้นที่ทำงาน HSA 74
4.37 ลักษณะผลที่ได้จากการหาแบบจำลองในโปรแกรม 75
4.38 ต้นไม้ตัดสินใจชุดข้อมูลพื้นที่ cleaning ตามเกณฑ์ควบคุมของโรงงาน 76
4.39 ต้นไม้ตัดสินใจชุดข้อมูลพื้นที่ HDE ตามเกณฑ์ควบคุมของโรงงาน 78
4.40 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 1 พื้นที่ cleaning 83
4.41 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 2 พื้นที่ cleaning 85
4.42 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 5 พื้นที่ cleaning 86
4.43 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 6 พื้นที่ cleaning 87

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่ หน้า

4.44 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 8 พื้นที่ cleaning 88
4.45 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 2 พื้นที่ HDE 90
4.46 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 7 พื้นที่ HDE 93
4.47 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 2 พื้นที่ HSA 97
4.48 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 6 พื้นที่ HSA 100
4.49 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 7 พื้นที่ HSA 102
4.50 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 8 พื้นที่ HSA 103

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

กระบวนการผลิตฮาร์ดดิสก์มีกระบวนการที่สำคัญอยู่หลายขั้นตอน แต่ทุกๆ ขั้นตอนในการผลิตและประกอบฮาร์ดดิสก์ที่มีคุณภาพนั้น ได้ดำเนินการอยู่ในห้องควบคุมสภาวะหรือที่เรียกกันว่า "ห้องคลีนรูม" (Clean room) โดยห้องคลีนรูมนี้จะดำเนินการประกอบชิ้นส่วนฮาร์ดดิสก์ตั้งแต่ขั้นตอนเริ่มต้นจนถึงขั้นตอนการประกอบชิ้นส่วนชิ้นสุดท้าย ดังนั้นการควบคุมและเฝ้าระวังห้องคลีนรูมให้ไม่มีสิ่งปลอมปนในอากาศจึงเป็นเรื่องสำคัญที่ไม่สามารถมองข้ามได้ เนื่องจากปริมาณของอนุภาคฝุ่นละอองนั้นอาจมีผลเสียต่อกระบวนการประกอบชิ้นส่วนภายในอุตสาหกรรม

การนำข้อมูลปริมาณอนุภาคฝุ่นละอองในอดีตมาวิเคราะห์เพื่อหารูปแบบของปัจจัยเสี่ยงที่เป็นสาเหตุให้เกิดปริมาณอนุภาคฝุ่นละอองที่เกินควบคุม (Out of control) ในอนาคตนั้นจึงเป็นสิ่งสำคัญและอาจเป็นประโยชน์ในการช่วยให้สามารถเฝ้าระวังและปรับปรุงกระบวนการได้อย่างถูกต้อง โดยวิทยานิพนธ์ฉบับนี้จะนำวิธีการเหมืองข้อมูลแบบการจำแนกข้อมูล (Classification) ด้วยเทคนิคต้นไม้การตัดสินใจ (Decision tree) และการควบคุมกระบวนการเชิงสถิติ (Statistical Process Control: SPC) ซึ่งเป็นเครื่องมือช่วยติดตามและปรับปรุงกระบวนการผลิตให้เป็นมาตรฐานและมีคุณภาพอย่างต่อเนื่องซึ่งจัดเป็นการควบคุมที่เน้นการป้องกันมากกว่าการตรวจจับ ด้วยแผนภูมิควบคุมชนิดผันแปร (Variable control chart) เป็นเครื่องมือทางสถิติที่สามารถระบุถึงการผันแปรของกระบวนการผลิตจากการกระจายตัวของข้อมูลในแผนภูมิควบคุมคุณภาพ

การนำความสามารถของการจำแนกประเภทข้อมูลด้วยต้นไม้การตัดสินใจและการควบคุมกระบวนการเชิงสถิติด้วยวิธีการใช้แผนภูมิควบคุม โดยอาศัยความสัมพันธ์ของข้อมูลในอดีตเพื่อใช้ในการวิเคราะห์ข้อมูลที่เป็นข้อมูลปริมาณอนุภาคฝุ่นละอองที่เกินควบคุมหรือเรียกว่าข้อมูลที่ผิดปกติ เพื่อสร้างระเบียบวิธีหรือตัวแบบในการหาข้อมูลที่ผิดปกติในกระบวนการผลิตเพื่อเป็นแนวทางในการปรับปรุงกระบวนการผลิตให้มีประสิทธิภาพยิ่งขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

1. ศึกษาลักษณะธรรมชาติของข้อมูลผิดปกติของข้อมูลปริมาณอนุภาคฝุ่นละอองในโรงงานอุตสาหกรรมการผลิตฮาร์ดดิสก์
2. สร้างระเบียบวิธีหรือตัวแบบที่เหมาะสมเพื่อใช้ในการตรวจหาปัจจัยที่ทำให้ปริมาณอนุภาคฝุ่นละอองผิดปกติ
3. เปรียบเทียบประสิทธิภาพของระเบียบวิธีที่ใช้ในการตรวจหาปัจจัยที่ทำให้เกิดอนุภาคฝุ่นละอองผิดปกติระหว่างวิธีการจำแนกข้อมูลด้วยต้นไม้การตัดสินใจและวิธีการจำแนกข้อมูลด้วยวิธีต้นไม้การตัดสินใจร่วมกับการควบคุมกระบวนการเชิงสถิติ

1.3 ทฤษฎีและแนวคิดที่ใช้ในการวิจัย

ในการควบคุมคุณภาพผลิตภัณฑ์นั้นจำเป็นต้องอาศัยกระบวนการเชิงควบคุมทางสถิติเข้ามาเพิ่มเติมความถูกต้องและเป็นมาตรฐาน วิทยานิพนธ์ฉบับนี้จึงนำแนวคิดการควบคุมกระบวนการเชิงสถิติมาใช้ในการควบคุมคุณภาพของกระบวนการและในขณะเดียวกันในการควบคุมกระบวนการผลิตนั้นย่อมเกิดสถานะที่ไม่สามารถควบคุมได้ซึ่งสถานะดังกล่าวก่อให้เกิดผลเสียต่อผลิตภัณฑ์ ดังนั้นจึงนำเทคนิคการจำแนกข้อมูลเข้ามาใช้ร่วมกันในการวิเคราะห์ข้อมูลเพื่อหาปัจจัยที่ก่อให้เกิดความเสียหายในกระบวนการผลิตภายใต้การควบคุมคุณภาพทางการผลิต

1.4 ขอบเขตการวิจัย

1. เป็นการประมวลผลแบบ off-line
2. เป็นกรณีศึกษาสำหรับข้อมูลปริมาณอนุภาคฝุ่นละอองในโรงงานอุตสาหกรรมการผลิตฮาร์ดดิสก์ของโรงงานฮิตาชิ โกลบอลสตอเรจ ประเทศไทยจำกัด ข้อมูลตัวอย่างระหว่างวันที่ 1 มีนาคม 2553 ถึง วันที่ 31 มีนาคม 2553

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.5 ขั้นตอนการดำเนินงานวิจัย

1. ศึกษาและเก็บรวบรวมข้อมูลปริมาณฝุ่นละอองจากระบบเฝ้าระวังและตรวจจับฝุ่นละออง
2. ศึกษาวิเคราะห์ความสัมพันธ์ของข้อมูลเบื้องต้นทางสถิติและหลักการควบคุมกระบวนการเชิงสถิติและเทคนิคการจำแนกข้อมูล
3. ออกแบบการทดลองเพื่อเปรียบประสิทธิภาพสำหรับการจำแนกข้อมูล
4. ดำเนินการจำแนกข้อมูลและตรวจหาข้อมูลผิดปกติด้วยการทดลองที่ออกแบบไว้
5. เปรียบเทียบประสิทธิภาพระเบียบวิธีที่ใช้ในวิเคราะห์ข้อมูล
6. วิเคราะห์ข้อมูลที่ได้จากการจำแนกข้อมูลและสรุปผลดำเนินการ
7. รายงานผลการดำเนินการประกอบด้วยการจัดทำรายงานฉบับสมบูรณ์ และจัดทำเอกสารที่เกี่ยวข้อง

1.6 ประโยชน์ที่คาดว่าจะได้รับ

1. ระเบียบวิธีหรือรูปแบบของการเกิดสภาวะนอกการควบคุม (out of control) ซึ่งจะเป็นประโยชน์ในการแก้ปัญหาการเกิดสภาวะนอกการควบคุมของโรงงานอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟของโรงงานฮิตาชิ โกลบอลสตอเรจ ประเทศไทยจำกัด
2. นำเทคนิคทางเหมืองข้อมูลและการควบคุมกระบวนการเชิงสถิติเพื่อปรับปรุงกระบวนการผลิตได้
3. ความรู้จากการวิเคราะห์ข้อมูลด้วยเทคนิคการจำแนกประเภทข้อมูลและการควบคุมกระบวนการเชิงสถิติเพื่อใช้ในการปรับปรุงกระบวนการผลิตฮาร์ดดิสก์เพื่อให้การผลิตมีประสิทธิภาพยิ่งขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึง งานวิจัยทางด้านการวิเคราะห์ข้อมูลด้วยเทคนิคการจำแนกข้อมูลร่วมกับการควบคุมกระบวนการเชิงสถิติและ โดยหัวข้อ 2.1 จะกล่าวถึง อุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟ หัวข้อที่ 2.2 จะกล่าวถึง การควบคุมคุณภาพเชิงสถิติด้วยเทคนิคประเภทต่างๆ หัวข้อที่ 2.3 กล่าวถึงการวิเคราะห์ข้อมูลเทคนิคเหมืองข้อมูลชนิดการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจ และหัวข้อที่ 2.4 จะกล่าวถึงงานวิจัยที่เกี่ยวข้องกับการควบคุมคุณภาพเชิงสถิติร่วมกับการวิเคราะห์ข้อมูลด้วยเทคนิคเหมืองข้อมูล

2.1 อุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟ

การผลิตฮาร์ดดิสก์ไดร์ฟให้มีคุณภาพนั้น การศึกษากระบวนการในแต่ละขั้นตอนอย่างละเอียดนั้นมีความสำคัญเนื่องจากทำให้สามารถวิเคราะห์ข้อมูลคุณภาพได้อย่างมีประสิทธิภาพ

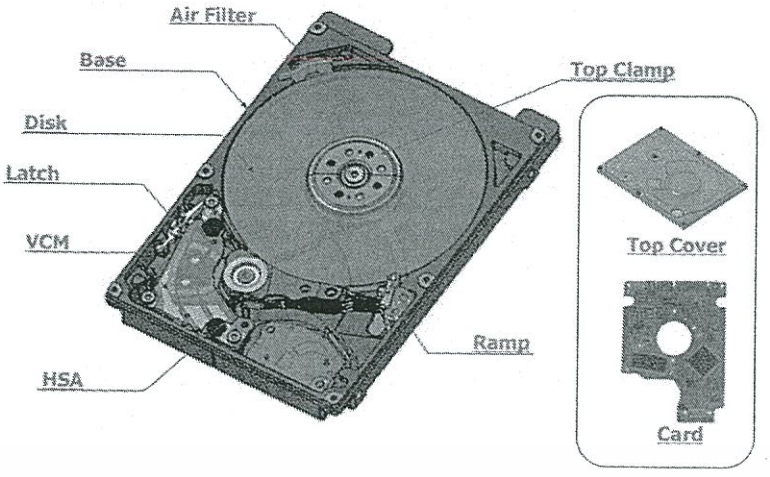
2.1.1 กระบวนการผลิตฮาร์ดดิสก์ไดร์ฟ (Hard Disk Drive Manufacturing Process)

ในการระบวนการผลิตฮาร์ดดิสก์ไดร์ฟนั้นจะประกอบด้วยการประกอบชิ้นส่วน (Hard disk part function) ต่างๆ ซึ่งเป็นองค์ประกอบของฮาร์ดดิสก์เข้าด้วยกัน ประกอบด้วย

1. Voice Coil Motor controller (VCM) เป็นชิ้นส่วนที่มีหน้าที่ควบคุมการหมุนของมอเตอร์และการเคลื่อนที่ของหัวอ่าน
2. Air Filter ใช้สำหรับการกรองอนุภาคที่เป็นอากาศระหว่างการดำเนินการ
3. Disk เป็นชั้นแม่เหล็กขนาดเล็กที่ใช้เขียนระเบียนข้อมูลไบนารี
4. Base เป็นฐานในการบรรจุองค์ประกอบทั้งหมดของฮาร์ดดิสก์
5. Latch เป็นสลักที่ใช้สำหรับยึด VCM กับดิสก์
6. Head Stack Assembly (HSA) เป็นหัวอ่านที่ใช้ในการอ่านเขียนสัญญาณแม่เหล็ก
7. Ramp เป็นส่วนประกอบที่ประกอบด้วยหัวเขียน-อ่าน ที่อยู่เหนือดิสก์
8. Card เป็นแผงวงจรควบคุมต่างๆ ของฮาร์ดดิสก์ไดร์ฟ
9. Top Clamp เป็นชิ้นส่วนอลูมิเนียมใช้สำหรับยึดแผ่นดิสก์กับมอเตอร์
10. Top Cover เป็นวัสดุสแตนเลสที่ใช้ป้องกันวัสดุจากภายนอกจะเข้ามาภายในของดิสก์

โดยส่วนประกอบของฮาร์ดดิสก์ไดร์ฟดังกล่าวนี้สามารถแสดงได้ดังรูปที่2.1[1]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.1 ส่วนประกอบของฮาร์ดดิสก์ไดร์ฟ

ในกระบวนการผลิตฮาร์ดดิสก์ไดร์ฟนั้นมีการทำงานที่ผสมผสานระหว่างการทำงานของเครื่องจักรกึ่งอัตโนมัติและแรงงานมนุษย์ กระบวนการผลิตนี้เป็น[2]การประกอบของ HSA, Disk, Motor, VCM, และชิ้นส่วนเล็กๆ อื่นๆ ลงบน Base แล้วจึงปิดส่วนประกอบทั้งหมดด้วย Top Cover กระบวนการดังกล่าวนั้นเรียกว่า "Hard disk assembly" (HDA) กระบวนการผลิตต้องดำเนินการภายในห้องควบคุมสภาวะหรือห้องคลีนรูม (Clean room) ซึ่งมีการควบคุมมาตรฐานปริมาณฝุ่นละอองในพื้นที่การทำงาน (Area) โดยมาตรฐานของปริมาณฝุ่นละอองที่เกิดขึ้นจะมีขนาดหรือจำนวนที่มากน้อยเพียงใดนั้นจะขึ้นอยู่กับความสำคัญของชิ้นส่วนหรือกระบวนการผลิตในแต่ละส่วน อันได้แก่กระบวนการผลิต 3 กระบวนการ ได้แก่

1. HSA (Head stack assembly Process) คือกระบวนการประกอบหัวอ่านและเขียนสัญญาณแม่เหล็ก HGA (Head gimbals assembly) กับชิ้นส่วนต่างๆ เช่น แผ่นวงจรไฟฟ้า (Flex cable) และตัวสร้างสนามแม่เหล็ก (Carriage) รวมถึงการทดสอบต่างๆ

2. HDE (Hard Disk Enclosure Process) คือกระบวนการประกอบ HSA กับชิ้นส่วนต่างๆ เข้าด้วยกัน (ยกเว้น card) เช่น Disk, VCM แล้วผ่านกระบวนการสร้าง Servo pattern ลงบนแผ่น Disk รวมถึงการตรวจสอบฝุ่นละอองและรอยรั่ว

3. HDD (Hard disk drive Process) คือกระบวนการขั้นตอนสุดท้ายในการผลิตฮาร์ดดิสก์ไดร์ฟ แบ่งการทำงานเป็น 3 ส่วน ได้แก่

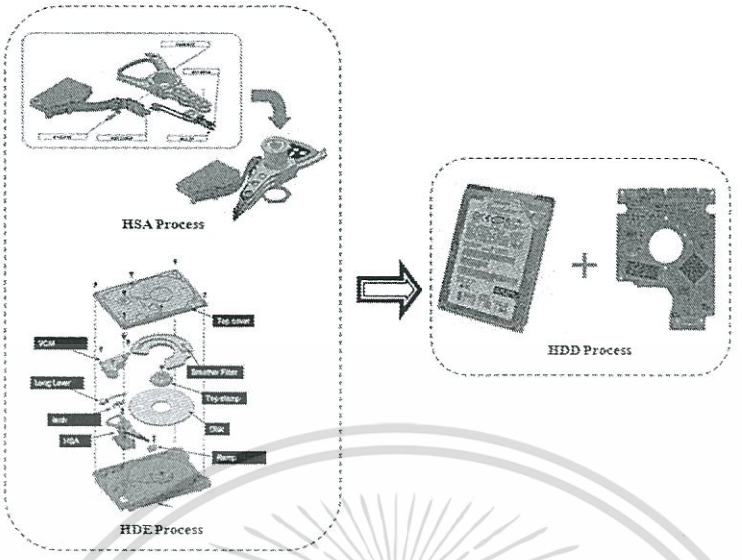
3.1 Final assembly คือการประกอบแผงวงจรควบคุม

3.2 Test คือการทดสอบการทำงานของฮาร์ดดิสก์ไดร์ฟในสภาวะการทำงานต่างๆ

3.3 Featuring and packaging คือการติดฉลากสินค้าและการใส่ข้อมูลจำเพาะบางอย่าง รวมถึงการบรรจุหีบห่อผลิตภัณฑ์ เพื่อเตรียมส่งมอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซึ่งในกระบวนการประกอบและผลิตฮาร์ดดิสก์ไดร์ฟนั้นปรากฏดังรูปที่ 2.2



รูปที่ 2.2 กระบวนการผลิตฮาร์ดดิสก์ไดร์ฟ

2.1.2 กระบวนการทำงานของห้องคลีนรูม ในอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟ

ห้องคลีนรูม [3] หมายถึงห้องที่ได้รับการปรับสภาวะอากาศ อุณหภูมิ ความชื้น ความดัน และติดตั้งระบบกรองอากาศเป็นพิเศษให้อากาศปราศจากฝุ่นในระดับจุลภาคหรือจุลชีพซึ่งมีขนาดเล็กมากไม่สามารถมองเห็นด้วยตาเปล่าขนาดเส้นผ่านศูนย์กลางไม่เกิน 0.5 ไมครอน จุลภาคดังกล่าว เช่น ควันน้ำมัน ควันบุหรี่ และควันอื่นๆ ฝุ่นในบรรยากาศที่ไม่ตก แบคทีเรียและไวรัสบางชนิด สำหรับห้องสะอาดที่ใช้ในอุตสาหกรรม (Industrial clean rooms) นั้นเป็นห้องสะอาดที่ใช้ในอุตสาหกรรมอิเล็กทรอนิกส์ นาโนอิเล็กทรอนิกส์ ออปติคอลไฟเบอร์ การพิมพ์ การผลิตฟิล์มถ่ายรูป คอมพิวเตอร์ ซึ่งจะหลักการจะเน้นที่การกำจัดอนุภาคในอากาศและการควบคุมสภาวะต่างๆ ภายในห้องให้เหมาะสมกับลักษณะงานกระบวนการผลิตและผลิตภัณฑ์แต่ละชนิด เช่นถ้าหากความชื้นสัมพัทธ์สูงเกินไปอาจทำให้ชิ้นส่วนของผลิตภัณฑ์เกิดสนิมแต่ถ้าความชื้นสัมพัทธ์ต่ำเกินไป อาจทำให้เกิดไฟฟ้าสถิตในชิ้นส่วนนั้นได้ ความดันในห้องสะอาดควรเป็นบวกเสมอ เพื่อป้องกันการปนเปื้อนจากอนุภาคภายนอกระดับความชื้นความสะอาดของห้องจะต่างกันขึ้นอยู่กับประเภทของอุตสาหกรรม สำหรับมาตรฐานที่ใช้เกี่ยวกับห้องสะอาดนั้นได้แก่ มาตรฐานการปนเปื้อนของพื้นที่ควบคุม (Standard for Contamination Control Area) ซึ่งนิยมใช้กันในอุตสาหกรรมต่างๆ รวมทั้งอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟนั้นได้แก่ ISO 14644-1 clean room standards, Clean rooms and Associated controlled environments, Part1: Classification of air cleanliness ซึ่งเป็นมาตรฐานสากลได้กำหนดระดับความสะอาดสูงสุดคือ ISO1 ซึ่งจะมีอนุภาคที่มีขนาดตั้งแต่ 0.1 ไมครอนขึ้นไปจำนวน 10 อนุภาคต่อลูกบาศก์เมตร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ความสัมพันธ์ของมาตรฐานห้องสะอาดกับพื้นที่การทำงานนั้นในอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟนั้นมีทั้งหมด 3 ระดับ ดังตารางที่ 2.1

ตารางที่ 2.1 มาตรฐานการปนเปื้อนของพื้นที่ควบคุมในอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟ

Clean Room Class	Process Area	Cleanliness Specification
10000	HSA and Cleaning	10,000 particles/ft^3 at 0.5μm 30,000 particles/ft^3 at 0.3μm
100	HDE	100 particles/ft^3 at 0.5μm 300 particles/ft^3 at 0.3μm
10	HDE Assembly Station	10 particles/ft^3 at 0.5μm 30 particles/ft^3 at 0.3μm

2.1.3 การควบคุมคุณภาพการผลิต

จุดมุ่งหมายสำคัญของการบริหารงานผลิต ได้แก่ การผลิตสินค้าที่มีคุณภาพ มีปริมาณตรงตามที่กำหนดส่งมอบทันตามกำหนดเวลาด้วยต้นทุนที่ต่ำ คำว่า "คุณภาพ" กำหนดในรูปของมาตรฐาน (Standard) หรือข้อกำหนด (Specification) แสดงถึงความถูกต้อง ตรงตามความต้องการของผู้บริโภค [4] แต่แนวคิดการจัดการด้านคุณภาพได้ขยายความรวมถึงการผลิตสินค้าและการให้บริการที่เป็นเลิศ (Excellence in products and services) สร้างความพึงพอใจให้ลูกค้าอย่างสมบูรณ์ (Total customer satisfaction) และเหนือความคาดหวังของลูกค้า (Exceeding customer expectation) แนวทางหนึ่งที่สามารถสร้างความพึงพอใจให้ลูกค้า คือการควบคุมคุณภาพระหว่างกระบวนการผลิตที่มีประสิทธิภาพ ซึ่งจะส่งผลให้การผลิตเป็นไปตามเป้าหมาย

ในอุตสาหกรรมการผลิตนั้นแม้ว่าผู้ควบคุมการผลิตจะดำเนินการดีเพียงใด ความแปรปรวนหรือการเบี่ยงเบนในคุณภาพผลิตภัณฑ์สามารถเกิดขึ้นได้ตลอดเวลา ความแปรปรวนที่เกิดขึ้นในกระบวนการผลิตจำแนกได้ 2 ประเภท ได้แก่ [5] ความแปรปรวนตามธรรมชาติ (Common causes) เกิดขึ้นเมื่อระดับความแปรปรวนน้อยและไม่ส่งผลกระทบต่อการลดลงของคุณภาพสินค้าที่ผลิต หรือกล่าวได้ว่ากระบวนการผลิตอยู่ในการควบคุม (In-control process) และประเภทที่สองความแปรปรวนสาเหตุผิดธรรมชาติ (Special causes) เกิดขึ้นเมื่อระดับความแปรปรวนมาก มีผลให้การกระจายตัวของข้อมูลอยู่ภายนอกการควบคุมหรือแสดงการกระจายตัวของข้อมูลอย่างเป็นรูปแบบ ซึ่งส่งผลกระทบต่อการลดลงของคุณภาพสินค้าที่ผลิต สินค้าที่ผลิตไม่เป็นไปตามข้อกำหนด ไม่ได้มาตรฐาน หรือกล่าวได้ว่ากระบวนการผลิตยังไม่อยู่ในการควบคุม (Out-of control process) ซึ่งอาจเกิดได้จากหลายสาเหตุ ได้แก่ เครื่องจักรอุปกรณ์เสื่อมสภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คุณภาพวัตถุดิบไม่สม่ำเสมอ หรือเกิดจากพนักงานไม่ปฏิบัติตามข้อกำหนด ความเหน็ดเหนื่อย เมื่อยล้า เป็นต้น

2.2 การควบคุมคุณภาพเชิงสถิติ (Statistical Quality Control)

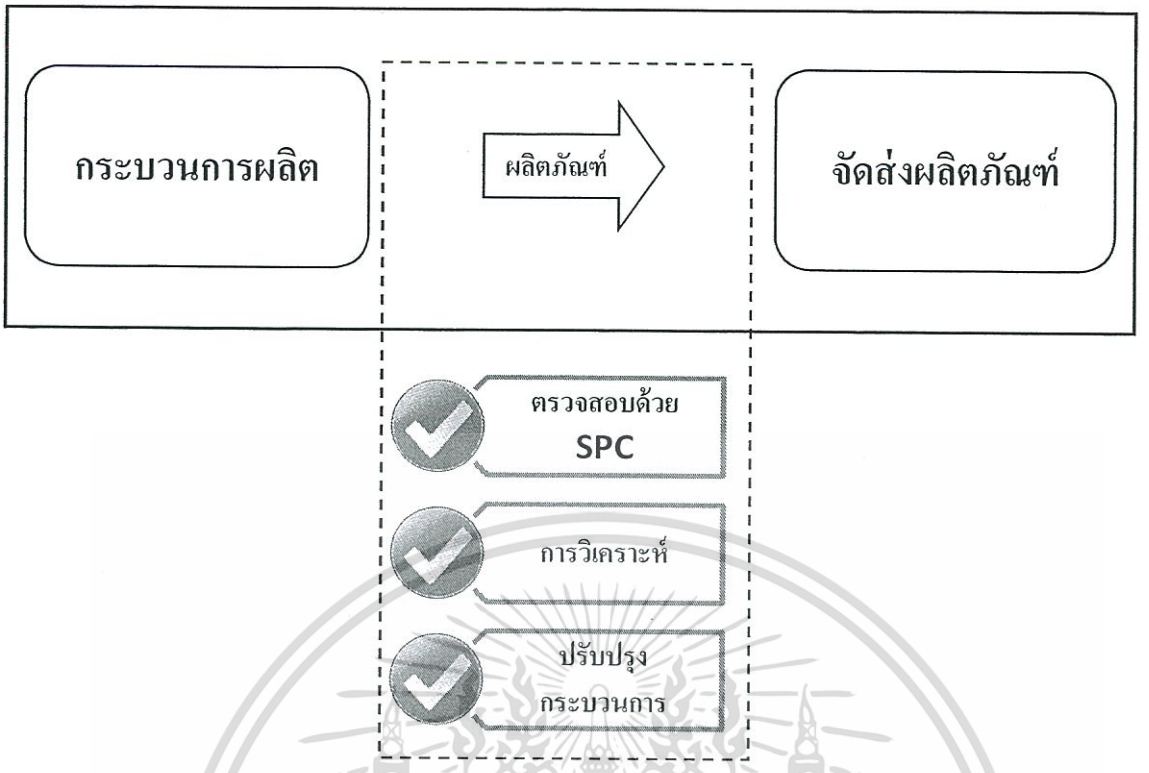
การควบคุมคุณภาพเชิงสถิติ (Statistical Quality Control; SQC) [6] เป็นเครื่องมือสำคัญในการรักษาระดับคุณภาพของสินค้าและบริการในกระบวนการผลิตให้ตรงตามมาตรฐานที่ผู้ผลิตและผู้บริโภคต้องการกล่าวคือ เป็นระดับมาตรฐานที่ทำให้ผู้บริโภคได้รับความพึงพอใจจากสินค้าและบริการสูงสุดและทำให้ได้รับผลกำไรสูงสุดในระยะยาว ซึ่งจะส่งผลให้กิจการสามารถดำรงอยู่ต่อไปได้ โดยการควบคุมคุณภาพดังกล่าวจะอาศัยวิธีการทางสถิติมาใช้ในการคำนวณและนำผลลัพธ์ที่ได้มาใช้ประกอบการตัดสินใจในเรื่องเกี่ยวกับคุณภาพของสินค้าในด้านต่างๆ เช่น การพัฒนาสินค้าให้ตรงตามมาตรฐานของผู้ผลิตเองตลอดจนพัฒนามาตรฐานสินค้า

เทคนิคหนึ่งที่ใช้สำหรับการกระบวนการเฝ้าระวังการผลิต [7] (Monitor manufacturing process) และให้ข้อเสนอแนะ ได้แก่ การควบคุมกระบวนการเชิงสถิติ (Statitical Process Control) ข้อเสนอแนะที่ได้จะใช้สำหรับบำรุงรักษาและปรับปรุงความสามารถของกระบวนการ

2.2.1 การควบคุมกระบวนการเชิงสถิติ (Statistical Process Control)

การควบคุมกระบวนการเชิงสถิติ (Statistical Process Control: SPC) เป็นเทคนิคของการควบคุมภาพที่สำคัญชนิดหนึ่งซึ่งใช้สำหรับควบคุมกระบวนการ โดยจะส่งสัญญาณเตือนเมื่อมีการปรับเปลี่ยนที่สำคัญ โดยใช้สัญญาณทางสถิติในการกำหนดแหล่งของการผันแปรเพื่อการปรับปรุงและพัฒนากระบวนการอย่างเป็นระบบ นำไปสู่เป้าหมายที่สำคัญได้แก่ เพิ่มผลการดำเนินงานให้สูงขึ้นและรักษาคุณภาพการผลิตให้อยู่ในระดับสูง ในปัจจุบันได้มีการนำไปใช้อย่างแพร่หลายในองค์กรต่างๆ เพื่อช่วยให้การบริหารงานผลิตเป็นไปตามเป้าหมาย ซึ่งกระบวนการควบคุมเชิงสถิติเน้นการป้องกัน (Prevention) มากกว่าการตรวจจับ (Detection) ดังแสดงในรูปที่ 2.3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.3 โครงสร้างแนวคิดการป้องกันในระหว่างกระบวนการผลิต

ดังนั้นการดำเนินการควบคุมกระบวนการเชิงสถิติ (SPC) ที่เหมาะสมต่อสัญญาณต่างๆ ที่ต่อเนื่องจะนำไปสู่การป้องกันการเกิดของเสีย (Waste) ลดความผันแปรที่เกิดระหว่างการผลิต เพิ่มการผลิตที่มีประสิทธิภาพ ลดความผันแปรหรือเพิ่มความสม่ำเสมอ (Consistency) ลดเศษทิ้ง (Scrap) และการแก้ไขซ้ำหรือทำซ้ำใหม่ (Rework) นำมาซึ่งคุณภาพของผลิตภัณฑ์ที่ตรงตามความต้องการ

โดยนิยมใช้แผนภูมิควบคุม (Control Chart) เพื่อติดตามการเปลี่ยนแปลงของกระบวนการผลิตว่าอยู่ในระดับปกติหรือไม่และเครื่องมือทางคุณภาพอื่นๆ นั้นก็สามารถนำมาใช้ในการรวบรวมข้อมูล นำเสนอข้อมูลและแสดงสาเหตุของปัญหาเป็นต้น ซึ่งเครื่องมือพื้นฐานสำหรับการควบคุมกระบวนการทางสถิติมีดังนี้ [8]

1. ใบตรวจสอบ (Check sheet)
2. แผนภูมิพาเรโต (Pareto chart)
3. ฮิสโตแกรม (Histogram)
4. แผนภาพการกระจาย (Scatter diagram)
5. แผนภูมิการไหลของกระบวนการ (Process flow chart)
6. แผนภาพสาเหตุและผล (Cause and effect diagram or Fish bone diagram)
7. แผนภูมิควบคุม (Control chart)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.2 แผนภูมิควบคุม (Control chart)

แผนภูมิควบคุม (Control chart) [16] เป็นเครื่องมือสำหรับการควบคุมที่ใช้สำหรับตรวจสอบค่าของตัวแปรที่ต้องการควบคุมคุณภาพว่าเกิดความแปรผันเกินจากขอบเขตที่กำหนดไว้หรือไม่ ซึ่งหากพบว่าเกินจากขอบเขตที่กำหนดไว้ ผู้วิเคราะห์จะต้องหาสาเหตุของความแปรผันและดำเนินการแก้ไขก่อนที่จะเกิดความเสียหายขึ้นกับผลิตภัณฑ์ โดยตรวจวัดค่าข้อมูลของผลิตภัณฑ์ใน 2 ลักษณะคือ

1. ค่าที่ได้จากการ ชั่ง ตวง วัด ที่มีค่าเป็นตัวเลข ซึ่งเรียกว่าตัวแปรปริมาณ (Quantity Variable) เช่น ความยาว ส่วนสูง ความหนา ความกว้างและปริมาณสารต่างๆ เป็นต้น

2. ค่าที่ได้จากการวัดในเชิงคุณลักษณะจะเป็นการพิจาณาภายนอก เช่น มีรอยตำหนิ – ไม่มีรอยตำหนิ สวย – ไม่สวย หรือการประเมิณผลลัพธ์สุดท้ายของผลิตภัณฑ์แต่ละชิ้นจากการวัดเชิงปริมาณแล้วสรุปผลว่าเป็นผลิตภัณฑ์ดีหรือไม่ดี ผลิตภัณฑ์เสียหรือไม่เสีย เป็นต้น แล้วนับจำนวนผลิตภัณฑ์ที่มีควมน่าสนใจ เช่น สนใจผลิตภัณฑ์ที่เสียก็จะนับค่าที่เป็นหน่วยนับ เรียกว่า ตัวแปรคุณลักษณะหรือตัวแปรคุณภาพ

แผนภูมิควบคุมคุณภาพ (Control chart) พัฒนาขึ้นโดย [17] Walter A. Shewhart ในปี ค.ศ.1924 จัดเป็น 1 ใน 7 เครื่องมือควบคุมกระบวนการเชิงสถิติ (SPC) ที่มีความสำคัญในการตรวจสอบการเปลี่ยนแปลงไปของกระบวนการผลิต หรือแผนภูมิควบคุมคือกราฟที่มีกลไกในการแยกความผันแปรจากสาเหตุผิดธรรมชาติ (Special causes) ออกจากสาเหตุธรรมชาติ (Common causes) ซึ่งลักษณะเฉพาะที่สำคัญของความผันแปรที่มีสาเหตุผิดธรรมชาติจะมีลักษณะที่ไม่สามารถคาดการณ์ตัวแบบได้เนื่องจากไม่มีความเสถียร อาจสรุปได้ว่าสรุปว่าแผนภูมิควบคุมอาศัยหลักการทางสถิติร่วมกับรูปแบบการกระจายตัวของข้อมูลเพื่อตรวจสอบว่ากระบวนการผลิตเป็นไปตามปกติหรืออยู่ในการควบคุมหรือไม่ เป็นเครื่องมือเตือนให้ผู้ควบคุมกระบวนการผลิตทราบถึงการเปลี่ยนแปลงไปหรือมีความผันแปรของกระบวนการผลิตในระดับภายนอกกการควบคุมที่กำหนดไว้

2.2.2.1 หลักการสร้างแผนภูมิควบคุมคุณภาพของชิวฮาร์ท

หลักการสร้างแผนภูมิควบคุมคุณภาพของชิววาร์ทจะสร้างบนพื้นฐานข้อมูลที่มีการแจกแจงปกติ (Normal Distribution) โดยรูปแบบของการแจกแจงแบบปกติขึ้นอยู่กับค่ากลางและลักษณะการกระจายของข้อมูล ในการประยุกต์ใช้การแจกแจงปกติกับการควบคุมคุณภาพนี้จะนำไปใช้กับการหาขอบเขตของการยอมรับผลิตภัณฑ์ภายใต้การกำหนดระดับความเชื่อมั่นทางสถิติ เมื่อทราบการแจกแจงข้อมูล

ให้ X คือ ค่าสังเกตหรือค่าที่ทำการศึกษาจากผลิตภัณฑ์

μ คือ ค่าเฉลี่ยของประชากร

σ คือ ค่าเบี่ยงเบนมาตรฐานของประชากร

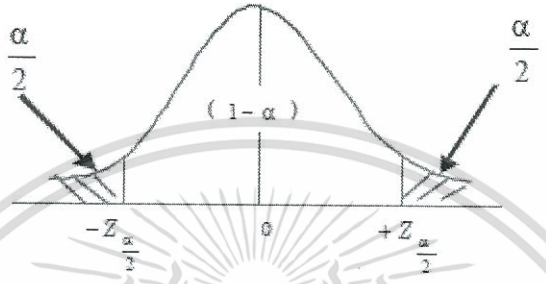
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถ้า X มีการแจกแจงแบบปกติ (Normal Distribution) จะได้ว่า

$$Z = \frac{X-\mu}{\sigma}$$

โดย Z เป็นค่ามาตรฐานและมีการแจกแจงแบบปกติ $(Z \sim N(0,1))$

ดังนั้น ที่ระดับความเชื่อมั่น $(1-\alpha)100\%$ จะได้ค่า Z เป็น $-Z_{\frac{\alpha}{2}}$ และ $+Z_{\frac{\alpha}{2}}$ ดังรูปที่ 2.4



รูปที่ 2.4 แสดงค่า Z และพื้นที่ใต้โค้งปกติ

โดยทั่วไปขอบเขตการยอมรับ (Specification limit) ที่มีความสัมพันธ์กับระดับความเชื่อมั่นนั้นมีทั้งหมด 4 ระดับ ได้แก่ เขตการยอมรับที่$\pm 1\sigma$,$\pm 2\sigma$, $\pm 3\sigma$ และ$\pm 4\sigma$ [12] ดังนั้นขอบเขตการยอมรับที่ใช้สำหรับขั้นตอนการควบคุมคุณภาพของงานวิจัยนี้ สามารถแสดงได้ดังตารางที่ 2.2

ตารางที่ 2.2 ระดับความเชื่อมั่นของการยอมรับระดับต่างๆ

ขอบเขตการยอมรับ (Spec. Limit)	ระดับความเชื่อมั่น (%)	ระดับความคลาดเคลื่อน (%)
$\pm 1\sigma$	68.27	31.73
$\pm 2\sigma$	95.45	4.55
$\pm 3\sigma$	99.73	0.270

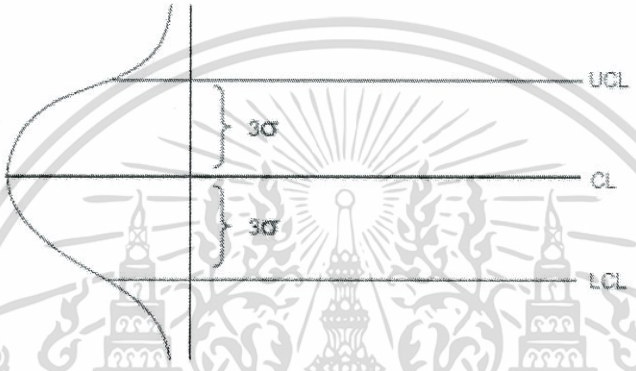
โดยทั่วไปนิยมใช้ระดับความเชื่อมั่น 99.73% เป็นขอบเขตของการควบคุมคุณภาพจริง หรือที่ค่า Z เท่ากับ ± 3 จะได้พื้นที่ใต้โค้งปกติจาก Z = -3 ถึง Z = +3 เท่ากับ 0.99 และใช้ระดับความเชื่อมั่น 95.45% เป็นขอบเขตของการควบคุมเตือน (Warning control) จะได้พื้นที่ใต้โค้งปกติ Z = -2 ถึง Z = +2 เท่ากับ 0.9545 สำหรับการสร้างแผนภูมิควบคุมจะประกอบด้วยองค์ประกอบอย่างน้อย 3 ส่วนได้แก่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. เส้นศูนย์กลาง หรือ μ_x (Central line; CL) เป็นค่าเฉลี่ยของกระบวนการผลิต
2. เส้นขอบเขตควบคุมบน หรือ $\mu_x + 3\sigma_x$ (Upper control limit; UCL) เป็นเส้นที่มีระยะห่างจากเส้นควบคุมเฉลี่ยเท่ากับ 3σ ทางค่ามาก
3. เส้นขอบเขตควบคุมล่าง หรือ $\mu_x - 3\sigma_x$ (Lower control limit; LCL) เป็นเส้นที่มีระยะห่างจากเส้นควบคุมเฉลี่ยเท่ากับ 3σ ทางค่าน้อย

เมื่อ μ_x คือ ค่าเฉลี่ยของข้อมูล ณ เวลาที่เก็บข้อมูลแต่ละครั้ง

σ_x คือ ส่วนเบี่ยงเบนเฉลี่ยของข้อมูล



รูปที่ 2.5 องค์ประกอบของแผนภูมิควบคุม

โดยทั่วไปช่วงขีดจำกัดของการควบคุมในการผลิตอยู่ที่ระดับ $\pm 3\sigma$ หรืออยู่ในขอบเขตจำกัดที่ $\mu_x \pm 3\sigma_x$ การกระจายตัวของข้อมูลในแผนภูมิควบคุมในลักษณะของคุณภาพอยู่ในแกนตั้งและให้ความถี่ เวลา หรือจำนวนครั้งที่จะตรวจสอบอยู่ในแกนนอน

2.2.2.2 ขั้นตอนในการสร้างแผนภูมิควบคุมคุณภาพของชิววาร์ท

1. เก็บรวบรวมข้อมูลที่จะทำการควบคุมหรือที่มีผลต่อคุณภาพของกระบวนการผลิต

2. เลือกแผนภูมิที่เหมาะสมกับข้อมูล โดยต้องศึกษาคุณสมบัติของแผนภูมิควบคุมคุณภาพแต่ละชนิดว่าใช้อย่างไร ใช้กับข้อมูลลักษณะใด คุณสมบัติที่กล่าวไว้ตรงกับข้อมูลที่มีอยู่หรือไม่ เช่น

i) ถ้าข้อมูลเป็นค่าวัดตัวแปรเชิงปริมาณเก็บข้อมูลในลักษณะเชิงเดี่ยวก็จะใช้แผนภูมิเชิงเดี่ยว (X-Chart)

ii) ถ้าเป็นข้อมูลที่วัดเชิงปริมาณเก็บข้อมูลในลักษณะเป็นล็อต (lots) ควรจะเลือกใช้แผนภูมิควบคุมคุณภาพที่ควบคุมค่ากลางและการกระจายควบคู่กันเสมอ เช่น ใช้แผนภูมิควบคุมเฉลี่ยคู่กับแผนภูมิควบคุมพิสัย หรือใช้แผนภูมิควบคุมเฉลี่ยคู่กับแผนภูมิควบคุมการกระจาย เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

iii) ถ้าข้อมูลที่เป็นหน่วยนับเก็บข้อมูลในลักษณะที่เป็นล็อตและข้อมูลเป็นข้อมูลคุณภาพสนใจจำนวนผลิตภัณฑ์ที่เสีย (Nonconforming or defectives) แล้วนำมาคิดเป็นสัดส่วนเสีย (Fraction nonconforming) ก็อาจจะใช้แผนภูมิควบคุม p หรือแผนภูมิควบคุม np-Chart

iv) ถ้าข้อมูลที่นับจำนวนรอยตำหนิหรือจำนวนเสีย (Count of nonconformities) ที่สนใจก็อาจใช้แผนภูมิควบคุม C-Chart หรือตำนวนรอยตำหนิหรือจำนวนเสียต่อหน่วย (Defect per unit or noconformities per unit) ก็จะใช้แผนภูมิควบคุม U-Chart เป็นต้น

3. คำนวณหาขอบเขตการควบคุมบนและล่างจากข้อมูลที่ได้

4. นำข้อมูลพล็อตลงบนแผนภูมิควบคุมคุณภาพแล้วพิจารณาว่ามีข้อมูลใดตกอยู่นอกขอบเขตการควบคุมหรือไม่ อธิบายและแปลความหมายของแผนภูมิที่ได้

5. วิเคราะห์หาสาเหตุและปรับปรุงกระบวนการผลิตหรือหาแหล่งที่ทำให้เกิดการแปรผันเมื่อมีข้อมูลตกอยู่นอกเขตการควบคุม

เนื่องจากการพิจารณาข้อมูลปริมาณอนุภาคฝุ่นละอองที่ทำการวิเคราะห์ด้านการควบคุมคุณภาพในวิทยานิพนธ์ฉบับนี้พบว่าข้อมูลที่ทำการรวบรวมมานั้นเป็นข้อมูลการวัดของของตัวแปรเชิงปริมาณ (Quantity Variable) ซึ่งเป็นข้อมูลค่าวัดแบบเชิงเดี่ยว ข้อมูลมีลักษณะการเก็บข้อมูลแบบต่อเนื่อง ดังนั้นแผนภูมิควบคุมที่ใช้ในงานวิจัยนี้ได้แก่ แผนภูมิควบคุมค่าวัดเชิงเดี่ยว (Individual control chart)

2.2.3 แผนภูมิควบคุมค่าวัดเชิงเดี่ยว (Individual control chart: X-chart)

แผนภูมิควบคุมค่าวัดเชิงเดี่ยว ใช้สำหรับข้อมูลที่เก็บรวบรวมค่าวัดแบบเชิงเดี่ยว เช่น ขนาดความยาว, น้ำหนัก, ความหนาแน่น เป็นต้น มีลักษณะการเก็บข้อมูลแบบต่อเนื่องหรืออาจเก็บข้อมูลในช่วงเวลาที่ห่างกันคงที่ในลักษณะข้อมูลโดดๆ นำข้อมูลที่ได้มาพิจารณาการแจกแจง ถ้าข้อมูล X มีการแจกแจงแบบปกติ ก็จะสร้างขอบเขตควบคุมคุณภาพตามหลักความเชื่อมั่นของการแจกแจงปกติ ค่าวัดแต่ละค่าในกระบวนการผลิตอยู่ภายใต้ขอบเขตการควบคุมคุณภาพที่กำหนดคือ ที่ระดับความเชื่อมั่น 99.73% จะถือว่ากระบวนการผลิตอยู่ภายใต้กระบวนการควบคุมทางสถิติและสามารถหาขอบเขตการควบคุมบนและขอบเขตการควบคุมล่างได้ดังนี้

$$\begin{aligned} UCL_X &= \mu + 3\sigma_x' \\ CL_X &= \mu \\ LCL_X &= \mu - 3\sigma_x' \end{aligned} \qquad (2.1)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ $\mu = \frac{\sum_{i=1}^{N} X_i}{N}$ เมื่อ μ คือค่าเฉลี่ยของข้อมูลประชากร

และ σ'_X คือส่วนเบี่ยงเบนมาตรฐานของข้อมูลประชากร $\sigma'_X = \sqrt{\frac{\sum_{i=1}^{N}(X_i-\mu)^2}{N}}$

จะประมาณค่า μ ด้วยค่าเฉลี่ยของข้อมูลตัวอย่าง $(\overline{X})$ และประมาณ σ'_X ด้วยค่าส่วนเบี่ยงเบนมาตรฐานของข้อมูลตัวอย่าง (S_X) เมื่อ $\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ และ $S_X = \sqrt{\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{n-1}}$

N คือ จำนวนข้อมูลที่เป็นประชากร (จำนวนผลิตภัณฑ์ทั้งหมด)

n คือ จำนวนข้อมูลที่สุ่มเป็นตัวอย่าง (จำนวนผลิตภัณฑ์ที่สุ่ม)

ดังนั้นสามารถสร้าง X-chart ได้ดังสมการ 2.2

$$\begin{aligned} UCL_X &= \overline{X} + 3S_X \\ CL_X &= \overline{X} \\ LCL_X &= \overline{X} - 3S_X \end{aligned} \qquad (2.2)$$

2.2.4 แผนภูมิค่าพิสัยเคลื่อนที่ (Moving range chart)

การกระจายของข้อมูลที่จะนำมาควบคุมนั้นจะไม่สามารถหาได้โดยตรงเพราะข้อมูลมีเพียงกลุ่มเดียวค่ากระจายก็จะมีค่าเดียว ณ เวลานั้นๆ จึงอาศัยค่าพิสัยเคลื่อนที่ (Moving Range) ในการควบคุมการกระจายได้ ดังตารางที่ 2.3

กำหนดให้ MR_i คือ ค่าพิสัยเคลื่อนที่

เมื่อ $MR_i = |X_i - X_{i-1}|$

และ $\overline{MR}$ คือ ค่าพิสัยเคลื่อนที่เฉลี่ย

จากสมการที่ 2.1 จะประมาณค่า μ ด้วย $\overline{X}$ และ σ'_X ด้วย $3\left(\frac{\overline{MR}}{d_2}\right)$ ได้ดังสมการที่ 2.3

สูตร X-chart ในรูปของค่าพิสัยเคลื่อนที่

$$\begin{aligned} UCL_X &= \overline{X} + 3\left(\frac{\overline{MR}}{d_2}\right) \\ CL_X &= \overline{X} \\ LCL_X &= \overline{X} - 3\left(\frac{\overline{MR}}{d_2}\right) \end{aligned} \qquad (2.3)$$

เมื่อ $d_2 = 1.128$ เป็นค่าคงที่จากตารางภาคผนวก Factor for control chart

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.3 วิธีคำนวณค่าพิสัยเคลื่อนที่ เมื่อ m = 2 และ m = 3

ค่าข้อมูล (X_i)	พิสัยเคลื่อนที่ (MR_i) (m = 2)	พิสัยเคลื่อนที่ (MR_i) (m = 3)
X_1	-	-
X_2	$X_2 - X_1$	-
X_3	$X_3 - X_2$	$X_3 - X_1$
X_4	$X_4 - X_3$	$X_4 - X_2$
.	.	$X_5 - X_3$
.	.	.
.	.	.
X_{N-2}	.	.
X_{N-1}	$X_{N-1} - X_{N-2}$	$X_{N-1} - X_{N-3}$
X_N	$X_N - X_{N-1}$	$X_N - X_{N-2}$
	$\overline{MR} = \sum MR_i$	$\overline{MR} = \sum MR_i$

ในการหาค่ากระจายซึ่งจะถือว่าค่านี้เทียบได้กับค่าพิสัยในแต่ละกลุ่มย่อย ดังนั้นจึงใช้สูตรของการหาขอบเขตควบคุมบนและขอบเขตควบคุมล่างด้วยสูตรเดียวกันกับสูตรของพิสัย และจากสูตร R-Chart ให้ $\overline{R} = \overline{MR}$ จะได้

$$\begin{aligned} UCL_{MR} &= D_4\overline{MR} \\ CL_{MR} &= \overline{MR} \\ LCL_{MR} &= D_3\overline{MR} \end{aligned} \qquad (2.4)$$

เมื่อ D_3, D_4 เป็นค่าคงที่จากตารางภาคผนวก Factor for control chart

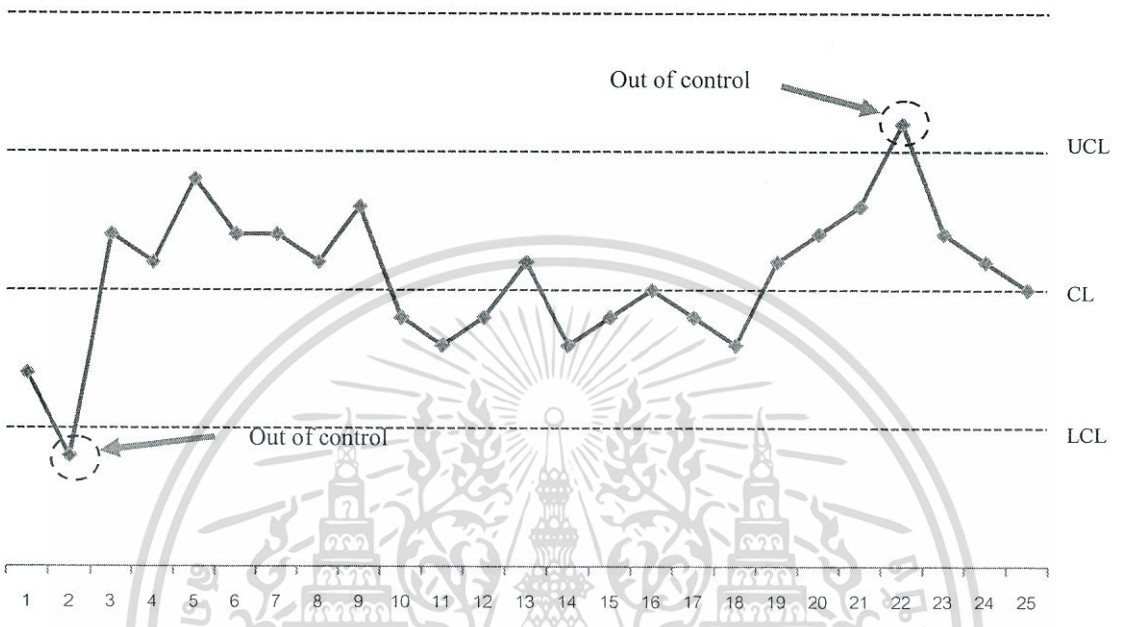
2.2.5 การตีความแผนภูมิควบคุมคุณภาพ

เมื่อนำข้อมูลมาสร้างแผนภูมิควบคุมคุณภาพแล้ว สามารถอ่านแผนภูมิแล้วแปลความหมายของแผนภูมิควบคุมเพื่อหาเหตุผลไปที่สภาวะของกระบวนการผลิตได้ เพราะอาการผิดปกติต่างๆ จะแสดงออกมาตามค่าข้อมูลทางสถิติดังกล่าว เมื่อพบความผิดปกติของกระบวนการผลิตแล้วจึงต้องหาแนวทางปรับปรุงกระบวนการเพื่อให้กระบวนการผลิตนั้นอยู่ภายใต้การควบคุม สำหรับรูปแบบของความผิดปกติของข้อมูลในลักษณะต่างๆ มีดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.5.1 เกิดจุดอยู่นอกขอบเขตการควบคุม

เมื่อมีจุดตกอยู่นอกขอบเขตการควบคุมบนหรือขอบเขตการควบคุมล่าง เรียกว่า จุดอยู่นอกการควบคุม (Out of control) แสดงว่ากระบวนการผลิตมีความผิดปกติผิดปกติ ดังรูปที่ 2.6



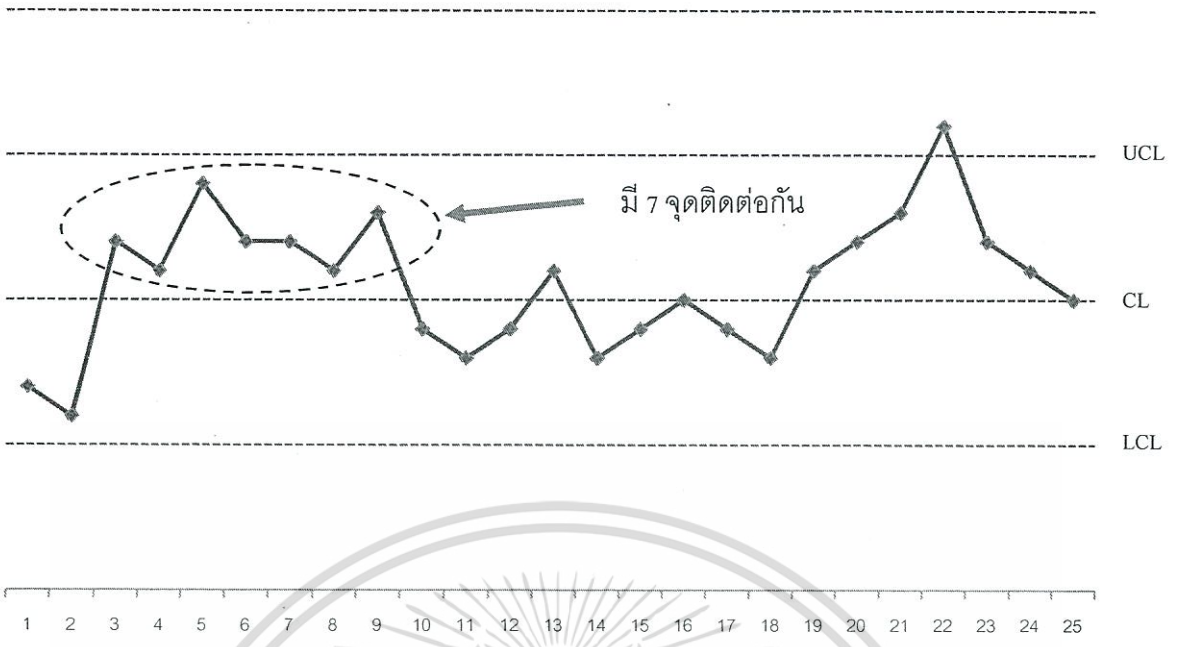
รูปที่ 2.6 การเกิดจุดอยู่นอกขอบเขตการควบคุม

2.2.5.2 เกิดการไม่สมดุล (Runs)

เมื่อมีจุดปรากฏติดต่อกันบนด้านใดด้านหนึ่งของเส้นกลาง เช่นมี 7 จุดติดต่อกันบนด้านใดด้านหนึ่งของเส้นกลาง แสดงว่ากระบวนการผลิตเกิดความผิดปกติขึ้นแล้ว ดังรูปที่ 2.7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

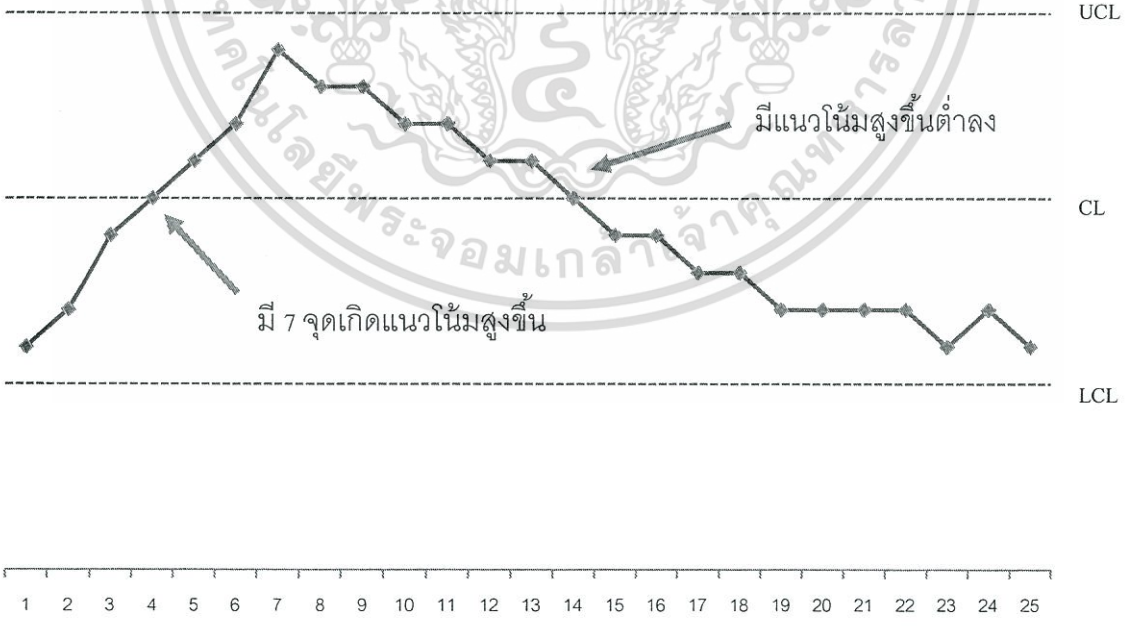
ห้องสมุดคณะเทคโนโลยีสารสนเทศ พระจอมเกล้าลาดกระบัง



รูปที่ 2.7 เกิดการไม่สมดุล (Runs)

2.2.5.3 เกิดแนวโน้ม (Trends)

เมื่อมีจุดที่ต่อเนื่องกันโดยมีแนวโน้มไปในทางใดทางหนึ่ง แสดงว่ากระบวนการผลิตกำลังผิดปกติ ดังรูปที่ 2.8

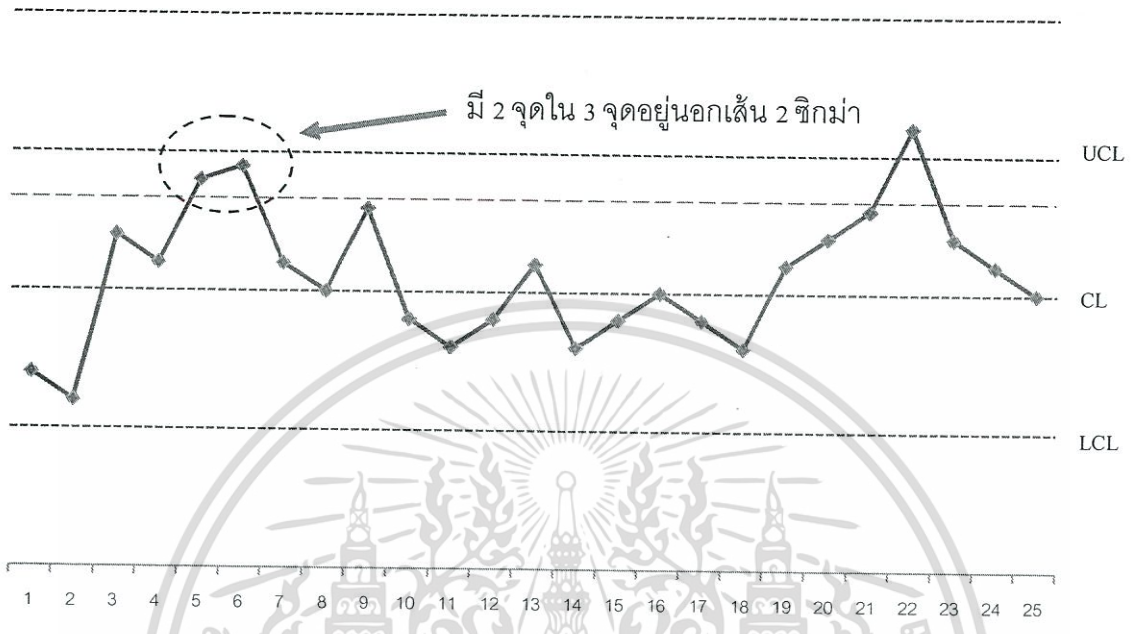


รูปที่ 2.8 การเกิดแนวโน้ม (Trends)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้
7167

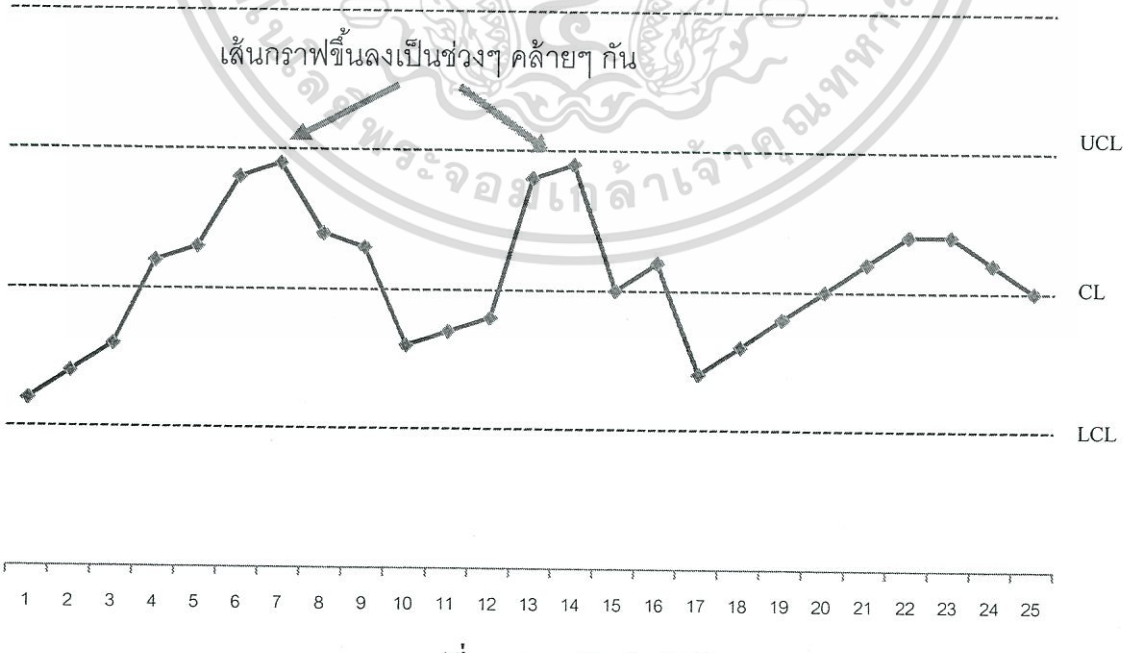
2.2.5.4 เกิดจุดอยู่นอกเขตควบคุมเตือน

การสร้างขอบเขตควบคุมเตือน ($\pm 2\sigma$) ถ้ามีจุด 2 จุดใน 3 จุด อยู่นอกเส้น $\pm 2\sigma$ หรืออยู่ใกล้เคียง $\pm 3\sigma$ แสดงว่ากระบวนการเริ่มมีความผิดปกติ ดังรูปที่ 2.9



รูปที่ 2.9 การเกิดจุดอยู่นอกเขตควบคุมเตือน

2.2.5.5 เกิดวัฏจักร (Cycle)



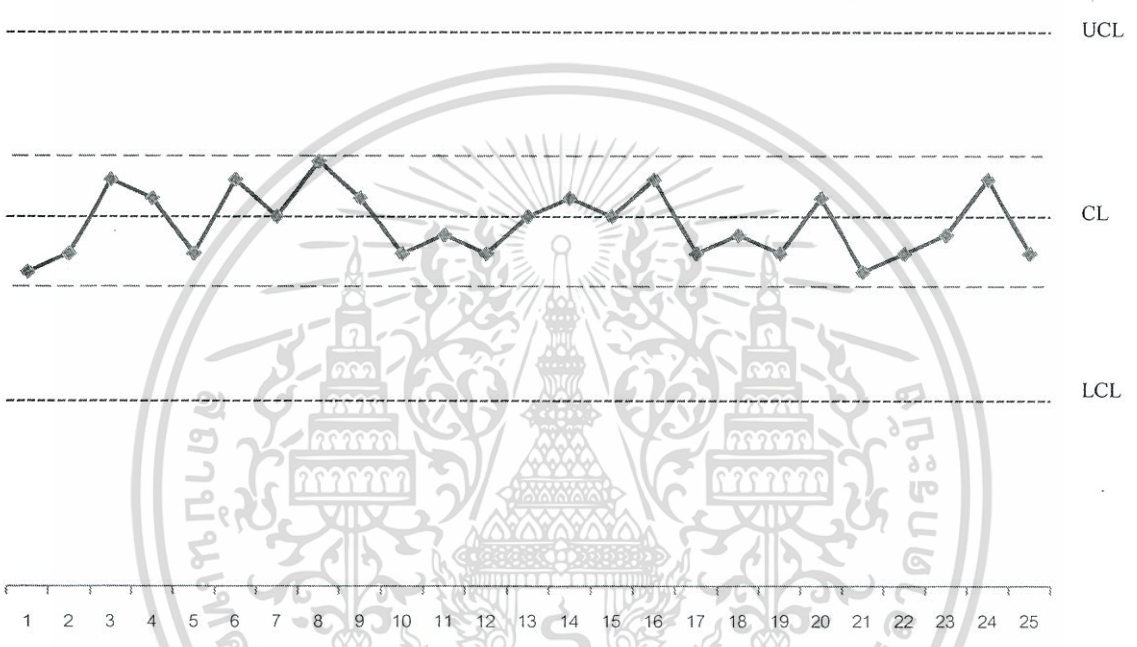
รูปที่ 2.10 การเกิดเกิดวัฏจักร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 2.10 การเปลี่ยนแปลงแบบขึ้นๆ ลงๆ ของกราฟเป็นรอบซึ่งสามารถทำนายการขึ้นลงในรอบต่อไปได้เป็นลักษณะที่ทำให้ทราบว่ากระบวนการผลิตเกิดความผิดปกติ

2.2.5.6 เกิดจุดใกล้เส้นกลาง

หากค่าทุกค่าของข้อมูลอยู่ภายใต้ $\pm 1.5\sigma$ อาจเกิดลักษณะที่เรียกว่า กระบวนการดีเยี่ยม แต่ในทางตรงข้ามกระบวนการอาจเกิดความผิดปกติจากการกำหนดขนาดของกลุ่มย่อย ข้อมูลมีการปะปนกันหรืออาจเกิดการปรุงแต่งข้อมูลมากเกินไป

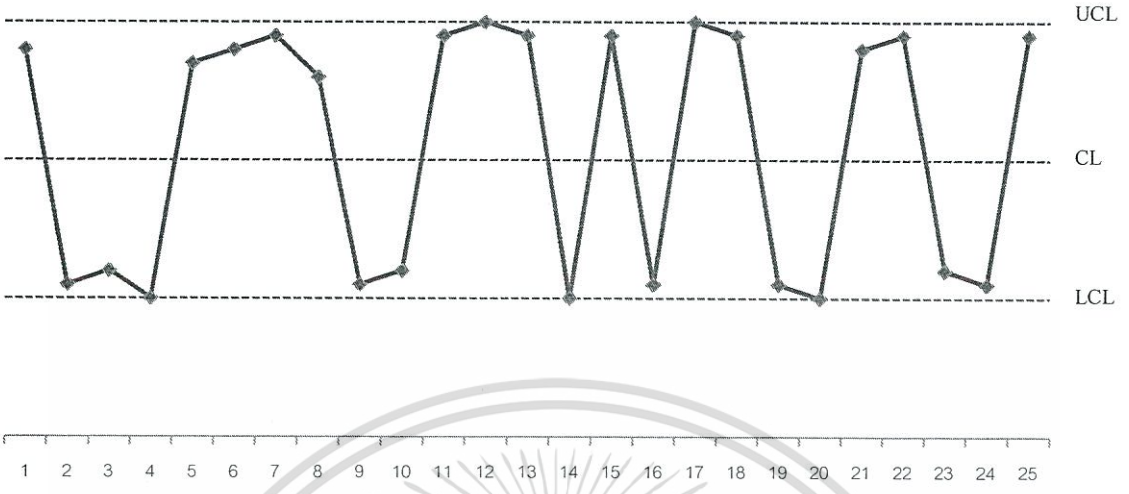


รูปที่ 2.11 การเกิดจุดใกล้เส้นกลาง

2.2.5.7 เกิดการผสม

จากรูปที่ 2.12 จะมีจุดใกล้เส้นควบคุมบนและใกล้เส้นควบคุมล่างเกือบทุกๆ จุด แสดงว่าเกิดความผิดปกติของกระบวนการผลิต ซึ่งสาเหตุอาจเกิดจากวัตถุดิบที่มีแหล่งคุณภาพต่างกัน หรือข้อมูลมาจาก 2 กระบวนการที่มีสภาวะแวดล้อมที่ต่างกันเป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.12 การเกิดการผสม

2.2.6 มาตรฐานการควบคุมคุณภาพ

ปัจจุบันอุตสาหกรรมต่างๆ จะมีมาตรฐานการควบคุมการผลิต โดยควบคุมคุณภาพในกรณีที่เกิดการรูปแบบต่างๆ ที่กล่าวมาข้างต้น โดยมาตรฐานที่กำหนด 2 แบบ ซึ่งในงานวิจัยนี้ จะทดสอบโดยการวิเคราะห์รูปแบบที่เหมาะสมทั้ง 2 แบบ ได้แก่

2.2.6.1 มาตรฐานแบบ 8 ข้อ [14]

2.2.6.1.1 มี 1 จุดขึ้นไปที่ตกอยู่นอกขอบเขตการควบคุม

2.2.6.1.2 มี 9 จุดติดต่อกันที่อยู่ข้างเดียวกันอาจจะเหนือขอบเขตควบคุมหรือใต้ขอบเขตควบคุมก็ได้

2.2.6.1.3 มี 6 จุดติดต่อกันที่มีแนวโน้มเพิ่มขึ้นหรือลดลง

2.2.6.1.4 มี 14 จุดต่อเนื่องกันที่มีแนวโน้มขึ้นหรือลง

2.2.6.1.5 มี 2 ใน 3 จุดที่ตกอยู่นอกเขตควบคุมเตือน ($\pm 2\sigma$) จากเส้นกลางของข้างเดียวกัน

2.2.6.1.6 มี 4 ใน 5 จุดที่ตกอยู่นอกเขตควบคุม $\pm 1\sigma$ จากเส้นกลางของข้างเดียวกัน

2.2.6.1.7 มี 15 จุดติดต่อกันที่อยู่เหนือหรือใต้เส้นกลางของขอบเขตควบคุม $\pm 1\sigma$ ข้างใดข้างหนึ่ง

2.2.6.1.8 มี 8 จุดติดต่อกันที่อยู่นอกเขตควบคุม $\pm 1\sigma$ จากเส้นกลางของข้างใดข้างหนึ่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการศึกษาวิจัยในส่วนของการควบคุมคุณภาพเชิงสถิติ (SPC) นั้น งานวิจัยหลายงานวิจัย [4][5] ได้ทำการควบคุมคุณภาพของปริมาณอนุภาคฝุ่นละอองด้วยแผนภูมิควบคุม (Control Chart) และใช้มาตรฐานสำหรับการควบคุมคุณภาพแบบ เกิดจุดอยู่นอกขอบเขตการควบคุม ที่ระดับความเชื่อมั่น 99.73 หรือที่ขอบเขตการยอมรับ $\pm 3\sigma$ เท่านั้น ดังนั้นการพิจารณาขอบเขตการยอมรับที่นำเสนอเพิ่มเข้ามาคือที่ระดับ $\pm 1\sigma$ และ $\pm 2\sigma$ และใช้มาตรฐานการควบคุมคุณภาพแบบ 8 ข้อ

โดยรายละเอียดของการควบคุมคุณภาพแต่ละพื้นที่การทำงาน ได้แก่ มาตรฐานการควบคุมคุณภาพแบบ 8 ข้อที่ใช้ในวิทยานิพนธ์จะกล่าวรายละเอียดและตัวอย่างในบทที่ 3

2.3 เทคนิคเหมืองข้อมูลชนิดการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจ

การจำแนกข้อมูล (Data Classification) [9] เป็นเทคนิคหนึ่งที่สำคัญในการสืบค้นความรู้ของฐานข้อมูลขนาดใหญ่ (Knowledge Discovery from very large Database: KDD) หรือเหมืองข้อมูล (Data Mining) เทคนิคการจำแนกข้อมูลเป็นกระบวนการสร้างโมเดลจัดการข้อมูลให้อยู่ในกลุ่มที่กำหนดมาจากกลุ่มตัวอย่างข้อมูลที่เรียกว่าข้อมูลสอนระบบ (training data) ที่แต่ละแถวของข้อมูลประกอบด้วยฟิลด์หรือแอททริบิวต์จำนวนมาก แอททริบิวต์นี้อาจเป็นค่าต่อเนื่อง (Continueous) หรือค่ากลุ่ม (Categorical) โดยจะมีแอททริบิวต์แบ่ง (Classification attribute) ซึ่งเป็นตัวบ่งชี้คลาสของข้อมูล จุดประสงค์ของการจำแนกข้อมูลคือการสร้างโมเดลการแยกแอททริบิวต์หนึ่งโดยขึ้นกับแอททริบิวต์อื่น โมเดลที่ได้จากการจำแนกข้อมูลจะทำให้สามารถพิจารณาคลาสในข้อมูลที่ยังไม่ได้แบ่งกลุ่มในอนาคตได้ เทคนิคการจำแนกข้อมูลนี้ได้นำไปประยุกต์ใช้ในหลายด้าน เช่น การจัดกลุ่มทางการตลาด, การตรวจสอบการผิดปกติ และการวิเคราะห์ทางการแพทย์ เป็นต้น

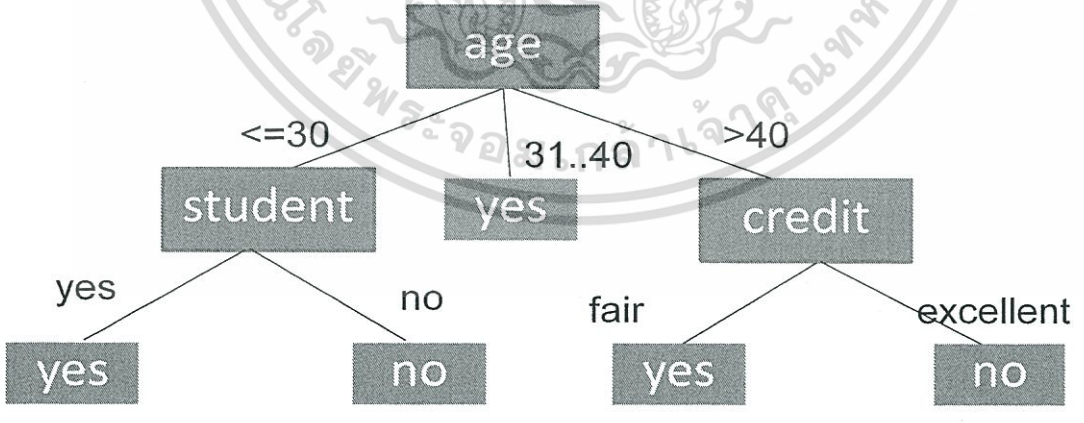
ตัวอย่างข้อมูลสองระบบที่แสดงการซื้อเครื่องคอมพิวเตอร์ โดยเราสนใจในการสร้างโมเดลคุณลักษณะของคนที่ซื้อเครื่องคอมพิวเตอร์ ตารางที่ 2.4 และ รูปที่ 2.13

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.4 ตัวอย่างข้อมูลการซื้อคอมพิวเตอร์

Age	Income	Student	Credit_rating	Buys_computer
≤ 30	High	No	fair	No
≤ 30	High	No	excellent	No
31 … 40	High	No	fair	Yes
>40	Medium	No	fair	Yes
>40	Low	Yes	fair	Yes
>40	Low	Yes	excellent	No
31 … 40	Low	Yes	excellent	Yes
≤ 30	Medium	No	fair	No
≤ 30	Low	Yes	fair	Yes
>40	Medium	Yes	fair	Yes
≤ 30	Medium	Yes	excellent	Yes
31 … 40	Medium	No	excellent	Yes
31 … 40	High	Yes	fair	Yes
>40	Medium	No	excellent	No

จากรูปการสร้างโมเดลการจำแนกข้อมูลสองระบบ ตารางที่ 2.4 ซึ่งประกอบด้วยอายุของผู้ซื้อ (age), เงินเดือน (income), การเป็นนักเรียน (student) และการจัดอันดับเครดิต (credit_rating) เป็นตัวแบ่งแยกว่าบุคคลนี้มีคุณสมบัติที่จะซื้อเครื่องคอมพิวเตอร์หรือไม่ และสามารถสร้างโมเดลต้นไม้ตัดสินใจได้ดังรูปที่ 2.13



รูปที่ 2.13 โมเดลต้นไม้ตัดสินใจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.1 ปัญหาการไม่สมดุลของข้อมูล (Imbalanced datasets Problem)

เนื่องจากปัญหาที่พิจาณาคือการเกิดปริมาณฝุ่นละอองเกินควบคุม (Out of control) ซึ่งจะพบว่าข้อมูลทั้ง 2 คลาส ได้แก่คลาสส่วนมาก (Majority class) คือคลาสที่มีปริมาณมากในชุดข้อมูล และคลาสส่วนน้อย (Minority class) คือคลาสที่มีปริมาณข้อมูลน้อยกว่าในชุดข้อมูลเดียวกัน โดยคลาสส่วนมากและคลาสส่วนน้อยมีปริมาณไม่สมดุลกันอย่างมาก เช่น ปริมาณคลาสของข้อมูล Failed มีน้อยเกินไป ในขณะที่คลาสข้อมูล Normal มีปริมาณมากกว่าคลาส Failed เป็นจำนวนมาก ทำให้ไม่สามารถหาต้นไม้ตัดสินใจหรือกฎปัจจัยที่ทำให้ปริมาณฝุ่นละอองเกินควบคุม ดังนั้นจึงต้องทำ Sampling ข้อมูลเพื่อให้ข้อมูลมีความสมดุล

2.3.1.1 SMOTE - Synthetic Minority Oversampling Technique

SMOTE [10][11] เป็นเทคนิคที่คิดค้นโดย Chawla, Hall and Kegelmeyer เมื่อปี ค.ศ.2002 โดยเสนอหลักการ Oversampling กับคลาสส่วนน้อย (Minority class) โดยสร้างข้อมูลสังเคราะห์จากข้อมูลคลาสส่วนน้อย

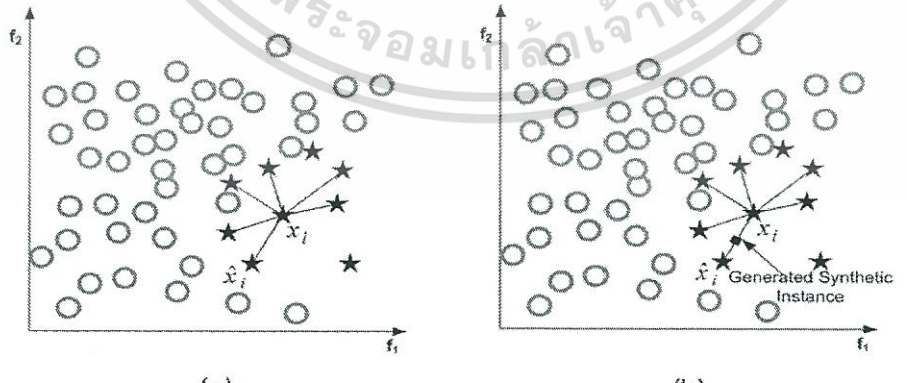
อัลกอริทึม SMOTE [19]สร้างข้อมูลจากข้อมูลที่ feature space ใกล้เคียงกันภายในคลาสส่วนน้อย โดยสับเซต $S_{min} \in S$ จะพิจารณาจาก k-nearest neighbors แต่ละข้อมูลคือ $x_i \in S_{min}$ โดยจะเลือกข้อมูลจำนวน k ข้อมูลใน S_{min} ซึ่งคำนวณระยะทางแบบยุคลิด จากนั้นสุ่มเลือกข้อมูลตัวหนึ่งใน k-nearest neighbors แล้วคูณ feature vector ที่สอดคล้องกันด้วยการสุ่มตัวเลขระหว่าง [0,1] และเพิ่มเวกเตอร์นี้เข้าไปใน x_i

$$x_{new} = x_i + (\hat{x}_i - x_i) \times \delta \tag{2.5}$$

โดย $x_i \in S_{min}$ คือ ข้อมูลส่วนน้อย (Minority instance) ที่พิจารณา

$\hat{x}_i$ คือ ข้อมูลค่าหนึ่งใน K-nearest neighbors ซึ่ง x_i: $\hat{x}_i \in S_{min}$

$\delta \in [0,1]$ คือ ตัวเลขสุ่ม

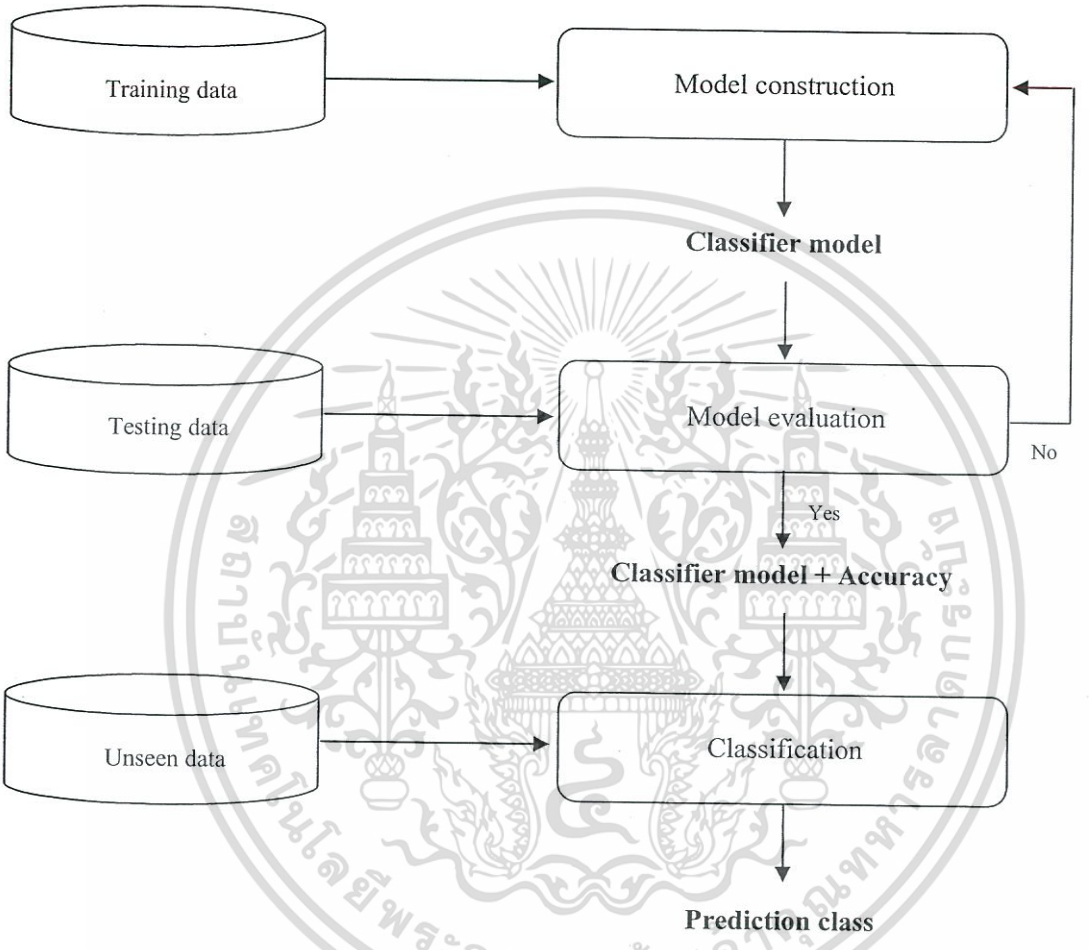


รูปที่ 2.14 (a) ตัวอย่าง K-nearest neighbors ของข้อมูลที่พิจารณา เมื่อ k=6
(b) ข้อมูลสังเคราะห์ (Synthetic Instance) ที่สร้างขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.2 กระบวนการจำแนกข้อมูล

การจำแนกข้อมูล (Classification) เป็นกระบวนการสร้างโมเดลจัดการข้อมูลให้อยู่ในกลุ่มที่กำหนดมาให้ ตัวอย่างเช่น จัดกลุ่มนักเรียนว่าดีมาก ดี ปานกลาง และไม่ดี โดยพิจารณาจากประวัติและผลการเรียนหรือแบ่งประเภทของลูกค้าว่าเชื่อถือได้หรือไม่โดยพิจาณาจากข้อมูลที่มีอยู่ กระบวนการจำแนกข้อมูลแบ่งออกเป็น 3 ขั้นตอนดังนี้



รูปที่ 2.14 กระบวนการจำแนกข้อมูล

2.3.2.1 Model construction (Learning)

เป็นขั้นตอนการสร้างโมเดลโดยอาศัยการเรียนรู้จากข้อมูลที่ได้กำหนดคลาสไว้เรียบร้อยแล้ว (Training data) ซึ่งโมเดลที่ได้อาจแสดงในรูปของต้นไม้ตัดสินใจคล้ายโครงสร้างต้นไม้ที่แต่ละโหนดแสดงแอททริบิวต์ แต่ละกิ่งแสดงผลในการทดสอบ และโหนดปลาย (Leaf node) แสดงคลาสที่กำหนดไว้ ซึ่งต้นไม้ตัดสินใจนี้ง่ายต่อการปรับเปลี่ยนเป็นกฎของการจำแนก (Classification rule)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.2.2 Model evaluation (Accuracy)

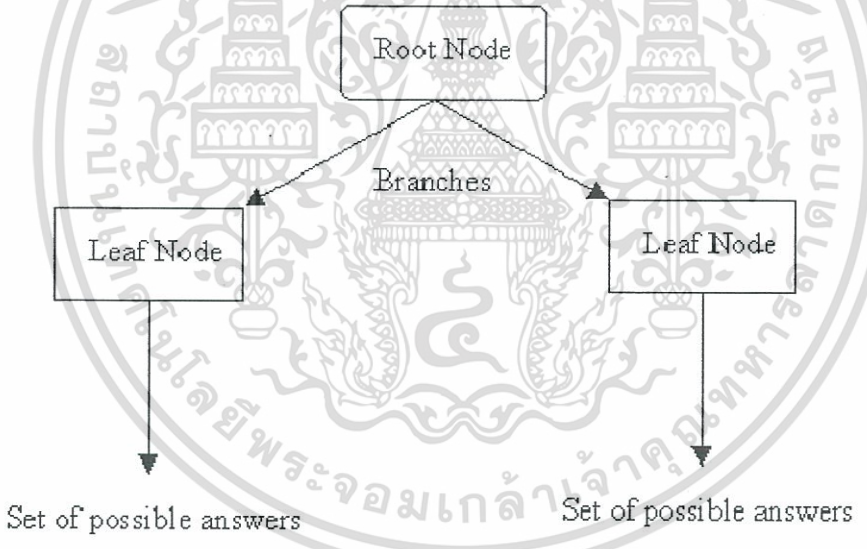
เป็นขั้นตอนประมาณความถูกต้องโดยอาศัยข้อมูลที่ใช้ทดสอบ (Testing data) ซึ่งคลาสที่แท้จริงของข้อมูลที่ใช้ทดสอบนี้จะถูกนำมาเปรียบเทียบกับคลาสที่หามาได้จากโมเดลเพื่อทดสอบความถูกต้อง

2.3.2.3 Model usage (Classification)

เป็นโมเดลสำหรับใช้กับข้อมูลที่ไม่เคยเห็นมาก่อน (Unseen data) โดยจะทำการกำหนดคลาสให้กับออปเจคใหม่ที่ได้มาหรือทำนายค่าที่ต้องการออกมา

2.3.3 การจำแนกข้อมูลด้วยต้นไม้ตัดสินใจ (Decision Tree Classification)

ต้นไม้ตัดสินใจ (Decision Tree) เป็นโครงสร้างที่ใช้แสดงกฎที่ได้จากเทคนิคการจำแนกข้อมูลโดยต้นไม้ตัดสินใจจะมีลักษณะคล้ายโครงสร้างของต้นไม้ โดยแต่ละโหนดจะแสดงคุณลักษณะ (attribute) แต่ละกิ่งแสดงเงื่อนไขในการทดสอบ และโหนดปลาย (leaf node) แสดงกลุ่มที่กำหนดไว้ โดยแสดงโครงสร้างของแผนภูมิต้นไม้ตัดสินใจดังรูปที่ 2.15



รูปที่ 2.15 โครงสร้างของแผนภูมิต้นไม้ตัดสินใจ

จากรูปที่ 2.15 โครงสร้างต้นไม้ตัดสินใจประกอบด้วย

Root node หรือ Decision node เป็นส่วนเงื่อนไขการตัดสินใจ

Leaf node คือ ค่าที่เป็นไปได้จากเงื่อนไขการตัดสินใจ

Branch คือ การเชื่อมต่อระหว่างโหนด (node)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แผนภูมิต้นไม้ตัดสินใจเป็นอัลกอริทึมเพื่อการจัดหมวดหมู่หรือการจำแนกข้อมูลอย่างหนึ่ง โดยการสำรวจรายการในฐานข้อมูลเพื่อแยกให้อยู่ในหมวดหมู่ที่ได้กำหนดไว้ล่วงหน้าแล้ว เช่น การแบ่งกลุ่มสินค้าเป็นกลุ่มเครื่องใช้, กลุ่มอาหารสด และกลุ่มอาหารแห้งเป็นต้น แผนภูมิชนิดนี้เป็นที่นิยมากเนื่องจากเป็นลักษณะที่เข้าใจง่าย ลักษณะการทำงานเหมือนการตั้งคำถามเพื่อต้องการคำตอบที่จะนำไปสู่ผลการจำแนกข้อมูลหรือผลการทำนาย (Prediction) ได้เร็วที่สุดและมีความแม่นยำสูง โดยแต่ละโหนดแทนคำถามเกี่ยวกับเรื่องหนึ่งๆ ที่นำไปใช้พิจารณาจำแนกข้อมูลหรือทำนายผล และแต่ละโหนดมีชุดคำตอบที่จะโยงไปสู่คำถามถัดไป ส่วนความแม่นยำแปรผันตรงกับข้อมูลที่ใช้ในการสอนแผนภูมิต้นไม้ตัดสินใจว่ามีความถูกต้องและมีปริมาณมากพอหรือไม่

ดังนั้นสิ่งสำคัญในการสร้างแผนภูมิต้นไม้ตัดสินใจคือ วิธีที่ใช้ในการเลือกหรือแบ่งแอททริบิวต์ (ชื่อโหนดและค่าที่เป็นไปได้ของโหนด) [15] ขณะสร้างแผนภูมิต้นไม้ตัดสินใจอย่างไรให้ได้โหนดที่ดีที่สุดมาสร้างเป็นแผนภูมิต้นไม้ตัดสินใจ เพราะจะทำให้แผนภูมิต้นไม้ตัดสินใจมีขนาดเล็กที่สุดซึ่งช่วยลดเวลาในการประมวลผลและมีความแม่นยำสูงสุด สำหรับงานวิจัยนี้เลือกใช้วิธีการสร้างต้นไม้ตัดสินใจ (Decision tree) ด้วยอัลกอริทึม C4.5

2.3.3.1 อัลกอริทึม C4.5

อัลกอริทึม C4.5 เป็นอัลกอริทึมที่สร้างต้นไม้ตัดสินใจที่พัฒนาโดย J.Ross Quinlan (1993) โดยนำอัลกอริทึม ID3 มาปรับปรุงให้มีความสามารถมากขึ้นโดยใช้วิธีการ Information Gain เพิ่มเติมการจัดรูปแบบข้อมูล ตัวเลข ข้อมูลที่ขาดหายและไม่สมบูรณ์และการพรุน (Prune) ด้วยการแทน Branch ที่ไม่ช่วยในการตัดสินใจด้วย Leaf node ที่ตัดสินใจได้ดีกว่า

- การเลือกแอททริบิวต์ทดสอบด้วย Information gain

ค่า Information gain เป็นค่าในการเลือกแอททริบิวต์การทดสอบที่แต่ละโหนดในต้นไม้ตัดสินใจ แอททริบิวต์ที่มีค่า Information gain สูงสุด (หรือค่าเอนโทรปีดีที่สุด) จะถูกเลือกให้เป็นแอททริบิวต์ทดสอบในเวลานั้น

ให้ S เป็นกลุ่มตัวอย่างของข้อมูล สมมติคลาสปลายทางมี m ค่า โดยการกำหนดเป็น m คลาสที่ต่างกัน, $C_i (\text{for } i = 1, \dots, m)$ ให้ s_i เป็นจำนวนคลาสข้อมูล S ในคลาส C_i ค่า Information gain จากกลุ่มตัวอย่างข้อมูลหาได้โดย

$$I(s_1, s_2, \dots, s_m) = -\sum_{i=1}^{m} p_i \log(p_i)$$

เมื่อ p_i คือความน่าจะเป็นของตัวอย่างข้อมูลในคลาส C_i และประมาณโดย s_i/s

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ให้แอททริบิวต์ A มีค่าที่แตกต่างกัน v ค่า $\{a_1, a_2, \ldots, a_v\}$ แอททริบิวต์ A สามารถแบ่งเป็นส่วน S ไปสู่ v ซับเซ็ต $\{S_1, S_2, \ldots, S_v\}$ เมื่อ S_j ประกอบด้วยตัวอย่างข้อมูลใน S ที่มีค่า a_j ใน A ถ้า A ได้รับเลือกเป็นแอททริบิวต์ทดสอบ (เป็นแอททริบิวต์ที่ดีที่สุดในการแบ่ง) แล้วค่าซับเซ็ตเหล่านี้จะแตกเป็นกิ่งจากโหนดที่ประกอบด้วยเซ็ต S ให้ s_{ij} เป็นจำนวนข้อมูลของคลาส C_i ในซับเซ็ต S_j ค่าเอนโทรปี (entropy) ที่ขึ้นกับการแบ่งส่วนซับเซ็ตโดย A กำหนดดังนี้

$$E(A) = \sum_{j=1}^{v} \frac{s_{1j} + \cdots + s_{mj}}{s} * I(s_{1j}, \ldots, s_{mj})$$

เทอม $\frac{s_{1j}+\cdots+s_{mj}}{s}$ เป็นค่าน้ำหนักของซับเซ็ตที่ j และเป็นจำนวนข้อมูลในซับเซ็ตที่ถูกแบ่งโดยจำนวนตัวอย่างข้อมูลรวมใน S โดยยิ่งค่าเอนโทรปีน้อย ค่าส่วนแบ่งซับเซ็ตจะยิ่งดีมาก สำหรับค่าซับเซ็ต S_j

$$I(s_{1j}, s_{2j}, \ldots, s_{mj}) = -\sum_{i=1}^{m} p_{ij} \log(p_{ij})$$

เมื่อ $p_{ij} = s_{ij}/|S_j|$ เป็นค่าความน่าจะเป็นของกลุ่มข้อมูล S_j ในคลาส C_i
ค่าที่แสดงถึงความแตกต่างบนแอททริบิวต์ A คือ

$$\text{Gain}(A) = I(S_1, S_2, \ldots, S_m) - E(A)$$

อัลกอริทึมจะคำนวณค่า Information gain สำหรับแต่ละแอททริบิวต์ แอททริบิวต์ที่มีค่า information gain สูงสุดจะถูกเลือกให้เป็นแอททริบิวต์ทดสอบในเซ็ต S จากนั้นจะสร้างโหนดและกิ่งที่แสดงค่าในแอททริบิวต์ และแบ่งข้อมูลต่อไปตามลำดับ

2.3.4 เครื่องมือสำหรับการสร้างต้นไม้ตัดสินใจ

เครื่องมือที่ใช้สำหรับการสร้างต้นไม้ตัดสินใจในงานวิจัยนี้เลือกใช้อัลกอริทึม C4.5 ซึ่งทำการสร้างผ่านโปรแกรม WEKA (Waikato Environment for Knowledge Analysis) โดย Ian H. Witten and Eibe Frank (2005) พัฒนาโดย the University of Waikato in New Zealand เพื่อเป็นเครื่องมือสำหรับการทำเหมืองข้อมูลโดยใช้หลักการของ Machine Learning ซึ่งมีรูปแบบการติดต่อกับผู้ใช้ (Graphic User Interface: GUI) ที่สะดวกในการเรียกใช้งานเรียกว่า Explore ทำงานบน Java VM และมีฟังก์ชั่นการทำงานต่างๆ เช่น Pre-processing, Visualizing data learning ฯลฯ ให้เลือกใช้ในงานต่างๆ เช่น Classification, Regression, Clustering, Association Rule เป็นต้น ซึ่งงานวิจัยนี้ใช้ J48 ซึ่งเป็นอัลกอริทึม C4.5

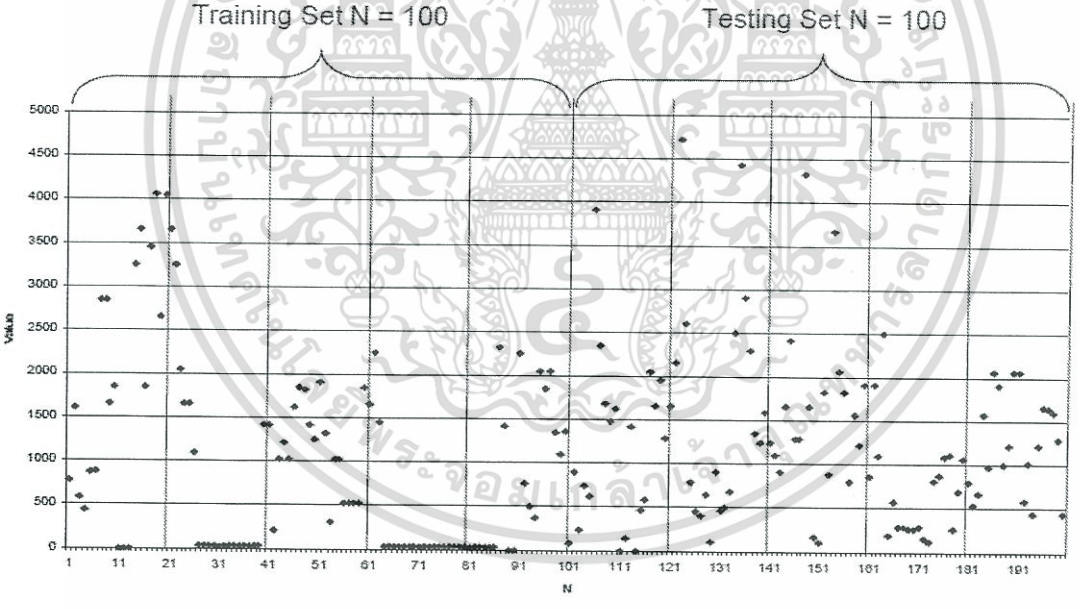
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.5 วิธี Cross-Validation

Cross-Validation [18] เป็นส่วนที่สำคัญในการใช้เพื่อแบ่งกลุ่มข้อมูลการเรียนรู้และกลุ่มข้อมูลทดสอบในการจำแนก ซึ่งถ้าเลือกรูปแบบการ Cross-Validation ไม่ดีจะทำให้ผลลัพธ์ที่ได้ออกมาจากการจำแนกมีความผิดพลาดในการจำแนกมาก ซึ่งการแบ่งกลุ่มข้อมูลเพื่อการเรียนรู้นั้นมีวิธีการหลายรูปแบบแตกต่างกัน ข้อดีและข้อเสียต่างกัน โดยการทำการแบ่งกลุ่มข้อมูลเพื่อเรียนรู้มีวิธีการดังนี้

2.3.5.1 Hold-out Cross-Validation

Hold-out Cross-Validation เป็นวิธีการดั้งเดิมในการแบ่งแยกกลุ่มข้อมูลซึ่งจะนำมาใช้ในการเรียนรู้ (Training) และทดสอบ (Testing) โดยจะทำการแบ่งออกเป็นอัตราส่วนร้อยละของข้อมูล เช่น แบ่งที่ร้อยละ 50 หมายถึง จะนำข้อมูลออกมาร้อยละ 50 มาใช้เป็นข้อมูลกลุ่มการเรียนรู้ (Training set) และนำอีกร้อยละ 50 มาเป็นข้อมูลกลุ่มการทดสอบ (Testing set) ตัวอย่างเช่น ข้อมูลจำนวน 200 ชุดข้อมูลจะถูกแบ่งออกเป็น 100 ชุด เพื่อเป็นกลุ่มข้อมูลเพื่อเรียนรู้ และอีกส่วนมาเป็นกลุ่มข้อมูลเพื่อทดสอบจำนวน 100 ชุด แสดงได้ดังรูปที่ 2.16



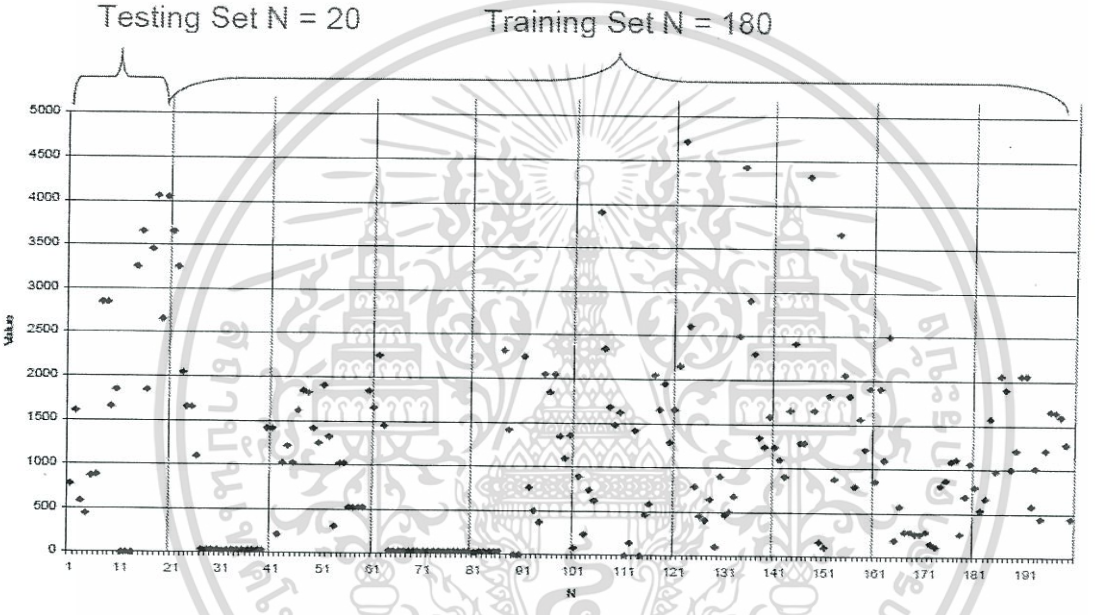
รูปที่ 2.16 การแบ่งกลุ่มข้อมูลเพื่อทดสอบ Cross-Validation แบบ Hold-out ที่ร้อยละ 50

แต่เนื่องจากวิธี Cross-Validation แบบ Hold-out มีข้อจำกัดคืออาจจะได้ผลการทดลองที่คลาดเคลื่อนได้ ถ้าเกิดการเลือกกลุ่มข้อมูลการเรียนรู้ (Training set) ไปเลือกข้อมูลที่ไม่สามารถครอบคลุมทุกกรณีที่เกิดขึ้นในข้อมูลทั้งหมด จะส่งผลให้ค่าความถูกต้องได้ผลลัพธ์ที่ไม่ดี จึงควรทำการสุ่มใหม่หลายๆ ครั้ง เพื่อจะหาค่าความถูกต้องเฉลี่ยของวิธีการนี้ แต่ก็มีข้อดีคือเหมาะกับข้อมูลขนาดใหญ่และสามารถจำแนกข้อมูลที่มีขนาดใหญ่ได้โดยใช้เวลาไม่นาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.5.2 k-Fold Cross-Validation

วิธีนี้จะทำโดยการแบ่งข้อมูลออกเป็น k_i, k_i =1,2,3,…,n กลุ่มแล้วให้ทำการใช้ 1 กลุ่มมาเป็นกลุ่มข้อมูลการทดสอบ (Testing set) ส่วนที่เหลือคือ k-1 นำมาใช้เป็นกลุ่มข้อมูลการเรียนรู้ (Training set) แล้วจะทำการวนป็นจำนวน k รอบ โดยจะเปลี่ยนกลุ่มทดสอบไปเรื่อยๆ ตามลำดับจนครบ โดยวิธีนี้จะทำให้จำนวนของข้อมูลทั้งหมดได้ผ่านขั้นตอนการจำแนกซึ่งแตกต่างจากการใช้ Cross-Validation แบบ Hold-out ที่กลุ่มการเรียนรู้จะไม่ได้นำมาทดสอบด้วย โดยส่วนใหญ่แล้วนิยมวช้ 10-Fold เนื่องจากได้ค่าความถูกต้องเป็นที่น่าพอใจและเวลาในการทำงาน (Time complexity) ไม่มากนัก สามารถแสดงตัวอย่างการทำงานได้ดังรูปที่ 2.17 และ รูปที่ 2.18

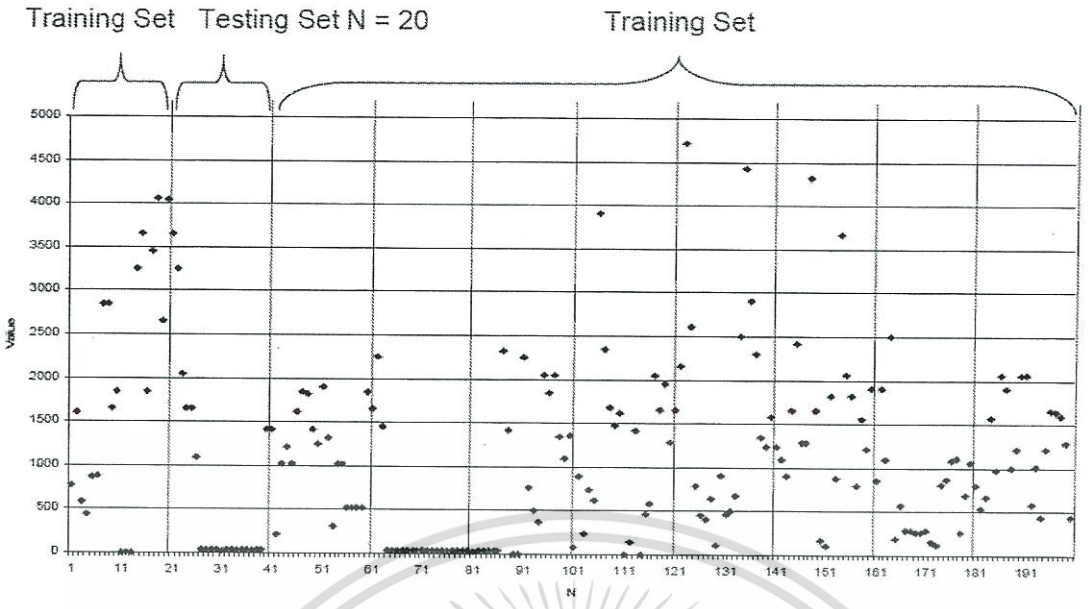


รูปที่ 2.17 ขั้นตอนที่ 1 ของการแบ่งกลุ่มข้อมูลเพื่อทดสอบ Cross-Validation แบบ 10-fold

โดยการ Cross-Validation แบบ k-fold ที่ k = 10 ในรูปที่ 2.10 จะเห็นว่า N = 200 ในรอบแรกจะใช้ข้อมูลจำนวน 20 ในการทดสอบแล้วให้ข้อมูลอีก 180 ข้อมูลเป็นกลุ่มการเรียนรู้

รูปที่ 2.18 เป็นการทำ 10-fold Cross-Validation รอบที่ 2 หลังจากนั้นจะวนทำจนกระทั่งครบทั้ง 10 กลุ่มของข้อมูล ซึ่งการทำ Cross-Validation แบบนี้เราจะสามารถให้ข้อมูลทั้งหมดเป็นกลุ่มข้อมูลการเรียนรู้ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.18 ขั้นตอนที่ 2 ของการแบ่งกลุ่มข้อมูลเพื่อทดสอบ Cross-Validation แบบ 10-fold

ซึ่งงานวิจัยนี้ได้เลือกการทดลองใช้ Cross-Validation แบบ k-Fold ที่ k = 10 ในการวิเคราะห์ข้อมูลปริมาณอนุภาคฝุ่นละอองเพื่อหากฎสำหรับการควบคุมคุณภาพปริมาณอนุภาคฝุ่นละออง โดยใช้การจำแนกข้อมูลอัลกอริทึม C4.5 (J48) สำหรับวิเคราะห์ข้อมูลในวิทยานิพนธ์จะกล่าวรายละเอียดและตัวอย่างในบทที่ 3

2.3.6 การวัดประสิทธิภาพของแบบจำลอง

การวัดประสิทธิภาพของแบบจำลองที่ได้จากการจำแนกข้อมูลด้วยเทคนิคต้นไม้ตัดสินใจอัลกอริทึม C4.5 ได้จากการนำแบบจำลองที่ได้จากชุดข้อมูลการเรียนรู้มาทดสอบด้วยชุดข้อมูลการทดสอบ ที่ให้ค่าความถูกต้องในการจำแนก (Accuracy), ค่าความครบถ้วน (Recall), ค่าความแม่นยำ (Precision) และค่าถ่วงดุล (F-measure) [16]

- ค่าความถูกต้องในการจำแนก (Accuracy) เป็นการวัดความถูกต้องของการจำแนกข้อมูลที่จำแนกได้ถูกต้องตรงกับความเป็นจริงเพียงใด

$$\text{ค่าความถูกต้องในการจำแนก} = \frac{\text{จำนวนข้อมูลที่จำแนกได้ถูกต้อง}}{\text{จำนวนข้อมูลทดสอบทั้งหมด}}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ค่าความครบถ้วน (Recall) เป็นค่าที่บอกว่าโมเดลทำนายว่าจริงนั้นเป็นอัตราส่วนเท่าไรของค่าจริงทั้งหมด โดยคำนวณจากค่าที่มีผลลัพธ์ถูกต้อง โดยพิจารณาจากข้อมูลของผลลัพธ์เดียวกัน

$$\text{ค่าความครบถ้วน} = \frac{\text{จำนวนข้อมูลที่จำแนกได้ถูกต้องของกลุ่มข้อมูลนั้น}}{\text{จำนวนข้อมูลที่อยู่ในกลุ่มนั้นจริง}}$$

- ค่าความแม่นยำ (Precision) เป็นค่าที่บอกว่าโมเดลทำนายว่าจริงนั้นถูกต้องเท่าไร โดยคำนวณจากค่าของข้อมูลที่มีผลลัพธ์ถูกต้องโดยพิจารณาจากจำนวนข้อมูลทั้งหมดที่ถูกจำแนกมีผลลัพธ์เดียวกัน

$$\text{ค่าความแม่นยำ} = \frac{\text{จำนวนข้อมูลที่จำแนกได้ถูกต้องของกลุ่มข้อมูลนั้น}}{\text{จำนวนข้อมูลที่ถูกจำแนกว่าเป็นกลุ่มนั้นทั้งหมด}}$$

- ค่าถ่วงดุล (F-measure) เป็นค่าที่แสดงความสัมพันธ์ระหว่างค่าความแม่นยำ (Precision) และ ค่าความระลึก (Recall) เพื่อหาความถูกต้องโดยรวม

$$\text{ค่าถ่วงดุล} = \frac{2 \times \text{ค่าความแม่นยำ} \times \text{ค่าความครบถ้วน}}{\text{ค่าความแม่นยำ} + \text{ค่าความครบถ้วน}}$$

2.4 งานวิจัยที่เกี่ยวข้อง

เนื่องจากการวิเคราะห์ข้อมูลในกระบวนการควบคุมกระบวนการทางสถิตินั้นมีหลายงานวิจัยที่ได้ทำการศึกษาเปรียบเทียบวิธีการทางเหมืองข้อมูลด้วยเทคนิคต่างๆ เพื่อพิจารณาเทคนิคที่เหมาะสมกับข้อมูลนั้นๆ และมีงานวิจัยที่นำเทคนิคเหมืองข้อมูลมาใช้ในการวิเคราะห์และทำนายข้อมูลผิดปกติ อีกทั้งยังเปรียบเทียบประสิทธิภาพของแต่ละเทคนิคในการตรวจหาข้อมูลผิดปกติ ซึ่งจากการศึกษางานวิจัยที่เกี่ยวข้องดังกล่าวมีดังนี้

2.4.1 การใช้เทคนิคเหมืองข้อมูลในการปรับปรุงคุณภาพการผลิต: กรณีศึกษาของโรงงานบรรจุ LCD Driver IC (Using Data Mining Technology to improve Manufacturing Quality – A Case Study of LCD Driver IC Packaging Industry) [4] มีวัตถุประสงค์ในการวิจัยเพื่อใช้เทคนิคเหมืองข้อมูลเพื่อปรับปรุงคุณภาพการผลิตของกรณีศึกษาโรงงานบรรจุแอลซีดี ไดรเวอร์ ไอซี (LCD Driver IC) โดยทำการศึกษากระบวนการทำงานของขั้นตอนการบรรจุแอลซีดี ไดรเวอร์ไอซี เพื่อค้นหาปัจจัยหลักในการแก้ปัญหาของกระบวนการผลิต ดังนั้นจึงได้นำเสนอเทคนิคเหมือง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลเพื่อใช้สำหรับหาวิเคราะข้อมูลเพื่อให้ทราบว่าสาเหตุความบกพร่องที่เกิดขึ้นแต่ละรูปแบบนั้นเกิดจากสาเหตุใด ซึ่งปัญหาทางด้านคุณภาพในการบรรจุไดรเวอร์ไอซี นั้นแสดงได้ดังตารางที่ 2.5

โดยการวิเคราะห์คุณภาพของงานวิจัยนี้ประกอบด้วย ฐานข้อมูลปัญหาของการควบคุมคุณภาพ (Quality Problems Database), การวิเคราะห์การผลิต (line analysis) และการวิเคราะห์ด้วยเหมืองข้อมูล (Data mining) ซึ่งมีขั้นตอนในการดำเนินการดังนี้

2.4.1.1 การรวบรวมข้อมูล (Data collection) โดยการรวบรวมข้อมูลการผลิตในสถานีปฏิบัติการของการบรรจุไดรเวอร์ไอซี ในช่วงระหว่าง เดือนมกราคม ค.ศ. 2002 ถึง เดือนธันวาคม ค.ศ.2003 จำนวน 9,752 ระเบียน

2.4.1.2 การเตรียมข้อมูล (Data preparation) ขั้นตอนนี้จะดำเนินการทั้งหมด 3 กิจกรรม เพื่อเตรียมข้อมูลให้เหมาะสมสำหรับการวิเคราะห์ได้แก่ การรวบรวมและคัดกรองข้อมูล (Data collecting and screening), การทำความสะอาดข้อมูล (Data scrubbing) และ การสร้างข้อมูลใหม่ (New data generating) จากการดำเนินการในกิจกรรมดังกล่าวพบว่า ข้อมูลปัญหาทางคุณภาพมี 15 ประเภท และตัวแปรสำหรับการจัดทำเหมืองข้อมูล 4 ตัวแปร ได้แก่ man จำนวน 17 ข้อมูล, machine จำนวน 13 ข้อมูล, material จำนวน 10 ข้อมูล และ method จำนวน 12 ข้อมูล ซึ่งแสดงตัวอย่างข้อมูลได้ดังตารางที่ 2.6

ตารางที่ 2.5 ปัญหาทางด้านคุณภาพของการบรรจุไดรเวอร์ไอซี

สถานีปฏิบัติการ (Station)	ปัญหา
Pre Mold Inspection	Inner Lead broken Inner Lead curved Inner Lead bent
Final Test	Short Open Function
Visual Monitor	Wrong stamp direction Tape jammed Chip has cracks
Rewind Inspection	Tape bent Denton tape Pot over resin

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.6 ตัวอย่างแอททริบิวต์และปัญหาคุณภาพ

Runcard	Member	Machine	Method	Material	Problem
TW010234-01	Wrong Maring Parameter	Marking sensor error	None	None	Repeat stamp
TW030046-02	Wrong Potting Parameter	ILB Pressure error	None	Inner lead is too thin	Inner lead is broken
...	...	...	...	...	...
TW310112-05	Wrong Maring Parameter	ILB Pressure error	Position of ILB is shifted	None	Side shipping
TW310067-01	Wrong Potting Parameter	Potting pressure error	None	Resin drop	Pot over resin

2.4.1.3 การใช้อัลกอริทึมต้นไม้ตัดสินใจ (Using decision tree algorithm) หลังจากเลือกข้อมูลปัญหาด้านคุณภาพและกำหนดตัวแปรลักษณะ (attribute variable) ได้แก่ คน (member), เครื่องจักร (machine), วัสดุ (material), วิธีการ (method) และ ตัวแปรจัดกลุ่ม (category variable) ได้แก่ ปัญหาคุณภาพ (quality problem) และใช้ข้อมูลดังกล่าวไปใช้ในกระบวนการเรียนรู้ด้วยอัลกอริทึม C4.5 ซึ่งตัวอย่างข้อมูลที่ใช้สำหรับการเรียนรู้สามารถแสดงได้ดังตารางที่ 2.7

ตารางที่ 2.7 ตัวอย่างข้อมูลปัญหาคุณภาพ

Id	Material	Machine	Method	Member	Problem
1	Glue itself is too thick	None	Abnormal Temporature	Abnormal marking parameter	bubble
2	Glue itself contain bubbles	None	Abnormal Temporature	None	bubble
3	Tape itself contracts	Abnormal marking pressure	None	Abnormal marking parameter	bubble
4	Tape itself contracts	None	Abnormal Temporature	Abnormal marking parameter	Pot over resin
5	Glue itself contain bubbles	None	None	Abnormal marking parameter	bubble
6	Glue itself is too thick	Abnormal curing temperature	Abnormal Temporature	None	Pot over resin
7	Tape itself contracts	Abnormal curing temperature	None	None	bubble

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

8	Glue itself contain bubbles	Abnormal curing temperature	Abnormal Temporature	Abnormal marking parameter	bubble
9	Glue itself is too thick	Abnormal curing temperature	None	Abnormal marking parameter	bubble
10	Tape itself contracts	Abnormal curing temperature	None	Abnormal marking parameter	bubble
11	Glue itself contain bubbles	Abnormal marking pressure	None	None	bubble
12	Tape itself contracts	Abnormal curing temperature	Abnormal Temporature	None	Pot over resin
13	Glue itself is too thick	Abnormal marking pressure	None	Abnormal marking parameter	Pot over resin
14	Glue itself is too thick	Abnormal marking pressure	Abnormal Temporature	None	Pot over resin

สำหรับการจำแนกข้อมูลในการทดลองของงานวิจัยนี้พบว่าตัวแปรวัสดุ (Material) มีค่าเกน (gain) สูงที่สุด และจากผลการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจ ผู้วิจัยได้นำผลการวิเคราะห์ข้อมูลไปปรับปรุงกระบวนการส่งผลต่อการควบคุมคุณภาพดังตารางที่ 2.8

ตารางที่ 2.8 ผลการปรับปรุงคุณภาพ

วัสดุ	ผลการปรับปรุง	การปรับปรุงดีขึ้น (%)	
		Decision tree	**Neural network**
Inner leads	ปริมาณบกพร่องลดลง	10.4	2.95
Chips	ปริมาณบกพร่องลดลง	11.6	4.6

จากตารางที่ 2.8 พบว่าเมื่อนำผลการวิเคราะห์คุณภาพด้วยอัลกอริทึมต้นไม้ตัดสินใจ (Decsion tree) และโครงข่ายประสาท (Neural Network) ไปปรับปรุงกระบวนการผลิตแล้วทำให้ลดปริมาณวัสดุที่เกิดการบกพร่องได้มากขึ้น ซึ่งจากผลการปรับปรุงพบว่า การปรับปรุงด้วยผลจากอัลกอริมทึมต้นไม้ตัดสินใจนั้นสามารถลดปริมาณวัสดุที่บกพร่องได้มากกว่าโครงข่ายประสาท

จากการศึกษางานวิจัยดังกล่าว สรุปได้ว่า การวิเคราะห์ปัญหาทางด้านคุณภาพของสินค้านั้นสามารถเพิ่มปริมาณสินค้าที่มีคุณภาพได้ ทั้งยังทำให้เข้าใจสาเหตุของการเกิดปัญหาในการผลิตก่อให้เกิดแนวทางแก้ไขปัญหา การผลิตมีประสิทธิภาพมากขึ้น สำหรับผลของการวิเคราะห์ปัญหาคุณภาพนั้นพบว่าอัลกอริทึมต้นไม้ตัดสินใจเป็นอัลกอริทึมที่เหมาะสมในการวิเคราะห์ทางด้าน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คุณภาพของสินค้าและการผลิตวัสดุคอมพิวเตอร์, การผลิตชิ้นส่วนอิเล็กทรอนิกส์, การบรรจุเซมิคอนดักเตอร์ และในอุตสาหกรรมการผลิตต่างๆ

2.4.2 ศึกษาระบบการปรับปรุงคุณภาพอย่างต่อเนื่องของการผลิตแอลอีดีด้วยเหมืองข้อมูล (Study on the Continuous Quality Improvement Systems of LED Packaging based on Data Mining) [5] มีวัตถุประสงค์ในการวิจัยคือการสร้างต้นแบบของระบบการปรับปรุงคุณภาพอย่างต่อเนื่องของการผลิตแอลอีดี โดยใช้การควบคุมคุณภาพด้วยการควบคุมกระบวนการเชิงสถิติและเหมืองข้อมูล

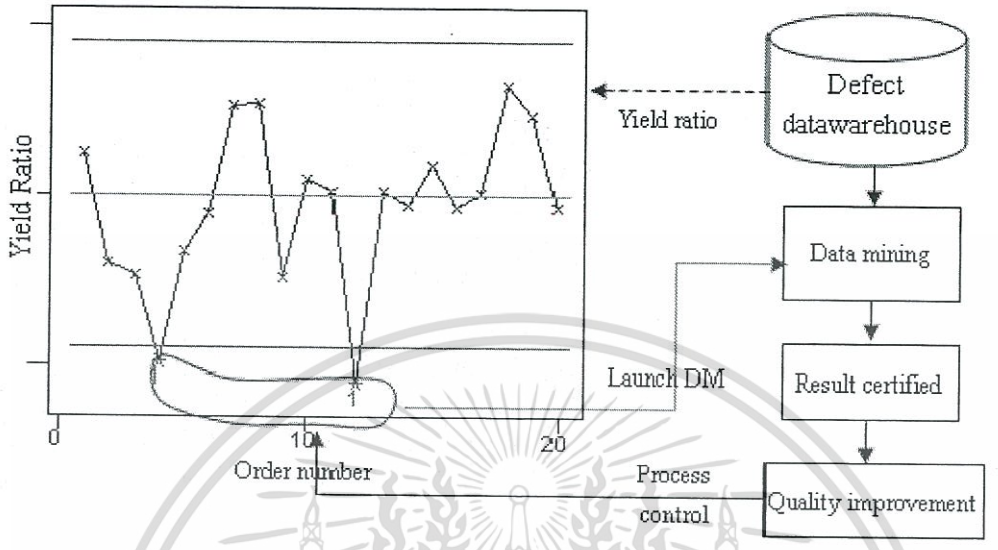
โดยเริ่มจากการวิเคราะห์ศึกษากระบวนการบรรจุแอลอีดี, การควบคุมกระบวนการเชิงสถิติ (Statistic Process Control: SPC) และเหมืองข้อมูล ซึ่งจากการศึกษาพบว่า ในกระบวนการบรรจุแอลอีดีนั้นประกอบด้วย 6 กระบวนการสำคัญได้แก่ การติด (mount), พันลวด (wire bond), การปั้น (molding), การเลื่อย (sawing), การทดสอบ (test) และ การบรรจุหีบห่อ (packing) และเมื่อศึกษาปัญหาทางด้านคุณภาพการบรรจะแอลอีดี พบว่าในกระบวนการที่ต่างกัน ดังนั้นจึงควบคุมกระบวนการด้วยมิติของ ผลิตภัณฑ์ (product), กระบวนการ (process), เวลา (time) และ คาบเวลา (period) การวิเคราะห์ข้อมูลทำด้วยเทคนิคต้นไม้ตัดสินใจ และการควบคุมคุณภาพทางสถิติใช้แผนภูมิควบคุมคุณภาพซึ่งใช้การควบคุมที่ระดับ $\pm 3\sigma$ ซึ่งปัญหาที่พบนั้นสามารถแสดงได้ดังตารางที่ 2.9

ตารางที่ 2.9 ข้อมูลปัญหาคุณภาพของการบรรจุแอลอีดี

สถานีปฏิบัติการ (Station)	ปัญหา/ตำหนิ
AM/AB	Wrong bonding Dice missing Line wrong Arc wrong
Molding	Pollute Bubble Peel off
Sawing	Shape
Final test	VF IV Hue

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ส่วนของการใช้กระบวนการควบคุมคุณภาพทางสถิติสำหรับงานวิจัยนี้ ผู้วิจัยได้ควบคุมคุณภาพโดยใช้แผนภูมิควบคุมเพื่อควบคุมสิ่งที่มีผลกระทบต่อลักษณะทางคุณภาพโดยอธิบายได้ดังรูปที่ 2.19



รูปที่ 2.19 โมเดลการควบคุมและปรับปรุงคุณภาพการผลิต

จากรูปที่ 2.19 สามารถอธิบายการทำงานได้ดังนี้

1. แผนภูมิควบคุมคุณภาพใช้สำหรับควบคุมสิ่งที่มีผลกระทบต่อคุณภาพ โดยถ้าหากไม่พบลักษณะที่อยู่นอกการควบคุมในแผนภูมิควบคุม ก็ไม่ต้องดำเนินการในขั้นตอนอื่น แต่ถ้าหากพบลักษณะที่อยู่นอกการควบคุมแล้วส่วนการทำงานเหมืองข้อมูลจะถูกนำมาใช้ทันที

2. เทคนิคเหมืองข้อมูลที่ใช้จะเป็นเทคนิคที่ใช้สำหรับการหากฎและความรู้จากฐานข้อมูลขนาดใหญ่ ซึ่งผลที่ได้อาจไม่จำเป็นต้องสมเหตุสมผล แต่ผลจากการวิเคราะห์ด้วยเหมืองข้อมูลจะต้องผ่านการตรวจสอบก่อนการนำไปใช้งาน

3. ผลจากการวิเคราะห์ด้วยเหมืองข้อมูลนั้น ปัจจัยที่ทำให้เกิดสถานะที่อยู่นอกเหนือการควบคุมสามารถได้รับการแก้ไขได้อย่างรวดเร็วและยังช่วยในการแก้ปัญหาทางคุณภาพได้อีกด้วย

4. แผนภูมิควบคุมคุณภาพนั้นยังคงใช้งานอย่างต่อเนื่องในระบบแม้การแก้ปัญหาทางคุณภาพจะได้รับการแก้ไขเรียบร้อยแล้ว

จากการศึกษางานวิจัยดังกล่าว สรุปได้ว่า การวิเคราะห์ปัญหาทางด้านคุณภาพของการบรรจุแอลอีดีนั้นพบว่าการควบคุมคุณภาพสามารถช่วยเพิ่มปริมาณผลิตภัณฑ์ที่ดีและสามารถช่วยให้การแก้ปัญหาทางคุณภาพนั้นสามารถทำได้อย่างรวดเร็วเนื่องด้วยการวิเคราะห์ด้วยเหมืองข้อมูล

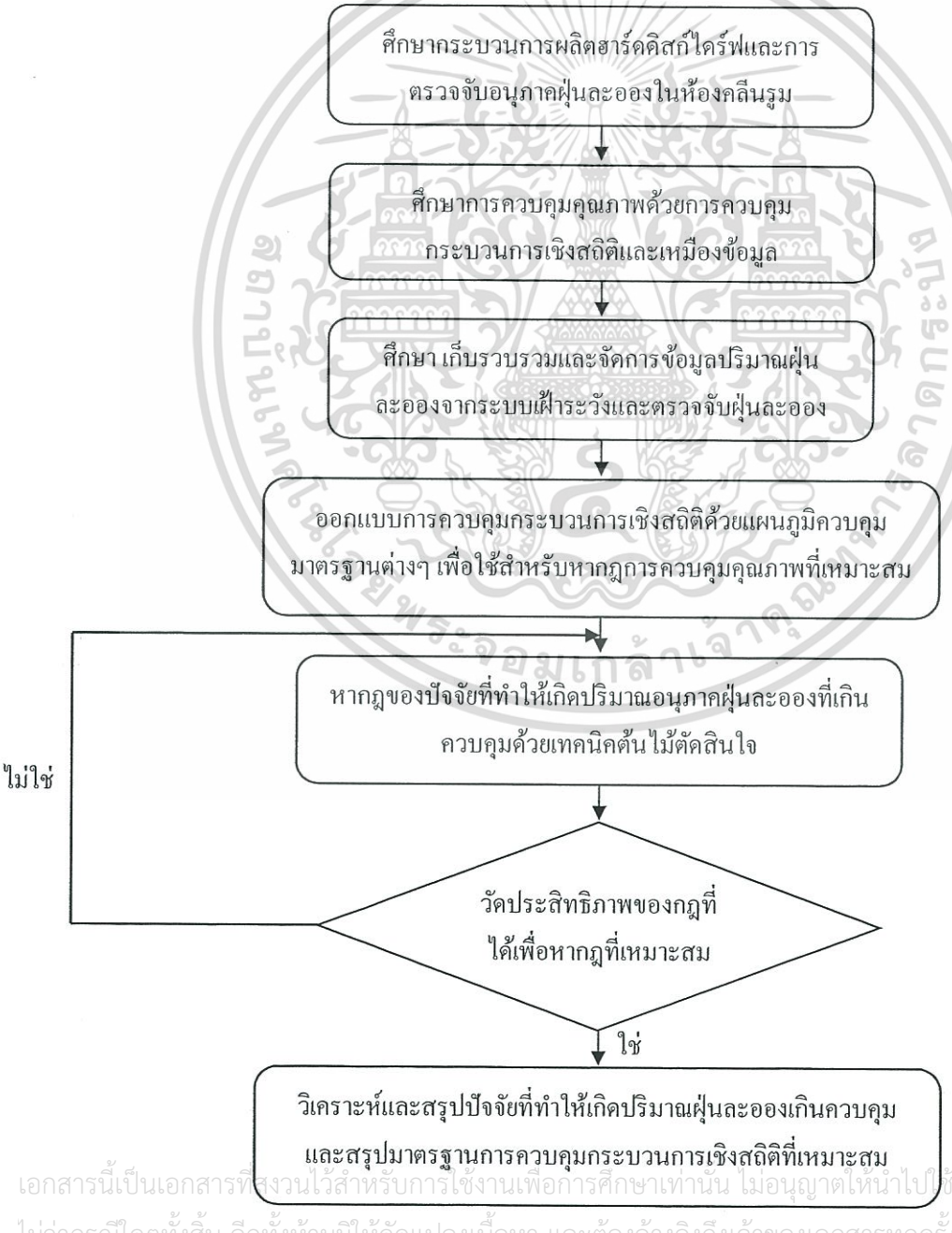
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.10 เปรียบเทียบแนวคิดเพื่อปรับปรุงงานวิจัยการควบคุมคุณภาพด้วยกระบวนการเชิงสถิติและเหมืองข้อมูล

ชื่องานวิจัย	ปี ค.ศ.	ผู้วิจัย	แนวคิดที่นำเสนอ	แนวคิดและการปรับปรุงสำหรับวิทยานิพนธ์
Using Data Mining Technology to improve Manufacturing Quality – A case Study of LCD Driver IC Packaging Industry	2006	Ruey-Shun Chen, Kun-Chieh Yeh, Chan-Chine Chang, H.H. Chien	- นำเสนอการหาสาเหตุของปัญหาทางด้านคุณภาพการผลิตในโรงงานอุตสาหกรรมด้วยการใช้เทคนิคเหมืองข้อมูล - ใช้เทคนิคการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจ	- งานวิจัยดังกล่าวนำเสนอเพียงวิธีการเทคนิคเหมืองข้อมูลเพียงอย่างเดียวซึ่งอาจไม่สามารถให้ผลที่ครอบคลุม สำหรับการควบคุมคุณภาพการผลิต - ดังนั้นนอกเหนือจากการใช้เหมืองข้อมูลในการวิเคราะห์เพื่อควบคุมคุณภาพแล้วจะทำการเพิ่มการควบคุมกระบวนการทางสถิติเข้าไป เพื่อให้การควบคุมคุณภาพมีประสิทธิภาพมากขึ้น - งานวิจัยดังกล่าวได้นำเสนอเทคนิคเหมืองข้อมูลด้วยต้นไม้ตัดสินใจซึ่งเป็นวิธีการที่เหมาะสม ดังนั้นวิทยานิพนธ์นี้จึงเลือกวิเคราะห์ข้อมูลด้วยเทคนิคการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจอัลกอริทึม C4.5
Study on the Continuous Quality Improvement Systems of LED Packaging based on Data Mining	2007	He Shu-guang, LI Li, Qi Er-shi	- เพื่อหาสาเหตุของปัญหาทางด้านคุณภาพการบรรจุแอลอีดีด้วยการใช้เหมืองข้อมูลร่วมกับการควบคุมกระบวนการเชิงสถิติ - ใช้เทคนิคการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจ	- งานวิจัยดังกล่าวนำเสนอวิธีการควบคุมคุณภาพทางสถิติที่ระดับการควบคุม $\pm 3\sigma$ ซึ่งอาจช้าสำหรับการดำเนินการแก้ไขเพื่อปรับปรุงคุณภาพ - ดังนั้นจึงเพิ่มเติมการควบคุมเตือน $\pm 2\sigma$ และการควบคุมระดับ $\pm 1\sigma$ เพื่อการควบคุมคุณภาพที่ดีขึ้น (ระดับการป้องกัน), หากฎมาตรฐานที่เหมาะสมสำหรับการควบคุมคุณภาพ และหารูปแบบของปัจจัยที่ทำให้เกิดปริมาณฝุ่นละอองที่เกินควบคุม

บทที่ 3
วิธีดำเนินการวิจัย

ในบทนี้จะกล่าวถึงการวิเคราะห์ข้อมูลด้วยเทคนิคเหมืองข้อมูลและการควบคุมกระบวนการเชิงสถิติ ประกอบด้วย 3.1 อธิบายลักษณะเบื้องต้นของข้อมูลที่ใช้เป็นกรณีศึกษาและรายละเอียดของการปรับปรุงข้อมูลเพื่อให้เหมาะสมสำหรับการวิเคราะห์และการสำรวจลักษณะของข้อมูลที่ใช้เป็นกรณีศึกษา หัวข้อ 3.2 กล่าวถึง การออกแบบมาตรฐานสำหรับการควบคุมกระบวนการเชิงสถิติของกรณีศึกษา และหัวข้อ 3.3 การหากฎของปัจจัยที่ทำให้ปริมาณอนุภาคฝุ่นละอองเกินควบคุมด้วยเหมืองข้อมูล



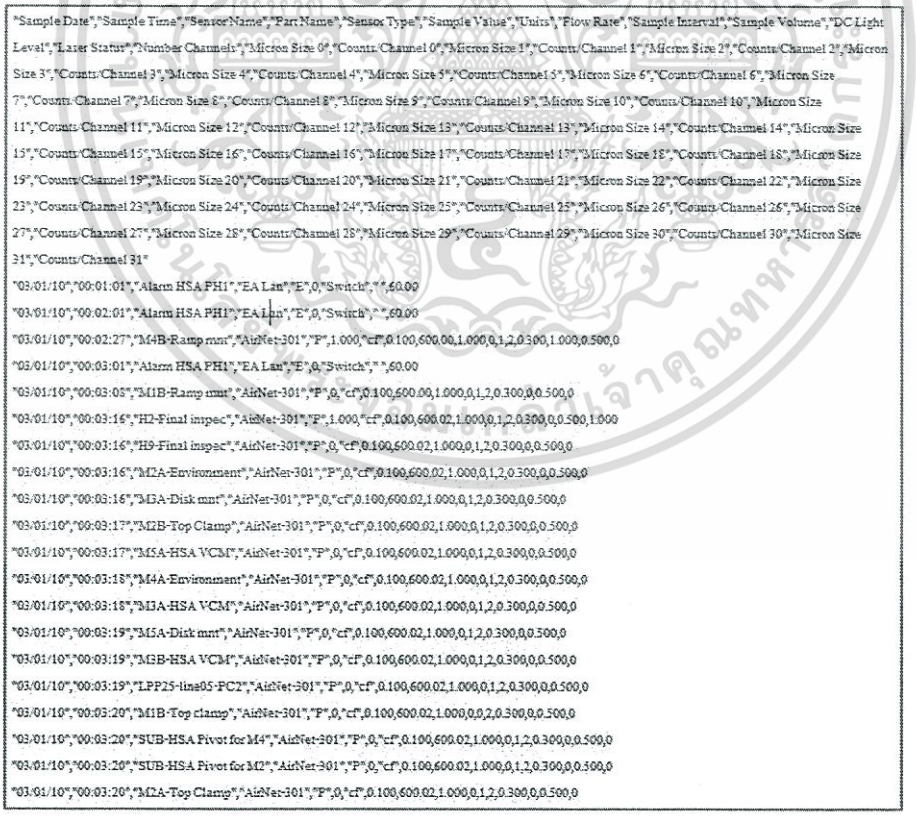
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1 การศึกษาและเก็บรวบรวมข้อมูลปริมาณฝุ่นละอองจากระบบเฝ้าระวังและตรวจจับฝุ่นละออง

สำหรับขั้นตอนการศึกษาและเก็บรวบรวมข้อมูลปริมาณอนุภาคฝุ่นละอองจากระบบเฝ้าระวังและตรวจจับฝุ่นละออง (Particle monitoring system) นั้นเป็นขั้นตอนเริ่มต้นของการทำความเข้าใจธรรมชาติของระบบและที่มาของปัญหาและศึกษากระบวนการของอุตสาหกรรมการผลิตฮาร์ดดิสก์ โดยเมื่อจัดเก็บข้อมูลปริมาณอนุภาคฝุ่นละอองจากระบบดังกล่าวแล้วจึงนำข้อมูลมาปรับปรุงเพื่อให้สอดคล้องกับการวิเคราะห์หาปัจจัยที่ทำให้เกิดปริมาณอนุภาคฝุ่นละอองเกินควบคุม

3.1.1 ลักษณะข้อมูลเบื้องต้นของระบบเฝ้าระวังและตรวจจับฝุ่นละออง

ข้อมูลปริมาณอนุภาคฝุ่นละอองในห้องคลีนรูมนั้นนำมาจากฐานข้อมูลบันทึกปริมาณฝุ่นละอองจากระบบเฝ้าระวังฝุ่นละอองหรือ Particle Monitoring System (PMS) ซึ่งเป็นโปรแกรมที่ใช้งานร่วมกับอุปกรณ์ที่ตรวจจับฝุ่นละอองในห้องคลีนรูมภายในโรงงานอุตสาหกรรมผลิตฮาร์ดดิสก์ไดร์ฟ ซึ่งบันทึกไว้เมื่อวันที่ 1 มีนาคม พ.ศ. 2553 ถึง วันที่ 31 มีนาคม พ.ศ. 2553 จำนวน 374,884 ข้อมูล โดยได้รับข้อมูลมาในรูปแบบของ Comma-Separated Value (*.csv) ดังรูปที่ 3.1



"Sample Date","Sample Time","SensorName","PartName","Sensor Type","Sample Value","Units","Flow Rate","Sample Interval","Sample Volume","DC Light Level","Laser Status","Number Channels","Micron Size 0","Counts/Channel 0","Micron Size 1","Counts/Channel 1","Micron Size 2","Counts/Channel 2","Micron Size 3","Counts/Channel 3","Micron Size 4","Counts/Channel 4","Micron Size 5","Counts/Channel 5","Micron Size 6","Counts/Channel 6","Micron Size 7","Counts/Channel 7","Micron Size 8","Counts/Channel 8","Micron Size 9","Counts/Channel 9","Micron Size 10","Counts/Channel 10","Micron Size 11","Counts/Channel 11","Micron Size 12","Counts/Channel 12","Micron Size 13","Counts/Channel 13","Micron Size 14","Counts/Channel 14","Micron Size 15","Counts/Channel 15","Micron Size 16","Counts/Channel 16","Micron Size 17","Counts/Channel 17","Micron Size 18","Counts/Channel 18","Micron Size 19","Counts/Channel 19","Micron Size 20","Counts/Channel 20","Micron Size 21","Counts/Channel 21","Micron Size 22","Counts/Channel 22","Micron Size 23","Counts/Channel 23","Micron Size 24","Counts/Channel 24","Micron Size 25","Counts/Channel 25","Micron Size 26","Counts/Channel 26","Micron Size 27","Counts/Channel 27","Micron Size 28","Counts/Channel 28","Micron Size 29","Counts/Channel 29","Micron Size 30","Counts/Channel 30","Micron Size 31","Counts/Channel 31"
"03/01/10","00:01:01","Alarm HSA PH1","EA Lan","E",0,"Switch","",60.00
"03/01/10","00:02:01","Alarm HSA PH1","EA Lan","E",0,"Switch","",60.00
"03/01/10","00:02:27","M4B-Ramp mnt","AirNet-301","F",1.000,"cf",0.100,600.00,1.000,0,1,2,0.300,1.000,0.500,0
"03/01/10","00:03:01","Alarm HSA PH1","EA Lan","E",0,"Switch","",60.00
"03/01/10","00:03:03","M1B-Ramp mnt","AirNet-301","F",0,"cf",0.100,600.00,1.000,0,1,2,0.300,0,0.500,0
"03/01/10","00:03:16","H2-Final inspec","AirNet-301","F",1.000,"cf",0.100,600.02,1.000,0,1,2,0.300,0,0.500,1.000
"03/01/10","00:03:16","H9-Final inspec","AirNet-301","F",0,"cf",0.100,600.02,1.000,0,1,2,0.300,0,0.500,0
"03/01/10","00:03:16","M2A-Environment","AirNet-301","F",0,"cf",0.100,600.02,1.000,0,1,2,0.300,0,0.500,0
"03/01/10","00:03:16","M3A-Disk mnt","AirNet-301","F",0,"cf",0.100,600.02,1.000,0,1,2,0.300,0,0.500,0
"03/01/10","00:03:17","M2B-Top Clamp","AirNet-301","F",0,"cf",0.100,600.02,1.000,0,1,2,0.300,0,0.500,0
"03/01/10","00:03:17","M5A-HSA VCM","AirNet-301","F",0,"cf",0.100,600.02,1.000,0,1,2,0.300,0,0.500,0
"03/01/10","00:03:18","M4A-Environment","AirNet-301","F",0,"cf",0.100,600.02,1.000,0,1,2,0.300,0,0.500,0
"03/01/10","00:03:18","M3A-HSA VCM","AirNet-301","F",0,"cf",0.100,600.02,1.000,0,1,2,0.300,0,0.500,0
"03/01/10","00:03:19","M5A-Disk mnt","AirNet-301","F",0,"cf",0.100,600.02,1.000,0,1,2,0.300,0,0.500,0
"03/01/10","00:03:19","M3B-HSA VCM","AirNet-301","F",0,"cf",0.100,600.02,1.000,0,1,2,0.300,0,0.500,0
"03/01/10","00:03:19","LPP25-line05-PC2","AirNet-301","F",0,"cf",0.100,600.02,1.000,0,1,2,0.300,0,0.500,0
"03/01/10","00:03:20","M1B-Top clamp","AirNet-301","F",0,"cf",0.100,600.02,1.000,0,0,2,0.300,0,0.500,0
"03/01/10","00:03:20","SUB-HSA Pivot for M4","AirNet-301","F",0,"cf",0.100,600.02,1.000,0,1,2,0.300,0,0.500,0
"03/01/10","00:03:20","SUB-HSA Pivot for M2","AirNet-301","F",0,"cf",0.100,600.02,1.000,0,1,2,0.300,0,0.500,0
"03/01/10","00:03:20","M2A-Top Clamp","AirNet-301","F",0,"cf",0.100,600.02,1.000,0,1,2,0.300,0,0.500,0

รูปภาพที่ 3.1 ลักษณะข้อมูลปริมาณอนุภาคฝุ่นละอองจากระบบเฝ้าระวังอนุภาคฝุ่นละออง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยในการวิจัยจะคัดเลือกเฉพาะคอลัมน์ที่ต้องนำมาใช้ในวิทยานิพนธ์นี้ ซึ่งจะนำข้อมูลที่เป็นค่าคงที่ออกไปโดยไม่นำข้อมูลดังกล่าวมาพิจารณาและนำมาจัดเรียงข้อมูลเพื่อให้เหมาะสมกับการวิเคราะห์ข้อมูลปริมาณอนุภาคฝุ่นละอองเพื่อหาปัจจัยเสี่ยงที่อาจทำให้ปริมาณอนุภาคฝุ่นละอองเกินการควบคุม โดยจัดรูปแบบของวันที่, เวลา กระบวนการและพื้นที่การทำงานเป็นกลุ่มและการจัดเรียงลำดับข้อมูลเพื่อการวิเคราะห์เรียงตามวัน เวลา และเซ็นเซอร์

ข้อมูลปริมาณอนุภาคฝุ่นละอองที่ได้เก็บรวบรวมจากระบบงานเฝ้าระวังอนุภาคฝุ่นละอองประกอบด้วยข้อมูลดังตารางที่ 3.1

ตารางที่3.1 ตารางอธิบายข้อมูลจากระบบงานเฝ้าระวังฝุ่นละออง

ชื่อคอลัมน์	ความหมายของข้อมูล
Sample Date	วันที่มีการเก็บข้อมูลตัวอย่างจากอุปกรณ์
Sample Time	เวลาที่มีการเก็บข้อมูลตัวอย่างจากอุปกรณ์
Sensor Name	ชื่อของเซ็นเซอร์ที่จับฝุ่นละออง
Part Name	ชื่อยี่ห้อและรุ่นของอุปกรณ์ในการตรวจจับฝุ่นละออง ได้แก่ AirNet-301
Sensor Type	ประเภทของตัวเซ็นเซอร์
Sample Value	จำนวนอนุภาคฝุ่นละอองที่พบในอากาศปริมาตร 1 ลูกบาศก์ฟุต
Units	หน่วยที่ใช้ในการวัดอัตราการไหลของปริมาณลม
Flow Rate	ค่าปริมาณการไหลของอากาศที่เหมาะสมกับเซ็นเซอร์
Sample Interval	เวลาที่ใช้ในการเก็บอากาศที่ผ่านเซ็นเซอร์เอาไว้ในท่อ (หน่วยเป็นวินาที)
Sample Volume	ปริมาณตัวอย่างที่ใช้ในการทำให้ข้อมูลอยู่ในรูปแบบมาตรฐาน
DC Light Level	ระดับแรงดันไฟฟ้ากระแสตรงในระหว่างที่มีอากาศผ่านเซ็นเซอร์
Laser Status	ค่าแสดงสถานะของเลเซอร์
Number Channels	จำนวนของ channel ในเซ็นเซอร์ 1 ตัว ที่เปิดใช้งาน
Micron Size 0	ขนาดของอนุภาคฝุ่นละอองที่ผ่านใน channel 0 มีขนาด 0.3 ไมครอน
Counts/Channel 0	จำนวนของอนุภาคที่ผ่าน channel 0 โดยต้องมีขนาด 0.3 – 0.49 ไมครอน
Micron Size 1	ขนาดของอนุภาคที่ผ่านใน channel 1 มีขนาด 0.5 ไมครอน
Counts/Channel 1	จำนวนของอนุภาคที่ผ่าน channel 1 โดยต้องมีขนาด 0.5 – 0.99 ไมครอน

3.1.2 การเตรียมข้อมูลสำหรับการวิเคราะห์ข้อมูล

หลังจากการรวบรวมและศึกษาความหมายของข้อมูลแล้วจะทำการเตรียมข้อมูลสำหรับการวิเคราะห์ โดยมีขั้นตอนดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.2.1 การคัดเลือกข้อมูล (Data Selection)

การคัดเลือกข้อมูลที่จะนำมาวิเคราะห์นั้นมีเป้าหมายคือ ปัจจัยที่ทำให้เกิดปริมาณอนุภาคฝุ่นละอองเกินควบคุม ซึ่งจะไม่คัดเลือกข้อมูลที่ทุกแถวมีค่าเดียวกันได้แก่ Part Name, Sensor Type, Units, Flow Rate, Sample Interval, Sample Volume, DC Light Level, Laser Status, Number Channels, Micron Size 0, Counts/Channel 0, Micron Size 1 และ Counts/Channel 1 โดยข้อมูลที่คัดเลือกได้แก่

- ข้อมูลที่บ่งปริมาณ (Quantitative data)
  - ข้อมูลปริมาณอนุภาคฝุ่นละออง (Sample Value)
  - ข้อมูลวันที่ (Sample date)
  - ข้อมูลเวลา (Sample time)
- ข้อมูลที่แบ่งตามประเภท (Categorical data)
  - ข้อมูลเซ็นเซอร์ (Sensor name) เช่น H2-Final inspec, M2A-Top clamp เป็นต้น

3.1.2.2 การจัดระเบียบข้อมูล (Data organization)

เนื่องจากข้อมูลที่นำมาวิเคราะห์นั้นเดิมมีการกระจายของข้อมูลอย่างไม่เป็นระเบียบ ดังนั้นจึงต้องมีการจัดเรียงข้อมูลโดย ทำการเรียงข้อมูลตาม เซ็นเซอร์, วัน และเวลา ตามลำดับ

Sample Da	Sample T	Sensor Name	Sample Value
3/1/2010	0:01:01	Alarm HSA PH1	0
3/1/2010	0:02:01	Alarm HSA PH1	0
3/1/2010	0:02:27	M4B-Ramp mnt	1
3/1/2010	0:03:01	Alarm HSA PH1	0
3/1/2010	0:03:08	M1B-Ramp mnt	0
3/1/2010	0:03:16	H2-Final inspec	1
3/1/2010	0:03:16	H9-Final inspec	0
3/1/2010	0:03:16	M2A-Environment	0
3/1/2010	0:03:16	M3A-Disk mnt	0
3/1/2010	0:03:17	M2B-Top Clamp	0
3/1/2010	0:03:17	M5A-HSA VCM	0
3/1/2010	0:03:18	M4A-Environment	0
3/1/2010	0:03:18	M3A-HSA VCM	0
3/1/2010	0:03:19	M5A-Disk mnt	0
3/1/2010	0:03:19	M3B-HSA VCM	0
3/1/2010	0:03:19	LPP25-line05-PC2	0
3/1/2010	0:03:20	M1B-Top clamp	0
3/1/2010	0:03:20	SUB-HSA Pivot for M	0
3/1/2010	0:03:20	SUB-HSA Pivot for M	0
3/1/2010	0:03:20	M2A-Top Clamp	0

Sample Date	Sample Time	Sensor Name	Sample Value
3/1/2010	0:03:29	Class 100 wipe	0
3/1/2010	0:13:29	Class 100 wipe	1
3/1/2010	0:23:29	Class 100 wipe	0
3/1/2010	0:33:29	Class 100 wipe	0
3/1/2010	0:43:28	Class 100 wipe	0
3/1/2010	0:53:28	Class 100 wipe	0
3/1/2010	1:03:29	Class 100 wipe	0
3/1/2010	1:13:28	Class 100 wipe	0
3/1/2010	1:23:29	Class 100 wipe	0
3/1/2010	1:33:29	Class 100 wipe	0
3/1/2010	1:43:28	Class 100 wipe	0
3/1/2010	1:53:28	Class 100 wipe	0
3/1/2010	2:03:29	Class 100 wipe	0
3/1/2010	2:13:29	Class 100 wipe	0
3/1/2010	2:23:29	Class 100 wipe	0
3/1/2010	2:33:29	Class 100 wipe	0
3/1/2010	2:43:28	Class 100 wipe	2
3/1/2010	2:53:28	Class 100 wipe	0
3/1/2010	3:03:29	Class 100 wipe	0
3/1/2010	3:13:29	Class 100 wipe	0
3/1/2010	3:23:29	Class 100 wipe	0
3/1/2010	3:33:29	Class 100 wipe	0
3/1/2010	3:43:28	Class 100 wipe	0

รูปภาพที่ 3.2 ตัวอย่างการจัดระเบียบข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.2.3 การทำความสะอาดข้อมูล (Data Cleaning)

เนื่องจากข้อมูลที่ได้มาเป็นข้อมูลที่อาจมีข้อมูลที่ไม่ถูกต้อง ซึ่งมีส่วนที่ไม่สามารถนำไปใช้ในการจำแนก (Classification) ได้จึงต้องมีการกลั่นกรองข้อมูลก่อน เช่นข้อมูลที่ขาดหาย (Missing Value) ข้อมูลชื่อเซ็นเซอร์ ข้อมูลวันที่ ข้อมูลเวลา

1. ข้อมูลที่ขาดหายไป (Missing Value) ได้ทำการตัดข้อมูลทั้งแถวที่ค่าในบางคอลัมน์ขาดหายไป เนื่องจากปริมาณข้อมูลที่ขาดหายไปนั้นมีปริมาณน้อย
2. ข้อมูลที่มีสิ่งปลอมปน (Noisy Data) ได้ทำการแก้ไขข้อมูลให้ถูกต้อง โดยแก้ไขรูปแบบวันที่ในการบันทึกให้อยู่ในรูปแบบเดียวกัน เพื่อง่ายต่อการจำแนกข้อมูล

3.1.2.4 การแปลงรูปแบบข้อมูล (Data transformation)

เมื่อจัดเรียงข้อมูลเรียบร้อยแล้ว ต่อไปจึงทำการแปลงรูปแบบข้อมูลให้อยู่ในรูปแบบที่เหมาะสมสำหรับการวิเคราะห์ข้อมูลและการทำเหมืองข้อมูล โดยการจัดเรียงข้อมูลนั้นจะจัดเรียงข้อมูล ในมิติที่สนใจศึกษาได้แก่

- มิติของวันในสัปดาห์ (Day)

เนื่องจากข้อมูลดิบที่แสดงวันที่ทำงาน (Sample Date) นั้นแสดงอยู่ในลักษณะ M/D/YYYY ซึ่งจากการรวบรวมข้อมูลตัวอย่างของกรณีศึกษาจำนวน 31 วัน ซึ่งทำให้มีค่าข้อมูลที่ไม่ซ้ำกันจำนวน 31 ข้อมูล ดังนั้นในการแปลงข้อมูลมิติของวันในสัปดาห์จึงแปลงข้อมูล ได้เป็น 7 ข้อมูลได้แก่

1) วันจันทร์ (1Mon)
2) วันอังคาร(2Tue)
3) วันพุธ (3Wed)
4) วันพฤหัสบดี (4Thu)
5) วันศุกร์ (5Fri)
6) วันเสาร์ (6Sat)
7) วันอาทิตย์ (7Sun)

- มิติของกะการทำงาน (Shift)

เนื่องจากข้อมูลดิบที่แสดงเวลาการทำงาน (Sample Time) นั้นแสดงอยู่ในลักษณะ HH:MM:SS หมายถึงเวลาจริงของการเก็บข้อมูลซึ่งข้อมูลดังกล่าวไม่เหมาะสมสำหรับการวิเคราะห์ ดังนั้นในการแปลงข้อมูลมิติของกะการทำงานจึงสามารถแปลงได้ดังนี้

1) กะที่หนึ่ง (1st shift) เริ่มตั้งแต่เวลา 08.00 – 15.59 น.
2) กะที่สอง (2nd shift) เริ่มตั้งแต่เวลา 16.00 – 23.59 น.
3) กะที่สาม (3rd shift) เริ่มตั้งแต่เวลา 00.00 – 07.59 น.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- มิติของพื้นที่การทำงาน (Area)

ตารางที่3.2 ตารางการแปลงข้อมูลตามมิติพื้นที่การทำงาน

พื้นที่ทำงาน (Area)	เซ็นเซอร์การทำงาน (Sensor Name)		
Cleaning	Cleaning Environment Ph1		
HDE	M4B-Ramp mnt	M1B-Ramp mnt	M15A-Ramp mnt
	M2A-Environment	M3A-Disk mnt	M1B-HSA VCM
	M2B-Top Clamp	M5A-HSA VCM	M3A-Ramp mnt
	M4A-Environment	M3A-HSA VCM	M4A-Hyper DB
	M5A-Disk mnt	M3B-HSA VCM	HDE Environment PH1
	LPP25-line05-PC2	M1B-Top clamp	M5B-Ramp mnt
	M2A-Top Clamp	M1A-Environment	SUB-Parts loading
	M5B-Top Clamp	M15A-Disk mnt	M4B-HSA VCM
	M15A-HSA VCM	M3B-Disk mnt	M4B-Top clamp
	M2A-HSA VCM	M2B-Disk mnt	M4A-Top clamp
	M15A-Top Clamp	M1B-Disk mnt	M2A-Hyper DB
	M4A-HSA VCM	M2A-Disk mnt	M4A-Disk mnt
	M2B-HSA VCM	LPP25-line01-PC3	M2B-Ramp mnt
	M4B-Disk mnt	M5B-Disk mnt	M4A-Ramp mnt
	M5A-Top Clamp	M15A-Top Cover	M2A-Ramp mnt
	Class 100 wipe	LPP25-line03-PC2	LPP25-line04-PC2
	M3A-Top Clamp	M3B-Top clamp	M3B-Ramp mnt
	M3A-Environment	M5A-Hyper DB	M3A-Hyper DB
	M5A-Ramp mnt	M5B-HSA VCM	LPP25-line02-PC4
	FA_Lab Environment	TEST PMS REPAIR 233.140	
HSA	H2-Final inspec	H9-Final inspec	SUB-HSA Pivot for M4
	SUB-HSA Pivot for M2	H1-Final inspec	H4-Final inspec
	H14-Final inspec	H12-Final inspec	H16-Final inspec
	H5-HSA stack	SUB-HSA Pivot for M3	HSA55-Final inspec
	H11-Final inspec	HSA-Environment PH1	HSA52-Final inspec
	HSA51-Final inspec	HSA50-Final Inspec	H5-Final inspec
	H5-Touch up	H6-Final inspec	H8-Final inspec
	H7-Final inspec	H3-Final inspec	HSA-Teardown29
	H15-Final inspec		

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตารางที่ 3.2 เนื่องจากข้อมูลดิบที่แสดงเซ็นเซอร์ที่จับอนุภาคฝุ่นละออง (Sensor Name) นั้นมีจำนวนมากจึงได้แปลงข้อมูลดังกล่าวให้อยู่ในมิติของพื้นที่การทำงาน (Area) ได้ดังนี้

1) Cleaning (Cleaning Process)
2) HDE (Hard Disk Enclosure Process)
3) HSA (Head stack assembly Process)

- มิติของกระบวนการทำงาน (Process)

เนื่องจากข้อมูลดิบที่แสดงเซ็นเซอร์ที่จับอนุภาคฝุ่นละออง (Sensor Name) นั้นมีจำนวนมากจึงได้แปลงข้อมูลดังกล่าวให้อยู่ในมิติของกระบวนการทำงานได้ดังนี้

1) Wipe เป็นกระบวนการเก็บชิ้นส่วนที่ทำความสะอาดแล้ว
2) Final inspection เป็นกระบวนการตรวจสอบขั้นสุดท้าย
3) HSA assembly เป็นกระบวนการบรรจุ HSA
4) Solder inspection เป็นกระบวนการตรวจสอบรอยบัดกรี
5) Tear down เป็นกระบวนการทำงาน re-work
6) Particle test เป็นกระบวนการตรวจสอบหาฝุ่นละอองของฮาร์ดดิสก์
7) Disk mounting เป็นกระบวนการบรรจุดิสก์
8) HSA/VCM mounting เป็นกระบวนการบรรจุ HSA หรือ VCM
9) Cleaning เป็นกระบวนการล้างทำความสะอาดชิ้นส่วน
10) Ramp mounting เป็นกระบวนการบรรจุหัวเขียนอ่านดิสก์
11) Top clamp mounting เป็นกระบวนการบรรจุชิ้นส่วนสำหรับยึดแผ่นดิสก์กับมอเตอร์
12) Top cover mounting เป็นกระบวนการปิดฮาร์ดดิสก์เพื่อป้องกันวัสดุจากภายนอก
13) Hyper disk balancer เป็นกระบวนการทดสอบความสมดุลของชุดแผ่นดิสก์
14) Pivot เป็นกระบวนการขันสกรูหรือแกนต่างๆ
15) Part loading conveyer เป็นกระบวนการเคลื่อนย้ายชิ้นส่วนต่างๆ จากห้องล้างเข้าสู่พื้นที่ HDE
16) Environment เป็นส่วนการตรวจจับสภาพแวดล้อมภายในพื้นที่การทำงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการเตรียมข้อมูลในขั้นตอน 3.1.2.1 – 3.1.2.4 นั้นทำให้มีข้อมูลใหม่ที่ใช้สำหรับการควบคุมกระบวนการเชิงสถิติเพื่อหาสถานะระหว่างการควบคุมคุณภาพ ได้แก่

1) วันที่เก็บข้อมูล (Day)
2) กะการทำงาน (Shift)
3) กระบวนการทำงาน (Process)
4) ปริมาณอนุภาคฝุ่นละออง (Particle)

Day	Shift	Process	Particle
1Mon	3rd shift	Cleaning	0
1Mon	3rd shift	Cleaning	2
1Mon	3rd shift	Cleaning	2
1Mon	3rd shift	Cleaning	8
1Mon	3rd shift	Cleaning	9
1Mon	3rd shift	Cleaning	49
1Mon	3rd shift	Cleaning	1386
1Mon	1st shift	Cleaning	2098
1Mon	1st shift	Cleaning	3005
1Mon	1st shift	Cleaning	5024
1Mon	1st shift	Cleaning	3670
1Mon	1st shift	Cleaning	2447
1Mon	1st shift	Cleaning	4175
1Mon	1st shift	Cleaning	2341
1Mon	1st shift	Cleaning	6641
1Mon	2nd shift	Cleaning	3507
1Mon	2nd shift	Cleaning	2566
1Mon	2nd shift	Cleaning	3503
1Mon	2nd shift	Cleaning	2189
1Mon	2nd shift	Cleaning	5084
1Mon	2nd shift	Cleaning	7200
1Mon	2nd shift	Cleaning	8422

รูปภาพที่ 3.3 ตัวอย่างข้อมูลตรวจจับอนุภาคฝุ่นละอองหลังการเตรียมข้อมูล

3.1.3 การสำรวจข้อมูลเบื้องต้น

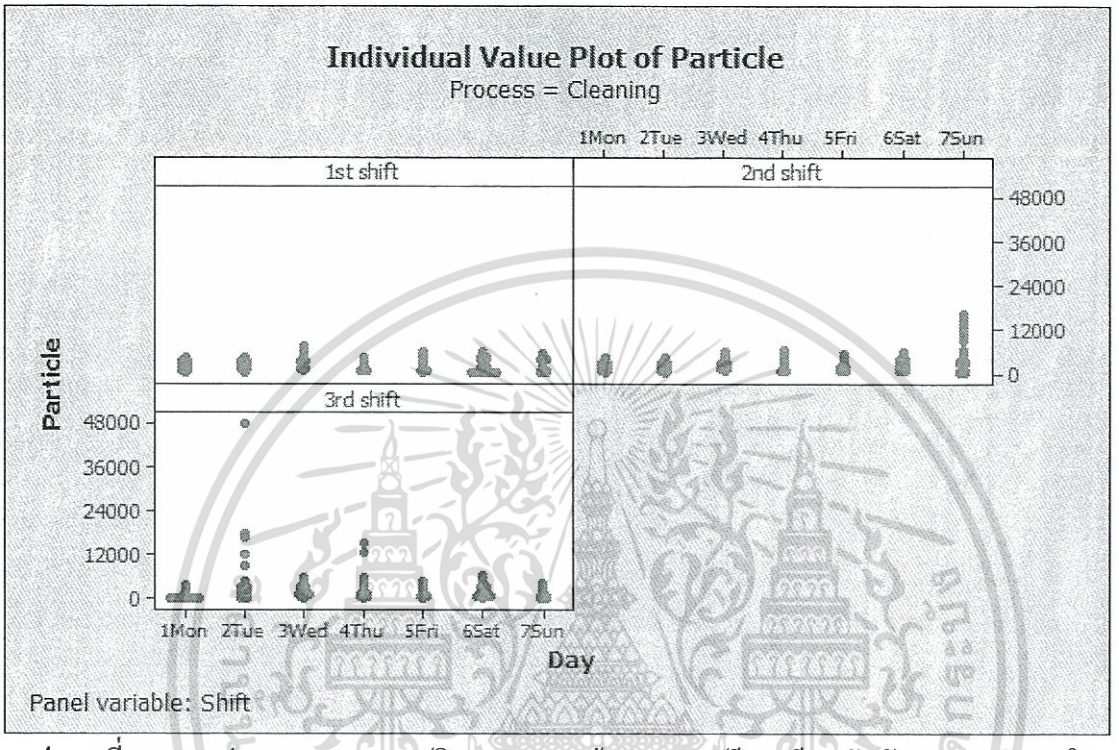
การสำรวจข้อมูลเบื้องต้นจะทำโดยนำข้อมูลปริมาณอนุภาคฝุ่นละอองมาเปรียบเทียบดูปริมาณที่เกิดขึ้นหรือการกระจายของข้อมูลในปัจจัยต่างๆ ได้แก่ วันทำงาน (Day), กะการทำงาน (Shift), พื้นที่การทำงาน (Area) และ กระบวนการทำงาน (Process) เพื่อสำรวจดูแนวโน้มเบื้องต้น ดังนี้

3.1.3.1 การสำรวจข้อมูลข้อมูลพื้นที่การทำงาน cleaning

- กราฟการกระจายของปริมาณอนุภาคฝุ่นละออง เมื่อวันเทียบกับกระบวนการทำงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เนื่องจากพื้นที่การทำงาน Cleaning มีกระบวนการทำงานเดียวคือ cleaning ซึ่งเมื่อสำรวจข้อมูลเบื้องต้น พบว่าปริมาณอนุภาคฝุ่นละอองที่พบในพื้นที่ทำงานดังกล่าวอาจมีปริมาณที่เพิ่มสูงในช่วงกะการทำงานที่ 3 ซึ่งเป็นช่วงเวลากลางคืนและช่วงกะที่ 2 ของวันอาทิตย์ สามารถแสดงกราฟการกระจายได้ดังรูปที่ 3.4



รูปภาพที่ 3.4 กราฟการกระจายของปริมาณอนุภาคฝุ่นละอองเปรียบเทียบวันกับกะการทำงานในพื้นที่การทำงาน cleaning

3.1.3.2 การสำรวจข้อมูลข้อมูลพื้นที่การทำงาน HDE

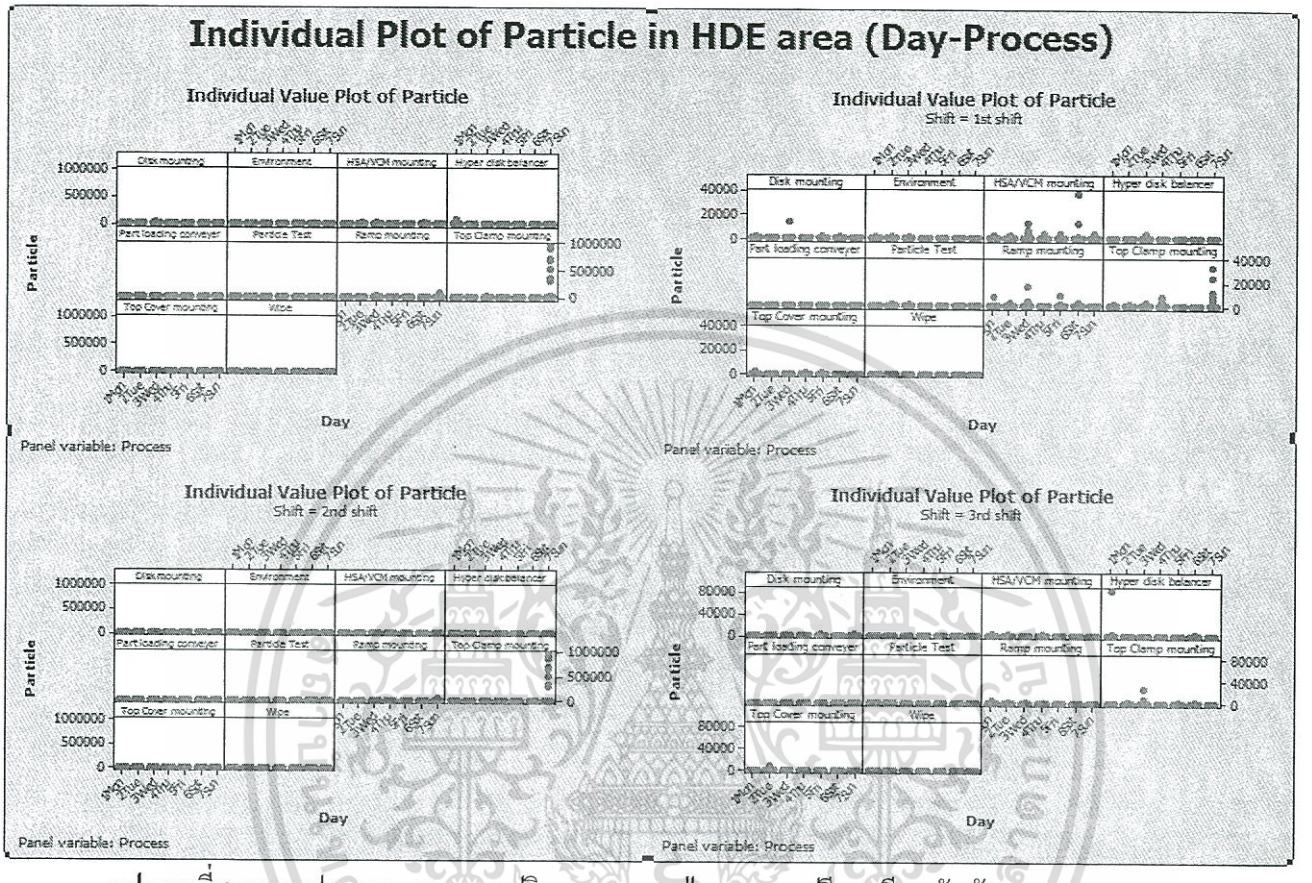
- กราฟการกระจายของปริมาณอนุภาคฝุ่นละออง เมื่อวันเทียบกับกระบวนการทำงาน

เมื่อเปรียบเทียบการกระจายของปริมาณอนุภาคฝุ่นละอองของวันต่างๆ ในแต่ละกระบวนการทำงาน พบว่า กระบวนการ Top Clamp mounting มีแนวโน้มที่อาจเกิดปริมาณอนุภาคฝุ่นละอองเกินควบคุม และเมื่อพิจารณาแยกเป็นกะการทำงานแต่ละกะการทำงาน พบว่า

1) กะการทำงานที่ 1 มีการกระจายของข้อมูลมากในกระบวนการ Disk mounting ในวันพุธ, กระบวนการ HSA/VCM mounting ในวันพุธและวันเสาร์, กระบวนการ Ramp mounting ในวันจันทร์ วันพุธและวันศุกร์, กระบวนการ Top clamp mounting ในวันศุกร์ ส่วนวันอื่นๆ นั้นการกระจายของข้อมูลเกิดในระดับต่ำ
2) กะการทำงานที่ 2 มีการกระจายของข้อมูลมากในกระบวนการ Top clamp mounting ในวันศุกร์ ส่วนวันอื่นๆ นั้นการกระจายของข้อมูลเกิดในระดับต่ำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) กะการทำงานที่ 3 มีการกระจายของข้อมูลมากในกระบวนการ Hyper disk balancer ในวันจันทร์, กระบวนการ Top clamp mounting ในวันพุธ ส่วนวันอื่นๆ นั้นการกระจายของข้อมูลเกิดในระดับต่ำ ซึ่งกราฟการกระจายทั้งหมดแสดงได้ดังรูปที่ 3.5

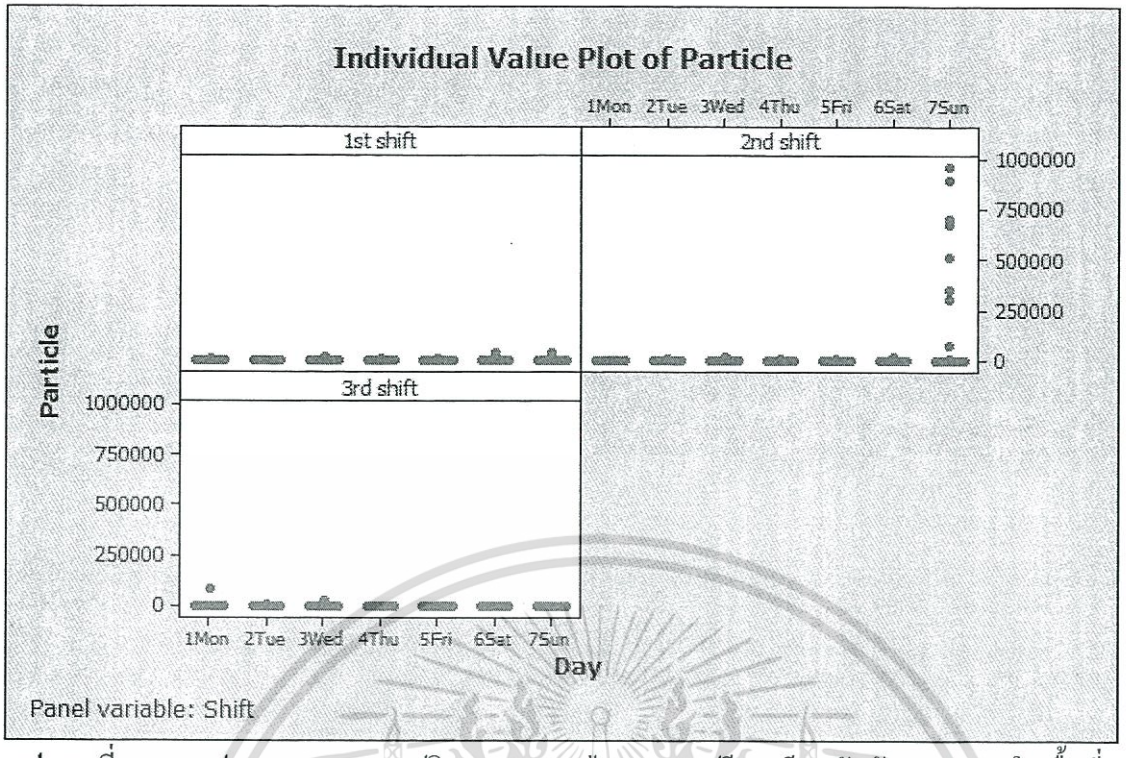


รูปภาพที่ 3.5 กราฟการกระจายของปริมาณอนุภาคฝุ่นละอองเปรียบเทียบวันกับกระบวนการทำงานในพื้นที่การทำงาน HDE

- กราฟการกระจายของปริมาณอนุภาคฝุ่นละออง เมื่อวันเทียบกับกะการทำงาน

เมื่อเปรียบเทียบการกระจายของปริมาณอนุภาคฝุ่นละอองของวันต่างๆ ในแต่ละกะการทำงาน พบว่า กะการทำงานที่ 2 ของวันอาทิตย์มีแนวโน้มที่อาจเกิดปริมาณอนุภาคฝุ่นละอองเกินควบคุมเนื่องจากมีการกระจายของข้อมูลมาก ส่วนในกะการทำงานที่ 3 ของวันจันทร์อาจเกิดปริมาณอนุภาคฝุ่นละอองเกินกำหนดได้เช่นกัน เนื่องจากมีการกระจายของข้อมูลที่ไม่เกาะกลุ่ม ซึ่งสามารถแสดงกราฟการกระจายของปริมาณฝุ่นละอองของกะต่างๆ ในแต่ละวันได้ดังรูปที่ 3.6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปภาพที่ 3.6 กราฟการกระจายของปริมาณอนุภาคฝุ่นละอองเปรียบเทียบวันกับกะทำงานในพื้นที่การทำงาน HDE

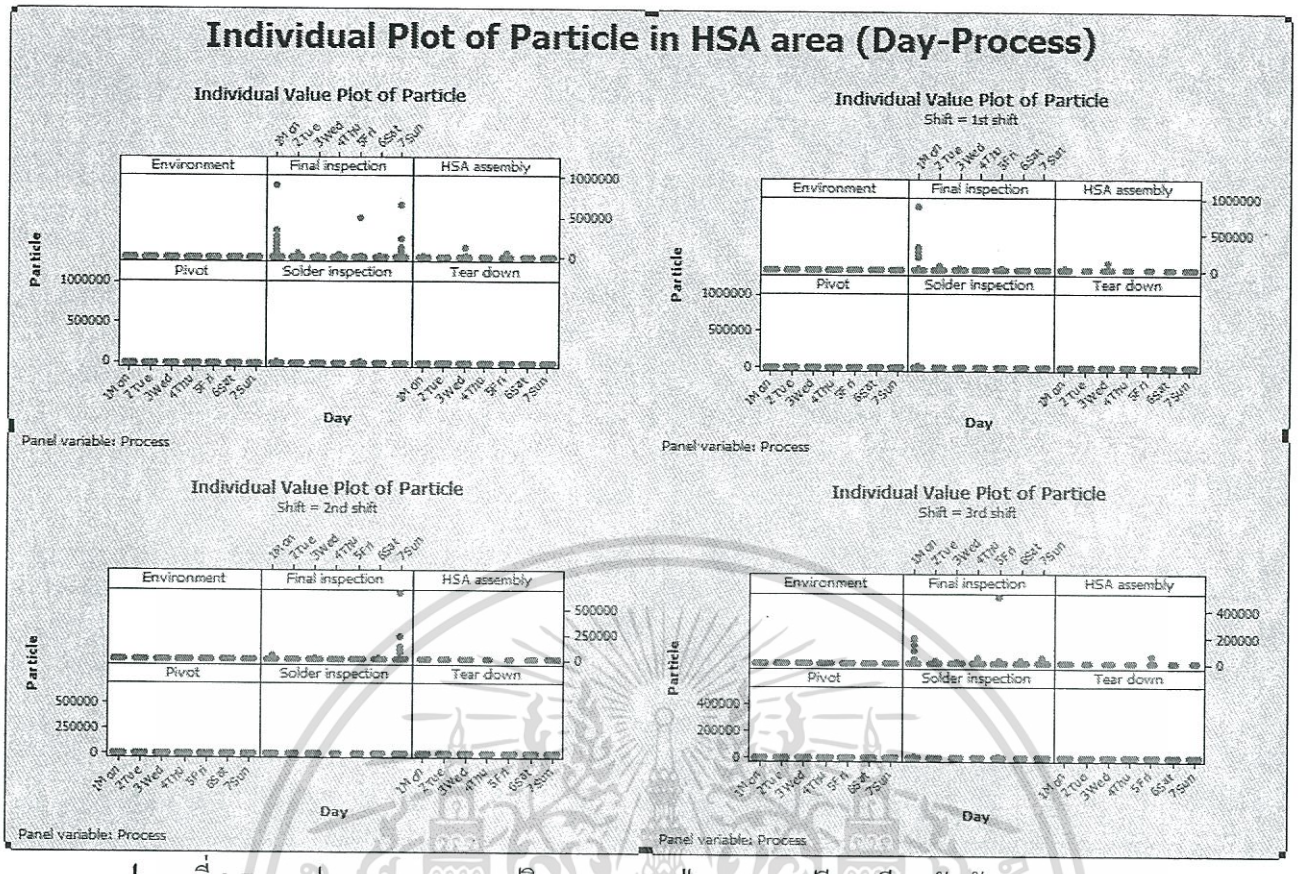
3.1.3.3 การสำรวจข้อมูลข้อมูลพื้นที่การทำงาน HSA

- กราฟการกระจายของปริมาณอนุภาคฝุ่นละออง เมื่อวันเทียบกับกระบวนการทำงาน

เมื่อเปรียบเทียบการกระจายของปริมาณอนุภาคฝุ่นละอองของวันต่างๆ ในแต่ละกระบวนการทำงาน พบว่า กระบวนการ Final inspection มีแนวโน้มที่อาจเกิดปริมาณอนุภาคฝุ่นละอองเกินควบคุมในวันจันทร์, วันศุกร์และวันอาทิตย์ เมื่อพิจารณาแยกเป็นกะการทำงานแต่ละกะการทำงาน พบว่า

1) กะการทำงานที่ 1 มีการกระจายของข้อมูลมากในกระบวนการ Final inspection ในวันจันทร์, กระบวนการ HSA assembly ในวันพุธ ส่วนวันอื่นๆ นั้นการกระจายของข้อมูลเกิดในระดับต่ำ
2) กะการทำงานที่ 2 มีการกระจายของข้อมูลมากในกระบวนการ Final inspection ในวันอาทิตย์ ส่วนวันอื่นๆ นั้นการกระจายของข้อมูลเกิดในระดับต่ำ
3) กะการทำงานที่ 3 มีการกระจายของข้อมูลมากในกระบวนการ Final inspection ในวันจันทร์และวันศุกร์, กระบวนการ HSA assembly ในวันศุกร์ ส่วนวันอื่นๆ นั้นการกระจายของข้อมูลเกิดในระดับต่ำ ซึ่งกราฟการกระจายทั้งหมดแสดงได้ดังรูปที่ 3.7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปภาพที่ 3.7 กราฟการกระจายของปริมาณอนุภาคฝุ่นละอองเปรียบเทียบวันกับกระบวนการทำงานในพื้นที่การทำงาน HSA

■ กราฟการกระจายของปริมาณอนุภาคฝุ่นละออง เมื่อวันเทียบกับกะการทำงาน

เมื่อเปรียบเทียบการกระจายของปริมาณอนุภาคฝุ่นละอองของวันต่างๆ ในแต่ละกะการทำงาน พบว่า

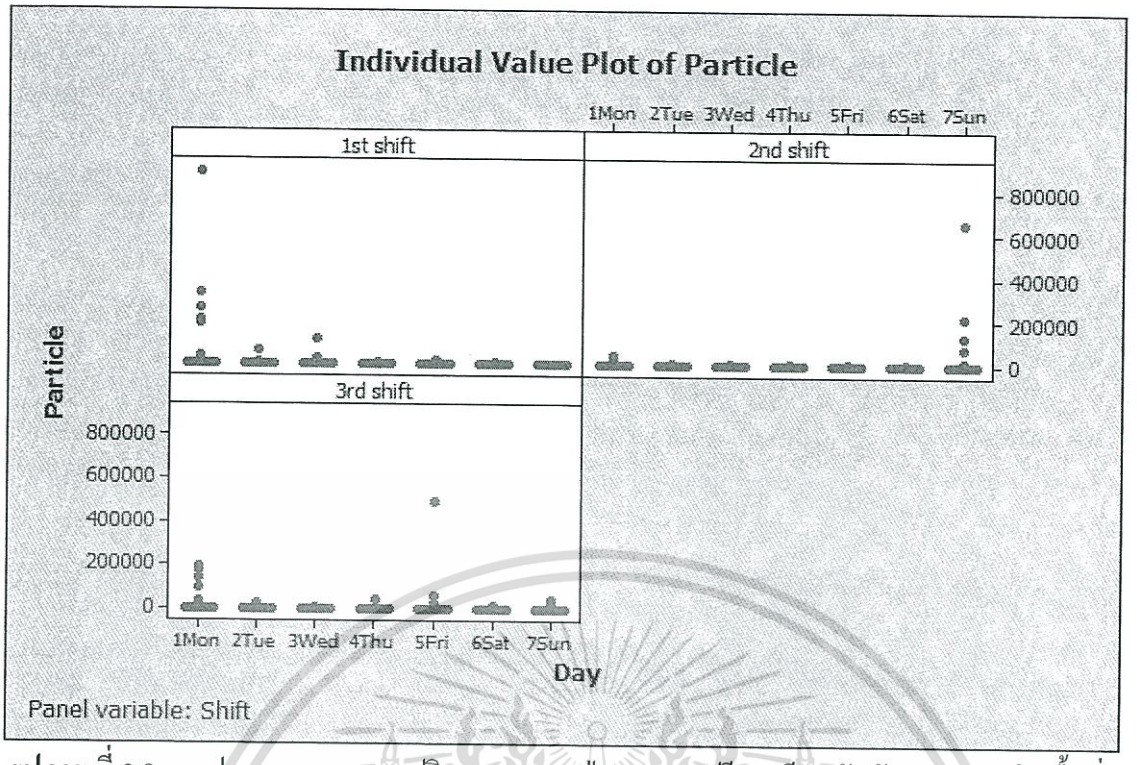
1) กะการทำงานที่ 1 ของวันจันทร์, วันอังคารและวันพุธ

2) กะการทำงานที่ 2 ของวันจันทร์และวันอาทิตย์

3) กะการทำงานที่ 3 ของวันจันทร์, วันพฤหัสบดี, วันศุกร์และวันอาทิตย์

มีแนวโน้มที่อาจเกิดปริมาณอนุภาคฝุ่นละอองเกินควบคุมเนื่องจากมีการกระจายของข้อมูลมาก เนื่องจากมีการกระจายของข้อมูลที่ไม่เกาะกลุ่ม ซึ่งสามารถแสดงกราฟการกระจายของปริมาณฝุ่นละอองของกะต่างๆ ในแต่ละวันได้ดังรูปที่ 3.8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปภาพที่ 3.8 กราฟการกระจายของปริมาณอนุภาคฝุ่นละอองเปรียบเทียบวันกับกะทำงานในพื้นที่การทำงาน HSA

3.2 การออกแบบมาตรฐานสำหรับการควบคุมกระบวนการเชิงสถิติของกรณีศึกษา

การออกแบบมาตรฐานสำหรับการควบคุมกระบวนการเชิงสถิติของปริมาณอนุภาคฝุ่นละออง เพื่อใช้สำหรับหามาตรฐานการควบคุมกระบวนของการผลิตฮาร์ดดิสก์ไดร์ฟอย่างเหมาะสม ประกอบด้วย 3.2.1 ชนิดของแผนภูมิควบคุมที่เหมาะสมสำหรับข้อมูลปริมาณอนุภาคฝุ่นละออง 3.2.2 มาตรฐานการควบคุมคุณภาพ

โดยวิทยานิพนธ์ฉบับนี้จะพิจารณาการควบคุมกระบวนการโดยมาตรฐานของโรงงานอุตสาหกรรมเดิมและมาตรฐานของการควบคุมกระบวนการเชิงสถิติจำนวน 8 มาตรฐาน เนื่องจากมาตรฐานการควบคุมเดิมเป็นมาตรฐานที่ถูกกำหนดขึ้นโดยไม่ได้อาศัยหลักการทางสถิติและยังมีการควบคุมในระดับเกินควบคุมเท่านั้น ไม่มีการควบคุมในระดับการควบคุมเตือนซึ่งสามารถช่วยให้ผู้ปฏิบัติงานสามารถดูแนวโน้มความผิดปกติล่วงหน้าและทำให้การป้องกันปัญหาได้ก่อนที่กระบวนการจะเกินควบคุม

3.2.1 ชนิดของแผนภูมิควบคุมที่เหมาะสมสำหรับข้อมูลปริมาณอนุภาคฝุ่นละออง

เนื่องจากแผนภูมิควบคุมที่ใช้ในการควบคุมกระบวนการทางคุณภาพนั้นมีอยู่หลายประเภท ซึ่งในการใช้งานแผนภูมิควบคุมให้มีประสิทธิภาพย่อมต้องมีการเลือกใช้แผนภูมิที่มีความเหมาะสมกับลักษณะของข้อมูล ซึ่งแผนภูมิควบคุมที่ใช้คือแผนภูมิควบคุมของชิวฮาร์ท (Shewhart) ซึ่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แผนภูมิควบคุมจะสร้างขึ้นโดยพิจารณาคุณลักษณะของตัวแปร โดยแบ่งเป็น 2 ประเภทใหญ่ๆ ได้แก่

3.2.1.1 แผนภูมิควบคุมตัวแปรเชิงปริมาณ (Control Chart for Variable) ได้แก่

- แผนภูมิควบคุมตัวแปรเชิงเดี่ยว (X-Chart or Individual Chart)
- แผนภูมิควบคุมเฉลี่ย ($\overline{X}$-Chart)
- แผนภูมิควบคุมพิสัย (R-Chart)
- แผนภูมิควบคุมการกระจาย (Sigma-Chart or S-Chart)

3.2.1.2 แผนภูมิควบคุมตัวแปรเชิงลักษณะ (Control Chart for Attributes) ได้แก่

- แผนภูมิควบคุมสัดส่วนของเสีย (P-Chart)
- แผนภูมิคุณภาพเสีย (nP-Chart)
- แผนภูมิข้อบกพร่องหรือรอยตำหนิ (C-Chart or U-Chart)

สำหรับวิทยานิพนธ์ฉบับนี้จากขั้นตอนการเตรียมข้อมูลดำเนินการเรียบร้อยแล้ว เมื่อถึงขั้นตอนการเลือกใช้แผนภูมิควบคุม โดยเมื่อพิจารณาข้อมูลปริมาณอนุภาคฝุ่นละอองที่ใช้นั้นในการวิเคราะห์นั้นพบว่า ข้อมูลเป็นข้อมูลตัวแปรเชิงปริมาณที่เก็บข้อมูลในลักษณะเชิงเดี่ยว ที่มีลักษณะการเก็บข้อมูลแบบต่อเนื่องซึ่งมีการเก็บข้อมูลในช่วงเวลาที่ห่างกันค่อนข้างคงที่ ในลักษณะข้อมูลโดดๆ ดังนั้นแผนภูมิที่เหมาะสมสำหรับการควบคุมคุณภาพได้แก่ แผนภูมิควบคุมตัวแปรเชิงเดี่ยว (X-Chart or Individual Chart) ซึ่งสามารถหาขอบเขตการควบคุมบนและขอบเขตการควบคุมล่างได้จากสมการ (3.1)

$$\begin{aligned} UCL_X &= \mu + 3\sigma_x' \\ CL_X &= \mu \\ LCL_X &= \mu - 3\sigma_x' \end{aligned} \tag{3.1}$$

สำหรับการกระจายของข้อมูลที่จะนำมาควบคุมนั้นไม่สามารถหาได้โดยตรงเพราะข้อมูลมีกลุ่มเดียวก็จะมีการกระจายค่าเดียว ณ เวลานั้นๆ จึงอาศัยค่าพิสัยเคลื่อนที่ (Moving Range) ในการควบคุมการกระจายได้จากสมการ (3.2)

$$\begin{aligned} UCL_{MR} &= D_4\overline{MR} \\ CL_{MR} &= \overline{MR} \\ LCL_{MR} &= D_3\overline{MR} \end{aligned} \tag{3.2}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.2 มาตรฐานการควบคุมคุณภาพ

สำหรับการควบคุมคุณภาพของกระบวนการผลิตฮาร์ดดิสก์ไดร์ฟนั้น เป็นการควบคุมเพื่อให้ผลิตภัณฑ์ที่ได้มีคุณภาพตามต้องการซึ่งในการควบคุมนั้นก็ต้องคอยควบคุมสิ่งที่ทำให้ผลิตภัณฑ์ไม่มีคุณภาพ ซึ่งสำหรับการผลิตฮาร์ดดิสก์นั้นได้แก่ ปริมาณอนุภาคฝุ่นละอองที่เกิดขึ้นในขณะการผลิตสินค้า โดยมีการควบคุมโดยการวัดปริมาณของอนุภาคฝุ่นละอองในอากาศภายในห้องควบคุม ซึ่งปริมาณที่วัดได้นั้นก็จะนำไปใช้เทียบกับปริมาณที่ใช้เป็นเกณฑ์ในการควบคุม ซึ่งเกณฑ์ที่ใช้ในการควบคุมจะแบ่งเป็น 2 ประเภทได้แก่ 1) เกณฑ์การควบคุมของโรงงานอุตสาหกรรม 2) เกณฑ์การควบคุมกระบวนการเชิงสถิติ

3.2.2.1 เกณฑ์การควบคุมของโรงงานอุตสาหกรรม (Industry standard) เป็นเกณฑ์การควบคุมปริมาณอนุภาคฝุ่นละอองที่ใช้ในโรงงานกรณีตัวอย่าง โดยกำหนดปริมาณอนุภาคฝุ่นละอองแตกต่างกันในแต่ละพื้นที่การทำงาน ซึ่งสามารถแสดงได้ดังตารางที่ 3.3

ตารางที่ 3.3 เกณฑ์การควบคุมปริมาณอนุภาคฝุ่นละอองโดยโรงงานอุตสาหกรรม

พื้นที่การทำงาน	ปริมาณควบคุม (อนุภาค)
HSA	20,000
HDE	100
Cleaning	500

3.2.2.2 เกณฑ์การควบคุมกระบวนการเชิงสถิติ เป็นเกณฑ์การควบคุมปริมาณอนุภาคฝุ่นละอองมาตรฐานที่วิทยานิพนธ์เลือกใช้ โดยจะใช้เกณฑ์การควบคุมดังกล่าวกับข้อมูลทุกพื้นที่การทำงาน ซึ่งประกอบด้วย

1) มี 1 จุดอยู่นอกเส้นควบคุม $\pm 3\sigma$
2) มี 9 จุดต่อเนื่องกันที่อยู่ด้านใดด้านหนึ่งของเส้นควบคุมกลาง (Central line)
3) มี 6 จุดต่อเนื่องกันที่มีแนวโน้มเพิ่มขึ้นหรือลดลงอย่างต่อเนื่อง
4) มี 14 จุดต่อเนื่องกันที่มีขึ้นลงสลับกัน
5) มี 2 ใน 3 จุดที่อยู่นอกเส้นควบคุมเตือน $\pm 2\sigma$ บนด้านเดียวกันของเส้นควบคุมกลาง
6) มี 4 ใน 5 จุดที่อยู่นอกเส้นควบคุมเตือน $\pm 1\sigma$ บนด้านเดียวกันของเส้นควบคุมกลาง
7) มี 15 จุดต่อเนื่องกันที่อยู่ในช่วง $\pm 1\sigma$
8) มี 8 จุดต่อเนื่องกันที่อยู่นอกเขตควบคุม $\pm 1\sigma$ จากเส้นควบคุมกลางของข้างใดข้างหนึ่ง

สำหรับวิทยานิพนธ์ฉบับนี้จะใช้แผนภูมิควบคุมและเกณฑ์มาตรฐานการควบคุมทั้ง 8 ข้อสำหรับการสร้างข้อมูลการควบคุมคุณภาพของปริมาณอนุภาคฝุ่นละอองเพื่อนำข้อมูลดังกล่าวไปใช้ในการค้นหาความรู้ที่ซ่อนอยู่จากการควบคุมคุณภาพด้วยเหมืองข้อมูลเทคนิคการจำแนกข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ด้วยต้นไม้ตัดสินใจ ซึ่งลักษณะข้อมูลที่ได้หลังจากการควบคุมคุณภาพด้วยแผนภูมิควบคุมมีลักษณะดังรูปที่ 3.9

Day	Shift	Process	Particle	Rule1	Rule2	Rule3	Rule4	Rule5	Rule6	Rule7	Rule8
1Mon	3rd shift	Wipe	0	normal	normal	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	1	normal	normal	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	0	normal	normal	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	0	normal	normal	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	0	normal	normal	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	0	normal	normal	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	0	normal	normal	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	0	normal	normal	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	0	normal	normal	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	0	normal	normal	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	0	normal	failed	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	0	normal	failed	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	0	normal	failed	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	0	normal	failed	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	0	normal	failed	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	0	normal	failed	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	2	failed	normal	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	0	normal	normal	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	0	normal	normal	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	0	normal	normal	normal	normal	normal	normal	normal	normal
1Mon	3rd shift	Wipe	0	normal	normal	normal	normal	normal	normal	normal	normal

รูปภาพที่ 3.9 ตัวอย่างข้อมูลที่ได้หลังจากการควบคุมคุณภาพด้วยแผนภูมิควบคุม

จากรูปที่ 3.9 เป็นข้อมูลหลังการควบคุมคุณภาพจากการควบคุมกระบวนการเชิงสถิติทั้ง 8 ข้อ ซึ่งประกอบด้วย ข้อมูลวัน, กะการทำงาน, กระบวนการทำงาน, ปริมาณอนุภาคฝุ่นละออง, ผลการควบคุมคุณภาพข้อที่ 1-8 ซึ่งผลที่ได้จากการควบคุมมี 2 แบบคือ ปริมาณอนุภาคฝุ่นละอองปกติ (normal) และปริมาณอนุภาคฝุ่นละอองเกินควบคุม (failed)

3.3 การหากฎของปัจจัยที่ทำให้ปริมาณอนุภาคฝุ่นละอองเกินควบคุมด้วยเหมืองข้อมูล

การออกแบบการทดลองสำหรับการหากฎของปัจจัยที่ทำให้ปริมาณอนุภาคฝุ่นละอองเกินควบคุมด้วยเหมืองข้อมูล ได้แก่การนำข้อมูลที่ได้ควบคุมคุณภาพเรียบร้อยแล้วตามข้อ 3.2 ไปหาความรู้ด้วยเทคนิคเหมืองข้อมูลด้วยแบบจำลองต้นไม้ตัดสินใจเพื่อให้ได้ปัจจัยที่มีผลต่อการเกิดปริมาณอนุภาคฝุ่นละอองที่เกินควบคุม โดยมีวิธีการดังต่อไปนี้

3.3.1 การแก้ปัญหาความไม่สมดุลของข้อมูล

เนื่องจากข้อมูลคลาสข้อมูลปริมาณฝุ่นละอองที่เกิดขึ้นอาจมีปริมาณน้อยในบางชุดข้อมูลที่จะทำการหาต้นแบบทำให้คลาสข้อมูลระหว่างคลาส Normal และ คลาส Failed ไม่สมดุลกัน ดังนั้นในข้อมูลบางข้อมูลจึงจำเป็นต้อง Sampling ข้อมูลก่อนโดยใช้วิธีการ SMOTE ตามที่ได้กล่าวไปใน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อ 2.3.1 ซึ่งจะสนใจคลาสส่วนน้อยเป็นหลัก โดยวิธีดังกล่าวนั้นมีประสิทธิภาพและไม่ทำให้เกิดปัญหา Over-fitting

3.3.2 การเลือกใช้เทคนิคเหมืองข้อมูล

การเลือกเทคนิคเหมืองข้อมูลที่เหมาะสมสำหรับการวิเคราะห์ข้อมูลนั้น งานวิจัยฉบับนี้ได้เลือกเทคนิคต้นไม้ตัดสินใจด้วยอัลกอริทึม C4.5 ซึ่งสามารถให้ผลลัพธ์เป็นต้นไม้ตัดสินใจ

จากการเตรียมข้อมูลทั้งหมดที่กล่าวมาแล้วนั้น สามารถสรุปข้อมูลนำเข้าสำหรับการจำแนกข้อมูลได้ดังตารางที่ 3.4 และชุดข้อมูลที่นำมาทดสอบนั้นแบ่งเป็น 3 ชุดตามมิติของพื้นที่การทำงาน

ตารางที่3.4 ปัจจัยนำเข้าของข้อมูลปริมาณอนุภาคฝุ่นละออง

ชื่อคุณสมบัติ	คำอธิบาย	ชนิดของข้อมูล
Day	วันที่ตรวจจับฝุ่นละออง	Categorical
Shift	กะการทำงาน	Categorical
Process	กระบวนการทำงาน	Categorical
Rule	ผลการควบคุมคุณภาพ	Class label

3.3.3 การแบ่งกลุ่มข้อมูลที่ใช้ทดสอบ

การแบ่งกลุ่มข้อมูลที่ใช้สำหรับการทดสอบนั้นใช้วิธี Cross-Validation เป็นขั้นตอนเพื่อใช้ในการแบ่งกลุ่มของข้อมูล โดยการแบ่งกลุ่มข้อมูลเพื่อทดสอบ 2 รูปแบบคือ

1) Hold-Out Cross Validation
2) k-Fold Cross-Validation

สำหรับงานวิจัยฉบับนี้เลือกวิธี K-Fold Cross-Validation สำหรับการแบ่งกลุ่มข้อมูลเพื่อทดสอบ ซึ่งกำหนดให้ k เท่ากับ 10 และสามารถแบ่งข้อมูลสำหรับการเรียนรู้และการทดสอบ

3.3.4 การวัดประสิทธิภาพของแบบจำลอง

การวัดประสิทธิภาพของแบบจำลองที่ได้จากการจำแนกข้อมูลด้วยเทคนิคต้นไม้ตัดสินใจอัลกอริทึม C4.5 ได้จากการนำแบบจำลองที่ได้จากชุดข้อมูลการเรียนรู้มาทดสอบด้วยชุดข้อมูลการทดสอบที่ให้ค่าความถูกต้องในการจำแนก (Accuracy), ค่าความครบถ้วน (Recall), ค่าความแม่นยำ (Precision) และค่าถ่วงดุล (F-measure) ดังแสดงราละเอียดของสมการในบทที่ 2 หัวข้อ 2.3.6

3.3.5 การวิเคราะห์และการประเมินผล

ขั้นตอนการวิเคราะห์และการแปลผลเป็นขั้นตอนการแปลความหมาย โดยในขั้นตอนนี้ผู้วิจัยได้ทำการวิเคราะห์ผล 3 ข้อ ได้แก่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1) วิเคราะห์กฎของการควบคุมกระบวนการที่เหมาะสมสำหรับข้อมูลแต่ละข้อมูล
2) วิเคราะห์ผลของต้นไม้ตัดสินใจและกฎที่ได้เพื่อสรุปปัจจัยที่ทำให้เกิดปริมาณฝุ่นละอองในแต่ละพื้นที่การทำงาน
3) วิเคราะห์เปรียบเทียบผลของการจำแนกข้อมูลระหว่างข้อมูลที่ควบคุมคุณภาพด้วยเกณฑ์ของโรงงานกับข้อมูลที่ควบคุมด้วยกระบวนการเชิงสถิติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
ผลการทดลองและวิเคราะห์เปรียบเทียบประสิทธิภาพของต้นไม้ตัดสินใจ

ในบทนี้จะกล่าวถึงผลการทดลองและการอภิปรายเปรียบเทียบผล ซึ่งแบ่งการพิจารณาเป็น 2 ส่วน ได้แก่ ส่วนของการควบคุมกระบวนการเชิงสถิติ และส่วนของกฎที่ได้จากการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจของข้อมูล 2 กลุ่ม ได้แก่ ข้อมูลควบคุมคุณภาพด้วยเกณฑ์ของโรงงานและข้อมูลที่ควบคุมคุณภาพด้วยการควบคุมกระบวนการเชิงสถิติทั้ง 8 กฎการควบคุม ซึ่งแต่ละกลุ่มข้อมูลประกอบด้วยข้อมูล 3 ชุด ตามมิติพื้นที่การทำงานในส่วนการผลิตฮาร์ดดิสก์ไดร์ฟ ได้แก่ ชุดข้อมูล cleaning, ชุดข้อมูล HDE และชุดข้อมูล HSA จากนั้นแปลงกฎที่ได้จากต้นไม้ตัดสินใจ เพื่อสรุปหาปัจจัยที่ทำให้เกิดปริมาณอนุภาคฝุ่นละอองที่เกินควบคุมและสรุปกฎที่เหมาะสมสำหรับการตรวจจับฝุ่นละอองเมื่อใช้การควบคุมกระบวนการเชิงสถิติ

4.1 ผลการทดลองและการอภิปรายผลการควบคุมกระบวนการเชิงสถิติของข้อมูลปริมาณอนุภาคฝุ่นละออง

สำหรับในหัวข้อนี้เป็นการนำข้อมูลปริมาณอนุภาคฝุ่นละอองทั้ง 3 ชุดที่เตรียมไว้ซึ่งแบ่งตามพื้นที่การทำงาน นำไปวิเคราะห์ด้วยแผนภูมิควบคุมด้วยเกณฑ์การควบคุม 8 เกณฑ์ ได้แก่

1) มี 1 จุดอยู่นอกเส้นควบคุม $\pm 3\sigma$
2) มี 9 จุดต่อเนื่องกันที่อยู่ด้านใดด้านหนึ่งของเส้นควบคุมกลาง (Central line)
3) มี 6 จุดต่อเนื่องกันที่มีแนวโน้มเพิ่มขึ้นหรือลดลงอย่างต่อเนื่อง
4) มี 14 จุดต่อเนื่องกันที่มีขึ้นลงสลับกัน
5) มี 2 ใน 3 จุดที่อยู่นอกเส้นควบคุมเตือน $\pm 2\sigma$ บนด้านเดียวกันของเส้นควบคุมกลาง
6) มี 4 ใน 5 จุดที่อยู่นอกเส้นควบคุมเตือน $\pm 1\sigma$ บนด้านเดียวกันของเส้นควบคุมกลาง
7) มี 15 จุดต่อเนื่องกันที่อยู่ในช่วง $\pm 1\sigma$
8) มี 8 จุดต่อเนื่องกันที่อยู่นอกเขตควบคุม $\pm 1\sigma$ จากเส้นควบคุมกลางของข้างใดข้างหนึ่ง

4.1.1 ผลการสร้างแผนภูมิควบคุมของชุดข้อมูล พื้นที่การทำงาน cleaning

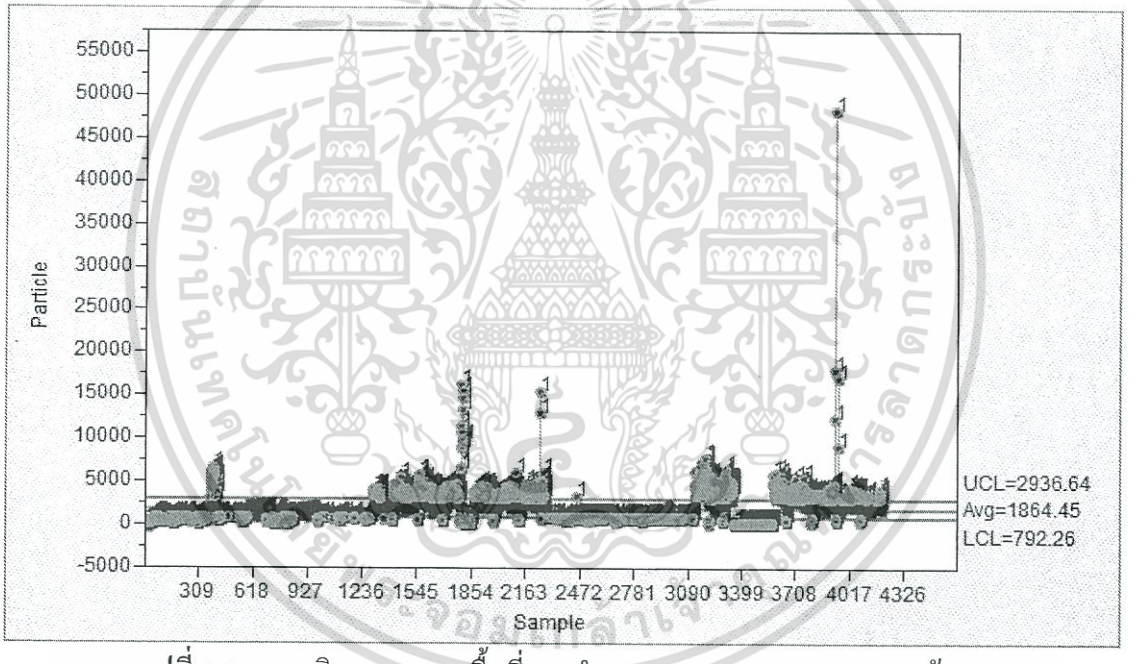
เมื่อสร้างแผนภูมิควบคุมด้วยเกณฑ์ทั้ง 8 เกณฑ์ ตามที่กำหนดแล้ว พบว่า สำหรับพื้นที่การทำงาน cleaning ซึ่งมีข้อมูลทั้งหมด 4,206 ข้อมูล พบว่าแผนภูมิควบคุมนี้มีจุดที่เกินควบคุมปริมาณมากในกฎข้อที่ 1, 2 ,5,6 และ 8 โดยแผนภูมิควบคุมค่าวัดเชิงเดี่ยว (Individual control chart) และ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

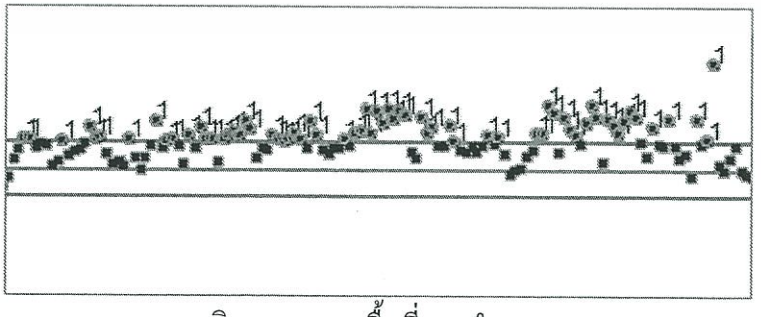
แผนภูมิค่าพิสัยเคลื่อนที่นั้นมีค่า มีค่าเฉลี่ยหรือเส้นควบคุมกลางเท่ากับ 1864.45 และ 403.279 ตามลำดับ ส่วนค่าเส้นควบคุมบนและเส้นควบคุมล่าง เท่ากับ 2937.00, 791.897 และ 1317.63, 0 ตามลำดับ และสามารถแสดงผลการทดลองได้ดังตาราง 4.1 และแสดงแผนภูมิควบคุมในรูปที่ 4.1-4.16

ตารางที่ 4.1 ผลการสร้างแผนภูมิควบคุมจากข้อมูลพื้นที่ cleaning

ข้อมูลพื้นที่ Cleaning											
ชนิดแผนภูมิ	ตำแหน่งเส้นควบคุม (ปริมาณฝุ่นละออง)			จำนวนข้อมูลเกินควบคุม (Records)							
	CL	UCL	LCL	กฎ 1	กฎ 2	กฎ 3	กฎ 4	กฎ 5	กฎ 6	กฎ 7	กฎ 8
I-chart	1864.45	2937.00	791.897	2043	3282	43	0	2844	3350	0	2652
MR-char	403.279	1317.63	0	175	997	4	10	-	-	-	-

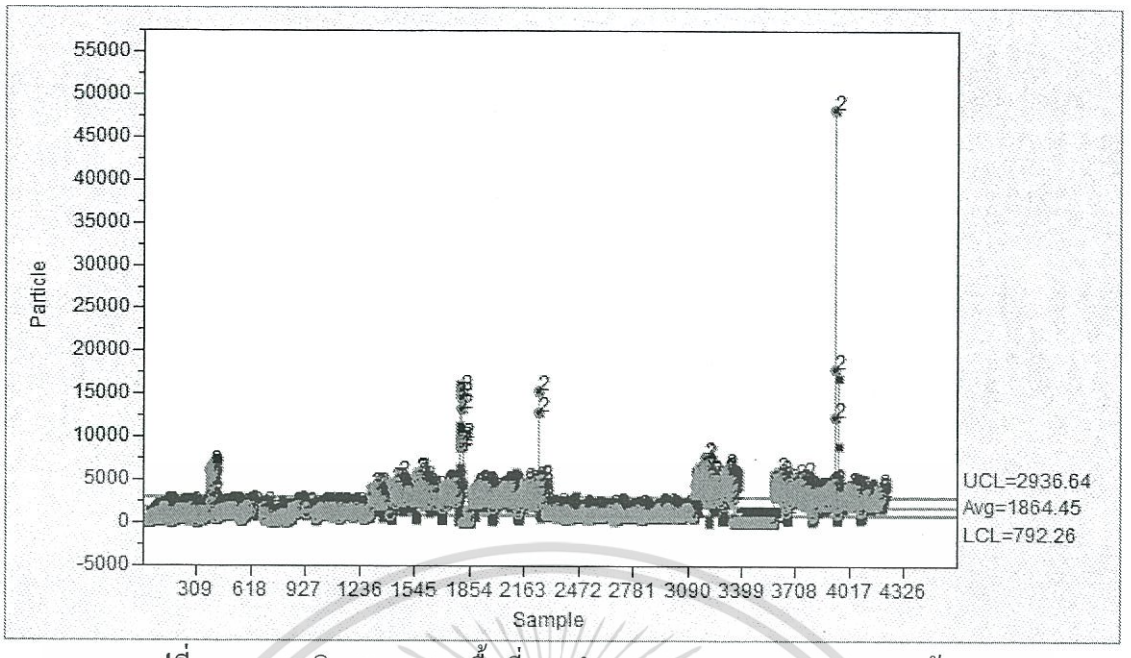


รูปที่ 4.1 แผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 1

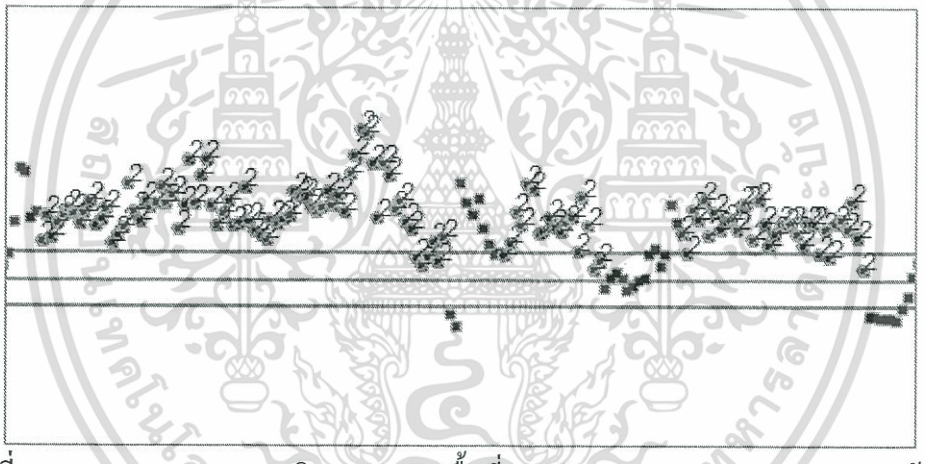


รูปที่ 4.2 ภาพขยายของแผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 1 ที่เรคคอร์ด 1989 ถึง 2120

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

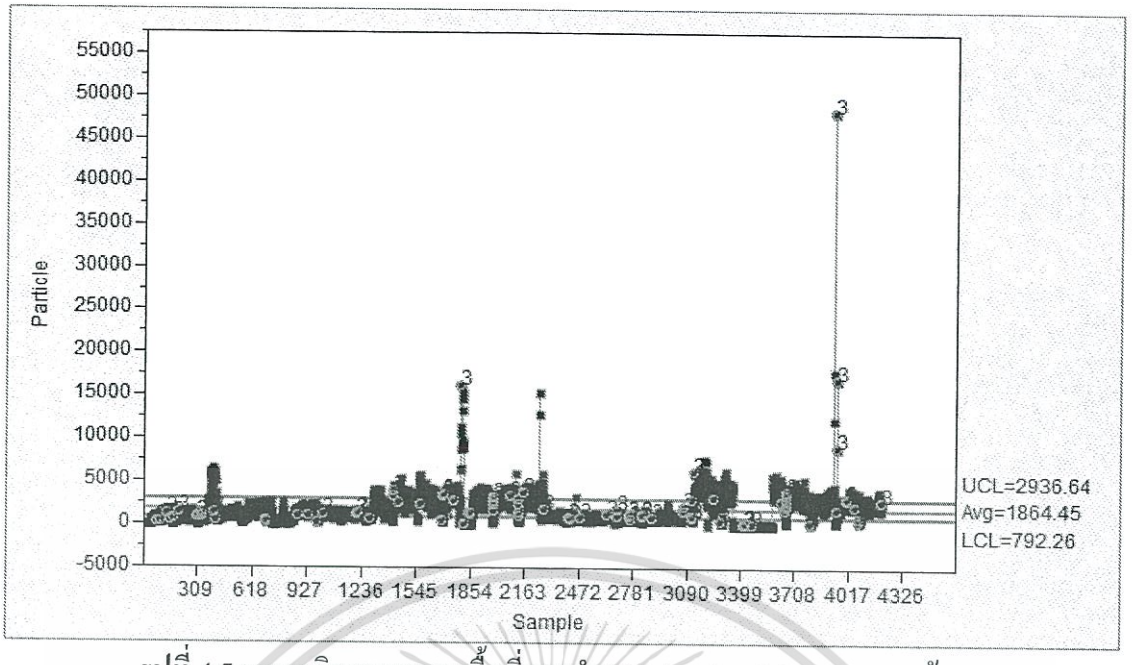


รูปที่ 4.3 แผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 2

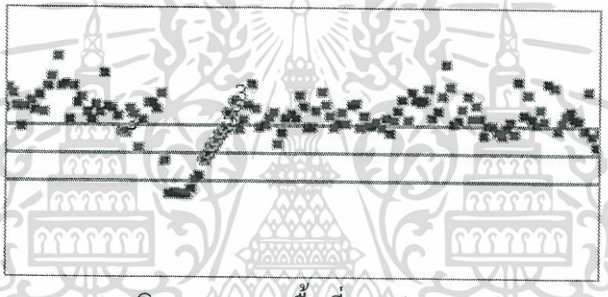


รูปที่ 4.4 ภาพขยายของแผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 2 ที่เรคคอร์ด 3131 ถึง 3294

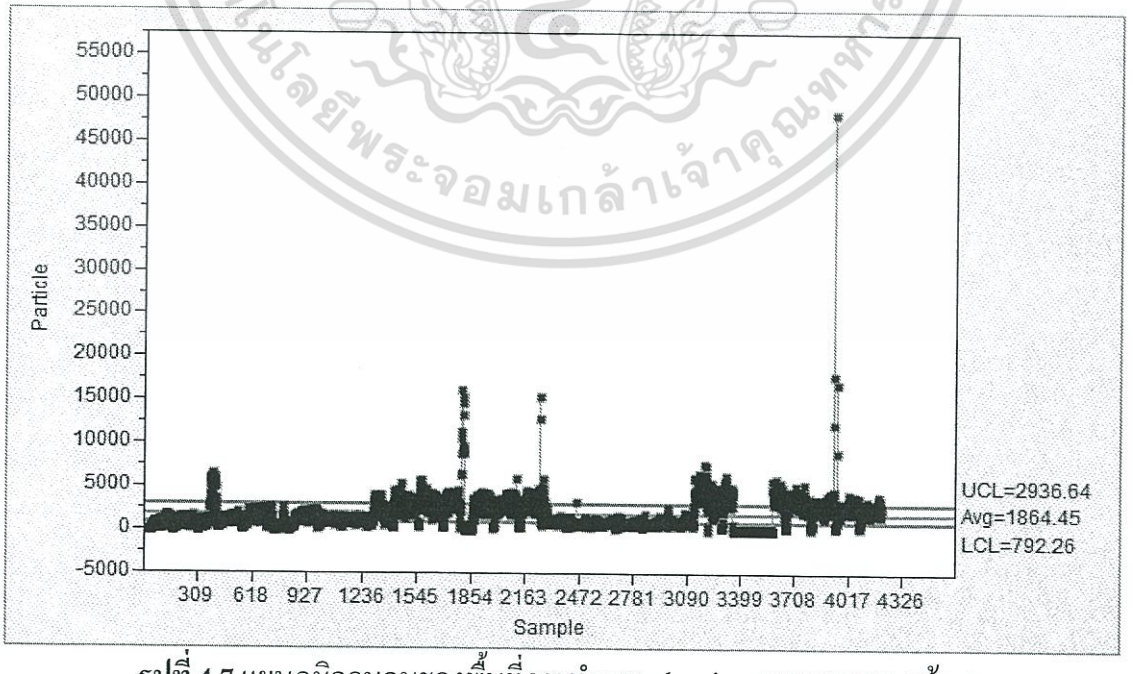
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.5 แผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 3

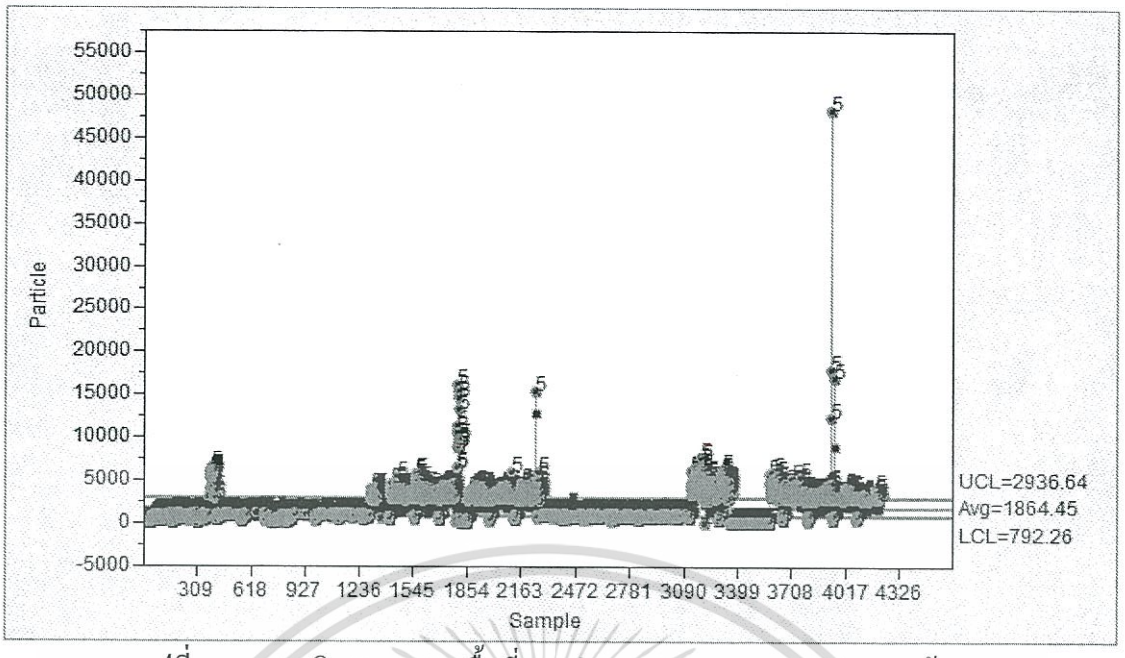


รูปที่ 4.6 ภาพขยายของแผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 3 ที่เรคคอร์ด 3618 ถึง 3749

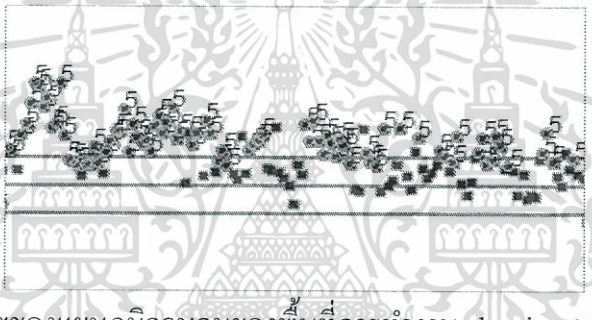


รูปที่ 4.7 แผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 4

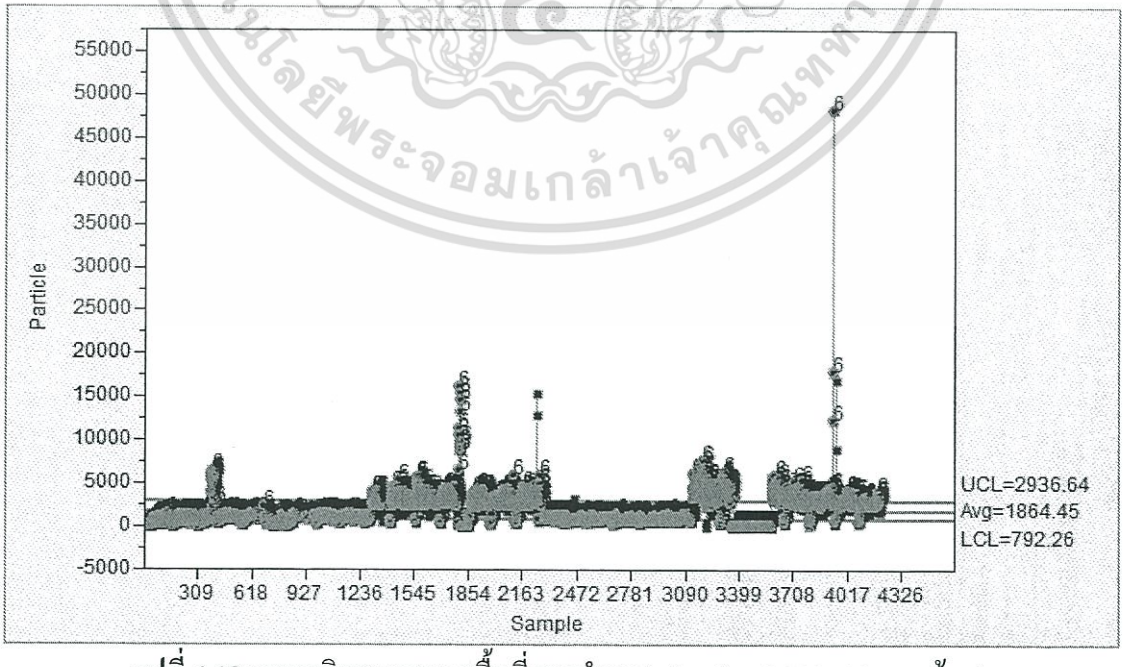
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.8 แผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 5

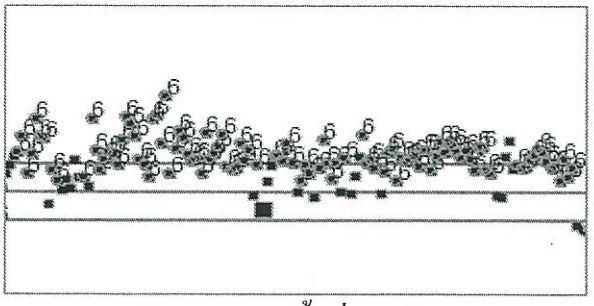


รูปที่ 4.9 ภาพขยายของแผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 5 ที่เรคคอร์ด 1557 ถึง 1691

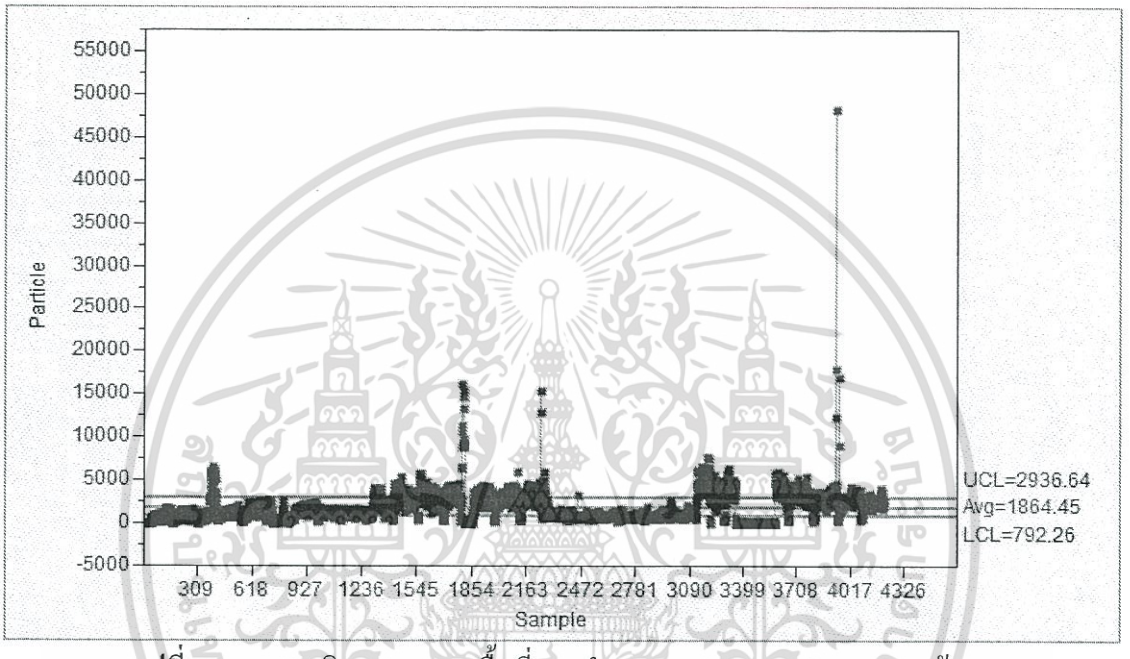


รูปที่ 4.10 แผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 6

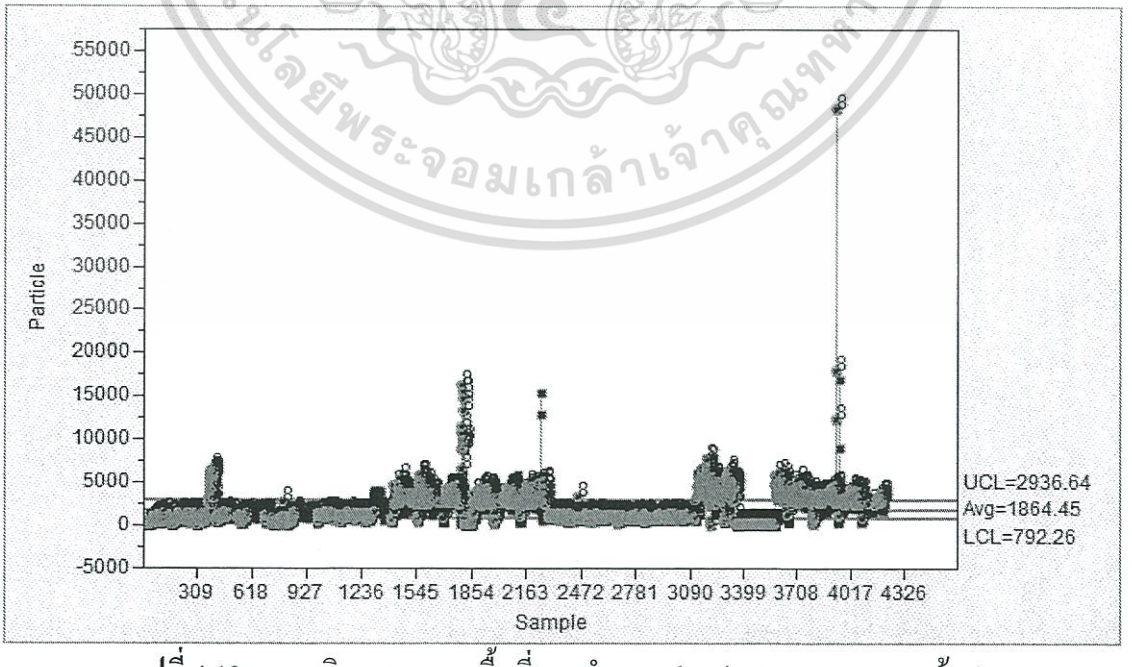
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.11 ภาพขยายของแผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 6 ที่เรคคอร์ด 1416 ถึง 1548

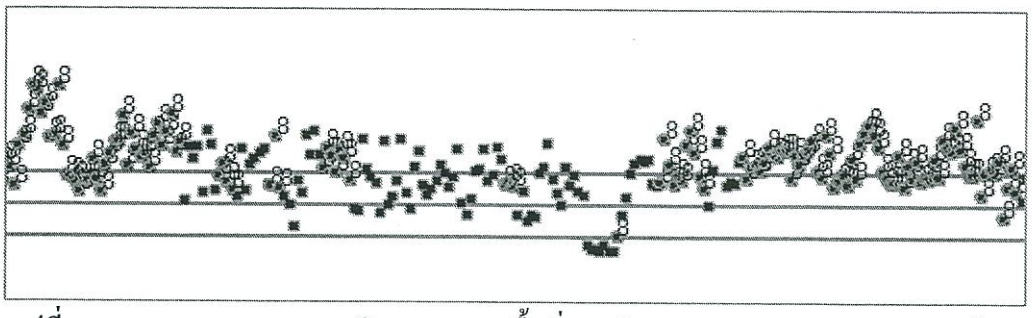


รูปที่ 4.12 แผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 7

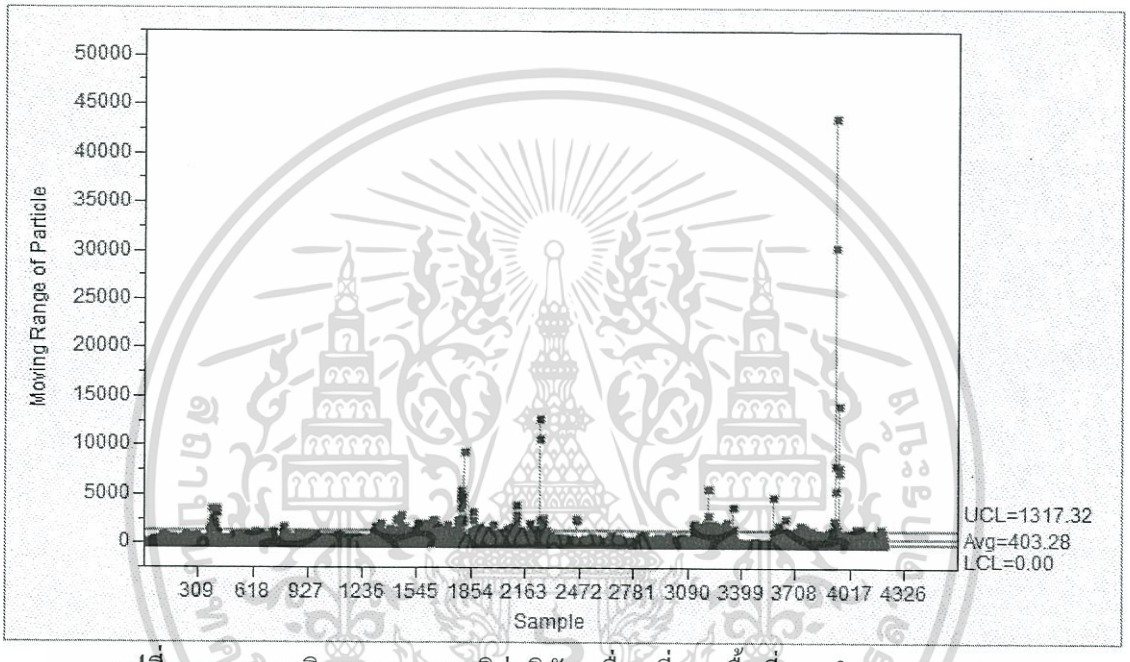


รูปที่ 4.13 แผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

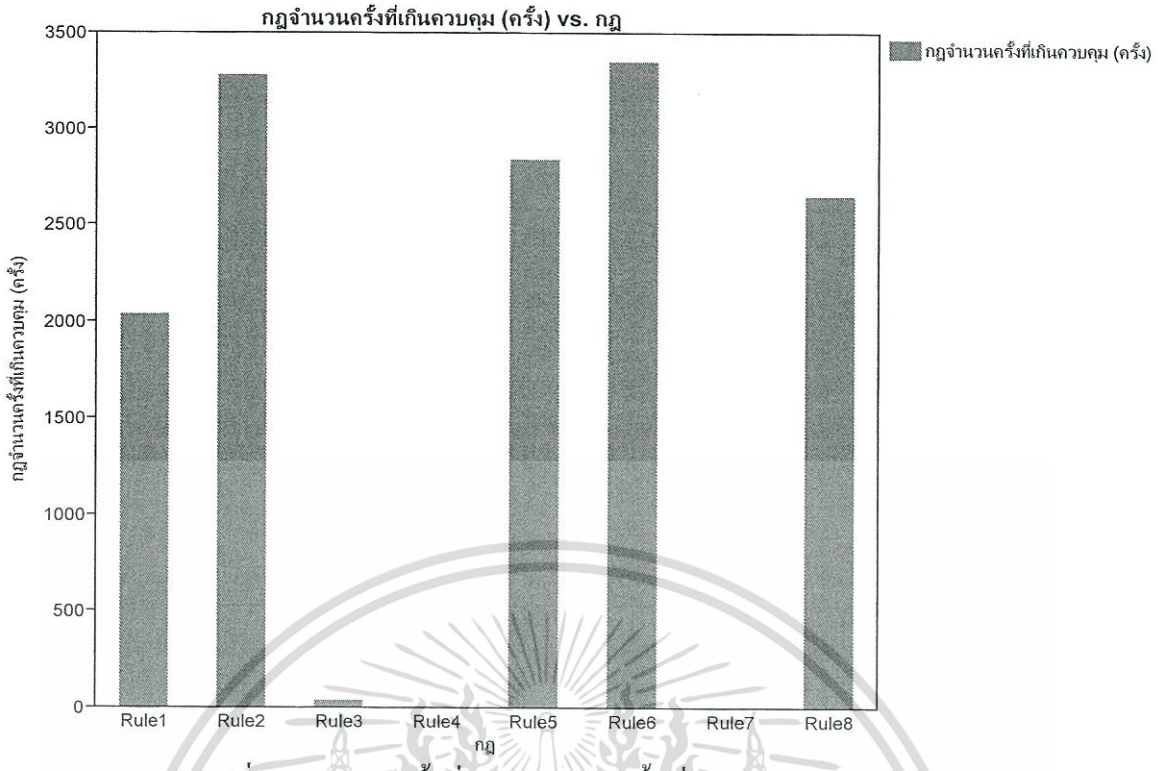


รูปที่ 4.14 ภาพขยายของแผนภูมิควบคุมของพื้นที่การทำงาน cleaning กฎการควบคุมข้อ 8 ที่เรคคอร์ด 1557 ถึง 1798



รูปที่ 4.15 แผนภูมิควบคุมแผนภูมิค่าพิสัยเคลื่อนที่ของพื้นที่การทำงาน cleaning

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.16 จำนวนครั้งที่เกินควบคุมในพื้นที่การทำงาน cleaning

จากตารางที่ 4.1 และรูปที่ 4.1-4.16 เมื่อพิจารณาแล้วพบว่า กฎที่ 1, 2, 5, 6 และ 8 มีจำนวนข้อมูลที่เกินควบคุมในปริมาณค่อนข้างมาก จึงสามารถอนุมานได้ว่ากฎดังกล่าวมีแนวโน้มที่อาจจะเหมาะสมในการใช้สำหรับควบคุมกระบวนการของพื้นที่การทำงาน cleaning ส่วนกฎควบคุมข้อที่ 4 และ 7 นั้นไม่มีจุดที่เกินควบคุมจากกฎดังกล่าวดังนั้นจึงอนุมานได้ว่ากฎข้อ 4 และ 7 ไม่เหมาะสมสำหรับการควบคุมปริมาณอนุภาคฝุ่นละอองของพื้นที่การทำงาน cleaning

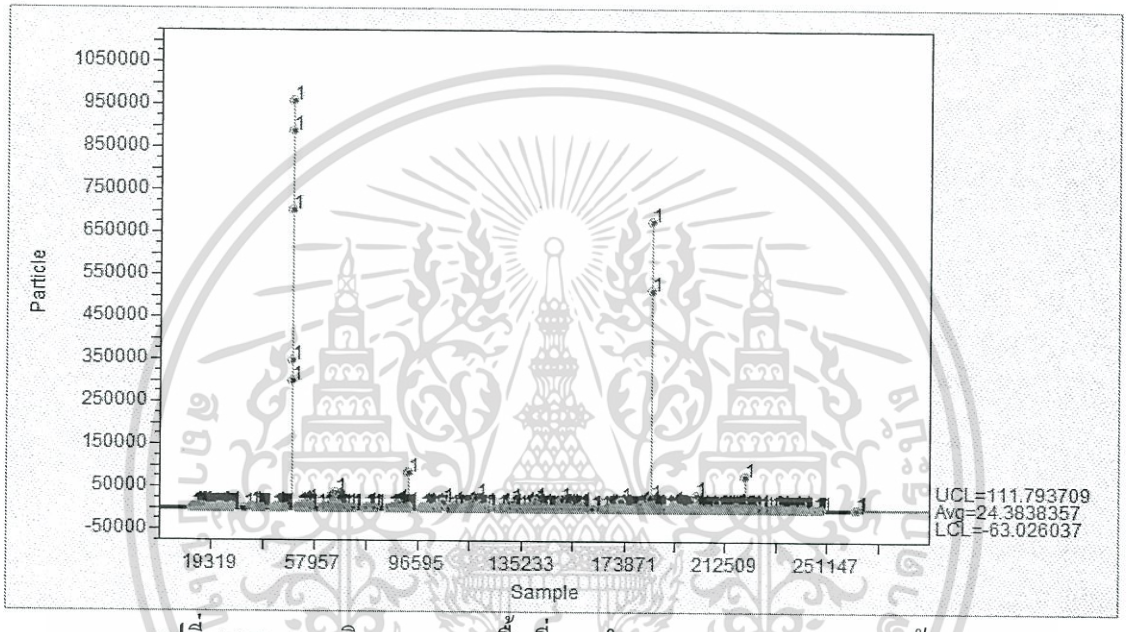
4.1.2 ผลการสร้างแผนภูมิควบคุมของชุดข้อมูล พื้นที่การทำงาน HDE

เมื่อสร้างแผนภูมิควบคุมด้วยเกณฑ์ทั้ง 8 เกณฑ์ ตามที่กำหนดแล้ว สำหรับพื้นที่การทำงาน HDE ซึ่งมีข้อมูลทั้งหมด 263,433 ข้อมูล พบว่าแผนภูมิควบคุมนี้มีจุดที่เกินควบคุมสำหรับกฎข้อที่ 1 - 8 โดยแผนภูมิควบคุมค่าวัดเชิงเดี่ยว (Individual control chart) และแผนภูมิค่าพิสัยเคลื่อนที่นั้นมีค่า มีค่าเฉลี่ยหรือเส้นควบคุมกลางเท่ากับ 24.3838 และ 32.8772 ตามลำดับ ส่วนค่าเส้นควบคุมบนและเส้นควบคุมล่าง เท่ากับ 111.823, -63.0554 และ107.419, 0 ตามลำดับ และสามารถแสดงผลการทดลองได้ดังตาราง 4.2 และแสดงแผนภูมิควบคุมในรูปที่ 4.17-4.26

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

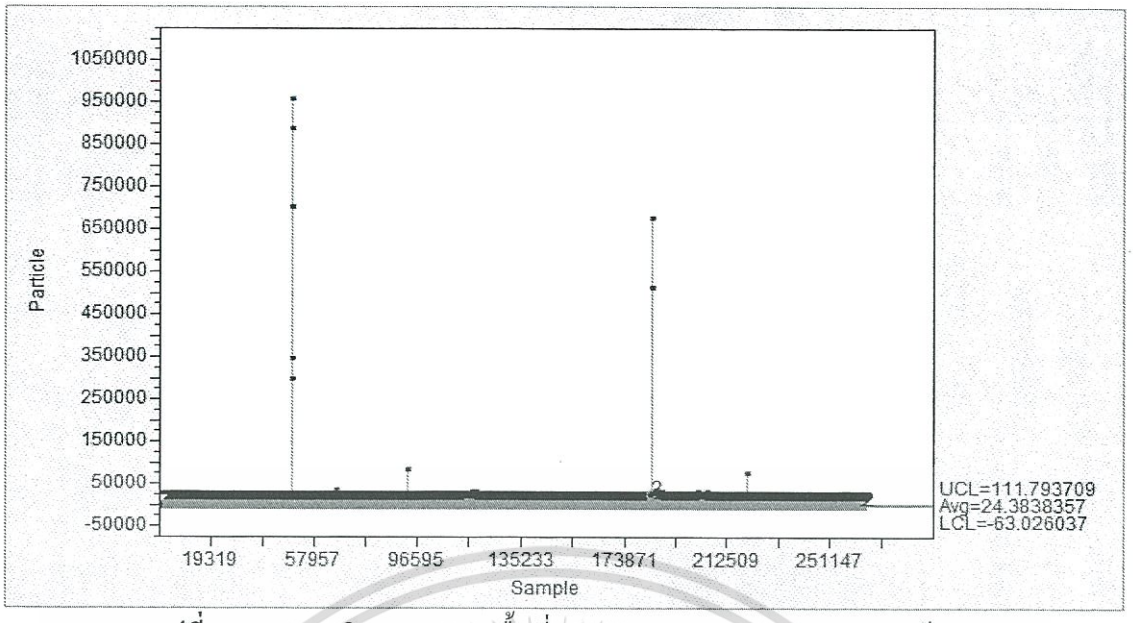
ตารางที่ 4.2 ผลการสร้างแผนภูมิควบคุมจากข้อมูลพื้นที่ HDE

ข้อมูลพื้นที่ HDE											
ชนิดแผนภูมิ	ตำแหน่งเส้นควบคุม (ปริมาณฝุ่นละออง)			จำนวนข้อมูลเกินควบคุม (Records)							
	CL	UCL	LCL	กฎ 1	กฎ 2	กฎ 3	กฎ 4	กฎ 5	กฎ 6	กฎ 7	กฎ 8
I-chart	24.3838	111.823	-63.0554	2176	228415	34	44	1220	983	233020	516
MR-char	32.8772	107.419	0	2757	234659	21	15	-	-	-	-



รูปที่ 4.17 แผนภูมิควบคุมของพื้นที่การทำงาน HDE กฎการควบคุมข้อ 1

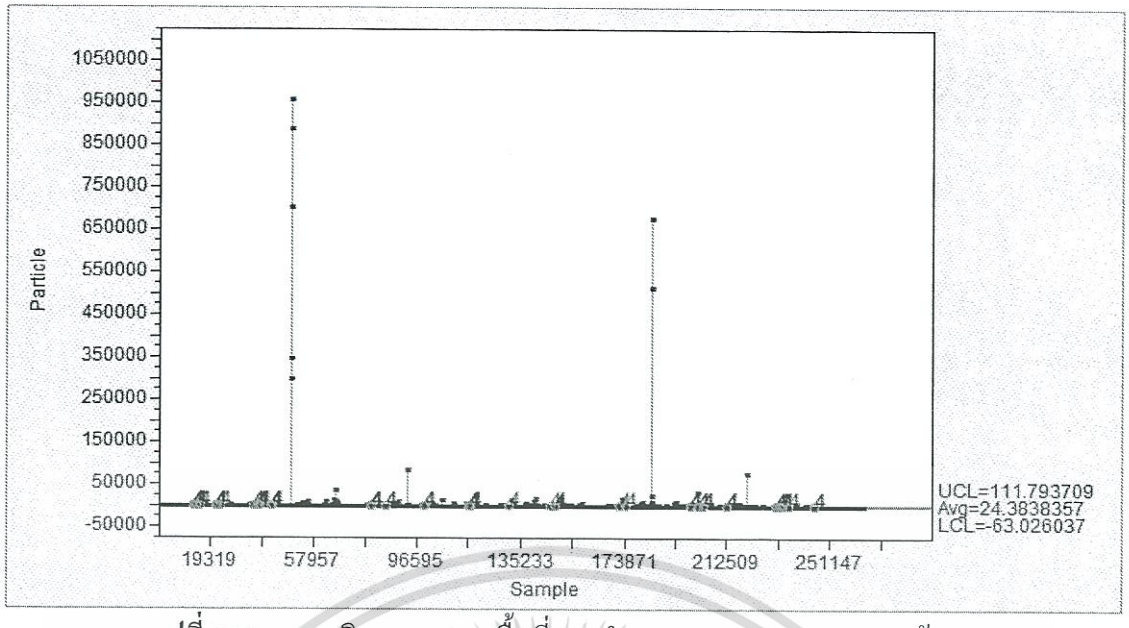
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



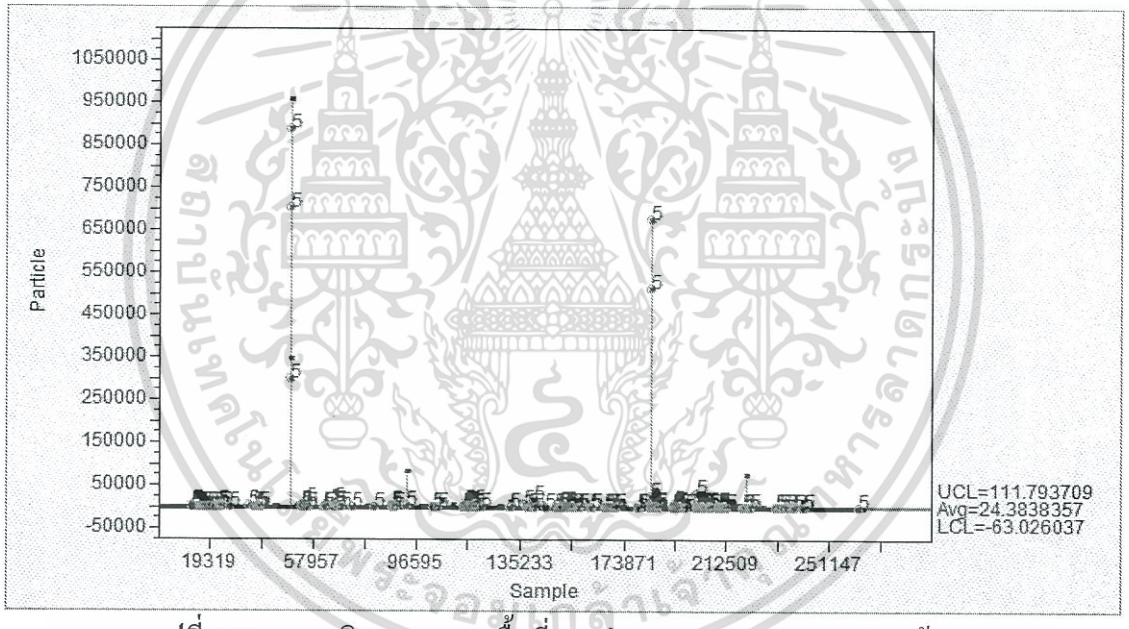
รูปที่ 4.18 แผนภูมิควบคุมของพื้นที่การทำงาน HDE กฎการควบคุมข้อ 2

รูปที่ 4.19 แผนภูมิควบคุมของพื้นที่การทำงาน HDE กฎการควบคุมข้อ 3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

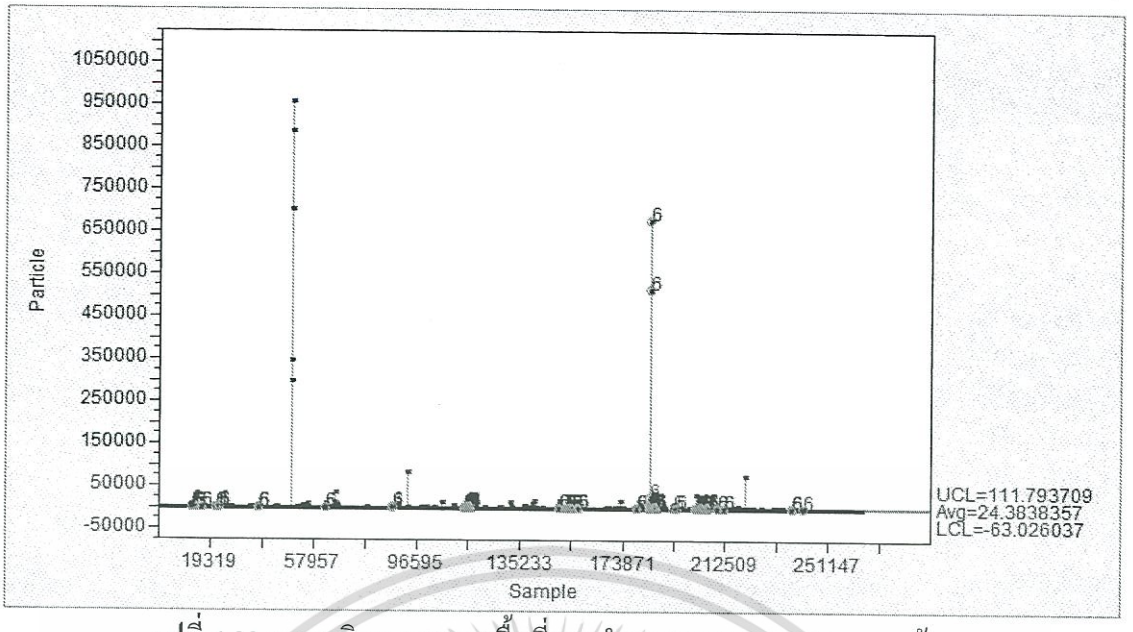


รูปที่ 4.20 แผนภูมิควบคุมของพื้นที่การทำงาน HDE กฎการควบคุมข้อ 4

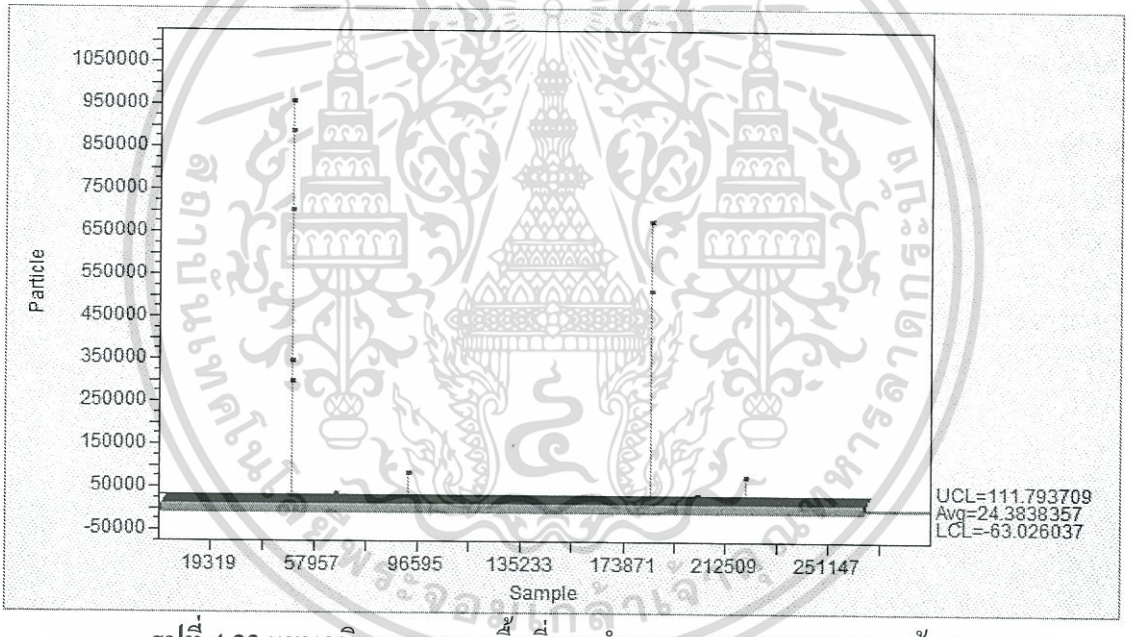


รูปที่ 4.21 แผนภูมิควบคุมของพื้นที่การทำงาน HDE กฎการควบคุมข้อ 5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

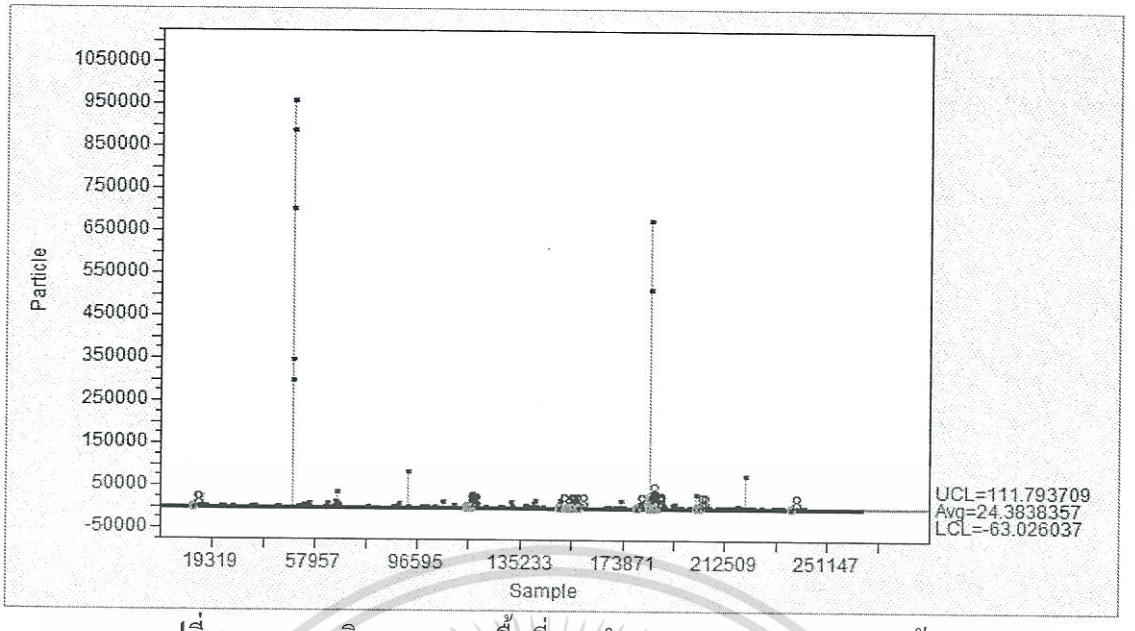


รูปที่ 4.22 แผนภูมิควบคุมของพื้นที่การทำงาน HDE กฎการควบคุมข้อ 6

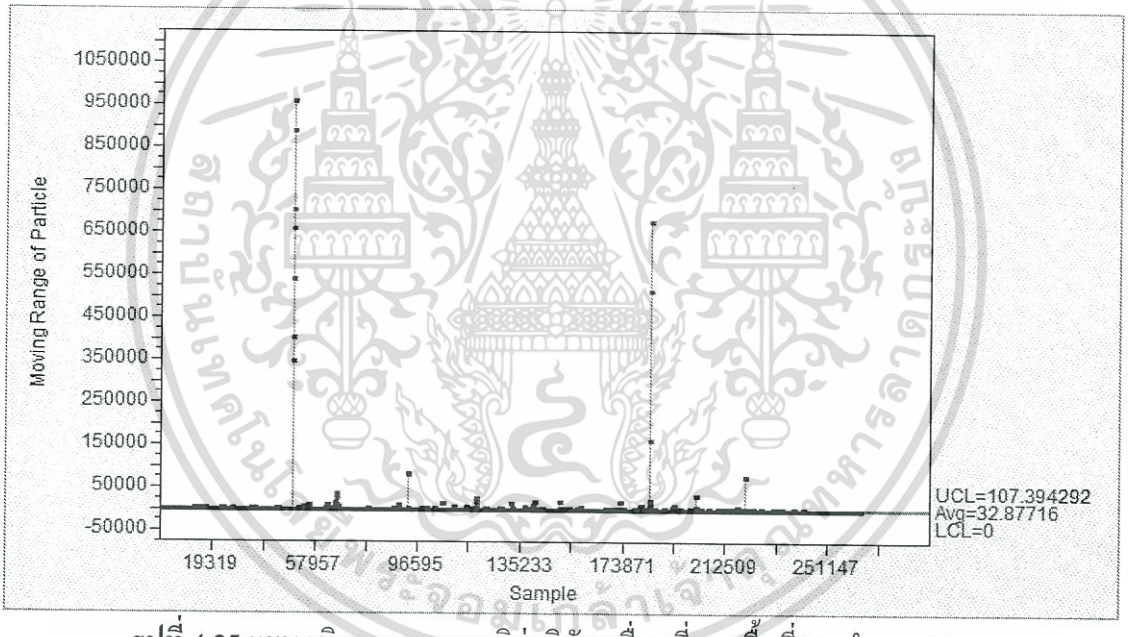


รูปที่ 4.23 แผนภูมิควบคุมของพื้นที่การทำงาน HDE กฎการควบคุมข้อ 7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

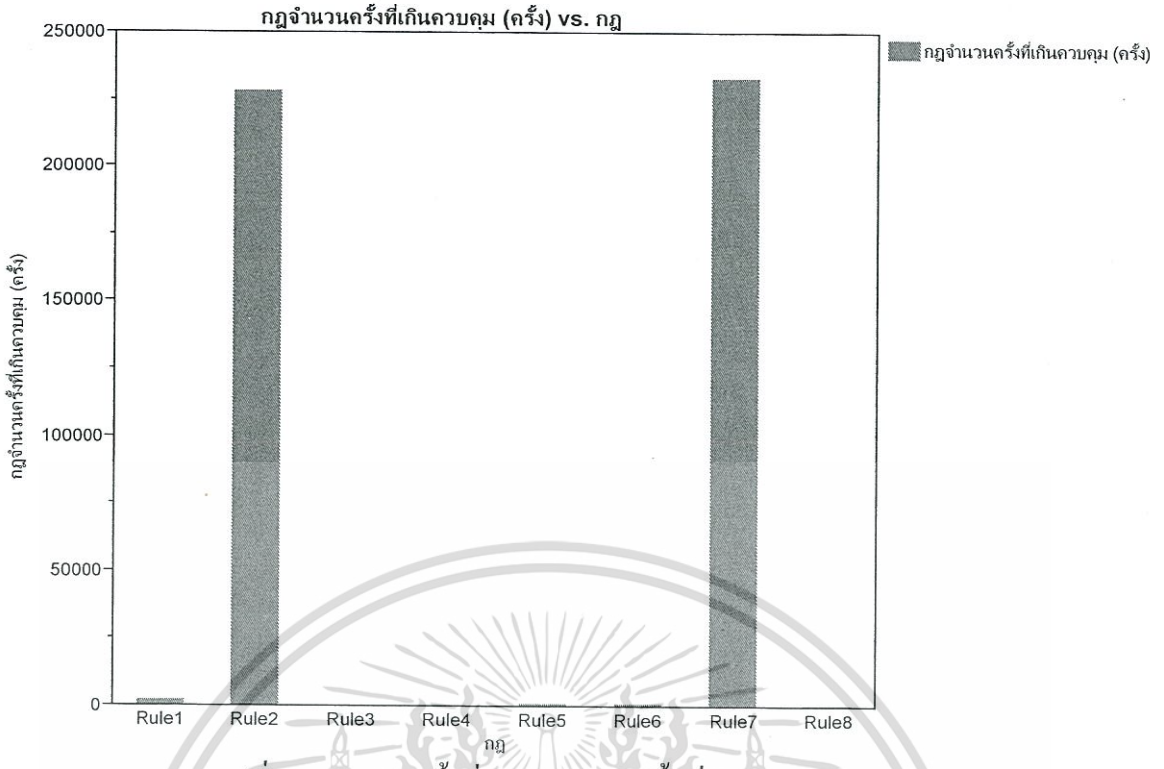


รูปที่ 4.24 แผนภูมิควบคุมของพื้นที่การทำงาน HDE กฎการควบคุมข้อ 8



รูปที่ 4.25 แผนภูมิควบคุมแผนภูมิค่าพิสัยเคลื่อนที่ของพื้นที่การทำงาน HDE

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.26 จำนวนครั้งที่เกินควบคุมในพื้นที่การทำงาน HDE

จากตารางที่ 4.2 และรูปที่ 4.17-4.26 เมื่อพิจารณาแล้วพบว่า กฎที่ 2 และ 7 มีจำนวนข้อมูลที่เกินควบคุมในปริมาณมาก จึงสามารถอนุมานได้ว่ากฎดังกล่าวมีแนวโน้มที่อาจจะเหมาะสมในการใช้สำหรับควบคุมกระบวนการของพื้นที่การทำงาน HDE

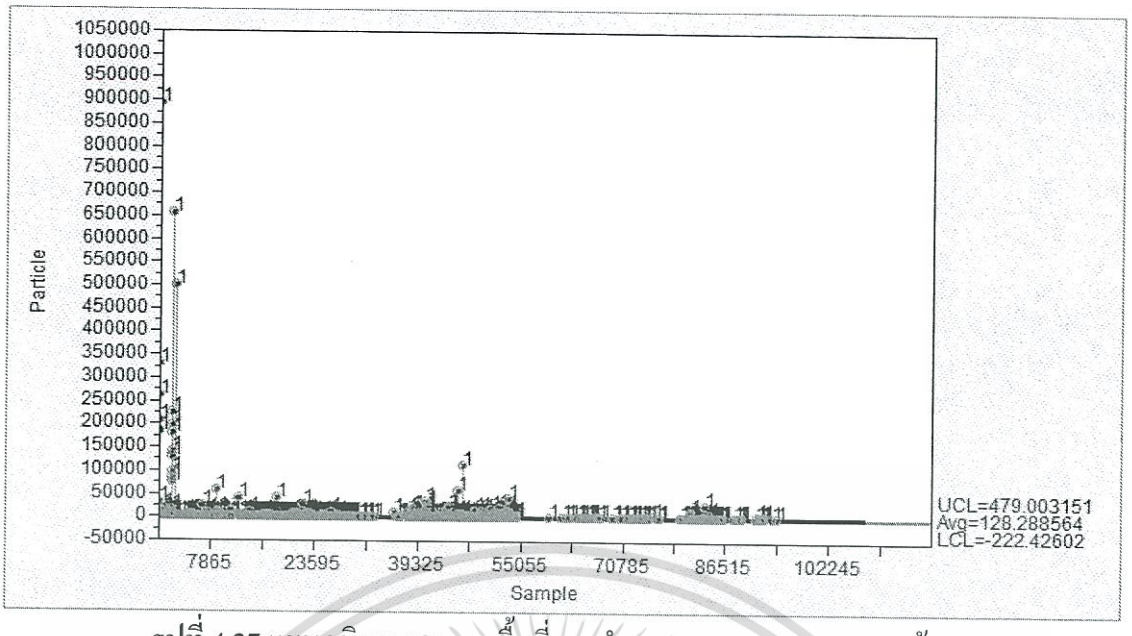
4.1.3 ผลการสร้างแผนภูมิควบคุมของชุดข้อมูล พื้นที่การทำงาน HSA

เมื่อสร้างแผนภูมิควบคุมด้วยเกณฑ์ทั้ง 8 เกณฑ์ ตามที่กำหนดแล้ว พบว่า สำหรับพื้นที่การทำงาน HSA ซึ่งมีข้อมูลทั้งหมด 107,245 ข้อมูล พบว่าแผนภูมิควบคุมนี้มีจุดที่เกินควบคุมสำหรับกฎข้อที่ 1 - 8 โดยแผนภูมิควบคุมค่าวัดเชิงเดี่ยว (Individual control chart) และแผนภูมิค่าพิสัยเคลื่อนที่นั้นมีค่า มีค่าเฉลี่ยหรือเส้นควบคุมกลางเท่ากับ 128.289 และ 131.913 ตามลำดับ ส่วนค่าเส้นควบคุมบนและเส้นควบคุมล่าง เท่ากับ 479.121, -222.544 และ 430.998, 0 ตามลำดับ และสามารถแสดงผลการทดลองได้ดังตาราง 4.3 และแสดงแผนภูมิควบคุมในรูปที่ 4.27-4.35

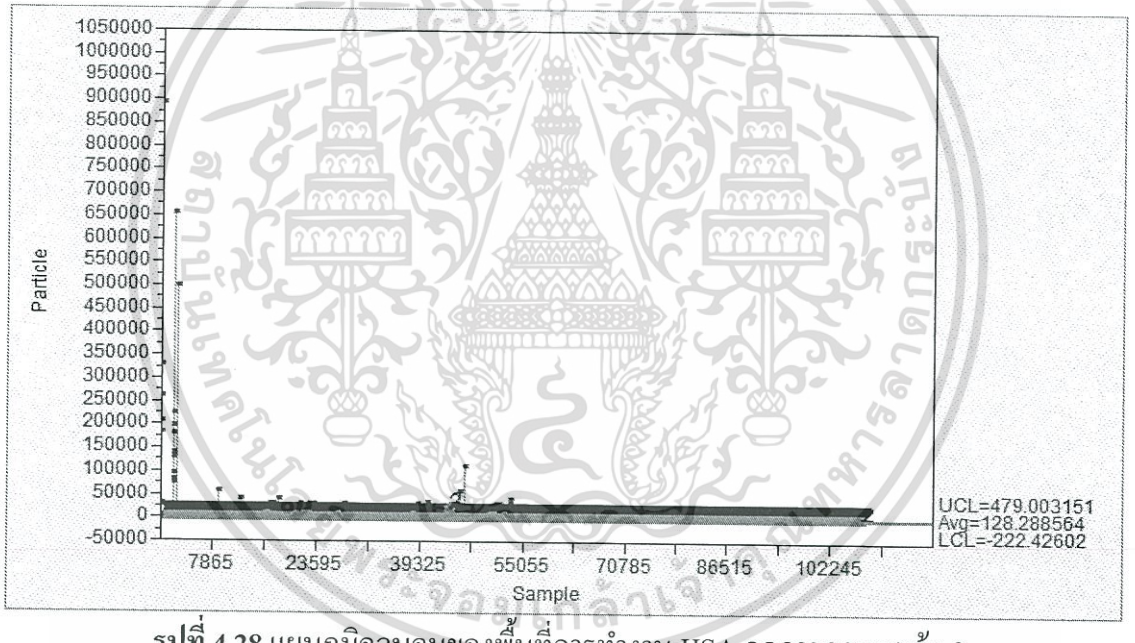
ตารางที่ 4.3 ผลการสร้างแผนภูมิควบคุมจากข้อมูลพื้นที่ HSA

ข้อมูลพื้นที่ HSA											
ชนิดแผนภูมิ	ตำแหน่งเส้นควบคุม (ปริมาณฝุ่นละออง)			จำนวนข้อมูลเกินควบคุม (Records)							
	CL	UCL	LCL	กฎ 1	กฎ 2	กฎ 3	กฎ 4	กฎ 5	กฎ 6	กฎ 7	กฎ 8
I-chart	128.289	479.121	-222.544	2787	69359	181	72	2610	36103	14684	24536
MR-char	131.913	430.998	0	2108	76689	119	95	-	-	-	-

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

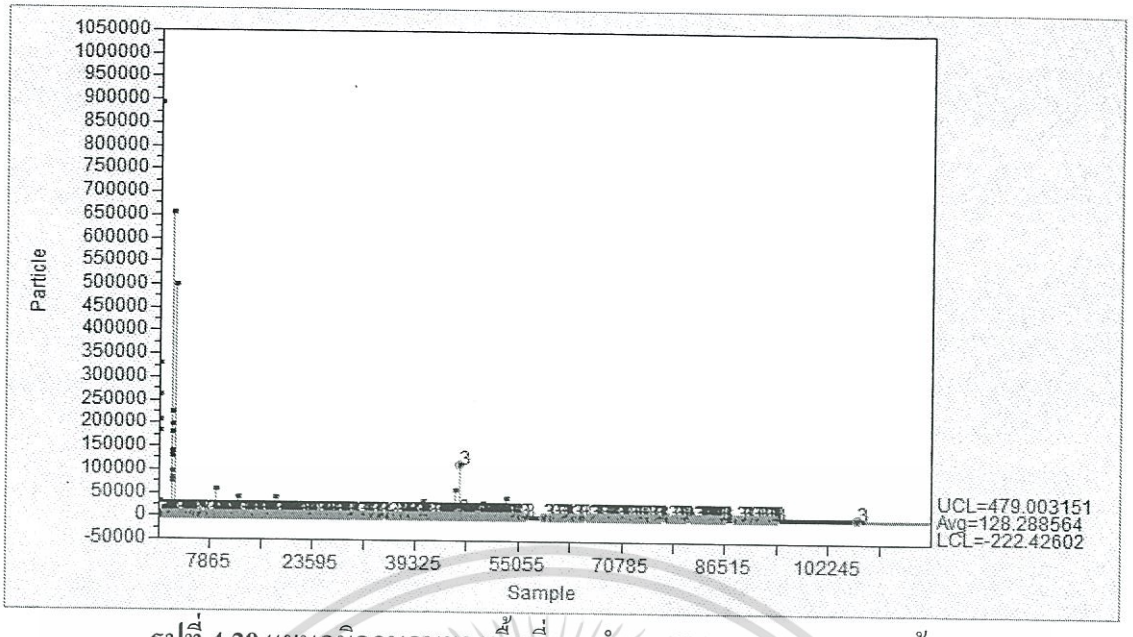


รูปที่ 4.27 แผนภูมิควบคุมของพื้นที่การทำงาน HSA กฎการควบคุมข้อ 1

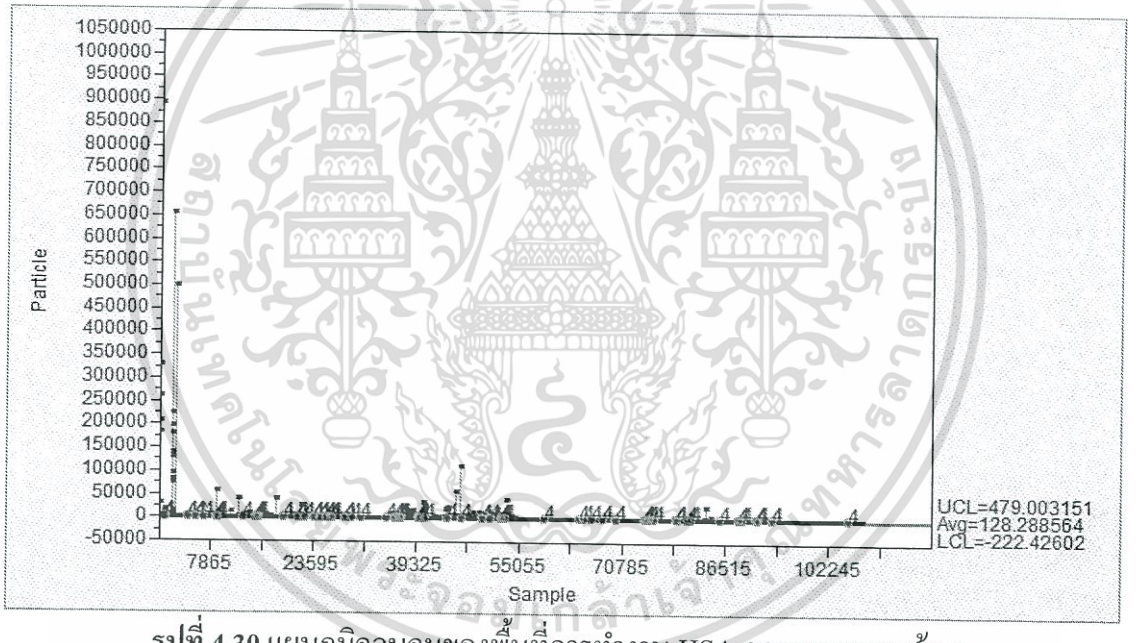


รูปที่ 4.28 แผนภูมิควบคุมของพื้นที่การทำงาน HSA กฎการควบคุมข้อ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

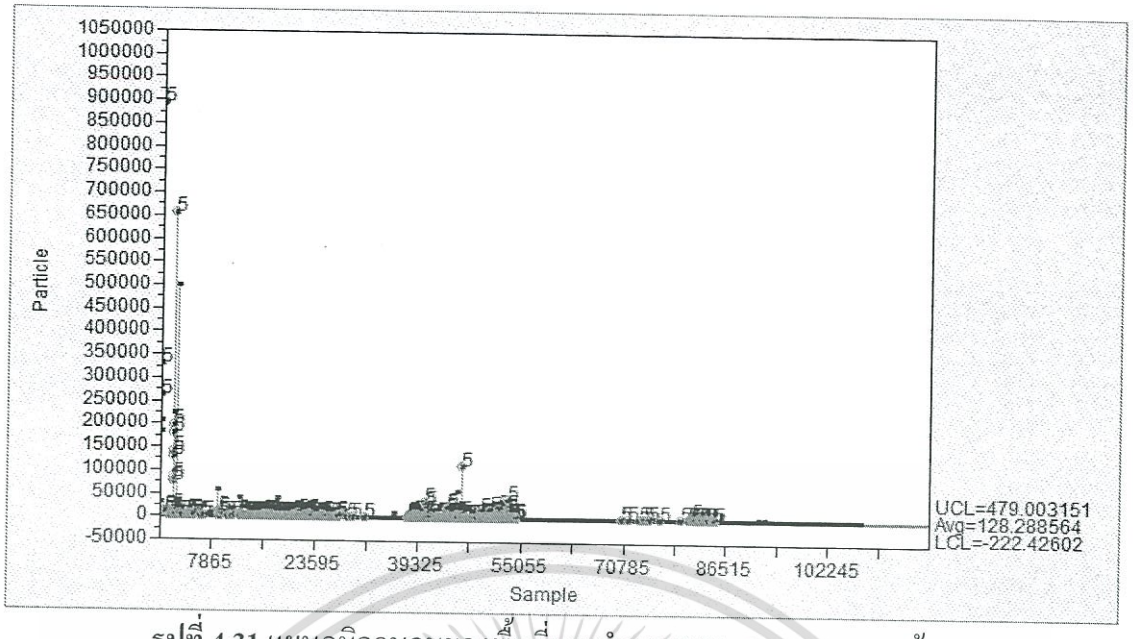


รูปที่ 4.29 แผนภูมิควบคุมของพื้นที่การทำงาน HSA กฎการควบคุมข้อ 3



รูปที่ 4.30 แผนภูมิควบคุมของพื้นที่การทำงาน HSA กฎการควบคุมข้อ 4

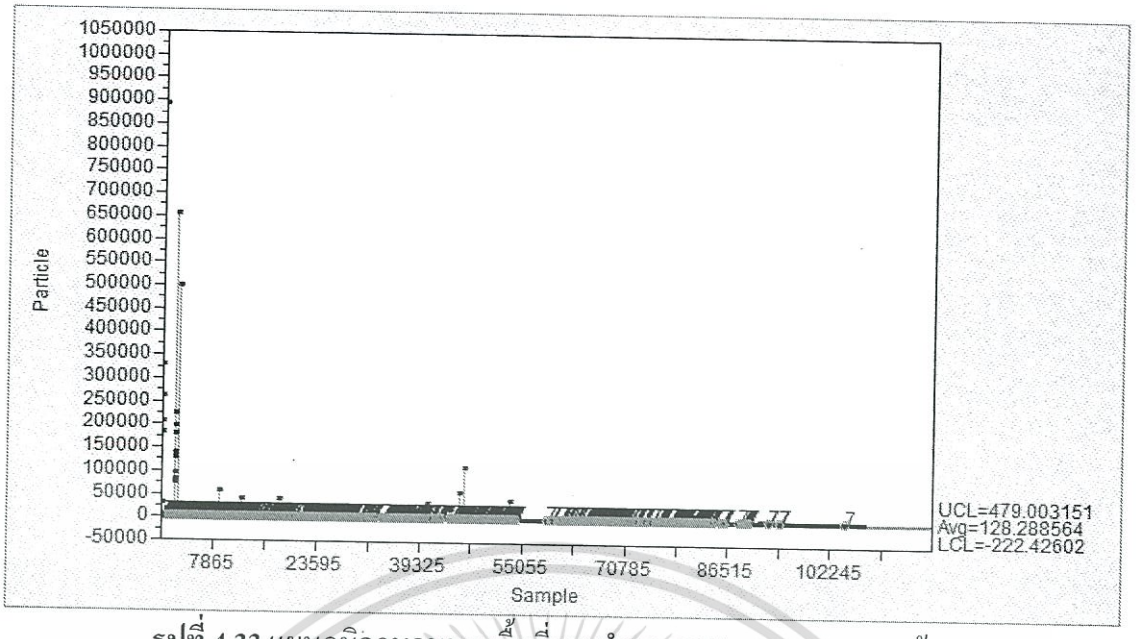
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



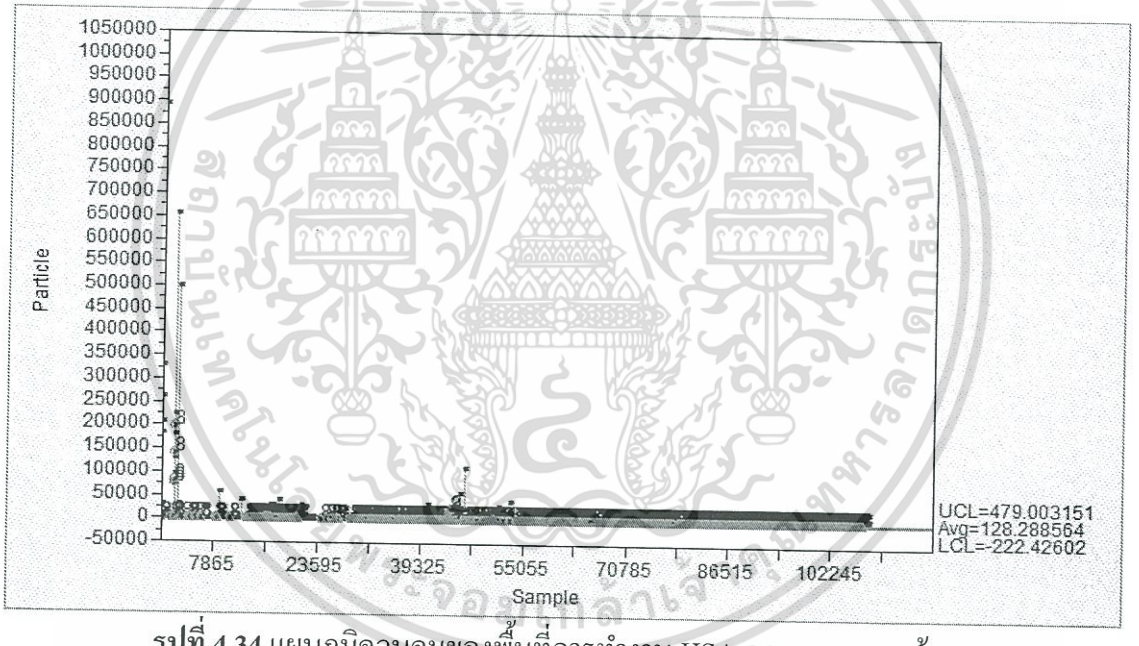
รูปที่ 4.31 แผนภูมิควบคุมของพื้นที่การทำงาน HSA กฎการควบคุมข้อ 5

รูปที่ 4.32 แผนภูมิควบคุมของพื้นที่การทำงาน HSA กฎการควบคุมข้อ 6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

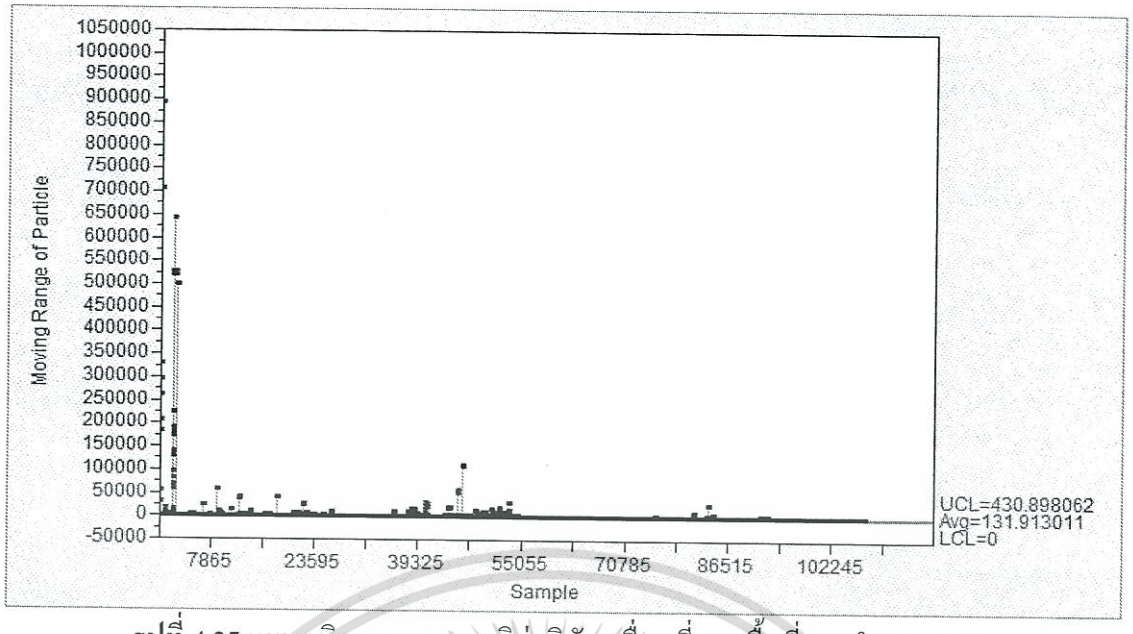


รูปที่ 4.33 แผนภูมิควบคุมของพื้นที่การทำงาน HSA กฎการควบคุมข้อ 7

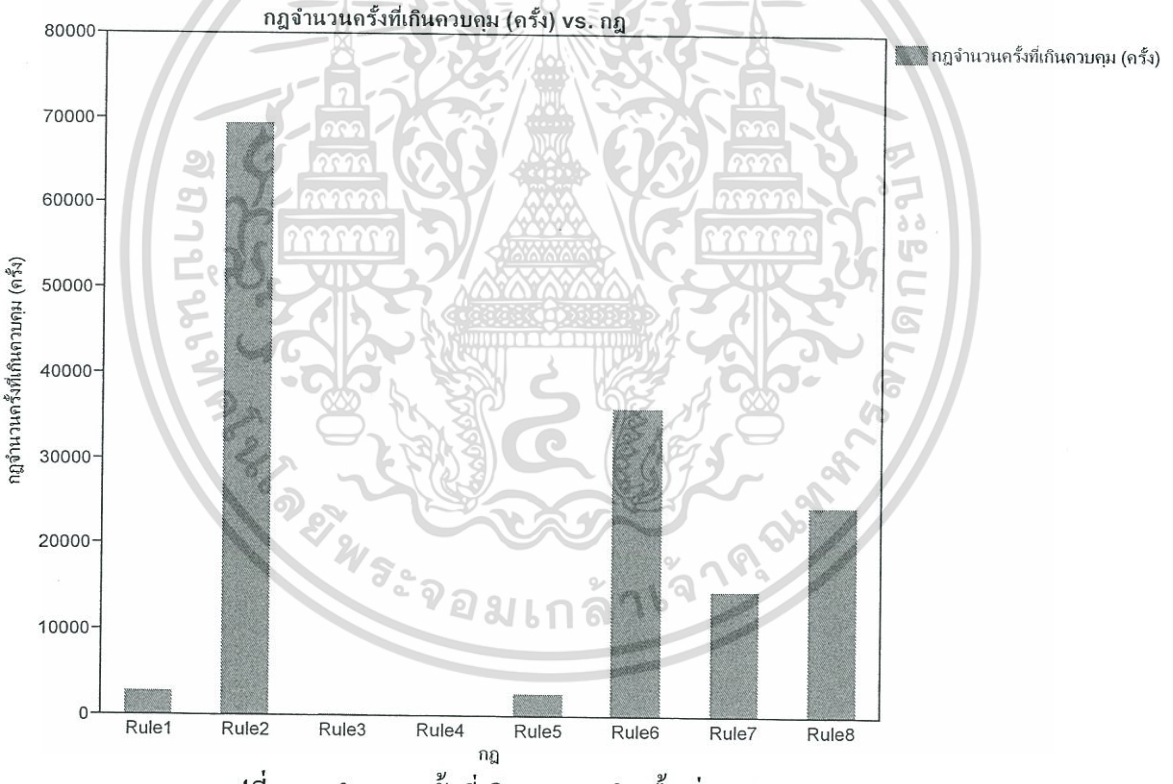


รูปที่ 4.34 แผนภูมิควบคุมของพื้นที่การทำงาน HSA กฎการควบคุมข้อ 8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.35 แผนภูมิควบคุมแผนภูมิค่าพิสัยเคลื่อนที่ของพื้นที่การทำงาน HSA



รูปที่ 4.36 จำนวนครั้งที่เกินควบคุมในพื้นที่การทำงาน HSA

จากตารางที่ 4.3 และรูปที่ 4.27-4.36 เมื่อพิจารณาแล้วพบว่า กฎที่ 2, 6, 7 และ 8 มีจำนวนข้อมูลที่เกินควบคุมในปริมาณมาก จึงสามารถอนุมานได้ว่ากฎดังกล่าวมีแนวโน้มที่อาจจะเหมาะสมในการใช้สำหรับควบคุมกระบวนการของพื้นที่การทำงาน HDE

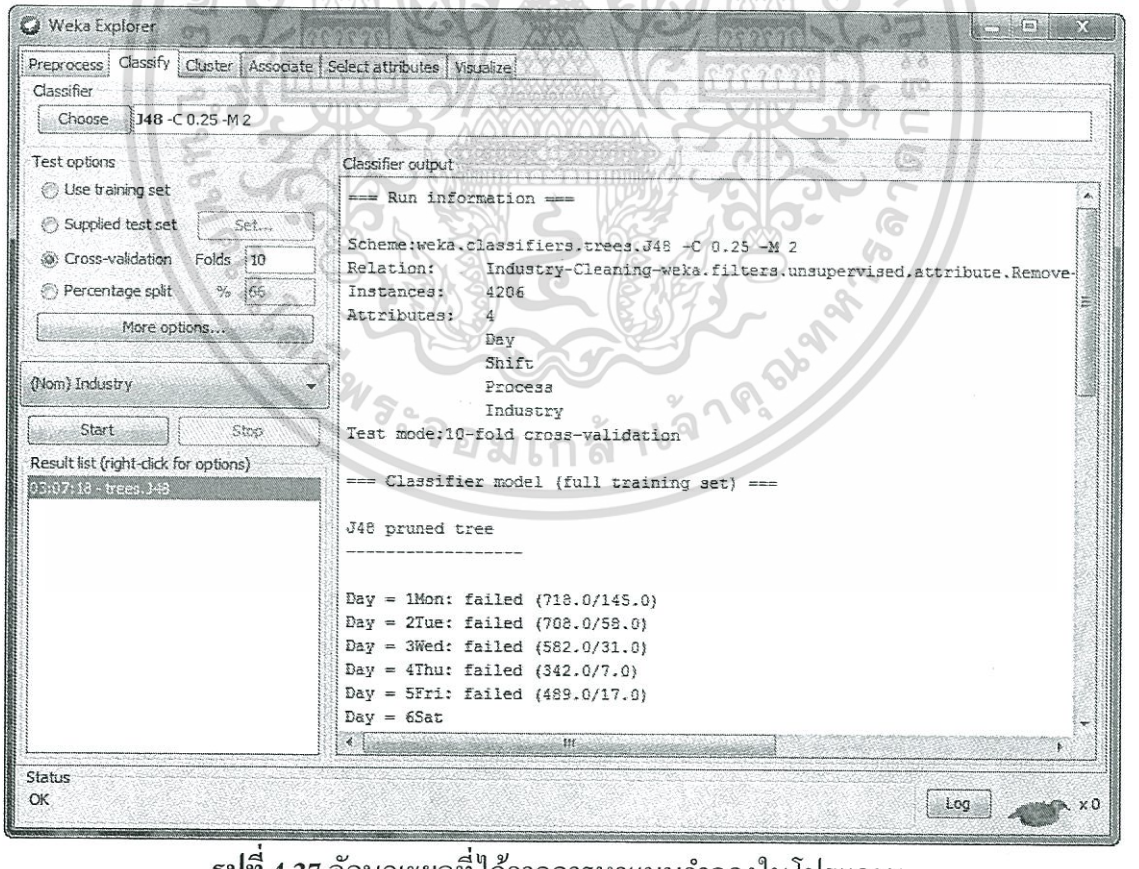
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 ผลการทดลองและการอภิปรายผลส่วนของกฎที่ได้จากการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจ

ส่วนผลการทดลองของการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจนั้นจะดำเนินการทดลองด้วยข้อมูล 2 กลุ่ม ได้แก่ กลุ่มข้อมูลการควบคุมอนุภาคฝุ่นละอองตามเกณฑ์ของโรงงานและกลุ่มข้อมูลการควบคุมอนุภาคฝุ่นละอองด้วยการควบคุมกระบวนการเชิงสถิติซึ่งนำชุดข้อมูลที่จัดระเบียบข้อมูลด้วยกระบวนการเชิงสถิติเฉพาะการควบคุมที่มีแนวโน้มว่าเหมาะสมในการควบคุมกระบวนการจากผลการทดลองที่ 4.1 ซึ่งสามารถแสดงรายละเอียดได้ดังนี้

4.2.1 ผลการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจของกลุ่มข้อมูลการควบคุมอนุภาคฝุ่นละอองตามเกณฑ์ของโรงงาน

หลังจากการสร้างข้อมูลควบคุมคุณภาพตามเกณฑ์การควบคุมของโรงงานในพื้นที่แต่ละพื้นที่และสร้างชุดข้อมูลแล้ว (ตามมิติพื้นที่ของการทำงาน) ทำให้มีข้อมูลการเรียนรู้จำนวน 3 ชุดข้อมูล นำมาวิเคราะห์กับแบบจำลองโดยนำชุดข้อมูลการเรียนรู้มาทดสอบทีละชุดกับอัลกอริทึมต้นไม้ตัดสินใจ C4.5 โดยยกตัวอย่างผลจากการวิเคราะห์ได้ดังรูปที่ 4.37



รูปที่ 4.37 ลักษณะผลที่ได้จากการหาแบบจำลองในโปรแกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.1.1 ผลการจำแนกข้อมูลของพื้นที่การทำงาน cleaning

ผลจากการใช้แบบจำลองของอัลกอริทึมต้นไม้ตัดสินใจ ทดสอบกับชุดข้อมูลให้ผลการทดสอบดังตาราง

ตารางที่ 4.4 จำนวนข้อมูลสำหรับการสร้างแบบจำลองต้นไม้ตัดสินใจชุดข้อมูลพื้นที่ cleaning

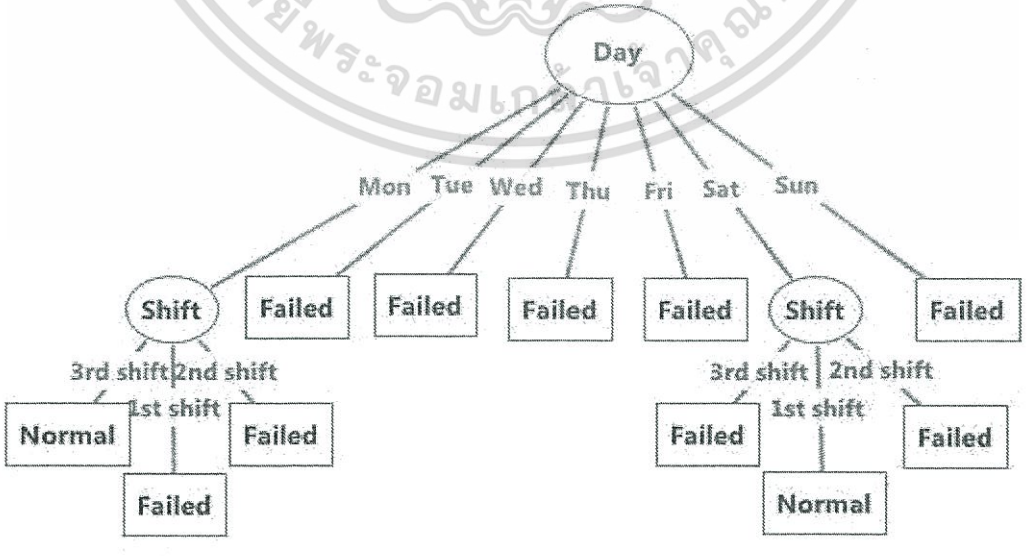
วิธีการจัดระเบียบข้อมูล	จำนวนข้อมูลพื้นที่ Cleaning (Records)			
	ก่อน SMOTE		หลัง SMOTE	
	Normal	Failed	Normal	Failed
เกณฑ์ของโรงงาน	601	3605	1202	3605

ตารางที่ 4.5 ค่าความถูกต้องและค่าเฉลี่ยความคลาดเคลื่อนจากการสร้างแบบจำลองต้นไม้ตัดสินใจชุดข้อมูลพื้นที่ cleaning

ชุดข้อมูล	ผลการทดสอบ (%)		
	Correctly Classified Instances	Incorrectly Classified Instances	RMSE
Cleaning	87.8511 %	12.1489 %	30.76

จากตารางที่ 4.5 แสดงว่าแบบจำลองต้นไม้ตัดสินใจของชุดข้อมูลพื้นที่ cleaning นั้นมีค่าความถูกต้อง 87.85% และมีค่าเฉลี่ยความคลาดเคลื่อน 12.15% โดยจากการจำแนกข้อมูลสามารถสร้างต้นไม้ตัดสินใจได้ดังรูป 4.38

- ต้นไม้ตัดสินใจข้อมูลพื้นที่การทำงาน cleaning



รูปที่ 4.38 ต้นไม้ตัดสินใจพื้นที่ cleaning ตามเกณฑ์การควบคุมของโรงงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูป 4.32 พบว่าชุดข้อมูลพื้นที่ cleaning ประกอบด้วย ขนาดของต้นไม้เท่ากับ 14 โหนด และลีฟโหนดมีขนาด 11 โหนด ซึ่งสามารถแปลงต้นไม้ตัดสินใจเป็นกฎในรูปแบบคำสั่ง IF-THEN ได้ดังตาราง 4.6

ตารางที่ 4.6 กฎจากต้นไม้ตัดสินใจข้อมูลพื้นที่ cleaning

กฎที่	เงื่อนไขกฎ
Rule1	IF DAY=1MON AND SHIFT = 3^{RD} SHIFT THEN PARTICLE = "NORMAL"
Rule2	IF DAY=1MON AND SHIFT = 1^{ST} SHIFT THEN PARTICLE = "FAILED"
Rule3	IF DAY=1MON AND SHIFT = 2^{ND} SHIFT THEN PARTICLE = "FAILED"
Rule4	IF DAY=2TUE THEN PARTICLE = "FAILED"
Rule5	IF DAY=3WED THEN PARTICLE = "FAILED"
Rule6	IF DAY=4THU THEN PARTICLE = "FAILED"
Rule7	IF DAY=5FRI THEN PARTICLE = "FAILED"
Rule8	IF DAY=6SAT AND SHIFT = 3^{RD} SHIFT THEN PARTICLE = "FAILED"
Rule9	IF DAY=6SAT AND SHIFT = 1^{ST} SHIFT THEN PARTICLE = "NORMAL"
Rule10	IF DAY=6SAT AND SHIFT = 2^{ND} SHIFT THEN PARTICLE = "FAILED"
Rule11	IF DAY=7SUN THEN PARTICLE = "FAILED"

4.2.1.2 ผลการจำแนกข้อมูลของพื้นที่การทำงาน HDE

ผลจากการใช้แบบจำลองของอัลกอริทึมต้นไม้ตัดสินใจ ทดสอบกับชุดข้อมูลให้ผลการทดสอบดังตาราง

ตารางที่ 4.7 จำนวนข้อมูลสำหรับการสร้างแบบจำลองต้นไม้ตัดสินใจพื้นที่ HDE

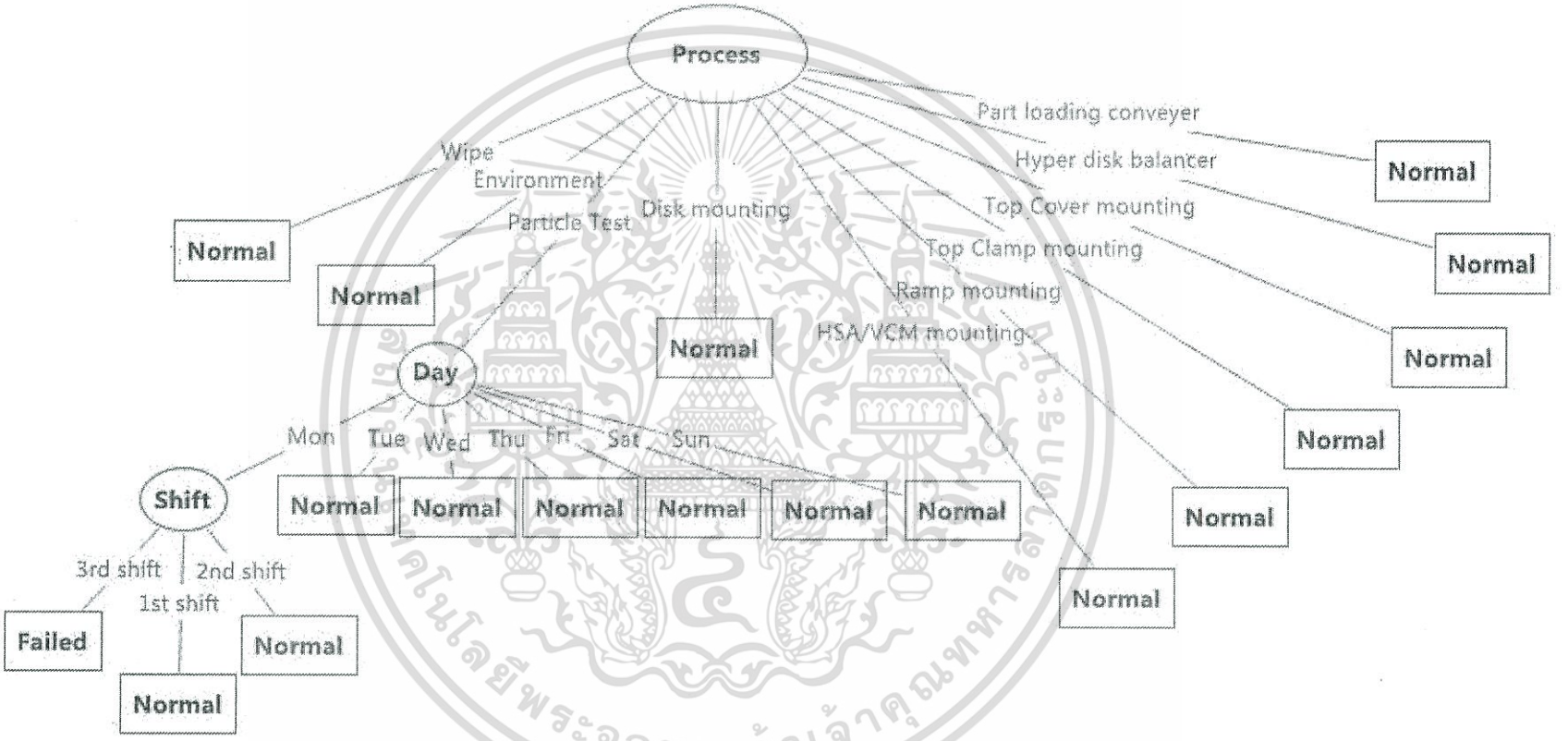
วิธีการจัดระเบียบข้อมูล	จำนวนข้อมูลพื้นที่ HDE (Records)			
	ก่อน SMOTE		หลัง SMOTE	
	Normal	Failed	Normal	Failed
เกณฑ์ของโรงงาน	261027	2406	261027	4812

ตารางที่ 4.8 ค่าความถูกต้องและค่าเฉลี่ยความคลาดเคลื่อนจากการสร้างแบบจำลองต้นไม้ตัดสินใจพื้นที่ HDE

ชุดข้อมูล	ผลการทดสอบ (%)		
	Correctly Classified Instances	Incorrectly Classified Instances	RMSE
HDE	98.5992	1.4008	11.18

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ต้นไม้ตัดสินใจข้อมูลพื้นที่การทำงาน HDE



รูปที่ 4.39 ต้นไม้ตัดสินใจพื้นที่ HDE ตามเกณฑ์การควบคุมของโรงงาน

จากรูป 4.39 พบว่าชุดข้อมูลพื้นที่ HDE ประกอบด้วย ขนาดของต้นไม้เท่ากับ 21 โหนด และลีฟโหนดมีขนาด 18 โหนด ซึ่งสามารถแปลงต้นไม้ตัดสินใจเป็นกฎในรูปแบบคำสั่ง IF-THEN ได้ดังตาราง 4.9

ตารางที่ 4.9 กฎจากต้นไม้ตัดสินใจข้อมูลพื้นที่ HDE

กฎที่	เงื่อนไขกฎ
Rule1	IF PROCESS=Wipe THEN PARTICLE = "NORMAL"
Rule2	IF PROCESS= Environment THEN PARTICLE = "NORMAL"
Rule3	IF PROCESS= Particle Test AND DAY = 1MON AND SHIFT = 3RD THEN PARTICLE = "FAILED"
Rule4	IF PROCESS= Particle Test AND DAY = 1MON AND SHIFT = 1ST THEN PARTICLE = "NORMAL"
Rule5	IF PROCESS= Particle Test AND DAY = 1MON AND SHIFT = 2ND THEN PARTICLE = "NORMAL"
Rule6	IF PROCESS= Particle Test AND DAY = 2TUE THEN PARTICLE = "NORMAL"
Rule7	IF PROCESS= Particle Test AND DAY = 3WED THEN PARTICLE = "NORMAL"
Rule8	IF PROCESS= Particle Test AND DAY = 4THU THEN PARTICLE = "NORMAL"
Rule9	IF PROCESS= Particle Test AND DAY = 5FRI THEN PARTICLE = "NORMAL"
Rule10	IF PROCESS= Particle Test AND DAY = 6SAT THEN PARTICLE = "NORMAL"
Rule11	IF PROCESS= Particle Test AND DAY = 7SUN THEN PARTICLE = "NORMAL"
Rule12	IF PROCESS= Disk mounting THEN PARTICLE = "NORMAL"
Rule13	IF PROCESS= HSA/VCM mounting THEN PARTICLE = "NORMAL"
Rule14	IF PROCESS= Ramp mounting THEN PARTICLE = "NORMAL"
Rule15	IF PROCESS= Top Clamp mounting THEN PARTICLE = "NORMAL"
Rule16	IF PROCESS= Top Cover mounting THEN PARTICLE = "NORMAL"
Rule17	IF PROCESS= Hyper disk balancer THEN PARTICLE = "NORMAL"
Rule18	IF PROCESS= Part loading conveyer THEN PARTICLE = "NORMAL"

4.2.1.3 ผลการจำแนกข้อมูลของพื้นที่การทำงาน HSA

ผลจากการใช้แบบจำลองของอัลกอริทึมต้นไม้ตัดสินใจ ทดสอบกับชุดข้อมูลให้ผลการทดสอบดังตาราง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.10 จำนวนข้อมูลสำหรับการสร้างแบบจำลองต้นไม้ตัดสินใจพื้นที่ HSA

วิธีการจัดระเบียบข้อมูล	จำนวนข้อมูลพื้นที่ HSA (Records)			
	ก่อน SMOTE		หลัง SMOTE	
	Normal	Failed	Normal	Failed
เกณฑ์ของโรงงาน	107214	31	107214	62

ตารางที่ 4.11 ค่าความถูกต้องและค่าเฉลี่ยความคลาดเคลื่อนจากการสร้างแบบจำลองต้นไม้ตัดสินใจพื้นที่ HSA

ชุดข้อมูล	ผลการทดสอบ (%)		
	Correctly Classified Instances	Incorrectly Classified Instances	RMSE
HSA	99.9711	0.0289	1.7

- ต้นไม้ตัดสินใจข้อมูลพื้นที่การทำงาน HSA

จากการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจของชุดข้อมูลพื้นที่การทำงาน HSA นั้นสร้างต้นไม้ตัดสินใจที่มีขนาด 1 โหนด จึงไม่สามารถสร้างกฎในรูปแบบคำสั่ง IF-THEN ได้ เนื่องจากข้อมูลคลาส Normal มีจำนวนมากถึง 99.97% และข้อมูลคลาส failed มีจำนวนเพียง 0.03% ซึ่งการทำให้ข้อมูลสมดุลนั้นก็ไม่สามารถสร้างกฎการจำแนกที่ถูกต้องได้ ดังนั้นจึงกล่าวได้ว่าข้อมูลปริมาณฝุ่นละอองที่จัดระเบียบตามเกณฑ์ของโรงงานมีแนวโน้มที่เกิด Normal และจากผลในตารางที่ 4.8 ค่าความถูกต้องคือ 99.9711% และจำแนกได้ต้นไม้ตัดสินใจ 1 โหนด

จากการจำแนกข้อมูลการควบคุมอนุภาคฝุ่นละอองตามเกณฑ์ของโรงงานทั้ง 3 ชุดข้อมูล ซึ่งแบ่งตามพื้นที่การทำงาน พบว่าการจำแนกชุดข้อมูลที่ควบคุมตามเกณฑ์ของโรงงานมีความถูกต้องอยู่ในระดับสูงซึ่งได้แก่ ข้อมูลพื้นที่การทำงาน HSA, HDE และ Cleaning ตามลำดับ และค่าเฉลี่ยความคลาดเคลื่อนต่ำที่สุดได้แก่ ข้อมูลพื้นที่การทำงาน HSA, HDE และ Cleaning ตามลำดับเช่นกัน โดยสามารถแสดงผลได้ดังตารางที่ 4.10 เนื่องจากข้อมูล HSA ไม่สามารถจำแนกเป็นต้นไม้ตัดสินใจที่มีขนาดของต้นไม้มากกว่า 1 โหนดได้ เพราะลักษณะข้อมูลไม่มีความสมดุล โดยมีคลาสข้อมูล Normal ถึง 99.97% ซึ่งกล่าวได้ว่าสำหรับพื้นที่การทำงาน HSA นี้ปริมาณฝุ่นละอองอยู่ในระดับปกติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.12 เปรียบเทียบค่าความถูกต้องและค่าเฉลี่ยความคลาดเคลื่อนจากการสร้างแบบจำลองต้นไม้ตัดสินใจของข้อมูลการควบคุมอนุภาคฝุ่นละอองตามเกณฑ์ของโรงงาน

ชุดข้อมูล	ผลการทดสอบ (%)		
	Correctly Classified Instances	Incorrectly Classified Instances	RMSE
Cleaning	87.8511	12.1489 %	30.76
HDE	98.5992	1.4008	11.18
HSA	99.9711	0.0289	1.7

ดังนั้นจึงสรุปได้ว่า การจำแนกข้อมูลด้วยต้นไม้ตัดสินใจของกลุ่มข้อมูลที่ควบคุมด้วยเกณฑ์ของโรงงานนั้นสามารถจำแนกได้กับข้อมูลชุดพื้นที่ Cleaning และ HDE และปัจจัยที่ทำให้เกิดปริมาณอนุภาคฝุ่นละอองเกินควบคุมได้แก่ วันในการทำงาน (Day) กระบวนการทำงาน (Process) และกะการทำงาน (Shift)

4.2.2 ผลการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจของกลุ่มข้อมูลการควบคุมอนุภาคฝุ่นละอองด้วยการควบคุมกระบวนการเชิงสถิติ

หลังจากการสร้างข้อมูลการควบคุมคุณภาพด้วยแผนภูมิควบคุมตามเกณฑ์การควบคุมทั้ง 8 เกณฑ์ของข้อมูลทั้ง 3 ชุด (ตามมิติของพื้นที่การทำงาน) ทำให้มีข้อมูลการเรียนรู้จำนวน 24 ชุดข้อมูล และเนื่องจากการวิเคราะห์เกณฑ์การควบคุมด้วยกระบวนการเชิงสถิติในหัวข้อ 4.1 พบว่าในแต่ละพื้นที่มีเกณฑ์บางข้อที่สามารถตรวจจับฝุ่นละอองได้ปริมาณมาก ดังนั้นจึงเลือกเฉพาะข้อมูลในกรณีดังกล่าวมาหาปัจจัยของการเกิดฝุ่นละอองและวิเคราะห์เกณฑ์การควบคุมกระบวนการเชิงสถิติต่อไป

4.2.2.1 ผลการจำแนกข้อมูลของพื้นที่การทำงาน cleaning

ผลจากการใช้แบบจำลองของอัลกอริทึมต้นไม้ตัดสินใจ ทดสอบกับชุดข้อมูลที่มีปริมาณเรคคอร์ดที่เกิด Out of control มาก จำนวน 5 ชุด ได้แก่ ชุดข้อมูลการเกิดฝุ่นละอองควบคุมด้วยกฎข้อ 1, 2, 5, 6 และ 8 ให้ผลการทดสอบดังตารางที่ 4.14 – 4.15

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.13 จำนวนข้อมูลสำหรับการสร้างแบบจำลองต้นไม้ตัดสินใจพื้นที่ cleaning

วิธีการจัดระเบียบข้อมูล	จำนวนข้อมูลพื้นที่ Cleaning (Records)			
	ก่อน SMOTE		หลัง SMOTE	
	Normal	Failed	Normal	Failed
เกณฑ์ของ SCP ข้อที่ 1 *	2043	2163	2043	2163
เกณฑ์ของ SCP ข้อที่ 2	924	3282	2772	3282
เกณฑ์ของ SCP ข้อที่ 5	1362	2844	2724	2844
เกณฑ์ของ SCP ข้อที่ 6	856	3350	2996	3350
เกณฑ์ของ SCP ข้อที่ 8	1554	2652	2641	2652

* เกณฑ์ของ SCP ข้อที่ 1 นั้นมีจำนวนข้อมูลที่สมดุลอยู่แล้ว ดังนั้นข้อมูลดังกล่าวจึงไม่ต้องทำ SMOTE

ตารางที่ 4.14 ค่าความถูกต้องและค่าเฉลี่ยความคลาดเคลื่อนจากการสร้างแบบจำลองต้นไม้ตัดสินใจพื้นที่ cleaning

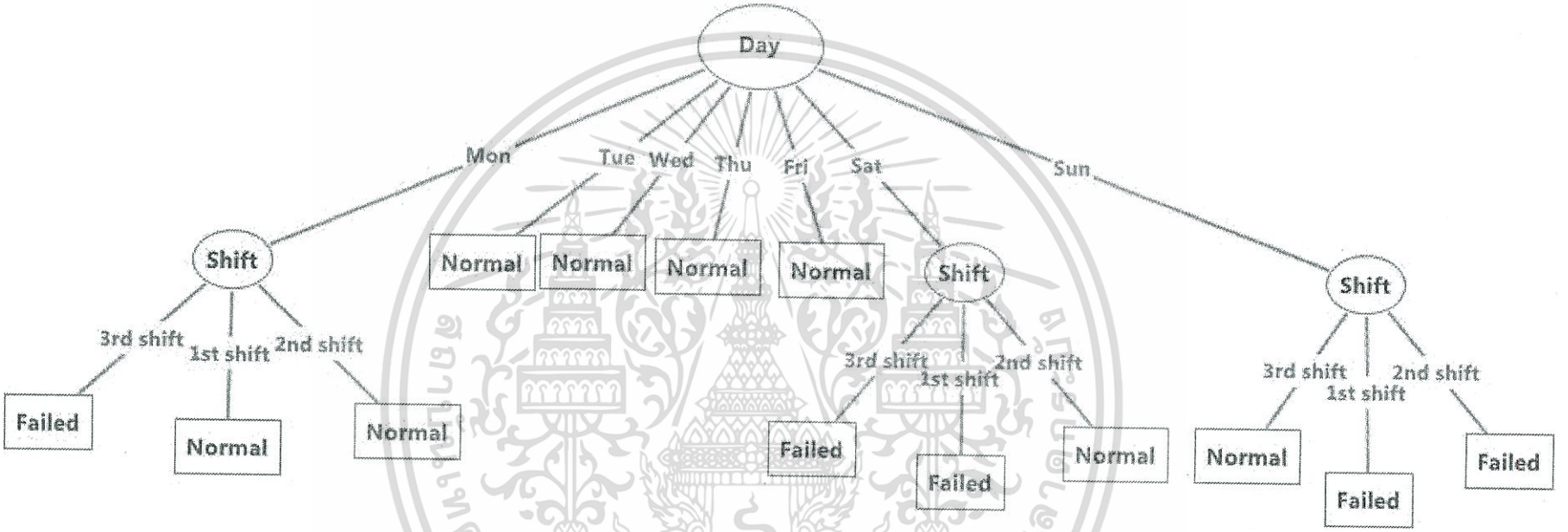
ชุดข้อมูล	ผลการทดสอบ (%)		
	Correctly Classified Instances	Incorrectly Classified Instances	RMSE
เกณฑ์ของ SCP ข้อที่ 1	64.6695	35.3305	47.33
เกณฑ์ของ SCP ข้อที่ 2	82.3257	17.6743	37.55
เกณฑ์ของ SCP ข้อที่ 5	74.2457	25.7543	41.62
เกณฑ์ของ SCP ข้อที่ 6	83.4541	16.5459	36.57
เกณฑ์ของ SCP ข้อที่ 8	68.0899	31.9101	44.15

ตารางที่ 4.15 ค่าประสิทธิภาพของแบบจำลองต้นไม้ตัดสินใจพื้นที่ cleaning

ชุดข้อมูล	ค่าประสิทธิภาพ		
	Precision	Recall	F-Measure
เกณฑ์ของ SCP ข้อที่ 1	0.669	0.647	0.63
เกณฑ์ของ SCP ข้อที่ 2	0.838	0.823	0.819
เกณฑ์ของ SCP ข้อที่ 5	0.266	0.788	0.742
เกณฑ์ของ SCP ข้อที่ 6	0.846	0.835	0.832
เกณฑ์ของ SCP ข้อที่ 8	0.704	0.681	0.671

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ต้นไม้ตัดสินในพื้นที่การทำงาน cleaning



รูปที่ 4.40 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 1 พื้นที่ cleaning

จากตารางที่ 4.14 พบว่า ความถูกต้องในการจำแนกของแบบจำลองต้นไม้ตัดสินใจอยู่ในเกณฑ์ปานกลาง – ดี ซึ่งเรียงจากมากไปหาน้อยได้แก่ เกณฑ์ของ SCP ข้อที่ 6, 2, 5, 8 และ 1 ตามลำดับ และเมื่อพิจารณาประสิทธิภาพของแบบจำลองในตารางที่ 4.15 พบว่า ประสิทธิภาพโดยรวมของแบบจำลองก็มีลักษณะเดียวกัน

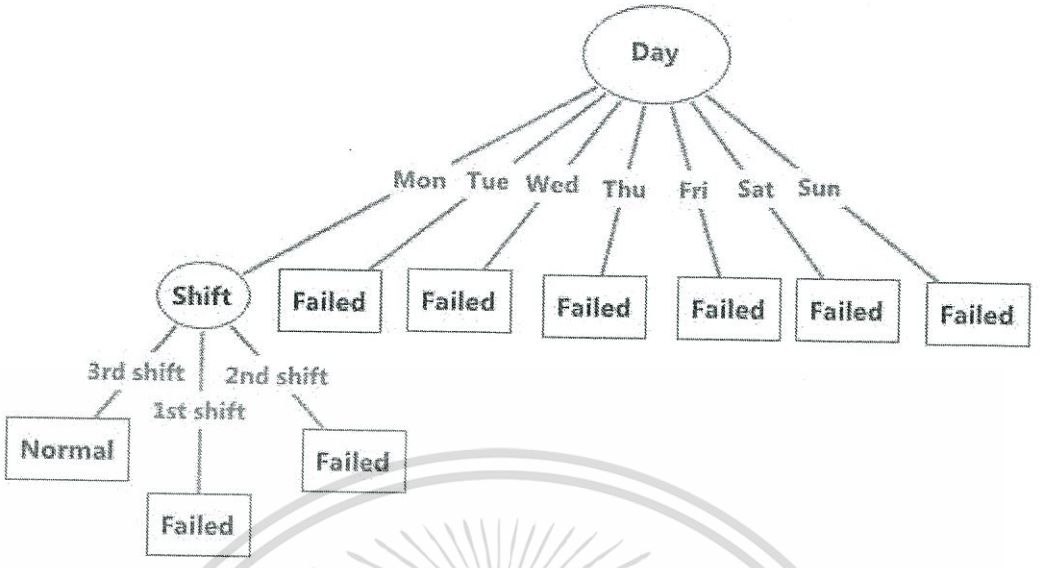
จากรูป 4.40 พบว่าชุดข้อมูลกฎข้อที่ 1 พื้นที่ cleaning ประกอบด้วย ขนาดของต้นไม้เท่ากับ 17 โหนด และลีฟโหนดมีขนาด 13 โหนด ซึ่งสามารถแปลงต้นไม้ตัดสินใจเป็นกฎในรูปแบบคำสั่ง IF-THEN ได้ดังตาราง 4.16

ตารางที่ 4.16 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 1 พื้นที่ cleaning

กฎที่	เงื่อนไขกฎ
Rule1	IF DAY=1MON AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule2	IF DAY=1MON AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule3	IF DAY=1MON AND SHIFT = 2ND SHIFT THEN PARTICLE = "NORMAL"
Rule4	IF DAY=2TUE THEN PARTICLE = "NORMAL"
Rule5	IF DAY=3WED THEN PARTICLE = "NORMAL"
Rule6	IF DAY=4THU THEN PARTICLE = "NORMAL"
Rule7	IF DAY=5FRI THEN PARTICLE = "NORMAL"
Rule8	IF DAY=6SAT AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule9	IF DAY=6SAT AND SHIFT = 1ST SHIFT THEN PARTICLE = "FAILED"
Rule10	IF DAY=6SAT AND SHIFT = 2ND SHIFT THEN PARTICLE = "NORMAL"
Rule11	IF DAY=7SUN AND SHIFT = 3RD SHIFT THEN PARTICLE = "NORMAL"
Rule12	IF DAY=7SUN AND SHIFT = 1ST SHIFT THEN PARTICLE = "FAILED"
Rule13	IF DAY=7SUN AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ต้นไม้ตัดสินในพื้นที่การทำงาน cleaning



รูปที่ 4.41 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 2 พื้นที่ cleaning

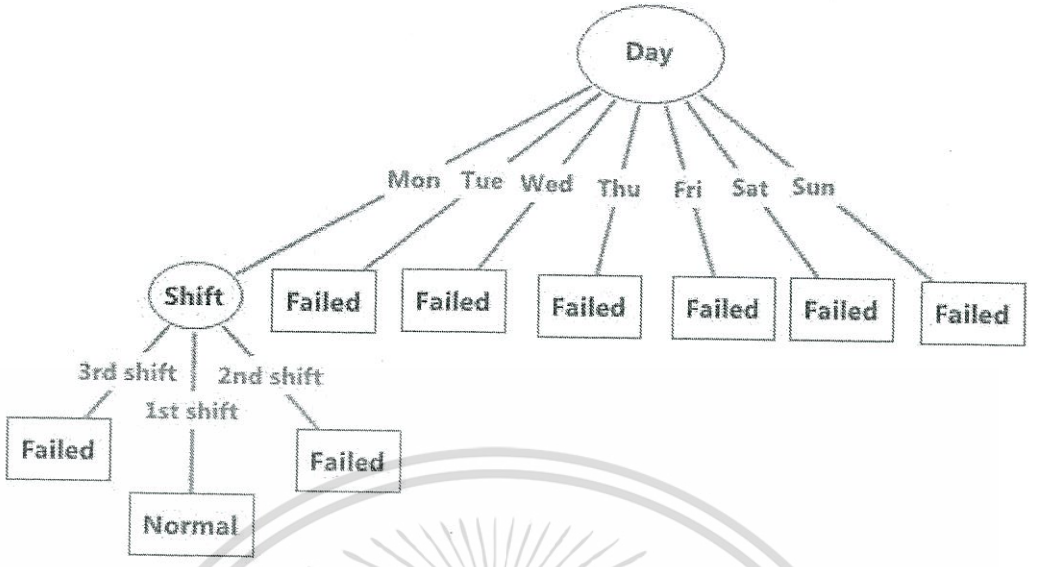
จากรูป 4.41 พบว่าชุดข้อมูลกฎข้อที่ 2 พื้นที่ cleaning ประกอบด้วย ขนาดของต้นไม้เท่ากับ 11 โหนด และลีฟโหนดมีขนาด 9 โหนด ซึ่งสามารถแปลงต้นไม้ตัดสินใจเป็นกฎในรูปแบบคำสั่ง IF-THEN ได้ดังตาราง 4.17

ตารางที่ 4.17 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 2 พื้นที่ cleaning

กฎที่	เงื่อนไขกฎ
Rule1	IF DAY = 1MON AND SHIFT= 3^{RD} AND THEN PARTICLE = "NORMAL"
Rule2	IF DAY = 1MON AND SHIFT= 1^{ST} AND THEN PARTICLE = "FAILED"
Rule3	IF DAY = 1MON AND SHIFT= 2^{ND} AND THEN PARTICLE = "FAILED"
Rule4	IF DAY = 2TUE THEN PARTICLE = "FAILED"
Rule5	IF DAY = 3WED THEN PARTICLE = "FAILED"
Rule6	IF DAY = 4THU THEN PARTICLE = "FAILED"
Rule7	IF DAY = 5FRI THEN PARTICLE = "FAILED"
Rule8	IF DAY = 6SAT THEN PARTICLE = "FAILED"
Rule9	IF DAY = 7SUN THEN PARTICLE = "FAILED"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ต้นไม้ตัดสินในพื้นที่การทำงาน cleaning



รูปที่ 4.42 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 5 พื้นที่ cleaning

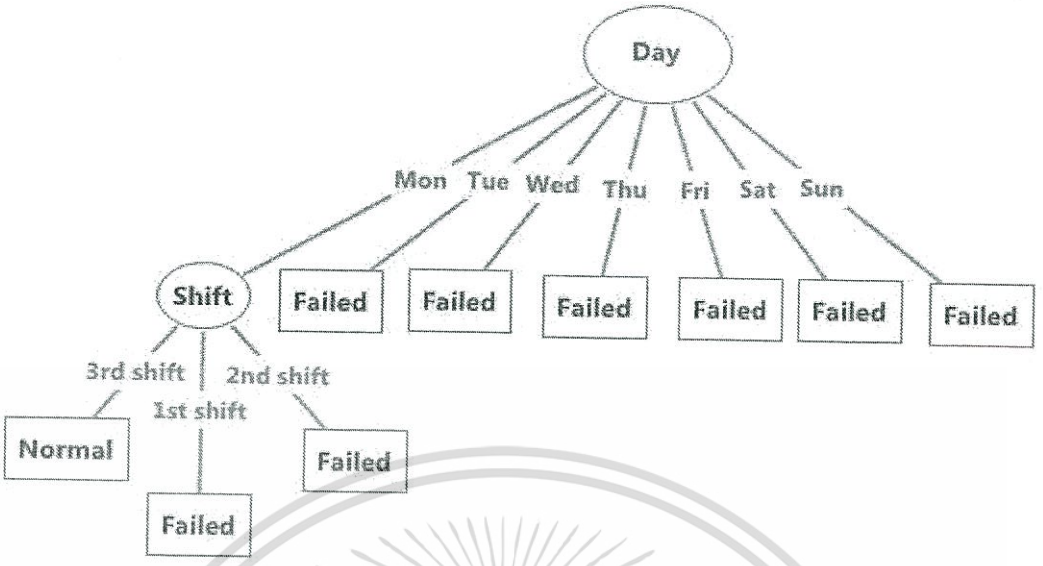
จากรูป 4.42 พบว่าชุดข้อมูลกฎข้อที่ 5 พื้นที่ cleaning ประกอบด้วย ขนาดของต้นไม้เท่ากับ 11 โหนด และลีฟโหนดมีขนาด 9 โหนด ซึ่งสามารถแปลงต้นไม้ตัดสินใจเป็นกฎในรูปแบบคำสั่ง IF-THEN ได้ดังตาราง 4.18

ตารางที่ 4.18 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 5 พื้นที่ cleaning

กฎที่	เงื่อนไขกฎ
Rule1	IF DAY = 1MON AND SHIFT= 3^{RD} AND THEN PARTICLE = "FAILED"
Rule2	IF DAY = 1MON AND SHIFT= 1^{ST} AND THEN PARTICLE = "NORMAL"
Rule3	IF DAY = 1MON AND SHIFT= 2^{ND} AND THEN PARTICLE = "FAILED"
Rule4	IF DAY = 2TUE THEN PARTICLE = "FAILED"
Rule5	IF DAY = 3WED THEN PARTICLE = "FAILED"
Rule6	IF DAY = 4THU THEN PARTICLE = "FAILED"
Rule7	IF DAY = 5FRI THEN PARTICLE = "FAILED"
Rule8	IF DAY = 6SAT THEN PARTICLE = "FAILED"
Rule9	IF DAY = 7SUN THEN PARTICLE = "FAILED"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ต้นไม้ตัดสินในพื้นที่การทำงาน cleaning



รูปที่ 4.43 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 6 พื้นที่ cleaning

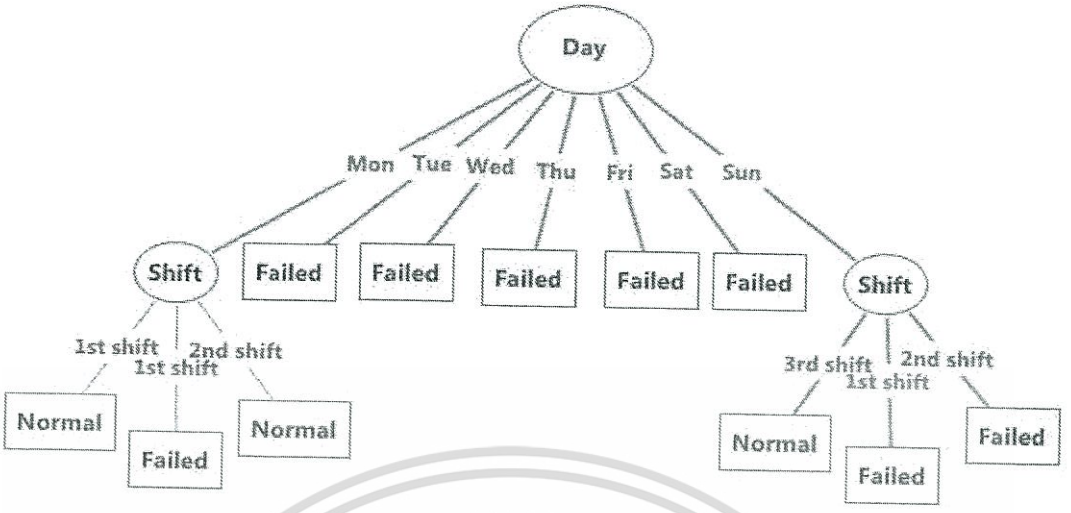
จากรูป 4.43 พบว่าชุดข้อมูลกฎข้อที่ 6 พื้นที่ cleaning ประกอบด้วย ขนาดของต้นไม้เท่ากับ 11 โหนด และลีฟโหนดมีขนาด 9 โหนด ซึ่งสามารถแปลงต้นไม้ตัดสินใจเป็นกฎในรูปแบบคำสั่ง IF-THEN ได้ดังตาราง 4.19

ตารางที่ 4.19 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 6 พื้นที่ cleaning

กฎที่	เงื่อนไขกฎ
Rule1	IF DAY = 1MON AND SHIFT= 3^{RD} AND THEN PARTICLE = "NORMAL"
Rule2	IF DAY = 1MON AND SHIFT= 1^{ST} AND THEN PARTICLE = "FAILED"
Rule3	IF DAY = 1MON AND SHIFT= 2^{ND} AND THEN PARTICLE = "FAILED"
Rule4	IF DAY = 2TUE THEN PARTICLE = "FAILED"
Rule5	IF DAY = 3WED THEN PARTICLE = "FAILED"
Rule6	IF DAY = 4THU THEN PARTICLE = "FAILED"
Rule7	IF DAY = 5FRI THEN PARTICLE = "FAILED"
Rule8	IF DAY = 6SAT THEN PARTICLE = "FAILED"
Rule9	IF DAY = 7SUN THEN PARTICLE = "FAILED"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ต้นไม้ตัดสินในพื้นที่การทำงาน cleaning



รูปที่ 4.44 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 8 พื้นที่ cleaning

จากรูป 4.44 พบว่าชุดข้อมูลกฎข้อที่ 8 พื้นที่ cleaning ประกอบด้วย ขนาดของต้นไม้เท่ากับ 14 โหนด และลีฟโหนดมีขนาด 11 โหนด ซึ่งสามารถแปลงต้นไม้ตัดสินใจเป็นกฎในรูปแบบคำสั่ง IF-THEN ได้ดังตาราง 4.20

ตารางที่ 4.20 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 8 พื้นที่ cleaning

กฎที่	เงื่อนไขกฎ
Rule1	IF DAY = 1MON AND SHIFT= 3^{RD} AND THEN PARTICLE = "NORMAL"
Rule2	IF DAY = 1MON AND SHIFT= 1^{ST} AND THEN PARTICLE = "FAILED"
Rule3	IF DAY = 1MON AND SHIFT= 2^{ND} AND THEN PARTICLE = "FAILED"
Rule4	IF DAY = 2TUE THEN PARTICLE = "FAILED"
Rule5	IF DAY = 3WED THEN PARTICLE = "FAILED"
Rule6	IF DAY = 4THU THEN PARTICLE = "FAILED"
Rule7	IF DAY = 5FRI THEN PARTICLE = "FAILED"
Rule8	IF DAY = 6SAT THEN PARTICLE = "FAILED"
Rule9	IF DAY = 7SUN AND SHIFT= 3^{RD} AND THEN PARTICLE = "NORMAL"
Rule10	IF DAY = 7SUN AND SHIFT= 1^{ST} AND THEN PARTICLE = "FAILED"
Rule11	IF DAY = 7SUN AND SHIFT= 2^{ND} AND THEN PARTICLE = "FAILED"

จากการสร้างแบบจำลองต้นไม้ตัดสินใจของชุดข้อมูลพื้นที่การทำงาน Cleaning ทั้ง 5 ชุดข้อมูลพบว่า ปัจจัยที่เกี่ยวข้องกับการเกิดฝุ่นละอองในพื้นที่การทำงาน cleaning ได้แก่ วันที่ทำงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(Day) และ กะการทำงาน (Shift) ซึ่งสำหรับการเลือกแบบจำลองและเกณฑ์การควบคุมกระบวนการเชิงสถิติที่เหมาะสมสำหรับพื้นที่ cleaning จะได้กล่าวต่อไป

4.2.2.2 ผลการจำแนกข้อมูลของชุดข้อมูลพื้นที่การทำงาน HDE

ผลจากการใช้แบบจำลองของอัลกอริทึมต้นไม้ตัดสินใจ ทดสอบกับชุดข้อมูลที่มีปริมาณเรคคอร์ดที่เกิด Out of control มาก จำนวน 2 ชุด ได้แก่ ชุดข้อมูลการเกิดฝุ่นละอองควบคุมด้วยกฎข้อ 2 และ 7 ให้ผลการทดสอบดังตารางที่ 4.22 – 4.23

ตารางที่ 4.21 จำนวนข้อมูลสำหรับการสร้างแบบจำลองต้นไม้ตัดสินใจพื้นที่ HDE

วิธีการจัดระเบียบข้อมูล	จำนวนข้อมูลพื้นที่ Cleaning (Records)			
	ก่อน Sampling		หลัง Sampling	
	Normal	Failed	Normal	Failed
เกณฑ์ของ SCP ข้อที่ 2*	35018	228415	35018	35018
เกณฑ์ของ SCP ข้อที่ 7*	30413	233020	30413	30413

* ใช้วิธี Under-sampling ข้อมูลเพื่อแก้ปัญหาข้อมูลไม่สมดุล

ตารางที่ 4.22 ค่าความถูกต้องและค่าเฉลี่ยความคลาดเคลื่อนจากการสร้างแบบจำลองต้นไม้ตัดสินใจพื้นที่ HDE

ชุดข้อมูล	ผลการทดสอบ (%)		
	Correctly Classified Instances	Incorrectly Classified Instances	RMSE
เกณฑ์ของ SCP ข้อที่ 2	62.872	37.128	47.92
เกณฑ์ของ SCP ข้อที่ 7	63.1473	36.8527	47.81

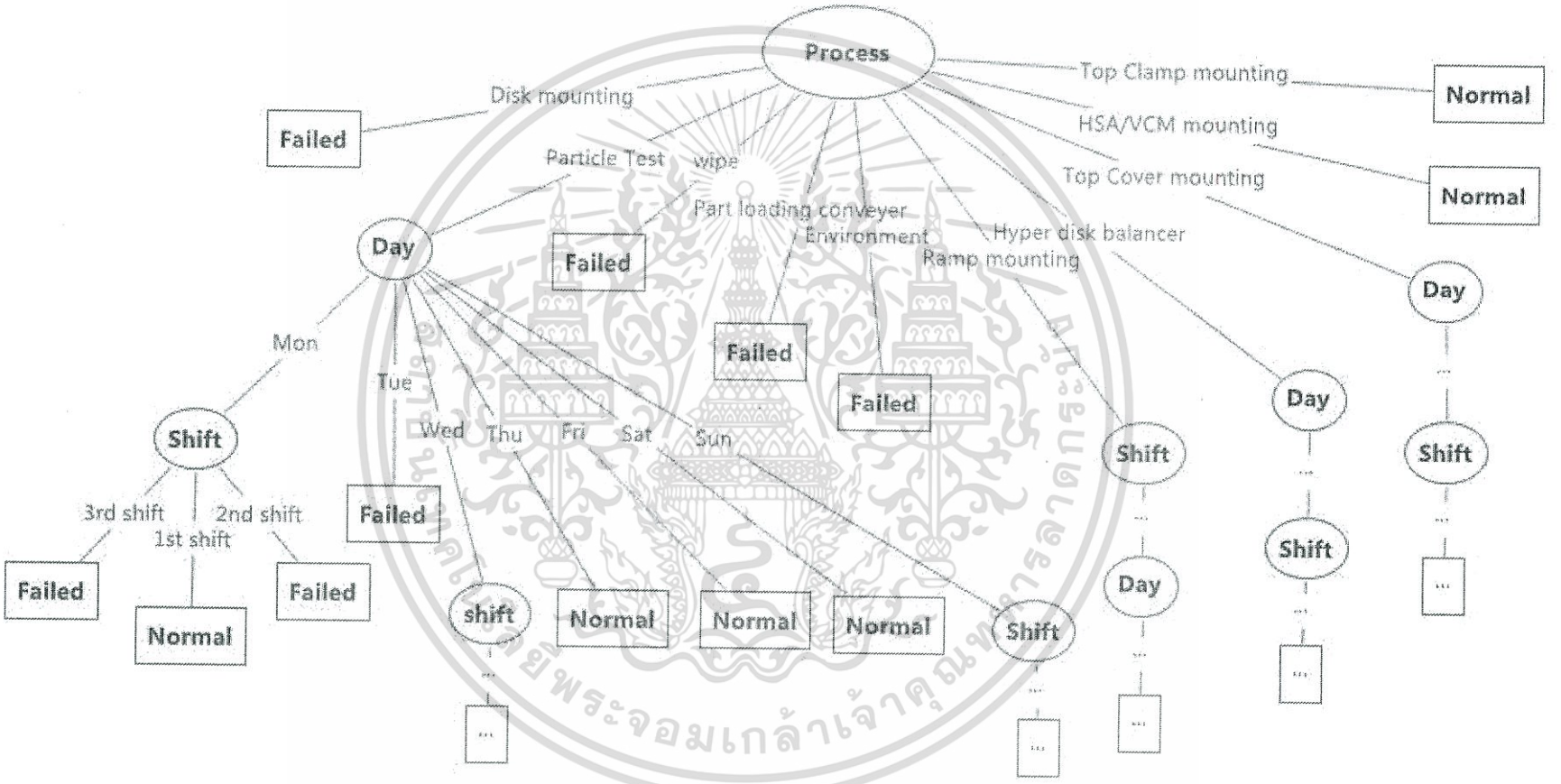
ตารางที่ 4.23 ค่าประสิทธิภาพของแบบจำลองต้นไม้ตัดสินใจพื้นที่ HDE

ชุดข้อมูล	ค่าประสิทธิภาพ		
	Precision	Recall	F-Measure
เกณฑ์ของ SCP ข้อที่ 2	0.63	0.629	0.628
เกณฑ์ของ SCP ข้อที่ 7	0.636	0.631	0.629

จากตารางที่ 4.22 พบว่า ความถูกต้องในการจำแนกของแบบจำลองต้นไม้ตัดสินใจอยู่ในเกณฑ์ดีมาก ซึ่งเรียงจากมากไปหาน้อยได้แก่ เกณฑ์ของ SCP ข้อที่ 7 และ 2 ตามลำดับ และเมื่อพิจารณาประสิทธิภาพของแบบจำลองในตารางที่ 4.23 พบว่า ประสิทธิภาพโดยรวมของแบบจำลองพบว่ามีประสิทธิภาพโดยรวมใกล้เคียงกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ต้นไม้ตัดสินในพื้นที่การทำงาน HDE



รูปที่ 4.45 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 2 พื้นที่ HDE

จากรูป 4.45 พบว่าชุดข้อมูลกฎข้อที่ 2 พื้นที่ HDE ประกอบด้วย ขนาดของต้นไม้เท่ากับ 69 โหนด และลีฟโหนดมีขนาด 54 โหนด ซึ่งสามารถแปลงต้นไม้ตัดสินใจเป็นกฎในรูปแบบคำสั่ง IF-THEN ได้ดังตาราง 4.24

ตารางที่ 4.24 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 2 พื้นที่ HDE

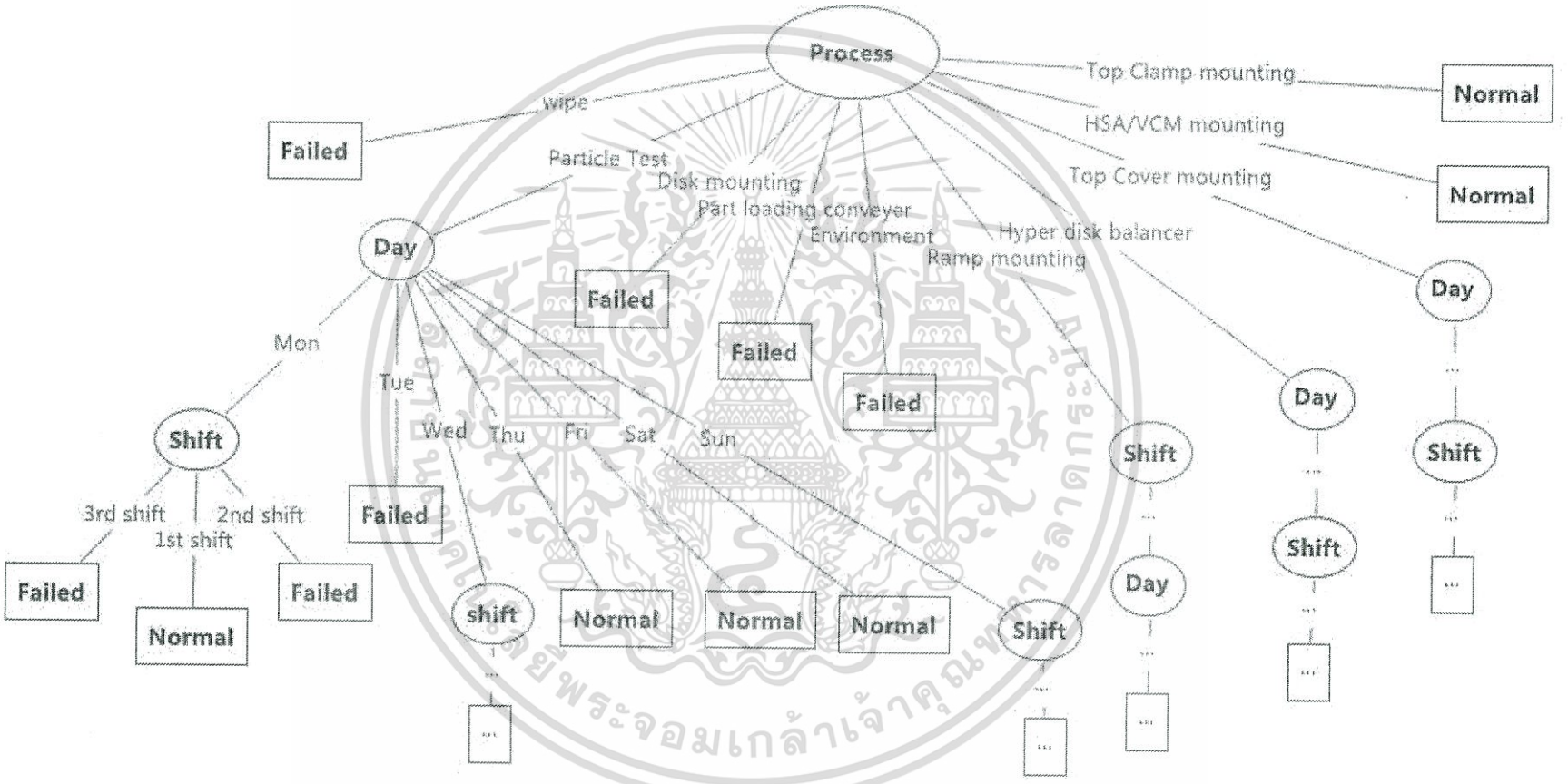
กฎที่	เงื่อนไขกฎ
Rule1	IF PROCESS = Disk mounting THEN PARTICLE = "FAILED"
Rule2	IF PROCESS = Ramp mounting AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule3	IF PROCESS = Ramp mounting AND SHIFT = 2ND SHIFT AND DAY = 1MON THEN PARTICLE = "NORMAL"
Rule4	IF PROCESS = Ramp mounting AND SHIFT = 2ND SHIFT AND DAY = 2TUE THEN PARTICLE = "FAILED"
Rule5	IF PROCESS = Ramp mounting AND SHIFT = 2ND SHIFT AND DAY = 3WED THEN PARTICLE = "FAILED"
Rule6	IF PROCESS = Ramp mounting AND SHIFT = 2ND SHIFT AND DAY = 4THU THEN PARTICLE = "FAILED"
Rule7	IF PROCESS = Ramp mounting AND SHIFT = 2ND SHIFT AND DAY = 5FRI THEN PARTICLE = "FAILED"
Rule8	IF PROCESS = Ramp mounting AND SHIFT = 2ND SHIFT AND DAY = 6SAT THEN PARTICLE = "FAILED"
Rule9	IF PROCESS = Ramp mounting AND SHIFT = 2ND SHIFT AND DAY = 7SUN THEN PARTICLE = "FAILED"
Rule10	IF PROCESS = Ramp mounting AND SHIFT = 3RD SHIFT THEN PARTICLE = "NORMAL"
Rule11	IF PROCESS = Particle Test AND DAY = 1MON AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule12	IF PROCESS = Particle Test AND DAY = 1MON AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule13	IF PROCESS = Particle Test AND DAY = 1MON AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule14	IF PROCESS = Particle Test AND DAY = 2TUE THEN PARTICLE = "FAILED"
Rule15	IF PROCESS = Particle Test AND DAY = 3WED AND SHIFT = 1ST SHIFT THEN PARTICLE = "FAILED"
Rule16	IF PROCESS = Particle Test AND DAY = 3WED AND SHIFT = 2ND SHIFT THEN PARTICLE = "NORMAL"
Rule17	IF PROCESS = Particle Test AND DAY = 3WED AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule18	IF PROCESS = Particle Test AND DAY = 4THU THEN PARTICLE = "NORMAL"
Rule19	IF PROCESS = Particle Test AND DAY = 5FRI THEN PARTICLE = "NORMAL"
Rule20	IF PROCESS = Particle Test AND DAY = 6SAT THEN PARTICLE = "NORMAL"
Rule21	IF PROCESS = Particle Test AND DAY = 7SUN AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule22	IF PROCESS = Particle Test AND DAY = 7SUN AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule23	IF PROCESS = Particle Test AND DAY = 7SUN AND SHIFT = 3RD SHIFT THEN PARTICLE = "NORMAL"
Rule24	IF PROCESS = Wipe THEN PARTICLE = "FAILED"
Rule25	IF PROCESS = HSA/VCM mounting THEN PARTICLE = "NORMAL"
Rule26	IF PROCESS = Environment THEN PARTICLE = "FAILED"
Rule27	IF PROCESS = Part loading conveyer THEN PARTICLE = "FAILED"
Rule28	IF PROCESS = Hyper disk balancer AND DAY = 1MON THEN PARTICLE = "FAILED"
Rule29	IF PROCESS = Hyper disk balancer AND DAY = 2TUE THEN PARTICLE = "NORMAL"
Rule30	IF PROCESS = Hyper disk balancer AND DAY = 3WED AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule31	IF PROCESS = Hyper disk balancer AND DAY = 3WED AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule32	IF PROCESS = Hyper disk balancer AND DAY = 3WED AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule33	IF PROCESS = Hyper disk balancer AND DAY = 4THU AND SHIFT = 1ST SHIFT THEN PARTICLE = "FAILED"
Rule34	IF PROCESS = Hyper disk balancer AND DAY = 4THU AND SHIFT = 2ND SHIFT THEN PARTICLE = "NORMAL"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Rule35	IF PROCESS = Hyper disk balancer AND DAY = 4THU AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule36	IF PROCESS = Hyper disk balancer AND DAY = 5FRI AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule37	IF PROCESS = Hyper disk balancer AND DAY = 5FRI AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule38	IF PROCESS = Hyper disk balancer AND DAY = 5FRI AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule39	IF PROCESS = Hyper disk balancer AND DAY = 6SAT THEN PARTICLE = "FAILED"
Rule40	IF PROCESS = Hyper disk balancer AND DAY = 7SUN THEN PARTICLE = "FAILED"
Rule41	IF PROCESS = Top Clamp mounting THEN PARTICLE = "NORMAL"
Rule42	IF PROCESS = Top Cover mounting AND DAY = 1MON THEN PARTICLE = "FAILED"
Rule43	IF PROCESS = Top Cover mounting AND DAY = 2TUE AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule44	IF PROCESS = Top Cover mounting AND DAY = 2TUE AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule45	IF PROCESS = Top Cover mounting AND DAY = 2TUE AND SHIFT = 3RD SHIFT THEN PARTICLE = "NORMAL"
Rule46	IF PROCESS = Top Cover mounting AND DAY = 3WED THEN PARTICLE = "FAILED"
Rule47	IF PROCESS = Top Cover mounting AND DAY = 4THU THEN PARTICLE = "FAILED"
Rule49	IF PROCESS = Top Cover mounting AND DAY = 5FRI AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule49	IF PROCESS = Top Cover mounting AND DAY = 5FRI AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule50	IF PROCESS = Top Cover mounting AND DAY = 5FRI AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule51	IF PROCESS = Top Cover mounting AND DAY = 6SAT AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule55	IF PROCESS = Top Cover mounting AND DAY = 6SAT AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule53	IF PROCESS = Top Cover mounting AND DAY = 6SAT AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule54	IF PROCESS = Top Cover mounting AND DAY = 7SUN THEN PARTICLE = "FAILED"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ต้นไม้ตัดสินในพื้นที่การทำงาน HDE



รูปที่ 4.46 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 7 พื้นที่ HDE

จากรูป 4.46 พบว่าชุดข้อมูลกฎข้อที่ 7 พื้นที่ HDE ประกอบด้วย ขนาดของต้นไม้เท่ากับ 79 โหนด และลีฟโหนดมีขนาด 62 โหนด ซึ่งสามารถแปลงต้นไม้ตัดสินใจเป็นกฎในรูปแบบคำสั่ง IF-THEN ได้ดังตาราง 4.25

ตารางที่ 4.25 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 7 พื้นที่ HDE

กฎที่	เงื่อนไขกฎ
Rule1	IF PROCESS = Wipe THEN PARTICLE = "FAILED"
Rule2	IF PROCESS = Environment THEN PARTICLE = "FAILED"
Rule3	IF PROCESS = Disk mounting THEN PARTICLE = "FAILED"
Rule4	IF PROCESS = Part loading conveyer THEN PARTICLE = "FAILED"
Rule5	IF PROCESS = Ramp mounting AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule6	IF PROCESS = Ramp mounting AND SHIFT = 2ND SHIFT AND DAY = 1MON THEN PARTICLE = "NORMAL"
Rule7	IF PROCESS = Ramp mounting AND SHIFT = 2ND SHIFT AND DAY = 2TUE THEN PARTICLE = "FAILED"
Rule8	IF PROCESS = Ramp mounting AND SHIFT = 2ND SHIFT AND DAY = 3WED THEN PARTICLE = "NORMAL"
Rule9	IF PROCESS = Ramp mounting AND SHIFT = 2ND SHIFT AND DAY = 4THU THEN PARTICLE = "FAILED"
Rule10	IF PROCESS = Ramp mounting AND SHIFT = 2ND SHIFT AND DAY = 5FRI THEN PARTICLE = "FAILED"
Rule11	IF PROCESS = Ramp mounting AND SHIFT = 2ND SHIFT AND DAY = 6SAT THEN PARTICLE = "NORMAL"
Rule12	IF PROCESS = Ramp mounting AND SHIFT = 2ND SHIFT AND DAY = 7SUN THEN PARTICLE = "FAILED"
Rule13	IF PROCESS = Ramp mounting AND SHIFT = 1ST SHIFT AND DAY = 1MON THEN PARTICLE = "NORMAL"
Rule14	IF PROCESS = Ramp mounting AND SHIFT = 1ST SHIFT AND DAY = 2TUE THEN PARTICLE = "NORMAL"
Rule15	IF PROCESS = Ramp mounting AND SHIFT = 1ST SHIFT AND DAY = 3WED THEN PARTICLE = "NORMAL"
Rule16	IF PROCESS = Ramp mounting AND SHIFT = 1ST SHIFT AND DAY = 4THU THEN PARTICLE = "FAILED"
Rule17	IF PROCESS = Ramp mounting AND SHIFT = 1ST SHIFT AND DAY = 5FRI THEN PARTICLE = "NORMAL"
Rule18	IF PROCESS = Ramp mounting AND SHIFT = 1ST SHIFT AND DAY = 6SAT THEN PARTICLE = "NORMAL"
Rule19	IF PROCESS = Ramp mounting AND SHIFT = 1ST SHIFT AND DAY = 7SUN THEN PARTICLE = "NORMAL"
Rule20	IF PROCESS = Top Clamp mounting THEN PARTICLE = "NORMAL"
Rule21	IF PROCESS = Particle Test AND DAY = 1MON AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule22	IF PROCESS = Particle Test AND DAY = 1MON AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule23	IF PROCESS = Particle Test AND DAY = 1MON AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule24	IF PROCESS = Particle Test AND DAY = 2TUE THEN PARTICLE = "FAILED"
Rule25	IF PROCESS = Particle Test AND DAY = 3WED AND SHIFT = 1ST SHIFT THEN PARTICLE = "FAILED"
Rule26	IF PROCESS = Particle Test AND DAY = 3WED AND SHIFT = 2ND SHIFT THEN PARTICLE = "NORMAL"
Rule27	IF PROCESS = Particle Test AND DAY = 3WED AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule28	IF PROCESS = Particle Test AND DAY = 4THU THEN PARTICLE = "FAILED"
Rule29	IF PROCESS = Particle Test AND DAY = 5FRI THEN PARTICLE = "NORMAL"
Rule30	IF PROCESS = Particle Test AND DAY = 6SAT THEN PARTICLE = "NORMAL"
Rule31	IF PROCESS = Particle Test AND DAY = 7SUN AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule32	IF PROCESS = Particle Test AND DAY = 7SUN AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule33	IF PROCESS = Particle Test AND DAY = 7SUN AND SHIFT = 3RD SHIFT THEN PARTICLE = "NORMAL"
Rule34	IF PROCESS = HSA/VCM mounting THEN PARTICLE = "NORMAL"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Rule35	IF PROCESS = Hyper disk balancer AND DAY = 1MON AND SHIFT = 1ST SHIFT THEN PARTICLE = "FAILED"
Rule36	IF PROCESS = Hyper disk balancer AND DAY = 1MON AND SHIFT = 2ND SHIFT THEN PARTICLE = "NORMAL"
Rule37	IF PROCESS = Hyper disk balancer AND DAY = 1MON AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule38	IF PROCESS = Hyper disk balancer AND DAY = 2TUE THEN PARTICLE = "NORMAL"
Rule39	IF PROCESS = Hyper disk balancer AND DAY = 3WED AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule40	IF PROCESS = Hyper disk balancer AND DAY = 3WED AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule41	IF PROCESS = Hyper disk balancer AND DAY = 3WED AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule42	IF PROCESS = Hyper disk balancer AND DAY = 4THU THEN PARTICLE = "NORMAL"
Rule43	IF PROCESS = Hyper disk balancer AND DAY = 5FRI AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule44	IF PROCESS = Hyper disk balancer AND DAY = 5FRI AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule45	IF PROCESS = Hyper disk balancer AND DAY = 5FRI AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule46	IF PROCESS = Hyper disk balancer AND DAY = 6SAT AND SHIFT = 1ST SHIFT THEN PARTICLE = "FAILED"
Rule47	IF PROCESS = Hyper disk balancer AND DAY = 6SAT AND SHIFT = 2ND SHIFT THEN PARTICLE = "NORMAL"
Rule48	IF PROCESS = Hyper disk balancer AND DAY = 6SAT AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule49	IF PROCESS = Hyper disk balancer AND DAY = 7SUN THEN PARTICLE = "FAILED"
Rule50	IF PROCESS = Top Cover mounting AND DAY = 1MON AND SHIFT = 1ST SHIFT THEN PARTICLE = "FAILED"
Rule51	IF PROCESS = Top Cover mounting AND DAY = 1MON AND SHIFT = 2ND SHIFT THEN PARTICLE = "NORMAL"
Rule52	IF PROCESS = Top Cover mounting AND DAY = 1MON AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule53	IF PROCESS = Top Cover mounting AND DAY = 2TUE AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule54	IF PROCESS = Top Cover mounting AND DAY = 2TUE AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule55	IF PROCESS = Top Cover mounting AND DAY = 2TUE AND SHIFT = 3RD SHIFT THEN PARTICLE = "NORMAL"
Rule56	IF PROCESS = Top Cover mounting AND DAY = 3WED AND SHIFT = 1ST SHIFT THEN PARTICLE = "FAILED"
Rule57	IF PROCESS = Top Cover mounting AND DAY = 3WED AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule58	IF PROCESS = Top Cover mounting AND DAY = 3WED AND SHIFT = 3RD SHIFT THEN PARTICLE = "NORMAL"
Rule59	IF PROCESS = Top Cover mounting AND DAY = 4THU THEN PARTICLE = "FAILED"
Rule60	IF PROCESS = Top Cover mounting AND DAY = 5FRI THEN PARTICLE = "NORMAL"
Rule61	IF PROCESS = Top Cover mounting AND DAY = 6SAT THEN PARTICLE = "NORMAL"
Rule62	IF PROCESS = Top Cover mounting AND DAY = 7SUN THEN PARTICLE = "FAILED"

จากการสร้างแบบจำลองต้นไม้ตัดสินใจของชุดข้อมูลพื้นที่การทำงาน HDE ทั้ง 2 ชุดข้อมูล พบว่า ปัจจัยที่เกี่ยวข้องกับการเกิดฝุ่นละอองในพื้นที่การทำงาน cleaning ได้แก่ กระบวนการทำงาน (Process) วันที่ทำงาน (Day) และ กะการทำงาน (Shift) ซึ่งสำหรับการเลือกแบบจำลองและเกณฑ์การควบคุมกระบวนการเชิงสถิติที่เหมาะสมสำหรับพื้นที่ HDE จะได้กล่าวต่อไป

4.2.2.3 ผลการจำแนกข้อมูลของชุดข้อมูลพื้นที่การทำงาน HSA

ผลจากการใช้แบบจำลองของอัลกอริทึมต้นไม้ตัดสินใจ ทดสอบกับชุดข้อมูลที่มีปริมาณเรคคอร์ดที่เกิด Out of control มาก จำนวน 4 ชุด ได้แก่ ชุดข้อมูลการเกิดฝุ่นละอองควบคุมด้วยกฎข้อ 2, 6, 7 และ 8 ให้ผลการทดสอบดังตารางที่ 4.27 – 4.28

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.26 จำนวนข้อมูลสำหรับการสร้างแบบจำลองต้นไม้ตัดสินใจพื้นที่ HSA

วิธีการจัดระเบียบข้อมูล	จำนวนข้อมูลพื้นที่ Cleaning (Records)			
	ก่อน SMOTE		หลัง SMOTE	
	Normal	Failed	Normal	Failed
เกณฑ์ของ SCP ข้อที่ 2	37886	69359	64406	69359
เกณฑ์ของ SCP ข้อที่ 6	71142	36103	71142	61375
เกณฑ์ของ SCP ข้อที่ 7	92561	14684	92561	88104
เกณฑ์ของ SCP ข้อที่ 8	82709	24536	82709	73608

ตารางที่ 4.27 ค่าความถูกต้องและค่าเฉลี่ยความคลาดเคลื่อนจากการสร้างแบบจำลองต้นไม้ตัดสินใจพื้นที่ HSA

ชุดข้อมูล	ผลการทดสอบ (%)		
	Correctly Classified Instances	Incorrectly Classified Instances	RMSE
เกณฑ์ของ SCP ข้อที่ 2	71.7549	28.2451	41.69
เกณฑ์ของ SCP ข้อที่ 6	82.5139	17.4861	36.82
เกณฑ์ของ SCP ข้อที่ 7	90.3125	9.6875	29.19
เกณฑ์ของ SCP ข้อที่ 8	89.3601	10.6399	29.97

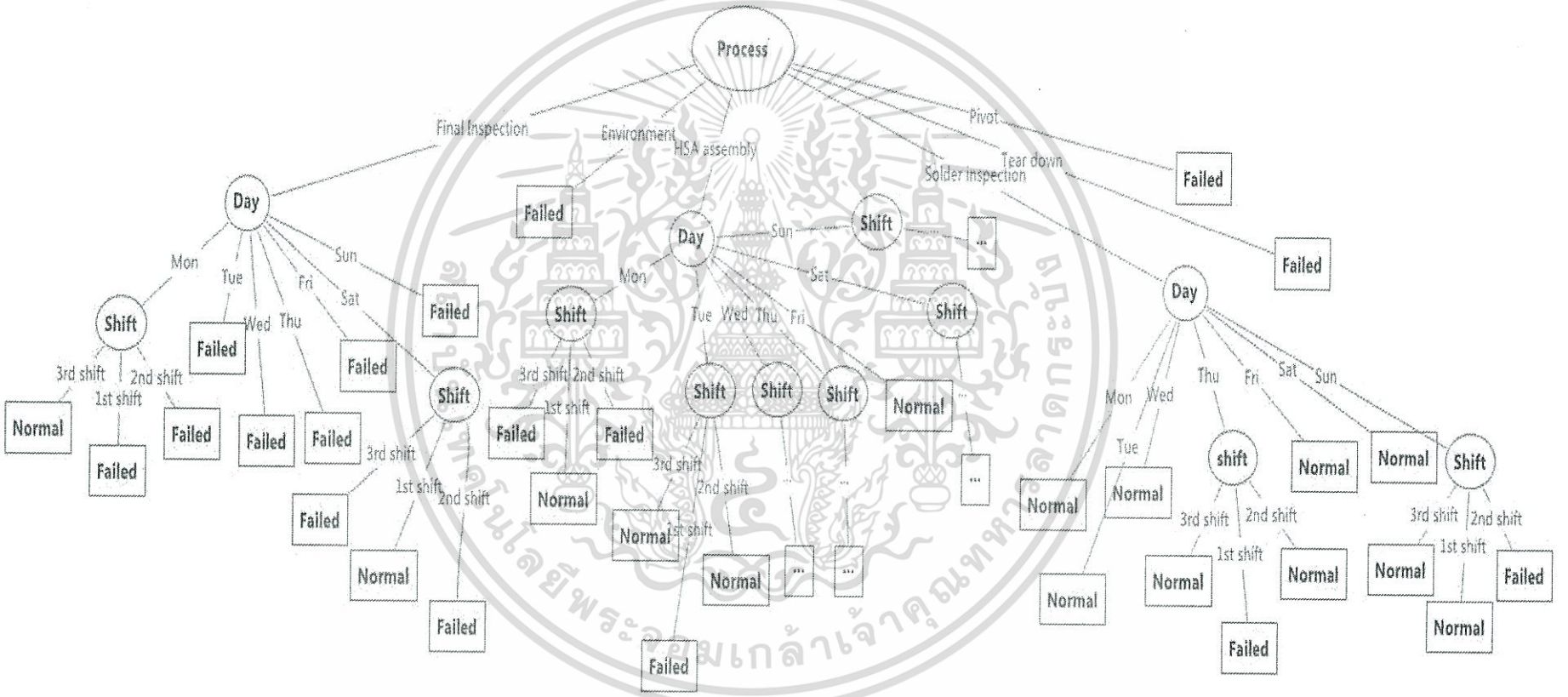
ตารางที่ 4.28 ค่าประสิทธิภาพของแบบจำลองต้นไม้ตัดสินใจพื้นที่ HSA

ชุดข้อมูล	ค่าประสิทธิภาพ		
	Precision	Recall	F-Measure
เกณฑ์ของ SCP ข้อที่ 2	0.753	0.718	0.704
เกณฑ์ของ SCP ข้อที่ 6	0.839	0.825	0.821
เกณฑ์ของ SCP ข้อที่ 7	0.908	0.903	0.903
เกณฑ์ของ SCP ข้อที่ 8	0.895	0.894	0.893

จากตารางที่ 4.27 พบว่า ความถูกต้องในการจำแนกของแบบจำลองต้นไม้ตัดสินใจอยู่ในเกณฑ์ดี – ดีมาก ซึ่งเรียงจากมากไปหาน้อยได้แก่ เกณฑ์ของ SCP ข้อที่ 7, 8, 6 และ 2 ตามลำดับ และเมื่อพิจารณาประสิทธิภาพของแบบจำลองในตารางที่ 4.28 พบว่า ประสิทธิภาพโดยรวมของแบบจำลองพบว่ามีประสิทธิภาพโดยรวมของแบบจำลองเรียงจากมากไปหาน้อยได้แก่ เกณฑ์ของ SCP ข้อที่ 7, 8, 6 และ 2 ตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ต้นไม้ตัดสินในพื้นที่การทำงาน HSA



รูปที่ 4.47 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 2 พื้นที่ HSA

จากรูป 4.47 พบว่าชุดข้อมูลกฎข้อที่ 2 พื้นที่ HSA ประกอบด้วย ขนาดของต้นไม้เท่ากับ 55 โหนด และลีฟโหนดมีขนาด 42 โหนด ซึ่งสามารถแปลงต้นไม้ตัดสินใจเป็นกฎในรูปแบบคำสั่ง IF-THEN ได้ดังตาราง 4.29

ตารางที่ 4.29 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 2 พื้นที่ HSA

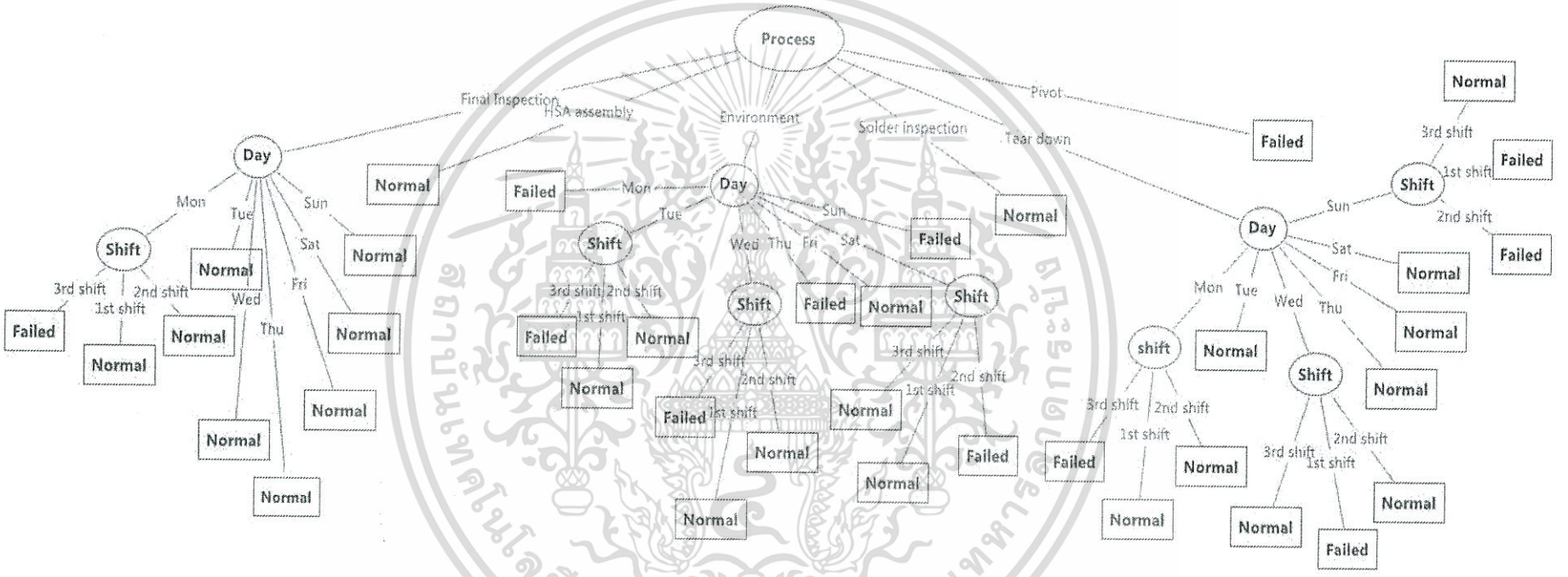
กฎที่	เงื่อนไขกฎ
Rule1	IF PROCESS = Final inspection AND DAY = 1MON AND SHIFT = 3RD SHIFT THEN PARTICLE = "NORMAL"
Rule2	IF PROCESS = Final inspection AND DAY = 1MON AND SHIFT = 1ST SHIFT THEN PARTICLE = "FAILED"
Rule3	IF PROCESS = Final inspection AND DAY = 1MON AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule4	IF PROCESS = Final inspection AND DAY = 2TUE THEN PARTICLE = "FAILED"
Rule5	IF PROCESS = Final inspection AND DAY = 3WED THEN PARTICLE = "FAILED"
Rule6	IF PROCESS = Final inspection AND DAY = 4THU THEN PARTICLE = "FAILED"
Rule7	IF PROCESS = Final inspection AND DAY = 5FRI THEN PARTICLE = "FAILED"
Rule8	IF PROCESS = Final inspection AND DAY = 6SAT AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule9	IF PROCESS = Final inspection AND DAY = 6SAT AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule10	IF PROCESS = Final inspection AND DAY = 6SAT AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule11	IF PROCESS = Final inspection AND DAY = 7SUN THEN PARTICLE = "FAILED"
Rule12	IF PROCESS = HSA assembly AND DAY = 1MON AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule13	IF PROCESS = HSA assembly AND DAY = 1MON AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule14	IF PROCESS = HSA assembly AND DAY = 1MON AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule15	IF PROCESS = HSA assembly AND DAY = 2TUE AND SHIFT = 3RD SHIFT THEN PARTICLE = "NORMAL"
Rule16	IF PROCESS = HSA assembly AND DAY = 2TUE AND SHIFT = 1ST SHIFT THEN PARTICLE = "FAILED"
Rule17	IF PROCESS = HSA assembly AND DAY = 2TUE AND SHIFT = 2ND SHIFT THEN PARTICLE = "NORMAL"
Rule18	IF PROCESS = HSA assembly AND DAY = 3WED THEN PARTICLE = "NORMAL"
Rule19	IF PROCESS = HSA assembly AND DAY = 4THU AND SHIFT = 3RD SHIFT THEN PARTICLE = "NORMAL"
Rule20	IF PROCESS = HSA assembly AND DAY = 4THU AND SHIFT = 1ST SHIFT THEN PARTICLE = "FAILED"
Rule21	IF PROCESS = HSA assembly AND DAY = 4THU AND SHIFT = 2ND SHIFT THEN PARTICLE = "NORMAL"
Rule22	IF PROCESS = HSA assembly AND DAY = 5FRI THEN PARTICLE = "NORMAL"
Rule23	IF PROCESS = HSA assembly AND DAY = 6SAT AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule24	IF PROCESS = HSA assembly AND DAY = 6SAT AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule25	IF PROCESS = HSA assembly AND DAY = 6SAT AND SHIFT = 2ND SHIFT THEN PARTICLE = "NORMAL"
Rule26	IF PROCESS = HSA assembly AND DAY = 7SUN AND SHIFT = 3RD SHIFT THEN PARTICLE = "NORMAL"
Rule27	IF PROCESS = HSA assembly AND DAY = 7SUN AND SHIFT = 1ST SHIFT THEN PARTICLE = "FAILED"
Rule28	IF PROCESS = HSA assembly AND DAY = 7SUN AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule29	IF PROCESS = Solder inspection AND DAY = 1MON THEN PARTICLE = "NORMAL"
Rule30	IF PROCESS = Solder inspection AND DAY = 2TUE THEN PARTICLE = "NORMAL"
Rule31	IF PROCESS = Solder inspection AND DAY = 3WED THEN PARTICLE = "NORMAL"
Rule32	IF PROCESS = Solder inspection AND DAY = 4THU AND SHIFT = 3RD SHIFT THEN PARTICLE = "NORMAL"
Rule33	IF PROCESS = Solder inspection AND DAY = 4THU AND SHIFT = 1ST SHIFT THEN PARTICLE = "FAILED"
Rule34	IF PROCESS = Solder inspection AND DAY = 4THU AND SHIFT = 2ND SHIFT THEN PARTICLE = "NORMAL"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Rule35	IF PROCESS = Solder inspection AND DAY = 5FRI THEN PARTICLE = "NORMAL"
Rule36	IF PROCESS = Solder inspection AND DAY = 6SAT THEN PARTICLE = "NORMAL"
Rule37	IF PROCESS = Solder inspection AND DAY = 7SUN AND SHIFT = 3^{RD} SHIFT THEN PARTICLE = "NORMAL"
Rule38	IF PROCESS = Solder inspection AND DAY = 7SUN AND SHIFT = 1^{ST} SHIFT THEN PARTICLE = "NORMAL"
Rule39	IF PROCESS = Solder inspection AND DAY = 7SUN AND SHIFT = 2^{ND} SHIFT THEN PARTICLE = "FAILED"
Rule40	IF PROCESS = Environment THEN PARTICLE = "FAILED"
Rule41	IF PROCESS = Tear down THEN PARTICLE = "FAILED"
Rule42	IF PROCESS = Pivot THEN PARTICLE = "FAILED"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ต้นไม้ตัดสินในพื้นที่การทำงาน HSA



รูปที่ 4.48 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 6 พื้นที่ HSA

จากรูป 4.48 พบว่าชุดข้อมูลกฎข้อที่ 6 พื้นที่ HSA ประกอบด้วย ขนาดของต้นไม้เท่ากับ 49 โหนด และลีฟโหนดมีขนาด 38 โหนด ซึ่งสามารถแปลงต้นไม้ตัดสินใจเป็นกฎในรูปแบบคำสั่ง IF-THEN ได้ดังตาราง 4.30

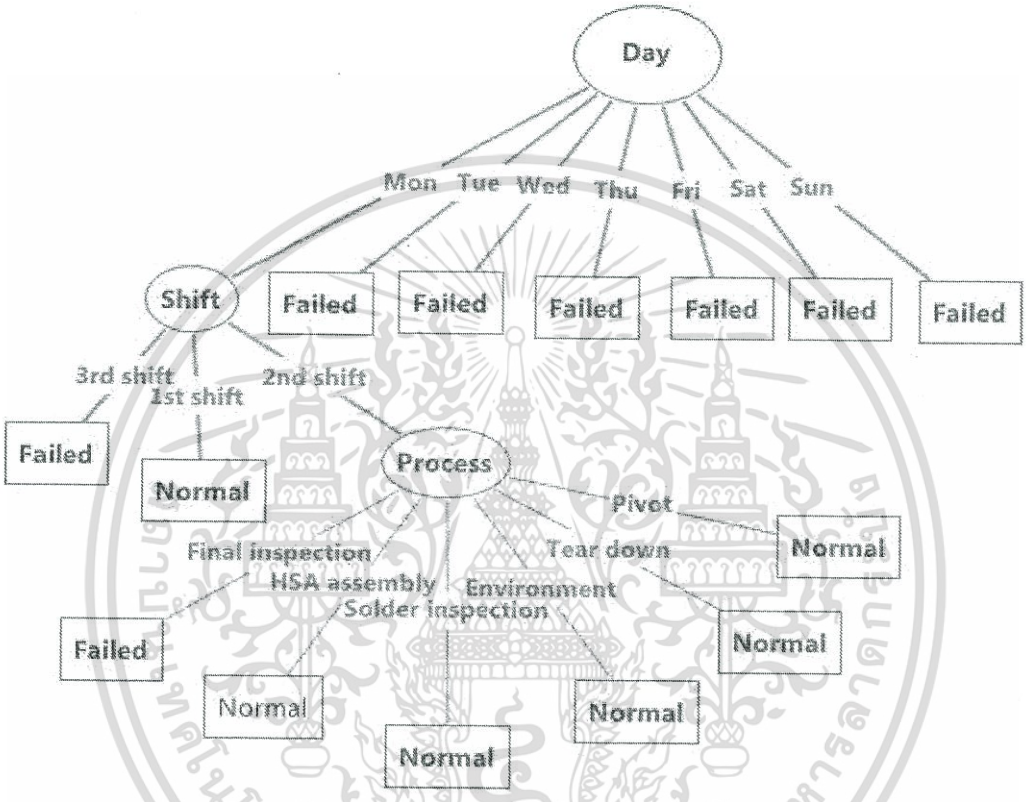
ตารางที่ 4.30 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 6 พื้นที่ HSA

กฎที่	เงื่อนไขกฎ
Rule1	IF PROCESS = Final inspection AND DAY = 1MON AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule2	IF PROCESS = Final inspection AND DAY = 1MON AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule3	IF PROCESS = Final inspection AND DAY = 1MON AND SHIFT = 2ND SHIFT THEN PARTICLE = "NORMAL"
Rule4	IF PROCESS = Final inspection AND DAY = 2TUE THEN PARTICLE = "NORMAL"
Rule5	IF PROCESS = Final inspection AND DAY = 3WED THEN PARTICLE = "NORMAL"
Rule6	IF PROCESS = Final inspection AND DAY = 4THU THEN PARTICLE = "NORMAL"
Rule7	IF PROCESS = Final inspection AND DAY = 5FRI THEN PARTICLE = "NORMAL"
Rule8	IF PROCESS = Final inspection AND DAY = 6SAT THEN PARTICLE = "NORMAL"
Rule9	IF PROCESS = Final inspection AND DAY = 7SUN THEN PARTICLE = "NORMAL"
Rule10	IF PROCESS = HSA assembly THEN PARTICLE = "NORMAL"
Rule11	IF PROCESS = Solder inspection THEN PARTICLE = "NORMAL"
Rule12	IF PROCESS = Environment AND DAY = 1MON THEN PARTICLE = "FAILED"
Rule13	IF PROCESS = Environment AND DAY = 2TUE AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule14	IF PROCESS = Environment AND DAY = 2TUE AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule15	IF PROCESS = Environment AND DAY = 2TUE AND SHIFT = 2ND SHIFT THEN PARTICLE = "NORMAL"
Rule16	IF PROCESS = Environment AND DAY = 3WED AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule17	IF PROCESS = Environment AND DAY = 3WED AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule18	IF PROCESS = Environment AND DAY = 3WED AND SHIFT = 2ND SHIFT THEN PARTICLE = "NORMAL"
Rule19	IF PROCESS = Environment AND DAY = 4THU THEN PARTICLE = "FAILED"
Rule20	IF PROCESS = Environment AND DAY = 5FRI THEN PARTICLE = "NORMAL"
Rule21	IF PROCESS = Environment AND DAY = 6SAT AND SHIFT = 3RD SHIFT THEN PARTICLE = "NORMAL"
Rule22	IF PROCESS = Environment AND DAY = 6SAT AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule23	IF PROCESS = Environment AND DAY = 6SAT AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule24	IF PROCESS = Environment AND DAY = 7SUN THEN PARTICLE = "FAILED"
Rule25	IF PROCESS = Tear down AND DAY = 1MON AND SHIFT = 3RD SHIFT THEN PARTICLE = "FAILED"
Rule26	IF PROCESS = Tear down AND DAY = 1MON AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"
Rule27	IF PROCESS = Tear down AND DAY = 1MON AND SHIFT = 2ND SHIFT THEN PARTICLE = "NORMAL"
Rule28	IF PROCESS = Tear down AND DAY = 2TUE THEN PARTICLE = "NORMAL"
Rule29	IF PROCESS = Tear down AND DAY = 3WED AND SHIFT = 3RD SHIFT THEN PARTICLE = "NORMAL"
Rule30	IF PROCESS = Tear down AND DAY = 3WED AND SHIFT = 1ST SHIFT THEN PARTICLE = "FAILED"
Rule31	IF PROCESS = Tear down AND DAY = 3WED AND SHIFT = 2ND SHIFT THEN PARTICLE = "NORMAL"
Rule32	IF PROCESS = Tear down AND DAY = 2THU THEN PARTICLE = "NORMAL"
Rule33	IF PROCESS = Tear down AND DAY = 5FRI THEN PARTICLE = "NORMAL"
Rule34	IF PROCESS = Tear down AND DAY = 6SAT THEN PARTICLE = "NORMAL"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Rule35	IF PROCESS = Tear down AND DAY = 7SUN AND SHIFT = 3RD SHIFT THEN PARTICLE = "NORMAL"
Rule36	IF PROCESS = Tear down AND DAY = 7SUN AND SHIFT = 1ST SHIFT THEN PARTICLE = "FAILED"
Rule37	IF PROCESS = Tear down AND DAY = 7SUN AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule38	IF PROCESS = Pivot THEN PARTICLE = "FAILED"

- ต้นไม้ตัดสินในพื้นที่การทำงาน HSA



รูปที่ 4.49 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 7 พื้นที่ HSA

จากรูป 4.49 พบว่าชุดข้อมูลกฎข้อที่ 7 พื้นที่ HSA ประกอบด้วย ขนาดของต้นไม้เท่ากับ 17 โหนด และลีฟโหนดมีขนาด 14 โหนด ซึ่งสามารถแปลงต้นไม้ตัดสินใจเป็นกฎในรูปแบบคำสั่ง IF-THEN ได้ดังตาราง 4.31

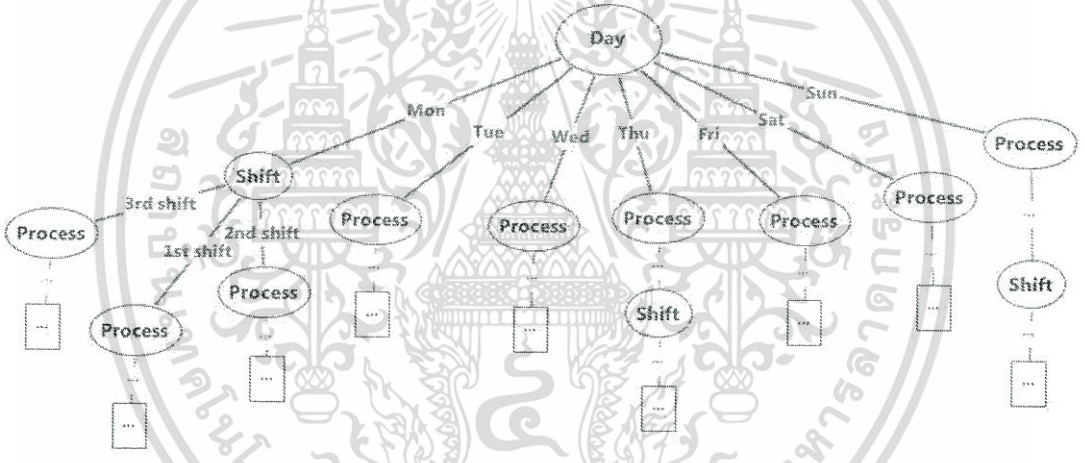
ตารางที่ 4.31 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 7 พื้นที่ HSA

กฎที่	เงื่อนไขกฎ
Rule1	IF PROCESS = Final inspection AND DAY = 1MON AND SHIFT = 3RD SHIFT THEN PARTICLE = "NORMAL"
Rule2	IF PROCESS = Final inspection AND DAY = 1MON AND SHIFT = 1ST SHIFT THEN PARTICLE = "NORMAL"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Rule3	IF PROCESS = Final inspection AND DAY = 1MON AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule4	IF PROCESS = Final inspection AND DAY = 2TUE THEN PARTICLE = "NORMAL"
Rule5	IF PROCESS = Final inspection AND DAY = 3WED THEN PARTICLE = "NORMAL"
Rule6	IF PROCESS = Final inspection AND DAY = 4THU THEN PARTICLE = "NORMAL"
Rule7	IF PROCESS = Final inspection AND DAY = 5FRI THEN PARTICLE = "NORMAL"
Rule8	IF PROCESS = Final inspection AND DAY = 6SAT THEN PARTICLE = "NORMAL"
Rule9	IF PROCESS = Final inspection AND DAY = 7SUN THEN PARTICLE = "NORMAL"
Rule10	IF PROCESS = HSA assembly THEN PARTICLE = "NORMAL"
Rule11	IF PROCESS = Solder inspection THEN PARTICLE = "NORMAL"
Rule12	IF PROCESS = Environment THEN PARTICLE = "NORMAL"
Rule13	IF PROCESS = Tear down THEN PARTICLE = "NORMAL"
Rule14	IF PROCESS = Pivot THEN PARTICLE = "NORMAL"

- ต้นไม้ตัดสินในพื้นที่การทำงาน HSA



รูปที่ 4.50 ต้นไม้ตัดสินใจของเกณฑ์ SPC ข้อที่ 8 พื้นที่ HSA

จากรูป 4.50 พบว่าชุดข้อมูลกฎข้อที่ 8 พื้นที่ HSA ประกอบด้วย ขนาดของต้นไม้เท่ากับ 71 โหนด และลีฟโหนดมีขนาด 58 โหนด ซึ่งสามารถแปลงต้นไม้ตัดสินใจเป็นกฎในรูปแบบคำสั่ง IF-THEN ได้ดังตาราง 4.32

ตารางที่ 4.32 กฎจากต้นไม้ตัดสินใจเกณฑ์การควบคุมข้อที่ 8 พื้นที่ HSA

กฎที่	เงื่อนไขกฎ
Rule1	IF DAY = 1MON AND SHIFT = 3RD SHIFT AND PROCESS = Final inspection THEN PARTICLE = "FAILED"
Rule2	IF DAY = 1MON AND SHIFT = 3RD SHIFT AND PROCESS = HSA assembly THEN PARTICLE = "NORMAL"
Rule3	IF DAY = 1MON AND SHIFT = 3RD SHIFT AND PROCESS = Solder inspection THEN PARTICLE = "NORMAL"
Rule4	IF DAY = 1MON AND SHIFT = 3RD SHIFT AND PROCESS = Environment THEN PARTICLE = "NORMAL"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Rule5	IF DAY = 1MON AND SHIFT = 3RD SHIFT AND PROCESS = Tear down THEN PARTICLE = "NORMAL"
Rule6	IF DAY = 1MON AND SHIFT = 3RD SHIFT AND PROCESS = Pivot THEN PARTICLE = "NORMAL"
Rule7	IF DAY = 1MON AND SHIFT = 1ST SHIFT AND PROCESS = Final inspection THEN PARTICLE = "NORMAL"
Rule8	IF DAY = 1MON AND SHIFT = 1ST SHIFT AND PROCESS = HSA assembly THEN PARTICLE = "NORMAL"
Rule9	IF DAY = 1MON AND SHIFT = 1ST SHIFT AND PROCESS = Solder inspection THEN PARTICLE = "NORMAL"
Rule10	IF DAY = 1MON AND SHIFT = 1ST SHIFT AND PROCESS = Environment THEN PARTICLE = "NORMAL"
Rule11	IF DAY = 1MON AND SHIFT = 1ST SHIFT AND PROCESS = Tear down THEN PARTICLE = "NORMAL"
Rule12	IF DAY = 1MON AND SHIFT = 1ST SHIFT AND PROCESS = Pivot THEN PARTICLE = "FAILED"
Rule13	IF DAY = 1MON AND SHIFT = 2ND SHIFT AND PROCESS = Final inspection THEN PARTICLE = "NORMAL"
Rule14	IF DAY = 1MON AND SHIFT = 2ND SHIFT AND PROCESS = HSA assembly THEN PARTICLE = "NORMAL"
Rule15	IF DAY = 1MON AND SHIFT = 2ND SHIFT AND PROCESS = Solder inspection THEN PARTICLE = "NORMAL"
Rule16	IF DAY = 1MON AND SHIFT = 2NDSHIFT AND PROCESS = Environment THEN PARTICLE = "NORMAL"
Rule17	IF DAY = 1MON AND SHIFT = 2ND SHIFT AND PROCESS = Tear down THEN PARTICLE = "NORMAL"
Rule18	IF DAY = 1MON AND SHIFT = 2ND SHIFT AND PROCESS = Pivot THEN PARTICLE = "FAILED"
Rule19	IF DAY = 2TUE AND PROCESS = Final inspection THEN PARTICLE = "NORMAL"
Rule20	IF DAY = 2TUE AND PROCESS = HSA assembly THEN PARTICLE = "NORMAL"
Rule21	IF DAY = 2TUE AND PROCESS = Solder inspection THEN PARTICLE = "NORMAL"
Rule22	IF DAY = 2TUE PROCESS = Environment THEN PARTICLE = "NORMAL"
Rule23	IF DAY = 2TUE PROCESS = Tear down THEN PARTICLE = "NORMAL"
Rule24	IF DAY = 2TUE PROCESS = Pivot THEN PARTICLE = "FAILED"
Rule25	IF DAY = 3WED AND PROCESS = Final inspection THEN PARTICLE = "NORMAL"
Rule26	IF DAY = 3WED AND PROCESS = HSA assembly THEN PARTICLE = "NORMAL"
Rule27	IF DAY = 3WED AND PROCESS = Solder inspection THEN PARTICLE = "NORMAL"
Rule28	IF DAY = 3WED PROCESS = Environment THEN PARTICLE = "NORMAL"
Rule29	IF DAY = 3WED PROCESS = Tear down THEN PARTICLE = "NORMAL"
Rule30	IF DAY = 3WED PROCESS = Pivot THEN PARTICLE = "FAILED"
Rule31	IF DAY = 4THU AND PROCESS = Final inspection THEN PARTICLE = "NORMAL"
Rule32	IF DAY = 4THU AND PROCESS = HSA assembly THEN PARTICLE = "NORMAL"
Rule33	IF DAY = 4THU AND PROCESS = Solder inspection THEN PARTICLE = "NORMAL"
Rule34	IF DAY = 4THU PROCESS = Environment AND SHIFT = 3RD SHIFT THEN PARTICLE = "NORMAL"
Rule35	IF DAY = 4THU PROCESS = Environment AND SHIFT = 1ST SHIFT THEN PARTICLE = "FAILED"
Rule36	IF DAY = 4THU PROCESS = Environment AND SHIFT = 2ND SHIFT THEN PARTICLE = "FAILED"
Rule37	IF DAY = 4THU AND PROCESS = Tear down THEN PARTICLE = "NORMAL"
Rule38	IF DAY = 4THU AND PROCESS = Pivot THEN PARTICLE = "FAILED"
Rule39	IF DAY = 5FRI AND PROCESS = Final inspection THEN PARTICLE = "NORMAL"
Rule40	IF DAY = 5FRI AND PROCESS = HSA assembly THEN PARTICLE = "NORMAL"
Rule41	IF DAY = 5FRI AND PROCESS = Solder inspection THEN PARTICLE = "NORMAL"
Rule42	IF DAY = 5FRI PROCESS = Environment THEN PARTICLE = "NORMAL"
Rule43	IF DAY = 5FRI PROCESS = Tear down THEN PARTICLE = "NORMAL"
Rule44	IF DAY = 5FRI PROCESS = Pivot THEN PARTICLE = "FAILED"
Rule45	IF DAY = 6SAT AND PROCESS = Final inspection THEN PARTICLE = "NORMAL"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Rule46	IF DAY = 6SAT AND PROCESS = HSA assembly THEN PARTICLE = "NORMAL"
Rule47	IF DAY = 6SAT AND PROCESS = Solder inspection THEN PARTICLE = "NORMAL"
Rule48	IF DAY = 6SAT PROCESS = Environment THEN PARTICLE = "NORMAL"
Rule49	IF DAY = 6SAT PROCESS = Tear down THEN PARTICLE = "NORMAL"
Rule50	IF DAY = 6SAT PROCESS = Pivot THEN PARTICLE = "FAILED"
Rule51	IF DAY = 7SUN AND PROCESS = Final inspection THEN PARTICLE = "NORMAL"
Rule52	IF DAY = 7SUN AND PROCESS = HSA assembly THEN PARTICLE = "NORMAL"
Rule53	IF DAY = 7SUN AND PROCESS = Solder inspection THEN PARTICLE = "NORMAL"
Rule54	IF DAY = 7SUN PROCESS = Environment AND SHIFT = 3^{RD} SHIFT THEN PARTICLE = "NORMAL"
Rule55	IF DAY = 7SUN PROCESS = Environment AND SHIFT = 1^{ST} SHIFT THEN PARTICLE = "NORMAL"
Rule56	IF DAY = 7SUN PROCESS = Environment AND SHIFT = 2^{ND} SHIFT THEN PARTICLE = "FAILED"
Rule57	IF DAY = 7SUN AND PROCESS = Tear down THEN PARTICLE = "NORMAL"
Rule58	IF DAY = 7SUN AND PROCESS = Pivot THEN PARTICLE = "FAILED"

จากการสร้างแบบจำลองต้นไม้ตัดสินใจของชุดข้อมูลพื้นที่การทำงาน HSA ทั้ง 4 ชุดข้อมูล พบว่า ปัจจัยที่เกี่ยวข้องกับการเกิดฝุ่นละอองในพื้นที่การทำงาน cleaning ได้แก่ กระบวนการทำงาน (Process) กะการทำงาน (Shift) และวันที่ทำงาน (Day) (Shift) ซึ่งสำหรับการเลือกแบบจำลองและเกณฑ์การควบคุมกระบวนการเชิงสถิติที่เหมาะสมสำหรับพื้นที่ HSA จะได้กล่าวต่อไป

4.2.3 ผลการวัดประสิทธิภาพของการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจ

การสร้างต้นไม้ตัดสินใจของข้อมูลการควบคุมคุณภาพปริมาณอนุภาคฝุ่นละออง จะวัดประสิทธิภาพเพื่อเลือกต้นไม้ตัดสินใจที่ดีที่สุดเพื่อนำกฎจากต้นไม้ตัดสินใจดังกล่าวเป็นแนวทางในการควบคุมคุณภาพของกระบวนการผลิตฮาร์ดดิสก์ไดร์ฟในห้องควบคุมปริมาณอนุภาคฝุ่นละออง (Clean room) โดยวิทยานิพนธ์ฉบับนี้จะเลือกใช้การวัดประสิทธิภาพ 4 ข้อได้แก่ ค่าความถูกต้องในการจำแนก (Correct Classified Instances), ค่าความครบถ้วน (Recall), ค่าความแม่นยำ (Precision) และ ค่าค่าถ่วงดุล (F-measure)

สำหรับการวัดค่าความถูกต้องในการจำแนก ค่าความครบถ้วน ค่าความแม่นยำและค่าถ่วงดุลในของต้นแบบการจำแนกข้อมูลปริมาณอนุภาคฝุ่นละอองด้วยต้นไม้ตัดสินใจ โดยแบ่งผลความถูกต้องในการจำแนกเป็น 4 ระดับ ได้แก่ ระดับดีมาก คือ $(0.8 \leq x \leq 1)$, ระดับดี คือ$(0.5 \leq x < 0.8)$, ระดับพอใช้ คือ$(0.2 \leq x < 0.5)$ และ ระดับต่ำ คือ $(x < 0.2)$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.33 เปรียบเทียบค่าความถูกต้องในการจำแนก, ค่าความครบถ้วน, ค่าความแม่นยำ และค่าถ่วงดุล ของต้นไม้ตัดสินใจพื้นที่การทำงาน cleaning

ตัวแบบต้นไม้ตัดสินใจ	เกณฑ์การวัดประสิทธิภาพ			
	ค่าความถูกต้องในการจำแนก (%)	ค่าความครบถ้วน (Recall)	ค่าความแม่นยำ (Precision)	ค่าถ่วงดุล (F-measure)
เกณฑ์ของโรงงาน	87.8511	0.879	0.88	0.879
เกณฑ์ SCP ข้อที่ 1	64.6695	0.647	0.669	0.63
เกณฑ์ SCP ข้อที่ 2	82.3257	0.823	0.838	0.819
เกณฑ์ SCP ข้อที่ 5	74.2457	0.742	0.788	0.73
เกณฑ์ SCP ข้อที่ 6	83.4541	0.835	0.846	0.832
เกณฑ์ SCP ข้อที่ 8	68.0899	0.681	0.704	0.671

จากตารางที่ 4.33 เมื่อพิจารณาเปรียบเทียบประสิทธิภาพการจำแนกของข้อมูลเกณฑ์ของโรงงานกับการควบคุมกระบวนการเชิงสถิติ พบว่า ประสิทธิภาพของข้อมูลเกณฑ์โรงงานมีประสิทธิภาพของกฎการจำแนกอยู่ในระดับดีมาก โดยมีความถูกต้อง ความครบถ้วน และความแม่นยำในการทำนายดีมาก สำหรับประสิทธิภาพของข้อมูลการควบคุมกระบวนการเชิงสถิติแต่ละเกณฑ์มีประสิทธิภาพอยู่ในระดับดีมากถึงพอใช้

ตารางที่ 4.34 เปรียบเทียบค่าความถูกต้องในการจำแนก, ค่าความครบถ้วน, ค่าความแม่นยำ และค่าถ่วงดุล ของต้นไม้ตัดสินใจพื้นที่การทำงาน HDE

ตัวแบบต้นไม้ตัดสินใจ	เกณฑ์การวัดประสิทธิภาพ			
	ค่าความถูกต้องในการจำแนก (%)	ค่าความครบถ้วน (Recall)	ค่าความแม่นยำ (Precision)	ค่าถ่วงดุล (F-measure)
เกณฑ์ของโรงงาน	98.5992	0.986	0.984	0.984
เกณฑ์ SCP ข้อที่ 2	62.872	0.629	0.63	0.628
เกณฑ์ SCP ข้อที่ 7	63.1473	0.631	0.636	0.629

จากตารางที่ 4.34 เมื่อพิจารณาเปรียบเทียบประสิทธิภาพการจำแนกของข้อมูลเกณฑ์ของโรงงานกับการควบคุมกระบวนการเชิงสถิติ พบว่า ประสิทธิภาพของข้อมูลเกณฑ์โรงงานมีประสิทธิภาพของกฎการจำแนกอยู่ในระดับดีมาก โดยมีความถูกต้อง ความครบถ้วน และความแม่นยำในการทำนายดีมาก สำหรับประสิทธิภาพของข้อมูลการควบคุมกระบวนการเชิงสถิติแต่ละเกณฑ์มีประสิทธิภาพอยู่ในระดับพอใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.35 เปรียบเทียบค่าความถูกต้องในการจำแนก, ค่าความครบถ้วน, ค่าความแม่นยำ และค่าถ่วงดุล ของต้นไม้ตัดสินใจพื้นที่การทำงาน HSA

ตัวแบบต้นไม้ตัดสินใจ	เกณฑ์การวัดประสิทธิภาพ			
	ค่าความถูกต้องในการจำแนก (%)	ค่าความครบถ้วน (Recall)	ค่าความแม่นยำ (Precision)	ค่าถ่วงดุล (F-measure)
เกณฑ์ของโรงงาน	-	-	-	-
เกณฑ์ SCP ข้อที่ 2	71.7549	0.718	0.753	0.704
เกณฑ์ SCP ข้อที่ 6	82.5139	0.825	0.839	0.821
เกณฑ์ SCP ข้อที่ 7	90.3125	0.903	0.908	0.903
เกณฑ์ SCP ข้อที่ 8	89.3601	0.894	0.895	0.893

จากตารางที่ 4.35 เนื่องจากไม่สามารถสร้างต้นไม้ตัดสินใจของข้อมูลเกณฑ์ของโรงงานได้ เนื่องจากข้อมูลมีความไม่สมดุลในระดับสูง เพราะคลาสของข้อมูล Normal มีปริมาณ 99.97% ซึ่งแสดงว่าเมื่อพิจารณาจากเกณฑ์ดังกล่าวให้ผลว่าพื้นที่ HSA มีปริมาณฝุ่นละอองเป็นปกติ และเมื่อประสิทธิภาพของข้อมูลการควบคุมกระบวนการเชิงสถิติแต่ละเกณฑ์มีประสิทธิภาพอยู่ในระดับดีมากถึงดี

4.2.4 การวิเคราะห์และประเมินผล

การวิเคราะห์และประเมินผลเป็นขั้นตอนการแปลความหมายโดยทำการวิเคราะห์ผล 3 ข้อ ได้แก่

4.2.4.1 วิเคราะห์กฎของการควบคุมกระบวนการที่เหมาะสมสำหรับข้อมูลแต่ละข้อมูล

- ข้อมูลพื้นที่การทำงาน cleaning

จากการวิเคราะห์ประสิทธิภาพของการจำแนก พบว่าข้อมูลที่จัดระเบียบด้วยการควบคุมกะบวนการเชิงสถิติ มีประสิทธิภาพการจำแนกอยู่ในระดับดี ดังนั้นสรุปว่าการควบคุมกระบวนการที่เหมาะสมสำหรับพื้นที่การทำงาน Cleaning คือเกณฑ์ข้อ 6, 2, 5, 8 และ 1 ตามลำดับ

- ข้อมูลพื้นที่การทำงาน HDE

จากการวิเคราะห์ประสิทธิภาพของการจำแนก พบว่าข้อมูลที่จัดระเบียบด้วยการควบคุมกะบวนการเชิงสถิติ มีประสิทธิภาพการจำแนกอยู่ในระดับดีมาก ดังนั้นสรุปว่าการควบคุมกระบวนการที่เหมาะสมสำหรับพื้นที่การทำงาน HDE คือเกณฑ์ข้อ 7 และ 2 ตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ข้อมูลพื้นที่การทำงาน HSA

จากการวิเคราะห์ประสิทธิภาพของการจำแนก พบว่าข้อมูลที่จัดระเบียบด้วยการควบคุมกระบวนการเชิงสถิติ มีประสิทธิภาพการจำแนกอยู่ในระดับดีและดีมาก จึงสามารถนำเกณฑ์ทั้ง 4 ข้อไปใช้ในงานตรวจจับฝุ่นละอองได้ ซึ่งสามารถเรียงลำดับเกณฑ์ที่มีประสิทธิภาพมากที่สุดได้แก่ ข้อ 7, 8, 6 และ 2 ตามลำดับ

4.2.4.2 วิเคราะห์กฎที่ได้เพื่อสรุปปัจจัยที่ทำให้เกิดปริมาณฝุ่นละอองในแต่ละพื้นที่การทำงาน

- ข้อมูลพื้นที่การทำงาน cleaning

จากการพิจารณากฎที่ได้จากต้นไม้ตัดสินใจของพื้นที่การทำงาน Cleaning พบว่าปัจจัยที่ทำให้ปริมาณอนุภาคฝุ่นละอองเกิดสภาวะเกินควบคุมประกอบด้วย วันที่ทำงาน (Day) และกะการทำงาน (Shift) ดังแสดงกฎของการการเกิดฝุ่นละอองในผลการทดลองหัวข้อที่ 4.2.2.1

- ข้อมูลพื้นที่การทำงาน HDE

จากการพิจารณากฎที่ได้จากต้นไม้ตัดสินใจของพื้นที่การทำงาน Cleaning พบว่าปัจจัยที่ทำให้ปริมาณอนุภาคฝุ่นละอองเกิดสภาวะเกินควบคุมประกอบด้วย กระบวนการทำงาน (Process) วันที่ทำงาน (Day) และกะการทำงาน (Shift) ซึ่งกระบวนการทำงานที่เป็นปัจจัยได้แก่ Wipe, Environment, Particle Test, Disk mounting, HSA/VCM mounting, Ramp mounting, Top Clamp mounting, Top Cover mounting, Hyper disk balancer และ Part loading conveyer ซึ่งแสดงกฎของการเกิดฝุ่นละอองในผลการทดลองหัวข้อที่ 4.2.2.2

- ข้อมูลพื้นที่การทำงาน HSA

จากการพิจารณากฎที่ได้จากต้นไม้ตัดสินใจของพื้นที่การทำงาน Cleaning พบว่าปัจจัยที่ทำให้ปริมาณอนุภาคฝุ่นละอองเกิดสภาวะเกินควบคุมประกอบด้วย กระบวนการทำงาน (Process) วันที่ทำงาน (Day) และกะการทำงาน (Shift) ซึ่งกระบวนการทำงานที่เป็นปัจจัยได้แก่ Final inspection, HSA assembly, Solder inspection, Environment, Tear down และ Pivot ซึ่งแสดงกฎของการเกิดฝุ่นละอองในผลการทดลองหัวข้อที่ 4.2.2.3

4.2.4.3 วิเคราะห์เปรียบเทียบประสิทธิภาพของการจำแนกข้อมูลระหว่างข้อมูลที่ควบคุมคุณภาพด้วยเกณฑ์ของโรงงานกับข้อมูลที่ควบคุมกระบวนการเชิงสถิติ

ประสิทธิภาพของการจำแนกของเกณฑ์จากโรงงานนั้นมีประสิทธิภาพสูง ทั้งในด้านความครบถ้วนและความแม่นยำ ในขณะเดียวกันประสิทธิภาพการจำแนกของเกณฑ์การควบคุมกระบวนการเชิงสถิติข้อต่างๆ ที่พิจารณาในข้อมูลนี้ก็มีประสิทธิภาพในระดับดีและดีมาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้นสำหรับข้อมูลพื้นที่ Cleaning, HDE และ HSA สามารถใช้เกณฑ์การควบคุมกระบวนการเชิงสถิติมาใช้ในการควบคุมกระบวนการผลิตเพิ่มเติมเพื่อให้สามารถตรวจจับฝุ่นละอองที่มีรูปแบบในการเกิดฝุ่นละอองที่หลากหลายมากขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
บทสรุปและข้อเสนอแนะ

5.1 บทสรุป

วิทยานิพนธ์ฉบับนี้ได้นำเสนอการควบคุมคุณภาพของปริมาณอนุภาคฝุ่นละอองในการผลิตฮาร์ดดิสก์ไดร์ฟ โดยใช้การควบคุมกระบวนการเชิงสถิติ (Statistic Process Control) ร่วมกับเหมืองข้อมูล (Data mining) โดยใช้กรณีศึกษาข้อมูลปริมาณอนุภาคฝุ่นละอองของอุตสาหกรรมผลิตฮาร์ดดิสก์ ซึ่งข้อมูลปริมาณอนุภาคฝุ่นละอองที่นำมาใช้ในการศึกษานั้นเป็นข้อมูลที่นำมาจากระบบตรวจจับฝุ่นละออง (Particle monitoring system) ในโรงงานอุตสาหกรรมผลิตฮาร์ดดิสก์ไดร์ฟจริง ซึ่งข้อมูลที่ได้มานั้นไม่เหมาะสมสำหรับการศึกษาจึงจำเป็นต้องมีการเตรียมข้อมูลก่อน ซึ่งสามารถแบ่งข้อมูลเป็นชุดทดสอบตามมิติพื้นที่การทำงานได้ 3 ชุดข้อมูลได้แก่ ชุดข้อมูลปริมาณอนุภาคฝุ่นละอองพื้นที่การทำงาน cleaning, ชุดข้อมูลปริมาณอนุภาคฝุ่นละอองพื้นที่การทำงาน HDE และ ชุดข้อมูลปริมาณอนุภาคฝุ่นละอองพื้นที่การทำงาน HSA

โดยจะพิจารณาเพื่อหาปัจจัยที่ทำให้เกิดปริมาณอนุภาคฝุ่นละอองที่เกินการควบคุมในกระบวนการผลิตฮาร์ดดิสก์ไดร์ฟ ซึ่งปัจจัยทั้งหมดที่พิจารณาได้แก่ วันที่ทำงาน (Day), กะการทำงาน (Shift) และกระบวนการทำงาน (Process) โดยใช้การควบคุมคุณภาพด้วยการควบคุมกระบวนการเชิงสถิติ (SPC) เพื่อสร้างคลาสข้อมูลการควบคุม โดยใช้การควบคุมชนิดแผนภูมิควบคุม (Control chart) ซึ่งกำหนดกฎการควบคุม 8 กฎ ที่ครอบคลุมการรูปแบบการควบคุมหลายๆ แบบ กฎการควบคุมทั้ง 8 กฎประกอบด้วย

1. มี 1 จุดอยู่นอกเส้นควบคุม $\pm 3\sigma$
2. มี 9 จุดต่อเนื่องกันที่อยู่ด้านใดด้านหนึ่งของเส้นควบคุมกลาง (Central line)
3. มี 6 จุดต่อเนื่องกันที่มีแนวโน้มเพิ่มขึ้นหรือลดลงอย่างต่อเนื่อง
4. มี 14 จุดต่อเนื่องกันที่มีขึ้นลงสลับกัน
5. มี 2 ใน 3 จุดที่อยู่นอกเส้นควบคุมเตือน $\pm 2\sigma$ บนด้านเดียวกันของเส้นควบคุมกลาง
6. มี 4 ใน 5 จุดที่อยู่นอกเส้นควบคุมเตือน $\pm 1\sigma$ บนด้านเดียวกันของเส้นควบคุมกลาง
7. มี 15 จุดต่อเนื่องกันที่อยู่ในช่วง $\pm 1\sigma$
8. มี 8 จุดต่อเนื่องกันที่อยู่นอกเขตควบคุม $\pm 1\sigma$ จากเส้นควบคุมกลางของข้างใดข้างหนึ่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยในการควบคุมคุณภาพเชิงสถิติของข้อมูลแต่ละชุดที่ควบคุมด้วยกฎแต่ละกฎนั้นจะกำหนดให้มีคลาสข้อมูลคือ Normal หมายถึง ปริมาณฝุ่นละอองปกติ (In control) และ Failed หมายถึง ปริมาณฝุ่นละอองเกินควบคุม (Out of control) เมื่อได้ข้อมูลการควบคุมคุณภาพจากกระบวนการควบคุมเชิงสถิติแล้วนำชุดข้อมูลนี้ไปทำการหาปัจจัยของการเกิดปริมาณอนุภาคฝุ่นละอองด้วยวิธีการเหมืองข้อมูลต่อไป

การหาปัจจัยของการเกิดปริมาณอนุภาคฝุ่นละอองที่เกินการควบคุมนั้น วิทยานิพนธ์ฉบับนี้ได้เลือกเทคนิคต้นไม้ตัดสินใจด้วยอัลกอริทึม C4.5 เนื่องจากเป็นวิธีการจำแนกข้อมูลที่มีประสิทธิภาพและเหมาะสมสำหรับการวิเคราะห์ร่วมกับข้อมูลปริมาณอนุภาคฝุ่นละอองของการผลิตฮาร์ดดิสก์ไดร์ฟ เนื่องจากเป็นวิธีการที่มีประสิทธิภาพและสามารถแสดงต้นไม้ตัดสินใจที่สามารถสร้างกฎได้ง่าย ซึ่งในขั้นตอนการสร้างต้นไม้ตัดสินใจนั้นมีขั้นตอนดังนี้

1. นำชุดข้อมูลการควบคุมคุณภาพแต่ละชุด ไปทำการสร้างต้นไม้ตัดสินใจด้วยอัลกอริทึม C4.5
2. สร้างต้นไม้ตัดสินใจด้วยการแบ่งข้อมูลการเรียนรู้และข้อมูลการทดสอบด้วยวิธี k-fold cross validation ที่มี k = 10
3. เปรียบเทียบข้อมูลประสิทธิภาพของต้นไม้ตัดสินใจของกฎการควบคุมแต่ละกฎที่ได้เพื่อเลือกต้นไม้ตัดสินใจที่มีประสิทธิภาพมากที่สุด
4. นำต้นไม้ตัดสินใจที่ได้จากการเปรียบเทียบประสิทธิภาพมาสรุปหากฎปัจจัยที่ทำให้ปริมาณอนุภาคฝุ่นละอองเกินควบคุม
5. สรุปกฎการควบคุมกระบวนการเชิงสถิติที่จะนำมาใช้สำหรับการควบคุมคุณภาพให้ดีขึ้นและเปรียบเทียบการแจกแจงข้อมูลของข้อมูลที่ควบคุมตามเกณฑ์ของโรงงานและกฎของการควบคุมกระบวนการเชิงสถิติ

5.1.1 การสรุปกฎของการควบคุมกระบวนการที่เหมาะสมสำหรับการควบคุมคุณภาพ

จากการสร้างข้อมูลควบคุมคุณภาพตามกระบวนการควบคุมเชิงสถิติและการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจ โดยการวัดประสิทธิภาพ ความครบถ้วนและความแม่นยำของต้นไม้ตัดสินใจนั้นสามารถสรุปกฎที่เหมาะสมสำหรับการควบคุมคุณภาพเพิ่มเติมได้ดังนี้

5.1.1.1 ข้อมูลพื้นที่ cleaning มีกฎที่เหมาะสมสำหรับการควบคุมคุณภาพคือ กฎข้อที่ 6, 2, 5, 8 และ 1

5.1.1.2 ข้อมูลพื้นที่ HDE มีกฎที่เหมาะสมสำหรับการควบคุมคุณภาพคือ กฎข้อที่ 7 และ 2

5.1.1.3 ข้อมูลพื้นที่ HSA มีกฎที่เหมาะสมสำหรับการควบคุมคุณภาพคือ กฎข้อที่ 7, 8, 6 และ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.1.2 การสรุปปัจจัยที่ทำให้เกิดปริมาณอนุภาคฝุ่นละอองที่เกินควบคุม

จากการสร้างต้นไม้ตัดสินใจที่สรุปต้นไม้ตัดสินใจที่ดีที่สุดในข้อ 5.1.1 นั้นสามารถนำข้อมูลจากต้นไม้ตัดสินใจมาสรุปปัจจัยที่ทำให้เกิดปริมาณอนุภาคฝุ่นละอองเกินควบคุมในรูปแบบของกฎ ซึ่งสามารถสรุปได้ดังนี้

- ปัจจัยที่ทำให้ปริมาณฝุ่นละอองเกินควบคุมในพื้นที่การทำงาน cleaning

ปัจจัยที่ทำให้ปริมาณอนุภาคฝุ่นละอองที่เกินควบคุมในพื้นที่การทำงาน Cleaning วันทำงาน (Day) และกะการทำงาน (Shift)

- ปัจจัยที่ทำให้ปริมาณฝุ่นละอองเกินควบคุมในพื้นที่การทำงาน HDE

ปัจจัยที่ทำให้ปริมาณอนุภาคฝุ่นละอองที่เกินควบคุมในพื้นที่การทำงาน HDE ได้แก่ กระบวนการทำงาน (Process) วันทำงาน (Day) และกะการทำงาน (Shift) กระบวนการทำงานที่เป็นปัจจัยได้แก่ Wipe, Environment, Particle Test, Disk mounting, HSA/VCM mounting, Ramp mounting, Top Clamp mounting, Top Cover mounting, Hyper disk balancer และ Part loading conveyer

- ปัจจัยที่ทำให้ปริมาณฝุ่นละอองเกินควบคุมในพื้นที่การทำงาน HSA

ปัจจัยที่ทำให้ปริมาณอนุภาคฝุ่นละอองที่เกินควบคุมในพื้นที่การทำงาน HSA ได้แก่ กระบวนการทำงาน (Process), วันทำงาน (Day) และ กะการทำงาน (Shift) ซึ่งสำหรับกระบวนการที่ต้องเฝ้าระวังได้แก่ กระบวนการ Final inspection, HSA assembly, Solder inspection, Environment, Tear down และ Pivot

5.1.3 การสรุปเปรียบเทียบประสิทธิภาพต้นไม้ตัดสินใจของข้อมูลที่ควบคุมคุณภาพตามเกณฑ์ของโรงงานและข้อมูลที่ควบคุมคุณภาพด้วยการควบคุมกระบวนการเชิงสถิติ (SPC)

ประสิทธิภาพของต้นไม้ตัดสินใจข้อมูลเกณฑ์ของโรงงานมีประสิทธิภาพดีและดีมาก และประสิทธิภาพของต้นไม้ตัดสินใจเกณฑ์การควบคุมกระบวนการเชิงสถิติก็มีประสิทธิภาพดีและดีมาก อีกทั้งยังสามารถหาปัจจัยการเกิดฝุ่นละอองได้ทุกข้อมูลพื้นที่การทำงาน

ดังนั้นจากผลการศึกษาของกรณีศึกษาดังกล่าว สามารถวิเคราะห์ได้ว่าเมื่อพิจารณาตามพื้นที่การทำงานแล้วสามารถนำกฎการควบคุมกระบวนการเชิงสถิติไปใช้ควบคุมคุณภาพการผลิตได้ ทำให้สามารถหาปัจจัยที่ทำให้เกิดปริมาณฝุ่นละอองเกินควบคุมได้ครอบคลุมและสามารถนำผลดังกล่าวไปควบคุมคุณภาพในอนาคตได้ดียิ่งขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 ข้อจำกัดและข้อเสนอแนะ

เนื่องจากการทดสอบเป็นการทดสอบด้วยกรณีศึกษาของ “ปริมาณอนุภาคฝุ่นละอองของอุตสาหกรรมผลิตฮาร์ดดิสก์ไดร์ฟ” เพียงกรณีศึกษาเดียว ซึ่งข้อมูลที่นำมาวิเคราะห์มีปริมาณน้อย และมาตรฐานในการควบคุมกระบวนการเชิงสถิติที่นำมาพิจารณาเป็นเกณฑ์พื้นฐาน ดังนั้นสำหรับงานวิเคราะห์ข้อมูลในอนาคตควรใช้ข้อมูลสำหรับการวิเคราะห์ที่มากขึ้นและปรับปรุงกฎการควบคุมกระบวนการเชิงสถิติจะพิจารณาเพิ่มเติมเพื่อให้เหมาะสมกับข้อมูลมากขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

[1] Training & Support Department, **"Hard Disk Drive Process Overview"**, Training Document. Hitachi Global Storage Technologies, April 2010.

[2] Manufacturing Environment Engineering Department, **"ESD and Contamination Control"**, Hitachi Global Storage Technologies, April 2009.

[3] R. W. Welker, **"Guideline for Design and Certification of Cleanroom Tooling"**, Clean room Tooling Document, Revised by Don DeLeo/Tim Karlberg, San Jose Site Contamination Control, August 8 2002.

[4] R.S. Chen, K.C. Yeh, C.C. Chang and H.H. Chien. **"Using Data Mining Technology to improve Manufacturing Quality – A Case Study of LCD Driver IC Packaging Industry"**, In Proceeding of the Seventh ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD'06), Las Vegas. USA, pp.115-119, 19-20 June, 2006.

[5] S.H. HE, L. LI and E.S. QI. **"Study on the Continuous Quality Improvement Systems of LED Packaging based on Data Mining"**, In the Wireless Communications, Networking and Mobile Computing, 2007 (WiCom2007), Shanghai, China, pp. 5625-5628, 21-25 September, 2007.

[6] M. Lazzaroni. **"A tool for quality controls in industrial process"**, In the IEEE Instrumentation and Measurement Technology Conference 2009 (I2MTC '09), Singapore, pp.68-73, 5-7 May, 2009.

[7] M.E. Camargo, W.P. Filho, A.I. dos Santos Dullius, J. Maciel and M. Pinto. **"Statistical Quality Control: A Case Study Research"**, In the 4th IEEE International Conference on Management of Innovation and Technology 2008 (ICMIT 2008), Bangkok, Thailand, pp. 746-750, 21-24 September, 2008.

[8] R. Srinivasu, G.S. Reddy and S.R. Rikkula. **"Utility of Quality Control Tools and Statistical Process Control to Improve the Productivity and Quality in an Industry"**, In the International Journal of Reviews in Computing, pp.15-20, 2009-2011.

[9] M. Inada and T. Terano. **"QC chart mining: extracting systematic error patterns from quality control charts"**, In IEEE International Conference on Systems, Man and Cybernetics (ICSCM2005), Vol. 4, pp. 3781-3787, 2005.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[10] N. V. Chawla, K.W. Bowyer, L. O. Hall and W. P. Kegelmeyer. **"SMOTE: Synthetic Minority Over-sampling Technique"**, In Journal of Artificial Intelligence Research, pp. 321 – 357, 2002.

[11] N. V. Chawla. **"Data Mining for Imbalanced Datasets: An Overview"**, In Data Mining and Knowledge Discovery Handbook, Springer US, USA, pp. 853-867, 2005.

[12] C.H. Zhang, L. Di and Z. An. **"Welding Quality Monitoring and Management System Based on Data Mining Technology"**, In Proceeding of the Second International Conference on Machine Learning and Cybernetics 2003, Xi-an, China, pp.13-17, 2003.

[13] A.M.M. Kamal. **"A Data Mining Approach for Improving Manufacturing Process Quality Control"**, In Proceeding of the 2nd International Conference on Next Generation Information Technology (ICNIT2011), Gyeongju, Korea, pp. 207-212, 2011.

[14] E.V. Ramana and P.Ravinder Reddy. **"Integration of Control Charts and Data Mining for Process Control and Quality Improvement"**, In the International Journal of Advances in Engineering & Technology, Vol. 1 Issue 6, pp.640, January, 2012.

[15] R.S. Chen, Y.C. Chen and C.C. Chen. **"Using Data Mining Technology to Design a Quality Control System for manufacturing Industry"**, In ADVANCES in COMMUNICATIONS, COMPUTERS, SYSTEMS, CIRCUITS and DEVICES (CSCC'10), Puerto De La Cruz, Tenerife, pp.272-276, November 30-December 2, 2010.

[16] David M. Levine, **"STATISTICS FOR SIX SIGMA GREEN BELTS WITH MINITAB AND JMP"**, Pearson Education, Inc., New Jersey, 2006.

[17] H.A. Shari, N. Khalid and N. Sahari. **"Statistical Process Control in Plastic Packaging Manufacturing: A Case Study"**, In 2009 International Conference on Electrical Engineering and Informatics, Selangor, Malaysia, pp.199-203, 5-7 August, 2009.

[18] E.S. Abouel Nasr and H. Al-Mubaid. **"Mining Process Control Data Using Machine Learning"**, In the International Conference on Computers & Industrial Engineering 2009 (CIE 2009), Troyes, France, pp.1434-1439, 2009.

[19] H. HE, IEEE and E.A. Garcia. **"Learning from Imbalanced Data"**, In the IEEE Transactions on Knowledge and Data Engineering, Vol.21, No.9, CA, USA, pp.1263-1284, September, 2009.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก

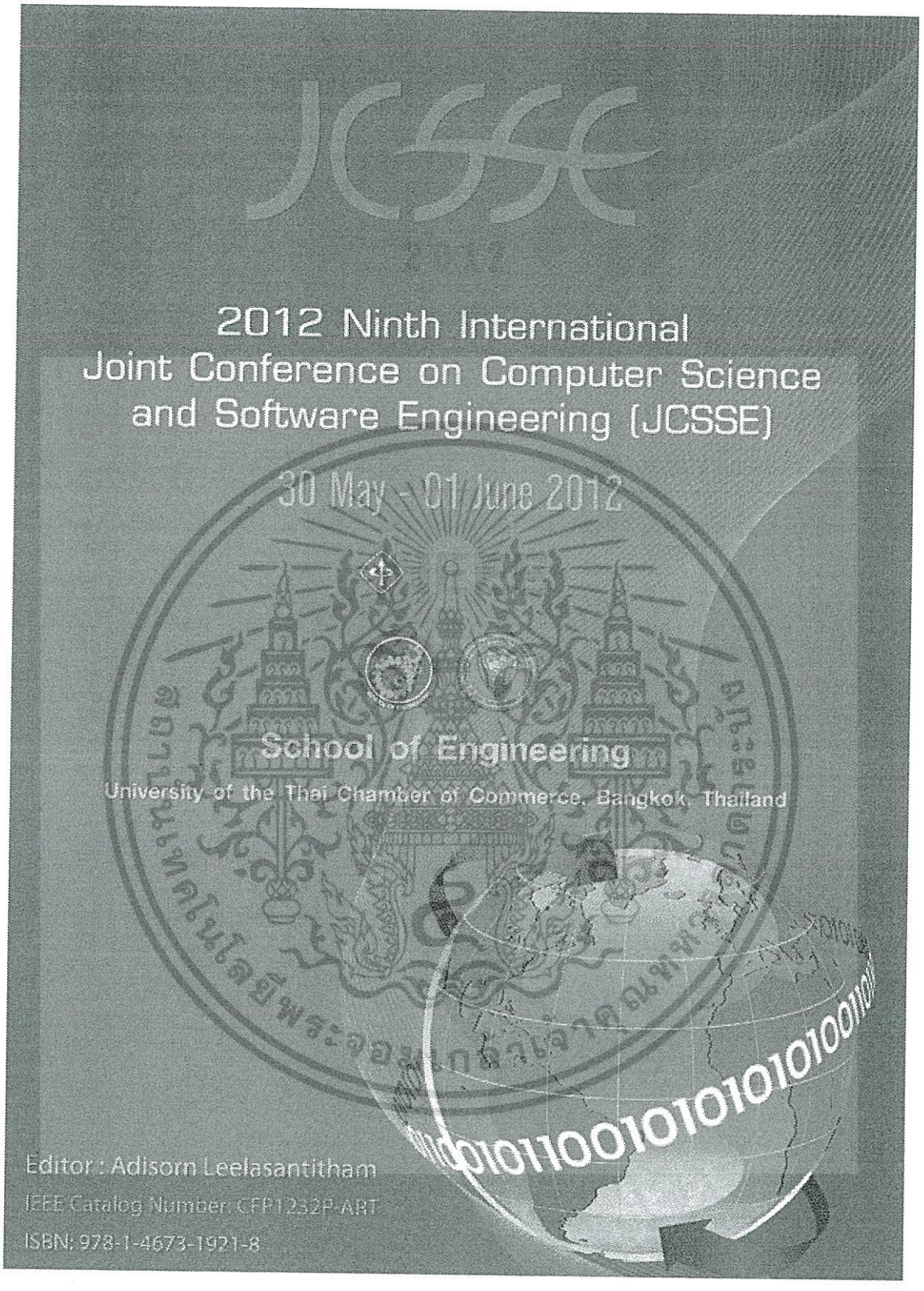
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก

บทความและผลงานวิจัยที่ได้รับการตีพิมพ์

1. Krithamarath Pheng-ubon and Ponrudee Netisopakul. "Development of Database support system for Particle monitoring system". **The 9th International Joint Conference on Computer Science and Software Engineering (JCSSE 2012)**, Thailand, 30 May - 1 June, 2012, pp.31-36
2. กฤษฎ์ฐมรัษฎร์ เพ็งอุบล และ พรฤดี เนติโสภากุล. "การวิเคราะห์ข้อมูลอนุภาคฝุ่นละอองในอุตสาหกรรมผลิตฮาร์ดดิสก์ โดยต้นไม้ตัดสินใจและทฤษฎีเบย์เชี่ยน". **2012 4th International Conference on Knowledge and Smart Technology (KST-2012),** Thailand, July 7-8, 2012, pp.22-28.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2012 Ninth International Joint Conference on Computer Science and Software Engineering (JCSSE)

การพัฒนาระบบฐานข้อมูลเพื่อสนับสนุนงานตรวจจับฝุ่นละออง

Development of Database support system for Particle monitoring system

กฤษฎ์ธรมวัษฎร์ เพ็งอุบล[1] และ ผศ.ดร.พรฤดี เนติโสภากุล[2]

[1]คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ

1 ซอยฉลองกรุง 1 แขวงลาดกระบัง เขตลาดกระบัง กรุงเทพมหานคร 10520 Email: s1066427@kmitl.ac.th

[2]คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ

1 ซอยฉลองกรุง 1 แขวงลาดกระบัง เขตลาดกระบัง กรุงเทพมหานคร 10520 Email: ponrudee@it.kmitl.ac.th

บทคัดย่อ

ระบบงานตรวจจับอนุภาคในอากาศภายในห้องควบคุมปริมาณอนุภาค (Clean room) นั้นเป็นส่วนงานที่สำคัญในการควบคุมและพัฒนาคุณภาพผลิตภัณฑ์ฮาร์ดดิสก์ไดร์ฟให้มีคุณภาพ บทความนี้นำเสนอวิธีการพัฒนาระบบฐานข้อมูลเพื่อสนับสนุนงานตรวจจับฝุ่นละอองของอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟ โดยการใช้ระบบฐานข้อมูลเพื่อแก้ปัญหาการสร้างรายงานสำหรับการวิเคราะห์ข้อมูลจากการสร้างด้วยมนุษย์เป็นการสร้างด้วยระบบฐานข้อมูลแบบอัตโนมัติทำให้มีการนำข้อมูลไปใช้งานได้อย่างหลากหลายและมีประสิทธิภาพในการทำงานมากขึ้น เช่นการนำข้อมูลไปใช้ในการสร้างโปรแกรมประยุกต์อื่นๆ หรือการนำข้อมูลไปวิเคราะห์ข้อมูลด้วยเหมืองข้อมูลเป็นต้น

คำสำคัญ: ห้องควบคุมปริมาณอนุภาคในอากาศ, ฮาร์ดดิสก์ไดร์ฟ, ระบบงานตรวจจับฝุ่นละออง, ระบบฐานข้อมูล

Abstract

Particle monitoring system is used in clean room where used to monitor values of particle. It's an important process in the hard disk drive manufacturing to control and development the quality of hard disk drive. This paper proposed a method to develop the database support system for particle monitoring system in hard disk drive manufacturing by using database system to solve the analytical reporting process that create by manual method. Thus, development of an automatic database support system leads to a wide range of information and the work is effectively finished such as create other applications using the information or make a data analysis by data mining technique.

Keywords: clean room, hard disk, particle monitoring system, database system

1. บทนำ

กระบวนการทำงานของอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟภายใต้สภาวะแวดล้อมการทำงานที่ต้องมีการควบคุมภายในห้องสะอาดนั้นมีความจำเป็นอย่างยิ่งเพื่อให้ได้ฮาร์ดดิสก์ไดร์ฟที่มีคุณภาพดังนั้นการเฝ้าระวังมิให้การเกิดฝุ่นละอองภายในห้องสะอาดมีปริมาณเกินกำหนดนั้นจึงมีความจำเป็นอย่างยิ่งและถือได้ว่าเป็นหัวใจหลักในการผลิตฮาร์ดดิสก์ไดร์ฟ แต่ในกระบวนการเฝ้าระวังดังกล่าวก็ไม่สามารถทำได้อย่างรวดเร็ว ทันที ในบางครั้งเพราะด้วยปริมาณผู้ปฏิบัติการที่จำกัดภายใต้การแข่งขันที่สูงขึ้นในอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟ ผู้ปฏิบัติงานจึงจำเป็นต้องวิเคราะห์ข้อมูลเพื่อแก้ไขปัญหาที่เกิดขึ้นแต่การนำข้อมูลมาวิเคราะห์ในปัจจุบันยังไม่สามารถทำได้อย่างทันท่วงทีเนื่องจากการสร้างรายงานเพื่อการวิเคราะห์ข้อมูลยังต้องทำด้วยมนุษย์ อันจะทำให้เกิดความล่าช้า ผิดพลาด และเสียเวลาในการดำเนินการได้ อีกทั้งยังอาจจะไม่ยืดหยุ่นสำหรับการดำเนินการในรูปแบบที่เพิ่มขึ้นและอาจมีการเปลี่ยนแปลงเพราะความสามารถในการดำเนินการของผู้ใช้มีจำกัดเพื่อเพิ่มศักยภาพการดำเนินงานของระบบงานเฝ้าระวังฝุ่นละออง จึงมีการพัฒนาระบบฐานข้อมูลขึ้นเพื่อให้ผู้ใช้สามารถนำข้อมูลไปใช้งานได้อย่างเหมาะสมกับการวิเคราะห์ข้อมูลที่ต้องการต่อไปในอนาคต

2. งานและทฤษฎีที่เกี่ยวข้อง

2.1 กระบวนการผลิตฮาร์ดดิสก์ไดร์ฟ (Hard Disk Drive Manufacturing Process)

ในการระบวนการผลิตฮาร์ดดิสก์ไดร์ฟนั้นจะประกอบด้วยการประกอบชิ้นส่วน (Hard Disk Part Function) ต่างๆ ซึ่งเป็นองค์ประกอบของฮาร์ดดิสก์เข้าด้วยกัน ประกอบด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. Voice Coil Motor controller (VCM) [1] เป็นชิ้นส่วนที่มีหน้าที่ควบคุมการหมุนของมอเตอร์และการเคลื่อนที่ของหัวอ่าน

2. Air Filter ใช้สำหรับการกรองอนุภาคที่เป็นอากาศ

3. Disk เป็นชั้นแม่เหล็กขนาดเล็กที่ใช้เขียนระเบียนข้อมูลไบนารี

4. Base เป็นฐานในการบรรจุองค์ประกอบทั้งหมดของฮาร์ดดิสก์

5. Latch เป็นสลักที่ใช้สำหรับยึด VCM กับดิสก์

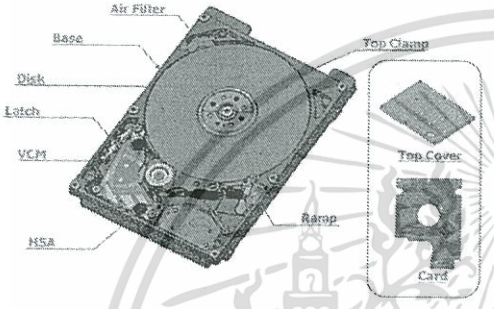
6. Head Stack Assembly (HSA) เป็นหัวอ่าน-เขียนสัญญาณแม่เหล็ก

7. Ramp เป็นส่วนประกอบที่ประกอบด้วยหัวเขียน-อ่าน ที่อยู่เหนือดิสก์

8. Card เป็นแผงวงจรควบคุมต่างๆ ของฮาร์ดดิสก์ไดร์ฟ

9. Top Clamp เป็นชิ้นส่วนอลูมิเนียมใช้สำหรับยึดแผ่นดิสก์กับมอเตอร์

10. Top Cover เป็นสแตนเลสที่ใช้ป้องกันวัสดุจากภายนอก



รูปที่ 1 ส่วนประกอบของฮาร์ดดิสก์ไดร์ฟ

กระบวนการผลิตฮาร์ดดิสก์ไดร์ฟมีการทำงานที่ผสมผสานระหว่างการทำงานของเครื่องจักรกึ่งอัตโนมัติและแรงงานมนุษย์ กระบวนการผลิตนี้เป็น[2] การประกอบของ HSA, Disk, Motor, VCM, และชิ้นส่วนเล็กๆ อื่นๆ ลงบน Base แล้วจึงปิดส่วนประกอบทั้งหมดด้วย Top Cover กระบวนการดังกล่าวนั้นเรียกว่า "Hard Disk Assembly" (HDA) กระบวนการผลิตต้องดำเนินการภายในห้องสะอาด (Clean Room) ซึ่งมีการควบคุมมาตรฐานปริมาณฝุ่นละอองในพื้นที่การทำงาน โดยมาตรฐานของปริมาณฝุ่นละอองที่เกิดขึ้นจะมีขนาดหรือจำนวนที่มากน้อยเพียงใดนั้นจะขึ้นอยู่กับความสำคัญของชิ้นส่วนหรือกระบวนการผลิตในแต่ละส่วน อันได้แก่กระบวนการผลิต 3 กระบวนการ ได้แก่

1. HSA (Head Stack Assembly) Process คือกระบวนการประกอบหัวอ่านและเขียนสัญญาณแม่เหล็ก HGA (Head Gimbals Assembly) กับชิ้นส่วน ต่างๆ เช่น แผ่นวงจรไฟฟ้า (Flex Cable) และตัวสร้างสนามแม่เหล็ก (Carriage) รวมถึงการทดสอบต่างๆ

2. HDE (Hard Disk Enclosure) Process คือกระบวนการประกอบ HSA กับชิ้นส่วนต่างๆ เข้าด้วยกัน (ยกเว้น card) เช่น Disk, VCM แล้วผ่านกระบวนการสร้าง Servo Pattern ลงบนแผ่น Disk รวมถึงการตรวจสอบฝุ่นละอองและรอยรั่ว

3. HDD (Hard Disk Drive) Process คือกระบวนการขั้นตอนสุดท้ายในการผลิตฮาร์ดดิสก์ไดร์ฟ แบ่งการทำงานเป็น 3 ส่วน ได้แก่

3.1 Final Assembly คือการประกอบแผงวงจรควบคุม

3.2 Test คือการทดสอบการทำงานของฮาร์ดดิสก์ไดร์ฟในสภาวะการทำงานต่างๆ

3.3 Featuring and Packaging คือการติดฉลากสินค้าและการใส่ข้อมูลจำเพาะบางอย่าง รวมถึงการบรรจุหีบห่อผลิตภัณฑ์ เพื่อเตรียมส่งมอบ

2.2 กระบวนการทำงานของห้องสะอาด (Clean room) ในอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟ

ห้องสะอาด[3] (Clean room) หมายถึงห้องที่ได้รับการปรับสภาวะอากาศ อุณหภูมิ ความชื้น ความดันและติดตั้งระบบกรองอากาศเป็นพิเศษ ให้อากาศปราศจากฝุ่นในระดับอนุภาคซึ่งมีขนาดเล็กมาก ไม่สามารถมองเห็นด้วยตาเปล่าขนาดเส้นผ่านศูนย์กลางไม่เกิน 0.5 ไมครอน อนุภาคดังกล่าว เช่น ควันน้ำมัน ควันบุหรี่และควันอื่นๆ สำหรับห้องสะอาดที่ใช้ในอุตสาหกรรมนั้นใช้ในอุตสาหกรรมอิเล็กทรอนิกส์ นาโนอิเล็กทรอนิกส์ ออปติคอลไฟเบอร์ การพิมพ์ การผลิตฟิล์มถ่ายรูป คอมพิวเตอร์ ซึ่งจะมีการกำจัดอนุภาคในอากาศและการควบคุมสภาวะต่างๆ ภายในห้องให้เหมาะสมกับลักษณะงาน กระบวนการผลิตและผลิตภัณฑ์แต่ละชนิด ความดันในห้องสะอาดควรเป็นบวกเสมอเพื่อป้องกันการปนเปื้อนจากอนุภาคภายนอก ระดับความชื้นความสะอาดของห้องจะต่างกันขึ้นอยู่กับประเภทของอุตสาหกรรม สำหรับมาตรฐานที่ใช้เกี่ยวกับห้องสะอาดนั้นได้แก่ มาตรฐานการปนเปื้อนของพื้นที่ควบคุม (Standard for Contamination Control Area) ซึ่งนิยมใช้กันในอุตสาหกรรมต่างๆ รวมทั้งอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟนั้นได้แก่ ISO 14644-1 clean room standards, 1999 (Clean rooms and Associated controlled environments, Part1: Classification of air cleanliness) ซึ่งเป็นมาตรฐานสากลได้กำหนดระดับความสะอาดสูงสุดคือ ISO1 ซึ่งจะมีอนุภาคที่มีขนาดตั้งแต่ 0.1 ไมครอนขึ้นไปจำนวน 10 อนุภาคต่อลูกบาศก์เมตร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 1 มาตรฐานการปนเปื้อนของพื้นที่ควบคุม
ในอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟ

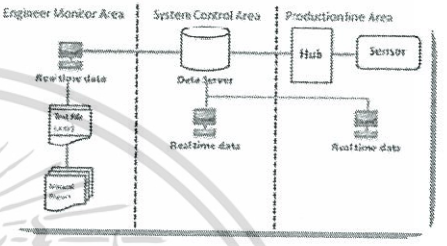
Clean Room Class	Process Area	Cleanliness Specification
10000	HSA and Cleaning	10,000 particles/ft^3 at 0.5μm 30,000 particles/ft^3 at 0.3μm
100	HDE	100 particles/ft^3 at 0.5μm 300 particles/ft^3 at 0.3μm
10	HDE Assembly Station	10 particles/ft^3 at 0.5μm 30 particles/ft^3 at 0.3μm

จากตารางที่ 1 มาตรฐานการปนเปื้อนของพื้นที่ควบคุมในอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟและพื้นควบคุมที่ที่กำหนด โดยที่ระดับ Clean room Class 10 เป็นมาตรฐานการควบคุมในพื้นที่ HDE Assembly Station ระดับ Clean room Class 100 เป็นมาตรฐานการควบคุมในพื้นที่ HDE Process และระดับ Clean room Class 10000 เป็นมาตรฐานการควบคุมในพื้นที่ HSA Process และ Cleaning Area

การประกอบฮาร์ดดิสก์ไดร์ฟต้องดำเนินการในห้องสะอาดที่ระดับคลาส 100 เพราะ [4] อากาศที่หมุนเวียนอยู่ภายในห้อง Main Assembly ต้องมีการหมุนเวียนที่ดี มีการดักฝุ่นและสารเคมีต่างๆในอากาศเพื่อกรองให้อากาศที่อยู่ภายในห้องมีความสะอาดระดับคลาส 100 ทำให้อากาศมีการถ่ายเทอยู่ตลอดเวลา ซึ่งการควบคุมความสะอาดนั้นจะดูแลพิเศษทั้งอากาศและอุปกรณ์ควบคุมความสะอาดต่างๆรวมถึงบุคลากร เพื่อให้ความสะอาดของห้องอยู่ในเกณฑ์ที่ยอมรับได้ในระดับคลาส 10

2.3 กระบวนการทำงานของระบบการตรวจจับฝุ่นละออง (Airborne Particle Monitoring System)

กระบวนการทำงานของระบบการตรวจจับฝุ่นละอองเริ่มต้นที่พื้นที่การผลิตโดยแต่ละสถานีการทำงานจะติดตั้งเครื่องตรวจจับฝุ่นละอองเอาไว้ในจุดต่างๆ ที่ได้มีการออกแบบในลักษณะที่ต่างกัน โดยเซ็นเซอร์จะดูดอากาศรอบๆ บริเวณท่อดูดในปริมาณ 1 ลูกบาศก์ฟุตเพื่อตรวจจับปริมาณฝุ่นละอองที่เกิดขึ้นในช่วงเวลาที่กำหนดแล้วตัวเซ็นเซอร์สำหรับฝุ่นละอองจะวิเคราะห์ขนาดและปริมาณของฝุ่นละอองที่เกิดขึ้นแล้วส่งข้อมูลผ่านระบบเครือข่ายโปรโตคอลไปยัง เครื่องเซิร์ฟเวอร์ที่เก็บข้อมูลหลัก โดยผู้ปฏิบัติงานภายในห้องสะอาดสามารถดูข้อมูลดังกล่าวได้ทันทีด้วยเครื่อง Real time monitoring ภายในห้องสะอาด ส่วนวิศวกรสามารถดูข้อมูลได้โดยวิธีการเชื่อมต่อเข้าสู่เครื่องเซิร์ฟเวอร์หลักและนำข้อมูลที่เก็บในเซิร์ฟเวอร์ลักษณะเท็กซ์ไฟล์ รูปแบบ Comma Separated Values (CSV) มาจัดเรียงและสร้างรายงานเชิงกราฟด้วยโปรแกรมไมโครซอฟท์เอ็กซ์เซลแสดงดังรูปที่ 3



รูปที่ 3 กระบวนการทำงานของระบบงานตรวจจับฝุ่นละออง

ด้วยขนาดของไฟล์ในแต่ละเดือนมีขนาดใหญ่และจำนวนข้อมูลมีปริมาณมาก จึงทำให้การค้นหาข้อมูลหรือการตรวจสอบทำได้ยากและเสียเวลาในการจัดเรียงข้อมูลใหม่ จึงมีการต่อยอดการพัฒนาระบบฐานข้อมูลเพื่อสนับสนุนงานตรวจจับฝุ่นละอองบนเว็บเพื่อให้การจัดเก็บและสร้างรายงานมีประสิทธิภาพ

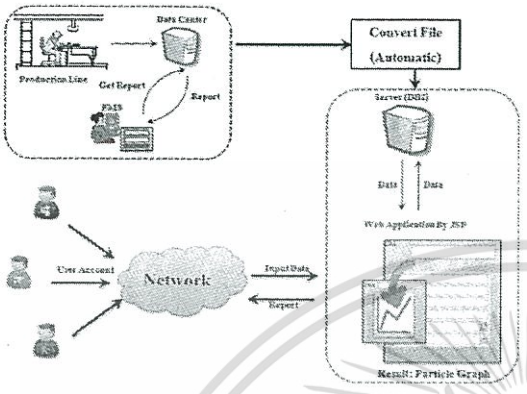
3. การพัฒนาระบบฐานข้อมูลเพื่อสนับสนุนงานตรวจจับฝุ่นละออง

ข้อมูลที่ได้รับจากโปรแกรมของอุปกรณ์ตรวจจับฝุ่นละอองในพื้นที่การผลิตอยู่ในรูปแบบที่ยังไม่ได้รับการประมวลผลตรงกับความต้องการขององค์กรจึงได้มีการพัฒนาระบบฐานข้อมูลขึ้นเพื่อนำข้อมูลจากอุปกรณ์ตรวจจับฝุ่นละอองมาใช้ในการประมวลผลและเป็นประโยชน์ต่อการใช้งานข้อมูลได้ง่ายขึ้น โดยแต่เดิมนั้นข้อมูลยังอยู่ในรูปข้อมูลที่เป็นข้อความ ทำให้ยากต่อการทำความเข้าใจจึงได้นำข้อมูลดังกล่าวมาจัดเก็บในรูปแบบระบบฐานข้อมูลและพัฒนาเว็บ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แอพพลิเคชั่นเพื่อทำให้การวิเคราะห์และการสืบค้นข้อมูลเป็นไปอย่างมีประสิทธิภาพ ครอบคลุมในทุกมิติที่ต้องการและเข้าถึงได้อย่างรวดเร็ว

3.1 หลักการทำงานของระบบฐานข้อมูลเพื่อสนับสนุนงานตรวจจับฝุ่นละออง

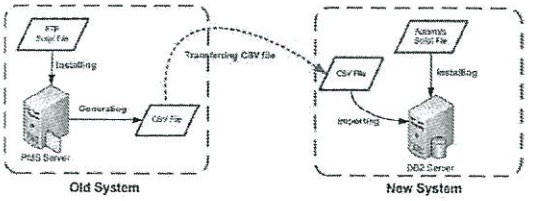


รูปที่ 4 สถาปัตยกรรมภาพรวมของระบบฐานข้อมูลเพื่อสนับสนุนงานตรวจจับฝุ่นละออง

หลักการทำงานของระบบฐานข้อมูลเพื่อสนับสนุนการตรวจจับฝุ่นละอองนั้นระบบที่สร้างขึ้นจะประกอบด้วยสองส่วนได้แก่ ส่วนแรกเป็นการย้ายข้อมูลที่ได้จากสายการผลิต (Production Line) ที่เก็บไว้ในเครื่องศูนย์ข้อมูลไปเก็บไว้ที่เครื่องเซิร์ฟเวอร์เครื่องใหม่จากนั้นจะแปลงข้อมูลจากเท็กซ์ไฟล์เป็นระบบฐานข้อมูล ส่วนที่สองเป็นการนำข้อมูลที่เก็บในฐานข้อมูลมาใช้ผ่านหน้าเว็บแอปพลิเคชั่น (Web Application) โดยสามารถออกรายงานข้อมูลเกี่ยวกับจำนวนฝุ่นละอองเป็นกราฟได้แบบออนไลน์ซึ่งผู้ใช้จะสามารถเข้าสู่ระบบได้โดยผ่านเครือข่ายเสมือนส่วนตัว (VPN: Virtual Private Network)

3.2 สถาปัตยกรรมส่วนการแปลงข้อมูล

การทำงานของระบบเริ่มต้นที่นำข้อมูลในรูปแบบข้อความที่เก็บในเซิร์ฟเวอร์ของระบบเดิมมาเก็บในรูปแบบระบบฐานข้อมูลในเครื่องเซิร์ฟเวอร์เพื่อสร้างโปรแกรมประยุกต์ที่ใช้งานข้อมูลจากฐานข้อมูลดังกล่าว ซึ่งในการทำงานของส่วนดังกล่าวนี้จำเป็นต้องทำโดยอัตโนมัติเพื่อลดเวลาและเพิ่มความสะดวกในการใช้งาน

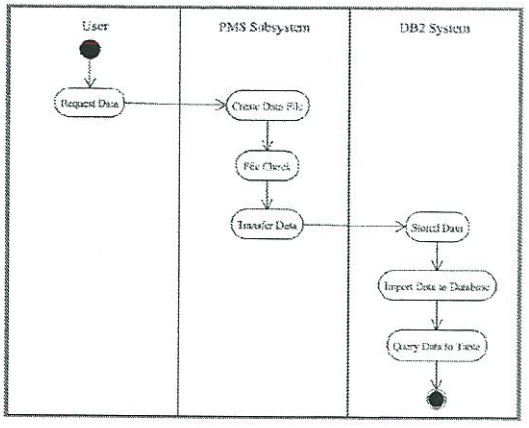


รูปที่ 5 สถาปัตยกรรมส่วนการแปลงข้อมูล

จากรูปที่ 5 การทำงานส่วนการแปลงข้อมูลจากระบบเดิมที่เป็นระบบของ Airborne Particle Monitoring System เป็นระบบ Database Support System for Particle Monitoring เริ่มต้นจากการนำข้อมูลแบบเท็กซ์ไฟล์ออกจากเครื่อง Particle Monitoring Server ซึ่งจะตั้งค่าการสร้างเท็กซ์ไฟล์จากโปรแกรมเฉพาะภายในเครื่องเซิร์ฟเวอร์และไฟล์ที่ได้จะถูกเก็บเอาไว้ในเครื่อง Particle Monitoring Server ขณะเดียวกันจะติดตั้ง FTP Script File ในการส่งข้อมูลนี้ไปยังเครื่องเซิร์ฟเวอร์ในระบบ Database Support System for Particle Monitoring โดยอัตโนมัติ การย้ายข้อมูลจากระบบเดิมไปยังเครื่องเซิร์ฟเวอร์ของระบบใหม่ด้วย FTP Script File จะระบุไอพีแอดเดรสของเครื่องต้นทางและปลายผ่านระบบเครือข่ายภายใน ส่วนระบบฐานข้อมูลนั้นเครื่องเซิร์ฟเวอร์จะติดตั้ง Automatic Script File ในการตั้งค่าการทำงานโดยอัตโนมัติต่างๆ และนำเข้าข้อมูลจากเท็กซ์ไฟล์เข้าสู่โปรแกรมระบบฐานข้อมูล DB2 และสั่งการทำงานภายในด้วย Deployment script เพื่อนำข้อมูลเข้าสู่ระบบฐานข้อมูล DB2 โดยได้ปรับข้อมูลจากเท็กซ์ไฟล์ลงในฐานข้อมูลที่ได้ออกแบบมาใหม่ตามความต้องการของระบบ

สำหรับแอคทิวิตี้ไดอะแกรม (Activity Diagram) ของส่วนงานการแปลงข้อมูลเมื่อผู้ใช้งานเรียกขอข้อมูลโดยการตั้งค่าใน Particle Monitoring Server แล้ว จะมีการสร้างไฟล์ข้อมูลขึ้นแล้วตรวจสอบว่าข้อมูลที่ได้ถูกต้องหรือไม่ จากนั้นจะย้ายข้อมูลไปยังเครื่องเซิร์ฟเวอร์ของระบบฐานข้อมูล แล้วจัดเก็บข้อมูลที่ย้ายเรียบร้อยในโฟลเดอร์ที่กำหนด จากนั้นนำข้อมูลจากไฟล์ข้อมูลดังกล่าวลงในระบบฐานข้อมูลที่สร้างไว้ในโปรแกรมระบบฐานข้อมูล DB2 โดยการนำข้อมูลดังกล่าวลงในตารางที่เตรียมไว้ ซึ่งสามารถเขียน Activity diagram ของส่วนการแปลงข้อมูลได้ดังรูปที่ 6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

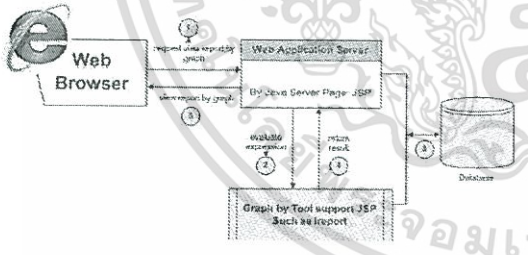


รูปที่ 6 Activity Diagram สำหรับส่วนงานการแปลงข้อมูล

3.3 สถาปัตยกรรมส่วนการจัดทำระบบรายงาน

การทำงานในส่วนจัดทำระบบรายงานข้อมูลของปริมาณฝุ่นละอองจะแสดงในลักษณะของกราฟ โดยผู้ใช้งานจะร้องขอรายงานของข้อมูลปริมาณฝุ่นละอองผ่านทาง Web browser โดยผ่าน Application Server พัฒนาโดยภาษา JSP ซึ่งการแสดงรายงานชนิดกราฟจะใช้เครื่องมือที่สนับสนุนการทำงานด้วย JSP เพื่อสร้างรายงานโดยจะสร้างรายงานจากข้อมูลที่เก็บอยู่ในฐานข้อมูลดังรูปที่ 7

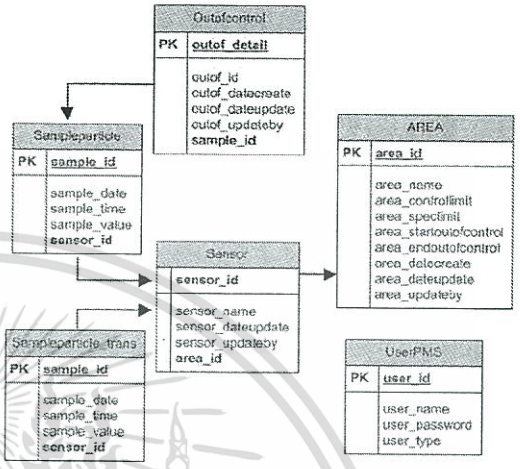
การพัฒนาแอพพลิเคชั่นใช้เทคโนโลยี J2EE, ภาษา JSP พัฒนาติดต่อกับฐานข้อมูล DB2 ทำงานบนเว็บเซิร์ฟเวอร์ apache Tomcat 6 และส่วนของอาวาไลบรารีสำหรับการสร้างรายงานรูปแบบต่างๆ ได้แก่ JFreeChart



รูปที่ 7 สถาปัตยกรรมของส่วนการจัดทำระบบรายงาน

การทำงานของระบบฐานข้อมูลเพื่อสนับสนุนงานตรวจจับฝุ่นละอองบนเว็บ เริ่มจากการเชื่อมต่อกับระบบฐานข้อมูลซึ่งมีการออกแบบการจัดเก็บข้อมูลปริมาณฝุ่นละอองตามวันและเวลาของแต่ละพื้นที่การทำงานเพื่อนำข้อมูลปริมาณฝุ่นละอองจากฐานข้อมูลมาใช้สำหรับประมวลผลและแสดงรายงานบนเว็บ

3.4 การออกแบบระบบฐานข้อมูลและตารางความสัมพันธ์



รูปที่ 8 โครงสร้างฐานข้อมูลของระบบการสร้างรายงาน

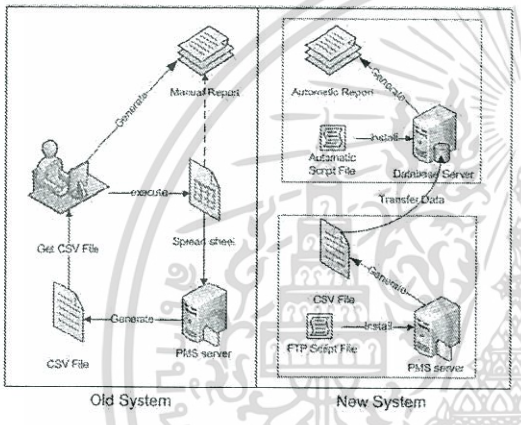
การออกแบบระบบฐานข้อมูลสำหรับระบบฐานข้อมูลประกอบด้วย ตารางสำหรับการจัดเก็บข้อมูลปริมาณฝุ่นละออง ดังนี้

1. ตาราง SENSOR เป็นตารางที่จัดเก็บอุปกรณ์เซ็นเซอร์ที่ใช้ในพื้นที่การผลิต
2. ตาราง AREA เป็นตารางที่ใช้สำหรับจัดเก็บพื้นที่การผลิต
3. ตาราง SAMPLEPARTICLE เป็นตารางที่ใช้จัดเก็บค่าปริมาณฝุ่นละอองที่ได้จากสถานการผลิต โดยจะมีการแปลงข้อมูลเข้าสู่ไฟล์ลงสู่ตารางนี้อัตโนมัติ
4. ตาราง SAMPLEPARTICLE_TRANS เป็นตารางที่ใช้จัดเก็บข้อมูลชั่วคราวเพื่อเตรียมจัดเก็บลงในตาราง SENSOR และตาราง SAMPLEPARTICLE
5. ตาราง OUTOFCONTROL เป็นตารางที่จัดเก็บข้อมูลฝุ่นละอองที่เกินมาตรฐานที่กำหนด
6. ตาราง USERPMS เป็นตารางที่ใช้จัดเก็บข้อมูลผู้ใช้งานระบบ

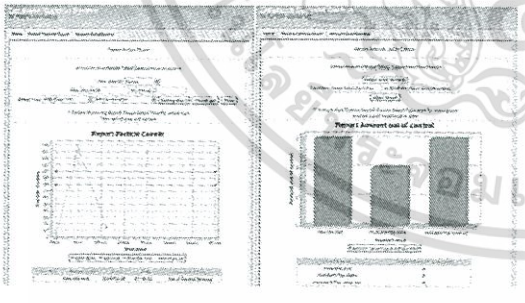
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.5 เปรียบเทียบการทำงานของระบบเดิมกับระบบใหม่

การทำงานของระบบเดิมนั้นผู้ใช้งานต้องดำเนินการแปลงข้อมูลให้อยู่ในรูปแบบที่สามารถใช้งานได้ด้วยตัวเอง ซึ่งทำให้เกิดความล่าช้าในการทำงานวิเคราะห์ข้อมูลที่มีปริมาณมากและอาจไม่สมตรงกับรูปแบบที่ต้องการ แต่ในการทำงานของระบบใหม่ผู้ใช้งานไม่ต้องดำเนินการแปลงข้อมูลด้วยตนเองและในการสร้างรายงานต่างๆนั้นระบบการทำงานเดิมผู้ใช้จะต้องนำข้อมูลเท็กซ์ไฟล์มาจัดเรียงด้วยโปรแกรมสำเร็จรูปอื่นๆ ด้วยตนเองซึ่งไม่สะดวกและอาจผิดพลาดได้ แต่ในระบบการจัดทำรายงานด้วยเว็บแอพพลิเคชั่นจะให้ผู้ใช้สามารถจัดทำรายงานได้สะดวกและอยู่ในรูปแบบที่ผู้ใช้งานต้องการ อีกทั้งยังมีความถูกต้องด้วย โดยเปรียบเทียบการทำงานของระบบเดิมกับระบบใหม่ได้ดังรูปที่ 10 และภาพหน้าจอการแสดงผลรายงานของเว็บแอพพลิเคชั่นของระบบฐานข้อมูลเพื่อสนับสนุนงานตรวจจับฝุ่นละอองได้ดังรูปที่ 11



รูปที่ 10 เปรียบเทียบการทำงานของระบบเดิมกับระบบใหม่



รูปที่ 11 หน้าจอการออกรายงานแบบกราฟของเว็บแอพพลิเคชั่น

4. สรุป

บทความฉบับนี้เสนอกระบวนการในการพัฒนาระบบฐานข้อมูลเพื่อสนับสนุนงานตรวจจับฝุ่นละอองของอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟ โดยนำเสนอปัญหาของการดำเนินการเดิมซึ่งข้อมูลที่ได้รับนั้นไม่อยู่ในรูปแบบที่เหมาะสมสำหรับการวิเคราะห์ข้อมูลและการจัดทำรายงานแบบกราฟ ซึ่งเสนอแนวทางแก้ปัญหาด้วยการนำข้อมูลจัดเก็บในรูปแบบของระบบฐานข้อมูลแบบอัตโนมัติเพื่อพัฒนาเว็บแอพพลิเคชั่นที่สามารถจัดทำรายงานในรูปแบบต่างๆ ที่ผู้ใช้ต้องการทั้งแบบกราฟและข้อความ ซึ่งทำให้ผู้ใช้งานสามารถใช้งานข้อมูลได้อย่างสะดวกและมีประสิทธิภาพ อีกทั้งข้อมูลที่จัดเก็บในระบบฐานข้อมูลยังสามารถนำไปใช้ในการวิเคราะห์ข้อมูลในรูปแบบอื่นๆ ได้ในอนาคต

5. กิตติกรรมประกาศ

งานชิ้นนี้เป็นส่วนหนึ่งของโครงการพัฒนาทรัพยากรมนุษย์ในอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟซึ่งได้รับทุนสนับสนุนจากศูนย์วิจัยร่วมเฉพาะทาง (Industry/University Cooperative Research Center, I/UCRC), สถาบันฮาร์ดดิสก์ไดร์ฟ (Hard disk Drive Institute, HDDI) ภายใต้การควบคุมของสำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ (สวทช.) กระทรวงวิทยาศาสตร์และเทคโนโลยี และยังได้รับการสนับสนุนจากวิทยาลัยร่วมด้านเทคโนโลยีการบันทึกข้อมูลและการประยุกต์ใช้งาน (Data Storage Technology and Applications) สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง โดยความร่วมมือของบริษัท อิคาชิ โกลบอลสตอเรจ เทคโนโลยี ประเทศไทย (HGST)

เอกสารอ้างอิง

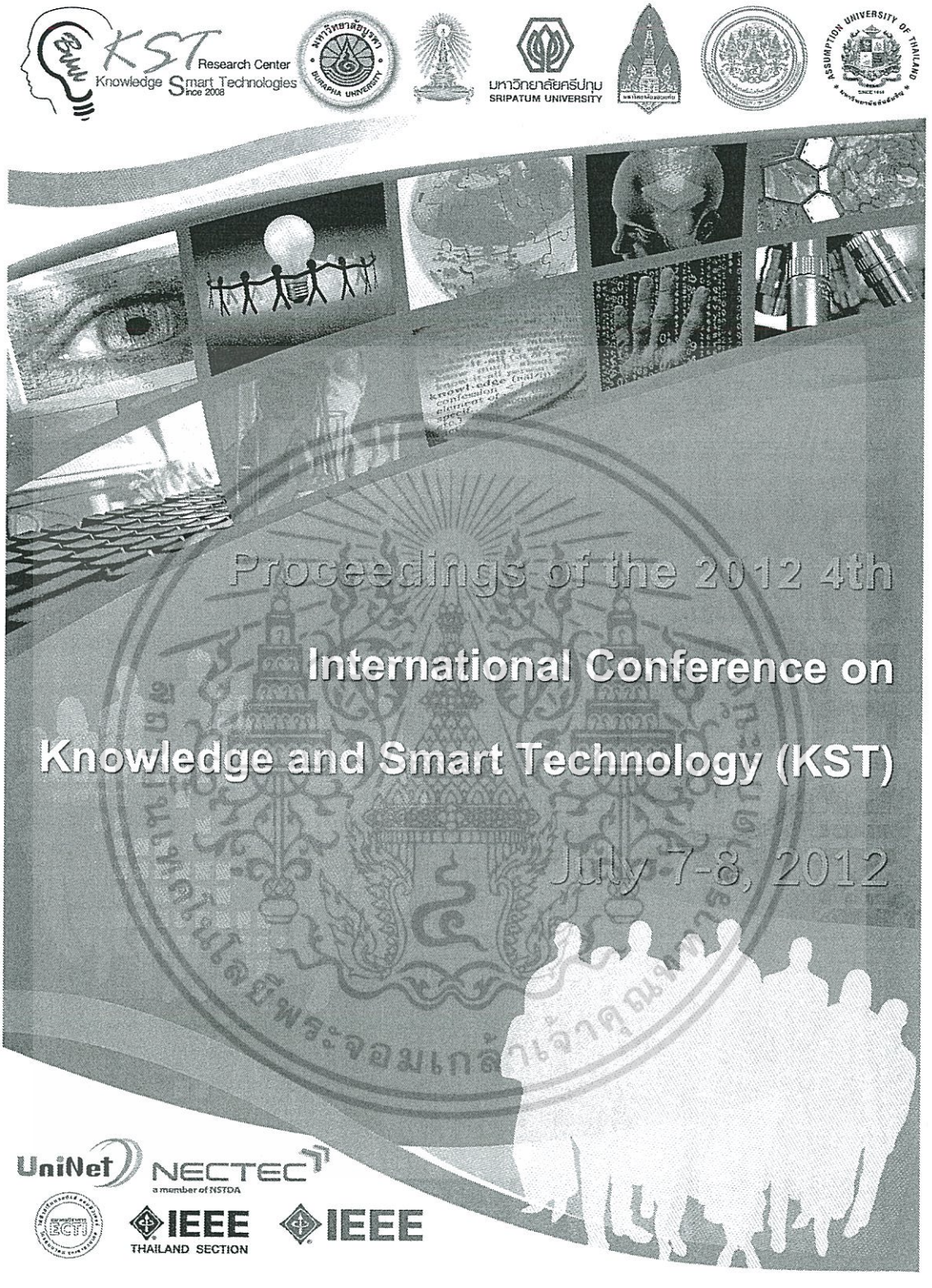
[1] IDEMA, "MATERIALS USED IN HARD DISK DRIVES", IDEMA STANDARDS Micro contamination, no. M2-98, pp.1-6, 1998.

[2] B.C. Yap, J. Turangan, "Field-induced Charging and FIM ESD Tests on GMR Heads in Hard Disk Assembly", ESD Association, pp. 159. 2001.

[3] F. Mrad, "The Characterization of a Clean Room Assembly Process", INDUSTRY APPLICATIONS, vol. 35, no. 2, MARCH/APRIL. 1999.

[4] Particle Measuring Systems, Inc, "Facility Net Operations Manual," Colorado, 2004.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2012 4th Conference on Knowledge and Smart Technology (KST)

การวิเคราะห์ข้อมูลอนุภาคฝุ่นละอองในอุตสาหกรรมผลิตฮาร์ดดิสก์ โดยต้นไม้ตัดสินใจและทฤษฎีเบย์เชี่ยน

Analysis of Particle data in Hard disk Manufacturing by Decision Tree Induction and Bayesian Theorem

กฤษฎ์ฐมรัษฎร์ เพ็งอุบล และ พรฤดี เนติโสภากุล
คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
แขวงลาดกระบัง เขตลาดกระบัง กรุงเทพมหานคร 10520
Email: s1066427@kmitl.ac.th, ponrudee@it.kmitl.ac.th

บทคัดย่อ

งานวิจัยนี้เสนอการวิเคราะห์ข้อมูลปัจจัยที่ทำให้ปริมาณอนุภาคฝุ่นละอองในอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟเกินปริมาณควบคุม โดยใช้เทคนิคการจำแนกข้อมูล (Data Classification) ด้วยหลักการต้นไม้ตัดสินใจ (Decision Tree) และทฤษฎีเบย์เชี่ยน (Bayesian Theorem) ของโรงงานผลิตฮาร์ดดิสก์แห่งหนึ่งและทำการวิเคราะห์ด้วยโปรแกรม Weka โดยพิจารณาเปรียบเทียบตัวจำแนกข้อมูลและสร้างตัวแบบที่เหมาะสมในการวิเคราะห์ข้อมูลดังกล่าว ซึ่งพบว่าเทคนิคการจำแนกทั้ง 2 ประเภทนั้นให้ค่าความถูกต้องในการจำแนกข้อมูลใกล้เคียงกันและวิธี Hold-out cross validation ที่เหมาะสมคือ ตัวแบบ 90:10 ค่าความถูกต้องคือ 99.29% ส่วนของเวลาในการสร้างตัวแบบนั้นทฤษฎีเบย์เชี่ยนใช้เวลาน้อยกว่าหลักการต้นไม้ตัดสินใจ โดยใช้เวลา 0.42 และ 0.77 วินาที ตามลำดับ

คำสำคัญ: การวิเคราะห์ข้อมูลอนุภาคฝุ่นละออง, ต้นไม้ตัดสินใจ, *ทฤษฎีเบย์เชี่ยน*

Abstract

This paper presents the factor analysis of out of control particle data in hard disk drive manufacturing using decision tree and Bayesian theorem. The result showed that decision tree and Bayesian theorem are equally good classification techniques because they have the same best accuracy of 99.29% using Hold-out cross validation 90:10. Times taken to build models for Bayesian theorem and decision tree are 0.42 and 0.77, respectively.

Key Words: Analysis of Particle data, Decision tree, Bayesian Theorem

1. บทนำ

ปริมาณอนุภาคของฝุ่นละอองภายในห้องควบคุมปริมาณอนุภาคของโรงงานอุตสาหกรรมผลิตฮาร์ดดิสก์ไดร์ฟนั้นมีความสำคัญต่อคุณภาพการผลิตของฮาร์ดดิสก์ไดร์ฟ ซึ่งข้อมูลปริมาณอนุภาคดังกล่าวนั้นจะมีการตรวจจับและเก็บข้อมูลด้วยเครื่องมือตรวจจับทางวิศวกรรมและเก็บข้อมูลปริมาณอนุภาคฝุ่นละอองไว้ด้วยโปรแกรมที่ผูกติดมากับอุปกรณ์ดังกล่าว ข้อมูลปริมาณอนุภาคฝุ่นละอองนั้นจะใช้เพื่อการควบคุมคุณภาพของการผลิตในขั้นตอนและพื้นที่การผลิตที่ต่างกัน โดยข้อมูลดังกล่าวนั้นเป็นข้อมูลที่มีปริมาณมากทำให้การวิเคราะห์หาปัจจัยที่เกี่ยวข้องในการเกิดปริมาณอนุภาคฝุ่นละอองที่เกินกว่าระดับควบคุมนั้นทำได้ยาก ดังนั้นเพื่อวิเคราะห์หาปัจจัยที่เกี่ยวข้องในการเกิดสภาวะเกินการควบคุม (Out of control) จึงเสนอเทคนิคการจำแนกข้อมูล (Data Classification) มาช่วยในการวิเคราะห์หาปัจจัยที่เกี่ยวข้อง โดยการสร้างต้นไม้ตัดสินใจ (Decision Tree) และทฤษฎีของเบย์เชี่ยน (Bayesian Theorem) เพื่อเปรียบเทียบประสิทธิภาพและความแม่นยำของตัวจำแนก เพื่อให้ได้ตัวจำแนกที่เหมาะสมในการประยุกต์ใช้ต่อไป

The research is supported by granted # HDD-01-52-02M.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ทฤษฎีที่และงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์ข้อมูลปริมาณอนุภาคฝุ่นละอองภายในห้องควบคุมจะอธิบายเป็น 2 ส่วนได้แก่ ส่วนของทฤษฎีที่เกี่ยวข้องในการวิเคราะห์ข้อมูลครั้งนี้ และงานวิจัยที่เกี่ยวข้องซึ่งเป็นงานวิจัยที่นำเทคนิคและวิธีการในส่วนแรกไปใช้ ซึ่งประกอบด้วย
1. กระบวนการทำงานของห้องควบคุมในอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟ 2. การจำแนกประเภทข้อมูล 3. จำแนกแบบต้นไม้ตัดสินใจ 4. ตัวจำแนกแบบเบย์อย่างง่าย และงานวิจัยที่เกี่ยวข้อง

2.1 กระบวนการทำงานของห้อง ควบคุมในอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟ

ห้องควบคุม [1] (Clean room) หมายถึงห้องที่ได้รับการปรับสภาวะอากาศ อุณหภูมิ ความชื้น ความดันและติดตั้งระบบกรองอากาศเป็นพิเศษ ให้อากาศปราศจากฝุ่นในระดับจุลภาคซึ่งมีขนาดเล็กมาก ไม่สามารถมองเห็นด้วยตาเปล่าขนาดเส้นผ่านศูนย์กลางไม่เกิน 0.5 ไมครอน ห้องควบคุมที่ใช้ในอุตสาหกรรมนั้นใช้ในอุตสาหกรรมอิเล็กทรอนิกส์ จะเน้นการกำจัดอนุภาคในอากาศและการควบคุมสภาวะต่างๆ ภายในห้องให้เหมาะสมกับลักษณะงานกระบวนการผลิตและผลิตภัณฑ์แต่ละชนิด สำหรับมาตรฐานที่ใช้ในห้องควบคุมของโรงงานผลิตฮาร์ดดิสก์ไดร์ฟได้แก่ มาตรฐานการปนเปื้อนของพื้นที่ควบคุม ซึ่งนิยมใช้กันในอุตสาหกรรมต่างๆ รวมทั้งอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟนั้นได้แก่ Clean Rooms - Federal Standard 209E

มาตรฐานการปน เปื้อนของพื้นที่ควบคุมในอุตสาหกรรม การผลิตฮาร์ดดิสก์ไดร์ฟและพื้นควบ คุมที่ที่กำหนด โดยที่ระดับ Class 10 เป็นมาตรฐานการควบคุมในพื้นที่ HDE Assembly Station ระดับ Class 100 เป็นมาตรฐานการควบคุมในพื้นที่ HDE Process และระดับ Class 10000 เป็นมาตรฐานการควบคุมในพื้นที่ HSA Process และ Cleaning Area ซึ่งแสดงได้ดังตารางที่ 1

2.2 การจำแนกประเภทข้อมูล (Data Classification)

การจำแนกข้อมูลเป็นเทคนิคหนึ่งในการเรียนรู้ของเครื่องจักร[2] (Machine Learning Technique) [3] ซึ่งเป็นกระบวนการสร้างตัวจำแนกจัดการข้อมูลให้อยู่ในรูปแบบที่กำหนด เพื่อแสดงให้เห็นถึงความแตกต่างระหว่างคลาสหรือกลุ่มของข้อมูลได้ และเพื่อทำนายว่าข้อมูลนี้ควรจัดอยู่ในกลุ่มของข้อมูลใด ตัวจำแนกที่ใช้จำแนกข้อมูลออกเป็นกลุ่มตามที่กำหนดไว้จะขึ้นอยู่กับการวิเคราะห์กลุ่มของข้อมูลการเรียนรู้ (Training data) โดยนำข้อมูลการเรียนรู้มาสอนให้ระบบเรียนรู้ว่ามีข้อมูลใดอยู่ในกลุ่มเดียวกันบ้างและผลที่ได้จากการเรียนรู้คือตัวจำแนกจัดประเภทข้อมูล (Classifier model) ซึ่งปรากฏได้หลายรูปแบบ เช่น Classification rule, ต้นไม้ตัดสินใจ (Decision tree) เป็นต้น และจะนำข้อมูลส่วนที่เหลือจากข้อมูลสอนระบบเป็นข้อมูลการทดสอบ (Testing data) โดยปรับปรุงตัวจำแนกจนกว่าจะได้ค่าความถูกต้องที่เหมาะสม

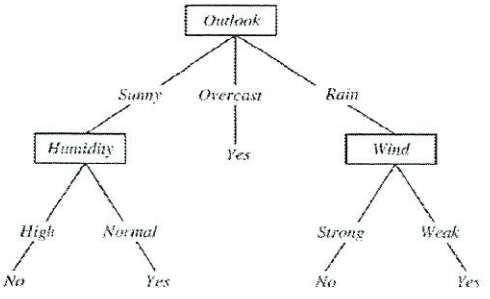
ตาราง 1 มาตรฐานการปนเปื้อนของพื้นที่ควบคุมในอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟ

Clean Room Class	Process Area	Cleanliness Specification
10000	HSA and Cleaning	10,000 particles/ft^3 at 0.5μm 30,000 particles/ft^3 at 0.3μm
100	HDE	100 particles/ft^3 at 0.5μm 300 particles/ft^3 at 0.3μm
10	HDE Assembly Station	10 particles/ft^3 at 0.5μm 30 particles/ft^3 at 0.3μm

2.3 ตัวจำแนกแบบต้นไม้ตัดสินใจ (Decision Tree)

ต้นไม้ตัดสินใจ [4] คือโครงสร้างต้นไม้ที่ประกอบด้วยโหนดตัดสินใจ (Decision node) เพื่อแสดงคุณลักษณะ (attribute) แล้วเชื่อมต่อกันด้วยสาขา (Branches) โดยขยายจากโหนดราก (Root node) มายังโหนดใบ (Leaf node) ค่าคุณลักษณะที่ใช้ในการเปรียบเทียบเงื่อนไขจะอยู่ในโหนดตัดสินใจซึ่งผลลัพธ์การเปรียบเทียบจะอยู่ในสาขาและในแต่ละสาขาจะนำไปสู่โหนดตัดสินใจอื่นหรือโหนดใบ ซึ่งโหนดใบก็จะเก็บผลลัพธ์ของการจำแนกไว้ดังรูปที่ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 1 แนวคิดของต้นไม้ตัดสินใจ [2]

2.4 ตัวจำแนกแบบเบย์อย่างง่าย (Naïve Bayesian)

ตัวจำแนกแบบเบย์อย่างง่ายเป็นขั้นตอนวิธีหนึ่งที่มีประสิทธิภาพในการจำแนกข้อมูล โดยการเรียนรู้ปัญหาที่เกิดขึ้นเพื่อนำมาสร้างเงื่อนไขในการจำแนกข้อมูลด้วยการใช้ทฤษฎีความน่าจะเป็น [6] เหมาะกับกรณีของกลุ่มตัวอย่างที่มีจำนวนมากและคุณลักษณะ (Attribute) ของตัวอย่างไม่ขึ้นต่อกันมีการจำแนกประเภทเบย์อย่างง่ายไปประยุกต์ใช้งานในด้านการจำแนกประเภทข้อความ การวินิจฉัยและพบว่าการใช้งานได้ดีไม่ต่างจากการจำแนกวิธีอื่น อีกทั้งเป็นวิธีการจำแนกข้อมูลที่มีประสิทธิภาพและมีอัลกอริทึมการทำงานไม่ซับซ้อน

สมมติฐานของการจำแนกแบบเบย์อย่างง่ายคือกำหนดให้คุณสมบัติแต่ละตัวไม่ขึ้นกันกับคุณสมบัติอื่นๆ ซึ่งจะวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระแต่ละตัวกับตัวแปรตาม เพื่อใช้สร้างเงื่อนไขความน่าจะเป็นสำหรับแต่ละความสัมพันธ์ โดยคำนวณได้จาก [7]

$$P(H|E) = [P(E|H) \times P(H)]/P(E)$$

เมื่อ P(H) คือความน่าจะเป็นที่จะเกิดเหตุการณ์ H และ

P(H|E) คือความน่าจะเป็นที่จะเกิดเหตุการณ์ H เมื่อเกิดเหตุการณ์ E

2.5 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่ผ่านมามักนำวิธีการจำแนกข้อมูลไปใช้กับข้อมูลต่างๆ ซึ่งมีความหลากหลายเพื่อเพิ่มประสิทธิภาพในการทำงานให้มากยิ่งขึ้น

งานวิจัยของ Ruey-Shun Chen [4] และคณะใช้เทคนิคการจำแนกข้อมูลแบบต้นไม้การตัดสินใจเพื่อเพิ่มคุณภาพการวิเคราะห์ข้อมูลสำหรับการค้นหาข้อบกพร่องในกระบวนการบรรจุวงจรรวม (Integrated Circuit) ของจอแสดงผลแบบแอลซีดี (LCD: liquid crystal display) เพื่อนำข้อมูลการวิเคราะห์ไปใช้ในการลดจำนวนข้อบกพร่องของชิ้นส่วน 2 ชนิด ได้แก่ inner leads และ chips โดยศึกษาข้อมูลจากกรณีตัวอย่างของโรงงานแห่งหนึ่งจำนวนข้อมูลที่ใช้ในวิเคราะห์คือ 9,752 ข้อมูล เป็นข้อมูลที่ใช้ในการเรียนรู้และทดสอบ และใช้เทคนิค C4.5 และ โครงข่ายประสาท (Neural network) พบว่า ชิ้นส่วน inner leads นั้น เทคนิค C4.5 ให้ประสิทธิภาพในการลดข้อบกพร่องสูงกว่าโครงข่ายประสาท คือ 10.4% และ 2.95% ตามลำดับ ชิ้นส่วน chips นั้น เทคนิค C4.5 ให้ประสิทธิภาพในการลดข้อบกพร่องสูงกว่าโครงข่ายประสาท คือ 11.6% และ 4.6% ตามลำดับ

งานวิจัยของ Emad S. Abouel Nasr และคณะ [7] ได้ประเมินผลเทคนิคการจำแนกข้อมูลที่เหมาะสมและรูปแบบของตัวแบบที่เหมาะสมสำหรับข้อมูลควบคุมกระบวนการ (Process control data) โดยพิจารณาศึกษาเทคนิคการจำแนกข้อมูล 6 เทคนิค ได้แก่ Naïve Bayes, Logistic, Support Vector Machine (SVM), Naïve Bayes Decision Tree Table hybrid algorithm (DTNB), Nearest-Neighbor with Generalization (NNge) และ Decision Trees PART สำหรับรูปแบบที่เหมาะสมนั้นได้เปรียบเทียบ 2 รูปแบบได้แก่ 5-fold cross validation และ 10-fold cross validation โดยทำการศึกษาด้วยชุดข้อมูลควบคุมกระบวนการของ Lavangnananda and Piyatumrong 2005, UCI Machine Learning Repository website ประกอบด้วย 600 ข้อมูล ซึ่งมี 60 คุณลักษณะ พบว่า รูปแบบ 5-fold cross validation นั้น SVM ให้ประสิทธิภาพที่ดีที่สุด คือ 99.0% อันดับที่ 2 คือ NNge 96% อันดับที่ 3 คือ Naïve Bayes 95.3% โดยเทคนิคที่ให้ประสิทธิภาพต่ำสุด คือ Logistic 82% รูปแบบ 10-fold cross validation นั้น SVM ให้ประสิทธิภาพที่ดีที่สุด คือ 99.2% อันดับที่ 2 คือ NNge 97% อันดับที่ 3 คือ Naïve Bayes 95.4% โดยเทคนิคที่ให้ประสิทธิภาพต่ำสุด คือ Logistic 84.4% แสดงว่า ทั้ง 5-fold cross validation และ 10-fold cross validation ของการจำแนกข้อมูลดังกล่าวนั้นให้ผลในแนวทางเดียวกัน

งานวิจัยของ R.S. Chen และคณะ [8] สร้างระบบควบคุมคุณภาพ (Quality control system) โดยใช้เทคนิคเหมืองข้อมูลเพื่อค้นหาเหตุผลความไม่สอดคล้องหลักใน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กระบวนการผลิตของโรงงานผลิตอุปกรณ์กึ่งตัวนำและเปรียบเทียบความถูกต้องของวิธีการจำแนกข้อมูลของวิธีต้นไม้ตัดสินใจกับทฤษฎีเบย์เซี่ยน เมื่อนำข้อมูลที่จำแนกไปปรับปรุงระบบแล้วพบว่า วิธีต้นไม้ตัดสินใจให้ประสิทธิภาพที่ดีกว่าทฤษฎีเบย์เซี่ยน โดยปรับปรุงผลด้านการควบคุมคุณภาพได้ดีขึ้น 12.5% และ 7.8% ตามลำดับ

3. วิธีการดำเนินการวิจัย

การวิจัยฉบับนี้เริ่มต้นตั้งแต่ขั้นตอนการเก็บรวบรวมข้อมูลงานวิจัยจนถึงกระบวนการเปรียบเทียบตัวจำแนกในปัจจัยต่างๆ เพื่อเปรียบเทียบประสิทธิภาพในการวิเคราะห์ข้อมูล โดยมีขั้นตอนดังต่อไปนี้

3.1 การจัดเก็บและการเตรียมข้อมูล

งานวิจัยนี้มุ่งเน้นในการเปรียบเทียบประสิทธิภาพของการใช้ต้นไม้ตัดสินใจ (Decision tree) และ เบย์เซี่ยน (Bayesian) เพื่อวิเคราะห์ปัจจัยที่มีผลต่อการเกิดปริมาณอนุภาคฝุ่นละอองที่เกินปริมาณการควบคุม โดยนำข้อมูลจากอุปกรณ์ตรวจจับภายในห้องควบคุมมาใช้ในการวิเคราะห์และพิจารณาควบคุมปริมาณอนุภาคฝุ่นละอองด้วยมาตรฐานการบ่นเปื้อนของพื้นที่ควบคุม ข้อมูลจัดเก็บในรูปแบบไฟล์ CSV ตั้งแต่วันที่ 1 มีนาคม – 31 มีนาคม 2553 จำนวนทั้งสิ้น 374,884 รายการ

3.2 การทำความสะอาดข้อมูล

เนื่องจากข้อมูลที่ได้มามีส่วนที่ไม่สามารถนำมาใช้ในการจำแนก (Classification) ได้ เช่นข้อมูลบางส่วนที่ไม่สามารถนำมาใช้ได้โดยตรง เช่น ข้อมูลวันที่ ข้อมูลเวลา เป็นข้อมูลเรียงกันเป็นจำนวนมาก หรือข้อมูลที่ไม่เกี่ยวข้องกับการจำแนกจึงไม่นำข้อมูลในส่วนนั้นมาทำการจำแนก

3.3 การจัดกลุ่มข้อมูล

ขั้นตอนนี้เป็นขั้นตอนการจัดกลุ่มของข้อมูลพื้นฐานเพื่อให้พร้อมสำหรับการจำแนก โดยได้ทำการแจกแจงดังนี้

3.3.1. ข้อมูลที่เป็นวันที่ จัดกลุ่มให้อยู่ในรูปแบบของวันในสัปดาห์ ตั้งแต่วันจันทร์ถึงอาทิตย์

3.3.2. ข้อมูลที่เป็นเวลา จัดกลุ่มข้อมูลเป็นกะการทำงาน โดยแบ่งเป็น 3 กะงาน ดังนี้

1. กะที่หนึ่ง (The first shift) เริ่มตั้งแต่เวลา 08.00 – 16.00 น.
2. กะที่สอง (The second shift) เริ่มตั้งแต่เวลา 16.00 – 24.00 น.
3. กะที่สาม (The third shift) เริ่มตั้งแต่เวลา 24.00 – 06.00 น.

3.3.3. ข้อมูลชื่ออุปกรณ์ตรวจจับอนุภาค สามารถจัดกลุ่มได้ 2 กลุ่ม ดังนี้

1. ข้อมูลพื้นที่การทำงาน (Area) โดยระบุตามโครงสร้างการทำงานว่าอุปกรณ์ตรวจจับอนุภาคอยู่ในพื้นที่การทำงานกลุ่มใด ได้แก่ HSA, HDE และ Cleaning
2. ข้อมูลกระบวนการทำงาน (Process) โดยการจัดกลุ่มว่าอุปกรณ์แต่ละตัวมีหน้าที่การทำงานในกระบวนการใด ซึ่งสามารถจัดกลุ่มได้ 16 กลุ่ม ดังนี้
   1) Wipe คือการเก็บชิ้นส่วนที่ทำความสะอาดแล้ว
   2) Cleaning คือการล้างทำความสะอาด
   3) Environment คือการตรวจสภาพแวดล้อม
   4) Final inspection คือการตรวจสอบขั้นสุดท้าย
   5) HSA assembly คือการบรรจุ HSA
   6) Solder inspection คือการตรวจสอบรอยบัดกรี
   7) Tear down คือการทำงาน re-work
   8) Particle test คือการตรวจสอบหาฝุ่นละอองของฮาร์ดดิสก์
   9) Disk mounting คือการบรรจุดิสก์
   10) HSA/VCM mounting คือการบรรจุ HSA หรือ VCM
   11) Ramp mounting คือการบรรจุหัวเขียนอ่านดิสก์
   12) Top clamp mounting คือการบรรจุชิ้นส่วนสำหรับยึดแผ่นดิสก์กับมอเตอร์
   13) Top cover mounting คือการปิดฮาร์ดดิสก์เพื่อป้องกันวัสดุจากภายนอก
   14) Hyper disk balancer คือการทดสอบความสมดุลของชุดแผ่นดิสก์
   15) Pivot คือการขันสกรูหรือแกนต่างๆ
   16) Part loading conveyer คือการเคลื่อนย้ายชิ้นส่วนต่างๆ จากห้องล้างเข้าสู่พื้นที่ HDE

3.3.4. ข้อมูลปริมาณฝุ่นละออง สามารถจัดกลุ่มได้เป็น 2 กลุ่ม คือ กลุ่มปริมาณอนุภาคฝุ่นละอองที่ไม่เกินกำหนดและกลุ่มปริมาณอนุภาคที่เกินกำหนด โดยพิจารณาจากเกณฑ์การควบคุม (Control limit) ในแต่ละพื้นที่การทำงาน ดังตารางที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 2 เกณฑ์การควบคุมปริมาณอนุภาคฝุ่นละออง

พื้นที่การทำงาน	ปริมาณควบคุม
HSA	20,000
HDE	100
Cleaning	500

ในการจัดกลุ่มของข้อมูลปริมาณอนุภาคฝุ่นละอองเพื่อให้ตัวจำแนกทราบว่าข้อมูลดังกล่าวมีส่วนที่เกี่ยวข้องกันและสามารถจัดกลุ่มเพื่อใช้ในการหาปัจจัยที่เกี่ยวข้องของปริมาณอนุภาคฝุ่นละอองเกินกว่าเกณฑ์ที่กำหนดได้ดังตารางที่ 3

ตาราง 3 จำนวนกลุ่มข้อมูลในการวิจัย

กลุ่มข้อมูล	จำนวนข้อมูล
กลุ่มข้อมูลไม่เกินควบคุม	368,842
กลุ่มข้อมูลเกินควบคุม	6,042
รวม	374,884

3.4 การแปลงข้อมูล

ขั้นตอนนี้เป็นขั้นตอนที่นำข้อมูลที่อยู่ในแบบข้อความออกมาเพื่อนำไปจำแนก (Classification) ด้วยเครื่องมือในการทำวิจัย โดยเครื่องมือ WEKA (Waikato Environment for Knowledge Analysis)

3.5 การ Cross-Validation

ขั้นตอนการ Cross-Validation เพื่อใช้ในการแบ่งกลุ่มของข้อมูล โดยเลือกใช้การแบ่งกลุ่มข้อมูลเพื่อทดสอบ 2 รูปแบบคือ

3.5.1. Hold-Out Cross Validation

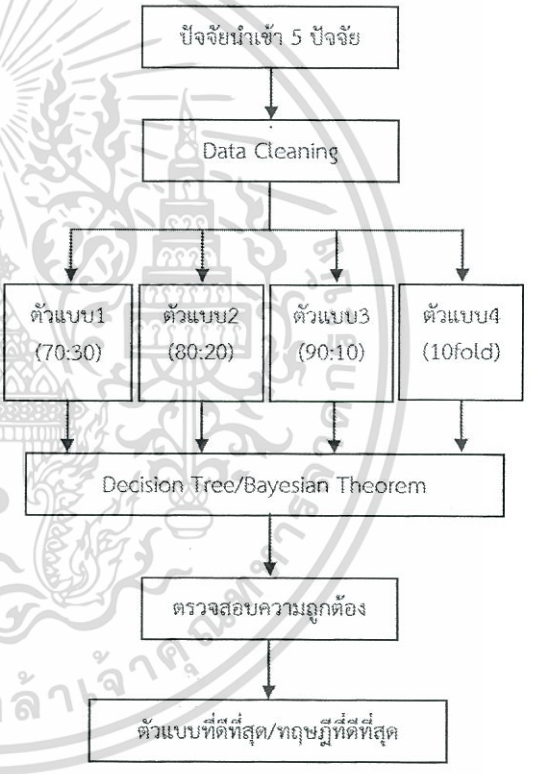
เป็นวิธีแบ่งกลุ่มข้อมูลที่จะนำมาใช้ในการเรียนรู้ (Training) และทดสอบ (Testing) แบบสุ่ม โดยทำการแบ่งเป็นอัตราส่วนร้อยละของข้อมูล เช่นแบ่งข้อมูลที่ร้อยละ 50 หมายถึงนำข้อมูลร้อยละ 50 เป็นกลุ่มการเรียนรู้ (Training set) และข้อมูลอีกร้อยละ 50 เป็นข้อมูลกลุ่มการทดสอบ (Testing set) ในการทดลองนี้แบ่งจำนวนกลุ่มของข้อมูลการเรียนเป็น 3 วิธี ได้แก่ ร้อยละ 70 80 และ 90 ซึ่งส่วนที่เหลือจะเป็นกลุ่มข้อมูลการทดสอบ คือ ร้อยละ 30 20 และ 10 ตามลำดับ

3.5.2. k-Fold Cross Validation

วิธีนี้จะทำโดยการแบ่งข้อมูลออกเป็น k_i, k_i=1,2,3,...,n กลุ่ม แล้วให้ใช้ 1 กลุ่มมาเป็นกลุ่มทดสอบ ส่วนที่เหลือคือ k-1 นำมาใช้เป็นกลุ่มข้อมูลการเรียนรู้แล้วจะทำการวนทำเป็นจำนวน k รอบ โดยจะเปลี่ยนกลุ่มทดสอบไปเรื่อยๆ ตามลำดับจนครบ ทำให้ข้อมูลทั้งหมดได้ผ่านการจำแนก ในการทดลองนี้ใช้ 10-Fold Cross Validation

3.6 การสร้างตัวแบบสำหรับการวิเคราะห์ข้อมูล

การสร้างตัวแบบสำหรับการจำแนกในการทดลองนี้สร้างตัวแบบโดยต้นไม้การตัดสินใจ (Decision tree) และทฤษฎีเบย์เซี่ยน (Bayesian Theorem) แสดงได้ดังรูปที่ 2



รูปที่ 2 ขั้นตอนการสร้างตัวแบบ

ซึ่งปัจจัยหรือข้อมูลนำเข้า (Input) ที่นำไปใช้ในการจำแนกข้อมูล ทั้ง 5 ปัจจัยแสดงดังตารางที่ 4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 4 ปัจจัยนำเข้าของข้อมูลปริมาณอนุภาคฝุ่นละออง

ชื่อข้อมูล	คำอธิบาย	ชนิดของข้อมูล
Day	วัน	Categorical
Shift	กะงาน	Categorical
Area	พื้นที่การทำงาน	Categorical
Process	กระบวนการทำงาน	Categorical
OutofControl	การเกินปริมาณควบคุม	Class label

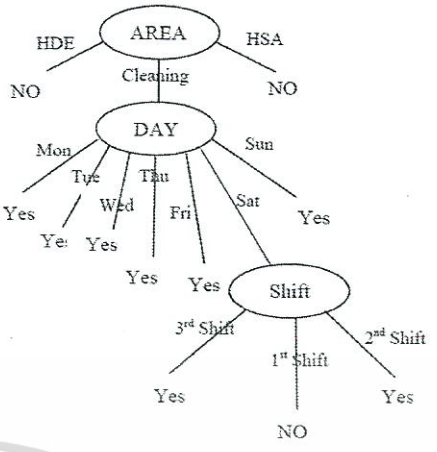
4. ผลการทดลอง

การวิเคราะห์ปัจจัยเสี่ยงในการเกิดปริมาณอนุภาคฝุ่นละอองที่เกินการควบคุมเป็นการนำเอาข้อมูลในส่วนการจับอนุภาคฝุ่นละอองของการผลิตเพื่อใช้สำหรับหารูปแบบของปัจจัยเสี่ยง การวิเคราะห์ปัจจัยจะทำโดยข้อมูลชุดเดียวกันโดยตัวแบบที่ใช้ในการศึกษาคือต้นไม้การตัดสินใจ C4.5 และตัวแบบ Naïve Bayes โดยวัดประสิทธิภาพด้วยการแบ่งชุดข้อมูลการเรียนรู้และชุดข้อมูลการทดสอบที่แตกต่างกันเพื่อหารูปแบบที่เหมาะสมที่สุดสำหรับการจำแนกข้อมูล พบว่าค่าความถูกต้องใกล้เคียงกัน ตัวแบบในการจำแนกที่เหมาะสมคือ ต้นไม้การตัดสินใจ C4.5 ซึ่งให้ผลการจำแนกที่ถูกต้องโดยเฉลี่ยมากกว่าตัวแบบเบย์เชียนในทุกๆ รูปแบบการ cross-validation แต่ก็ใช้เวลาในการจำแนกมากกว่าตัวแบบเบย์เชี่ยน ดังตารางที่ 5

ตาราง 5 ค่าความถูกต้องและเวลาที่ใช้ของแต่ละวิธีการจำแนก

Cross-Validation	Decision Tree (C4.5)		Naïve Bayesian	
	Accuracy	Time (sec)	Accuracy	Time (sec)
70:30	99.20%	0.91	99.19%	0.93
80:20	99.23%	0.73	99.23%	0.23
90:10	99.29%	0.77	99.29%	0.42
10-Fold	99.20%	0.87	99.18%	0.22

จากปัจจัยนำเข้าทั้งหมดสามารถจำแนกข้อมูลได้ตัวแบบด้วยวิธีต้นไม้การตัดสินใจ แสดงดังรูปที่ 3



รูปที่ 3 ผลการทดลองในรูปแบบต้นไม้ตัดสินใจ

จากรูปที่ 3 พบว่าปัจจัยที่ทำให้ปริมาณอนุภาคฝุ่นละอองเกินควบคุมได้แก่ พื้นที่การทำงาน, วัน และกะการทำงาน ตามลำดับ ซึ่งในพื้นที่การทำความสะอาด (Cleaning) มีความเสี่ยงในการเกิดปริมาณอนุภาคฝุ่นละอองเกินการควบคุมในทุกๆ วันและทุกกะการทำงานยกเว้นวันเสาร์เท่านั้นที่พบว่ามีเฉพาะกะการทำงานที่ 2 และ 3 ที่เป็นปัจจัยเสี่ยง สามารถสรุปเป็นกฎ ได้ดังต่อไปนี้

1. IF (AREA=HDE) Then Out of control = No
2. IF (AREA=HSA) Then Out of control = No
3. IF (AREA=Cleaning) and (DAY=Mon) Then Out of control = Yes
4. IF (AREA=Cleaning) and (DAY=Tue) Then Out of control = Yes
5. IF (AREA=Cleaning) and (DAY=Wed) Then Out of control = Yes
6. IF (AREA=Cleaning) and (DAY=Thu) Then Out of control = Yes
7. IF (AREA=Cleaning) and (DAY=Fri) Then Out of control = Yes
8. IF (AREA=Cleaning) and (DAY=Sat) and (Shift=1st Shift) Then Out of control = No
9. IF (AREA=Cleaning) and (DAY=Sat) and (Shift=2nd Shift) Then Out of control = Yes
10. IF (AREA=Cleaning) and (DAY=Sat) and (Shift=3rd Shift) Then Out of control = Yes

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

11. IF (AREA=Cleaning) and (DAY=Sun) Then Out of control = Yes

5. บทสรุปและข้อเสนอแนะ

งานวิจัยนี้นำเสนอการวิเคราะห์ปัจจัยที่ทำให้เกิดอนุภาคฝุ่นละอองในปริมาณที่เกินกำหนด มีวัตถุประสงค์เพื่อสร้างตัวแบบโดยวิธีต้นไม้ตัดสินใจกับทฤษฎีเบย์เซี่ยน การเปรียบเทียบประสิทธิภาพของทฤษฎีเบย์เซี่ยนและต้นไม้ตัดสินใจให้ค่าความถูกต้องใกล้เคียงกัน และมีค่าความถูกต้องสูงสุดเท่ากันคือ 99.29% ด้วยตัวแบบการ cross-validation 90:10 ทั้งวิธีต้นไม้การตัดสินใจและเบย์เซี่ยน เนื่องจากชนิดข้อมูลที่ต้องการจำแนกประเภท (Class label) มี 2 กลุ่มได้แก่ ปริมาณที่เกินควบคุมและปริมาณที่ไม่เกินควบคุม เมื่อคำนึงถึงข้อได้เปรียบและข้อเสียเปรียบของเทคนิคทั้งสองพบว่า วิธีต้นไม้การตัดสินใจนั้นใช้เวลาในการจำแนกมากกว่าแต่ก็มีสามารถนำตัวแบบไปใช้งานได้ง่ายกว่าวิธีการแบบเบย์เซี่ยน ในขณะที่วิธีแบบเบย์เซี่ยนไม่มีโมเดลที่สามารถนำไปใช้ได้ง่ายแต่ต้องคำนวณจากค่าความน่าจะเป็นเท่านั้น

โดยในการวิจัยในอนาคตจะนำการควบคุมกระบวนการเชิงสถิติมาประยุกต์ใช้ร่วมกับการจำแนกประเภทข้อมูลเพื่อให้ได้ผลการจำแนกที่เหมาะสมกับการควบคุมกระบวนการผลิตสำหรับโรงงานอุตสาหกรรมมากขึ้นและสามารถนำไปใช้ในการแก้ไขปรับปรุงกระบวนการผลิตเพื่อให้ได้ผลิตภัณฑ์ที่ดีขึ้นในอนาคต

6. กิตติกรรมประกาศ

งานชิ้นนี้เป็นส่วนหนึ่งของโครงการพัฒนาทรัพยากรมนุษย์ในอุตสาหกรรมการผลิตฮาร์ดดิสก์ไดร์ฟซึ่งได้รับทุนสนับสนุนจากศูนย์วิจัยร่วมเฉพาะทาง (Industry/University Cooperative Research Center, I/UCRC), สถาบันฮาร์ดดิสก์ไดร์ฟ (Hard disk Drive Institute, HDDI) ภายใต้การควบคุมของสำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ (สวทช.) กระทรวงวิทยาศาสตร์และเทคโนโลยี และยังได้รับการสนับสนุนจากวิทยาลัยร่วมด้านเทคโนโลยีการบันทึกข้อมูลและการประยุกต์ใช้งาน (Data Storage Technology and Applications) สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง โดยความร่วมมือของบริษัท ฮิตาชิ โกลบอลสตอเรจ เทคโนโลยี ประเทศไทย (HGST)

7. เอกสารอ้างอิง

[1] Mrad F., "The Characterization of a Clean Room Assembly Process", INDUSTRY APPLICATIONS, vol. 35, no. 2, MARCH/APRIL. 1999.

[2] T. Mitchell," Machine Learning", McGraw-Hill Companies. Inc., New York, 1997.

[3] Berry, M. and G. Linoff, "Data Mining Techniques for Marketing Sales and Customer Support", John Wiley and Sons, New York, 1997.

[4] R.S. Chen, K.C. Yeh, C.C. Chang and H.H. Chien, "Using Data Mining Technology to improve Manufacturing Quality- A Case Study of LCD Driver IC Packaging Industry", In: Proceedings of the Seventh International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing., 2006.

[5] Maimon, O. and L. Rokach., "The Data Mining and Knowledge Discovery Handbook", Springer, New York, 2005.

[6]Bretthorse, G. Larry., "Bayesian Spectrum Analysis and Parameter Estimation in Lecture", Springer-Verlag vol2, New York, 1998.

[7] Abouel Nasr, E.S., Al-Mubaid, H., "Mining Process Control Data Using Machine Learning", In: Proceedings of International Conference on Computers & Industrial Engineering, Egypt, 2009.

[8] R.S. Chen, Y.C. Chen, C.C. Chang and C.C. Chen, "Using Data Mining Technology to Design an Quality Control System for Manufacturing Industry", In: Proceedings of the European conference of systems, and European conference of circuits technology and devices, and European conference of communications, and European conference on Computer science., 2010.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

นางสาว กฤษฎ์ฐมรัษฎร์ เพ็งอุบล เกิดเมื่อวันที่ 24 ตุลาคม พ.ศ.2528 จบการศึกษา วิทยาศาสตรบัณฑิต สาขาวิทยาการคอมพิวเตอร์ คณะศิลปศาสตร์และวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ ในปีการศึกษา 2550

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้