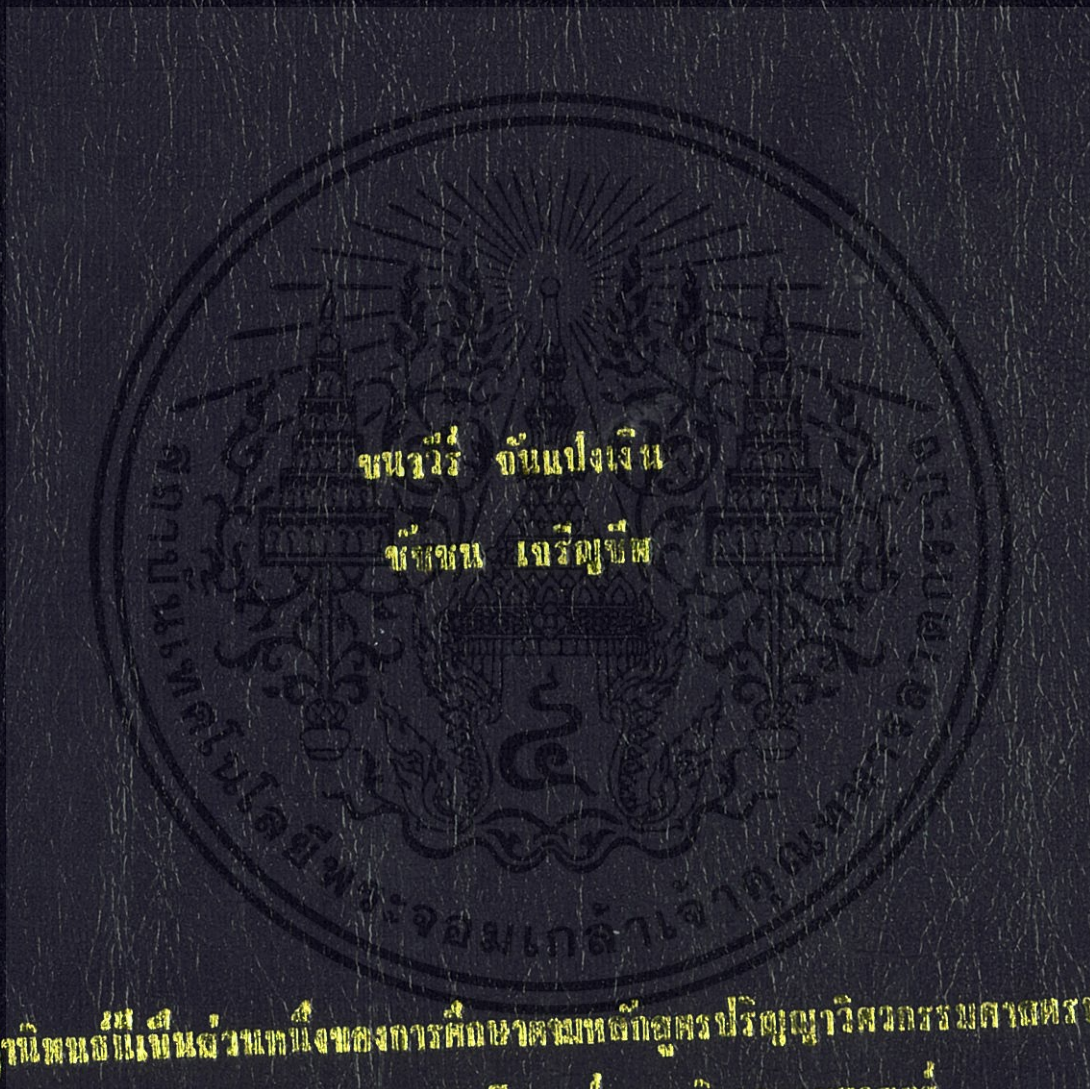


ระบบตรวจสอบการปลอมแปลงเอกสารรับรองความถูกต้อง
TRANSCRIPT INTEGRITY VERIFICATION SYSTEM



ปริญญาโท สาขาวิชาวิศวกรรมคอมพิวเตอร์
มหาวิทยาลัยราชภัฏบุรีรัมย์
ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2557

ระบบตรวจสอบการปลอมแปลงเอกสารรับรองผลการศึกษา
TRANSCRIPT INTEGRITY VERIFICATION SYSTEM



ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปีการศึกษา 2557

ปริญญาานิพนธ์ปีการศึกษา 2557

ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง ระบบตรวจสอบการปลอมแปลงเอกสารรับรองผลการศึกษา

TRANSCRIPT INTEGRITY VERIFICATION SYSTEM

ผู้จัดทำ

1. นายชนาวีร์ จันแปงเงิน รหัสนักศึกษา 54010275

2. นายชัชชน เจริญชีพ รหัสนักศึกษา 54010298



อาจารย์ที่ปรึกษา

(อาจารย์ธนัญชัย ตริภาค)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระบบตรวจสอบการปลอมแปลงเอกสารรับรองผลการศึกษา

นายชนาวีร์	เงินแปงเงิน	54010275
นายชัชชน	เจริญชีพ	54010298
อาจารย์ธัญชัย	ตรีภาค	อาจารย์ที่ปรึกษา
ปีการศึกษา 2557		

บทคัดย่อ

เพื่อป้องกันปัญหาการปลอมแปลงเอกสารทางการศึกษา จึงมีขั้นตอนในการตรวจสอบความถูกต้องของเอกสาร แต่ขั้นตอนดังกล่าวจำเป็นต้องใช้คนในการตรวจสอบ และใช้เวลานาน

โครงการนี้เป็นโครงการที่ออกแบบและพัฒนาระบบเพื่อตรวจสอบภาพเอกสารรับรองผลการศึกษาโดยอัตโนมัติ โดยเพิ่มการพิมพ์ค่าแฮชของข้อมูลลงบนเอกสารเอกสาร ในขั้นตอนการตรวจสอบจะทำการแปลงภาพเอกสารเป็นตัวอักษรโดยใช้ OCR Library และตรวจสอบค่าแฮชดังกล่าว ว่ามีค่าตรงกับค่าที่พิมพ์ไว้บนเอกสาร

ระบบตรวจสอบการปลอมแปลงเอกสารรับรองผลการศึกษามีความสามารถตรวจสอบการปลอมแปลงเอกสารได้แม่นยำ โดยใช้เวลาในการตรวจสอบเฉลี่ย 25 วินาที/ฉบับ โดยสามารถตรวจสอบได้เฉพาะไฟล์ภาพเอกสารรับรองผลการศึกษาของสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังรูปแบบภาษาอังกฤษที่มีความละเอียดของไฟล์ภาพสแกน 360 dpi ขึ้นไป ระบบนี้เหมาะสำหรับผู้ที่ทำหน้าที่ในการตรวจสอบเอกสารรับรองผลการศึกษา ระบบนี้สามารถนำไปต่อยอดเพื่อให้สามารถตรวจสอบเอกสารรับรองอื่นได้ต่อไปได้ในอนาคต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Transcript Integrity Verification System

Mr. Chanawee Chanpaeng 54010275

Mr. Shutchachon Jaroencheep 54010298

Dr. Thanunchai Threepak Advisor

Academic Year 2014

ABSTRACT

To protect forging transcripts, Academy has a procedure for verifying transcript integrity but that procedure has human error and take a long time.

This project is designed and developed for transcript integrity verification. To accomplish this, system use Optical Character Recognition to convert data on transcript image to text data and pass it to hashing function for generate codes and then compare similarity with pregenerated codes which are on transcript to verify integrity.

Transcript integrity verification system can verify integrity with high accuracy and use time around 25 seconds to process. But this system can verify only king mongkut's institute of technology Ladkrabang transcript english format which has image's pixel density 360 dpi. This system suit for who has responsible to verify transcript integrity and can improve for verify other certificate in the future.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ข้าพเจ้าขอขอบคุณอาจารย์ รัตนัญชัช ตริภาค ที่ให้คำแนะนำและคำปรึกษาเสมอมาตลอดการทำโครงการชิ้นนี้ โดยเฉพาะอย่างยิ่งในคำแนะนำในเรื่องกระบวนการคิดอย่างเป็นระบบที่สามารถนำไปประยุกต์ใช้ได้กับเรื่องต่างๆในชีวิตประจำวัน และข้าพเจ้าขอขอบคุณอาจารย์ในภาควิชาวิศวกรรมคอมพิวเตอร์ทุก ๆ ท่านที่อบรมสั่งสอนข้าพเจ้ามาตลอด

ขอขอบคุณเพื่อนทุกคนในภาควิชาคอมพิวเตอร์ที่คอยให้คำปรึกษาในด้านเทคนิค กระบวนการทำโครงการและช่วยเหลือในการทำกรทดลองต่างๆของโครงการ

สุดท้ายนี้ข้าพเจ้าขอขอบคุณบิดามารดาที่ให้โอกาสทางการศึกษา และคอยให้กำลังใจ ช่วยทำให้ข้าพเจ้ามีทุกวันนี้ ข้าพเจ้าขอขอบพระคุณมา ณ ที่นี้ด้วย

ชนาวีร์ จันแปงเงิน
ชัชชน เจริญชีพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อ	I
ABSTRACT	II
สารบัญ	IV
สารบัญตาราง	VII
สารบัญรูป.....	IX

บทที่ 1 บทนำ.....	1
1.1 ความสำคัญและที่มาของโครงการ	1
1.2 วัตถุประสงค์ของโครงการ	2
1.3 ขอบเขตของโครงการ.....	3
1.4 วิธีการดำเนินงาน	3
1.5 ประโยชน์ที่คาดว่าจะได้รับ	3
1.6 ส่วนประกอบของปริญญานิพนธ์.....	3

บทที่ 2 ทฤษฎีที่เกี่ยวข้อง.....	5
2.1 Optical Character Recognition (OCR)	5
2.1.1 แบบมองอักษรที่ตรวจจับในลักษณะเมตริกซ์.....	5
2.1.2 แบบมองอักษรที่ตรวจจับในลักษณะ Vector	6
2.2 การทำงานของ Tesseract OCR	6
2.3 รูปแบบเอกสารรับรองผลการศึกษาของสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหาร ลาดกระบัง	8
2.4 ทฤษฎี Hashing	10

บทที่ 3 การออกแบบและพัฒนา.....	13
3.1 การทำงานของระบบ	14
3.2 วิธีการตรวจสอบ	16
3.3 โครงสร้างของระบบ	20

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่ ใช้งานด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น ผู้อ่านที่มีข้อผิดพลาดหรือข้อสงสัย กรุณาแจ้งให้ทราบเพื่อปรับปรุงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

บทที่ 4 การทดลองและผลการทดลอง	23
4.1 การทดลองนำเอกสารรับรองผลการศึกษาเข้าโปรแกรม OCR เพื่อหาประเภทไฟล์ที่รองรับ.....	24
4.1.1 จุดประสงค์การทดลอง.....	24
4.1.2 เครื่องมือและอุปกรณ์ที่ใช้ในการทดลอง.....	24
4.1.3 วิธีการทดลอง.....	24
4.1.4 ผลการทดลอง.....	25
4.1.5 สรุปผลการทดลอง.....	25
4.2 การทดลองนำเอกสารรับรองผลการศึกษาเข้าโปรแกรม OCR เพื่อหาค่า DPI ที่เหมาะสม	25
4.2.1 จุดประสงค์การทดลอง.....	25
4.2.2 เครื่องมือและอุปกรณ์ที่ใช้ในการทดลอง.....	25
4.2.3 วิธีการทดลอง.....	25
4.2.4 ผลการทดลอง.....	26
4.2.5 สรุปผลการทดลอง.....	28
4.3 การทดสอบประสิทธิภาพ เมื่อใช้งานส่วนการปรับปรุงข้อมูล OCR โดยระบบ	29
4.3.1 จุดประสงค์การทดลอง.....	29
4.3.2 เครื่องมือและอุปกรณ์ที่ใช้ในการทดลอง.....	29
4.3.3 วิธีการทดลอง.....	29
4.3.4 ผลการทดลอง.....	29
4.3.5 สรุปผลการทดลอง.....	32
4.4 การทดลองความถูกต้องการทำงานของระบบ	32
4.4.1 จุดประสงค์การทดลอง.....	32
4.4.2 เครื่องมือและอุปกรณ์ที่ใช้ในการทดลอง.....	32
4.4.3 วิธีการทดลอง.....	33
4.4.4 ผลการทดลอง.....	33
4.4.5 สรุปผลการทดลอง.....	34

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่ ใช้งานด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
บทที่ 5 บทสรุปและข้อเสนอแนะ	35
5.1 บทสรุป.....	35
5.2 ปัญหาอุปสรรค และแนวทางการแก้ไข.....	36
5.2.1 ความละเอียดของไฟล์ภาพ	36
5.2.2 การทำการแก้ไขโดยอัตโนมัติ	36
5.3 แนวทางการพัฒนาต่อ	36
5.3.1 ส่วนการแก้ไขข้อมูล.....	36
5.3.2 ส่วนการเก็บผลลัพธ์การทำงาน	37
5.3.3 ส่วนปรับแต่งเพื่อตรวจสอบเอกสารรับรองอื่นๆ	39
5.3.4 ปรับแต่งส่วนเพิ่มเอกสารรับรองผลการศึกษา	40
บรรณานุกรม.....	41

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

	หน้า
ตาราง 2.1 แสดงความแตกต่างของฟังก์ชันแฮช.....	10
ตาราง 4.1 ผลการทำกระบวนการ OCR ของไฟล์แต่ละประเภท.....	25
ตาราง 4.2 ค่าความถูกต้องในการตรวจสอบและเวลาที่ใช้ในการประมวลผล.....	26
ตาราง 4.3 ค่าความถูกต้อง ต่อ เวลาทั้งหมดที่ใช้ในการตรวจสอบ.....	27
ตาราง 4.4 ค่าความถูกต้องในการตรวจสอบและเวลาที่ใช้ในการประมวลผล กรณีปิดการใช้งาน ส่วนการปรับปรุงข้อมูล OCR โดยระบบ.....	30
ตาราง 4.5 ค่าความถูกต้องในการตรวจสอบและเวลาที่ใช้ในการประมวลผล กรณีที่เปิดการใช้งาน ส่วนการปรับปรุงข้อมูล OCR โดยระบบ.....	30
ตาราง 4.6 ความถูกต้องในการทำงานแต่ละส่วนของระบบ.....	33
ตาราง 5.1 รูปแบบตารางสำหรับเก็บข้อมูลการแก้ไขจากผู้ใช้.....	37
ตาราง 5.2 รูปแบบตารางสำหรับเก็บข้อมูลการใช้งานของผู้ใช้.....	38

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

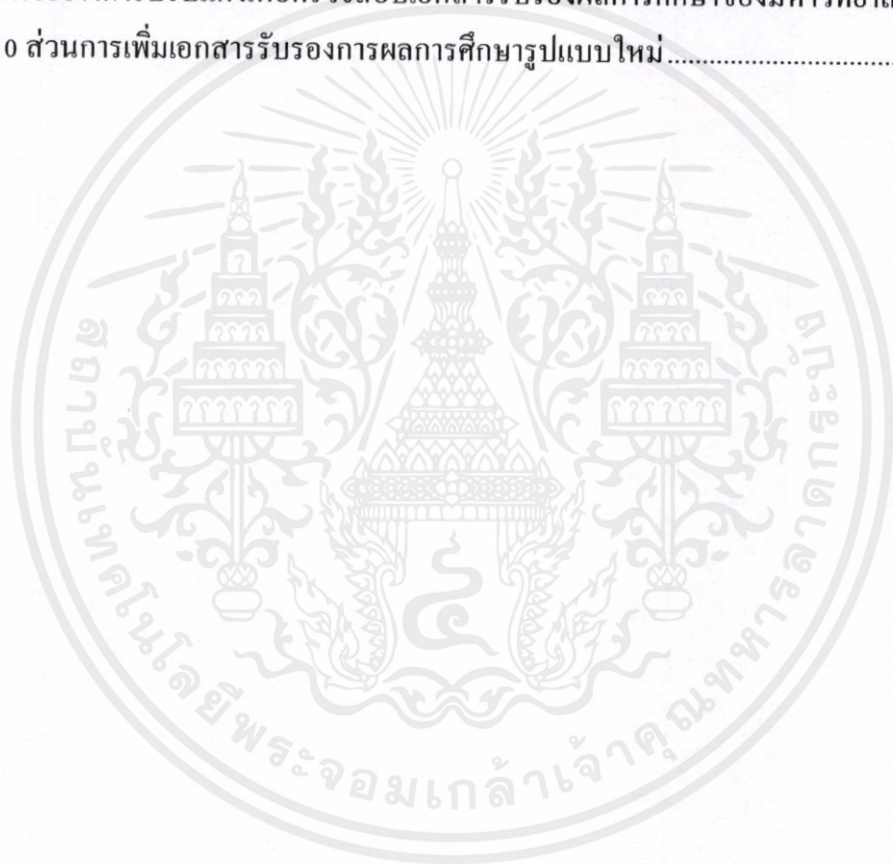
หน้า

รูป 1.1 ตัวอย่างผู้ที่รับทำวุฒิการศึกษาปลอม	1
รูป 1.2 ตัวอย่างบุคคลที่สนใจทำวุฒิการศึกษาปลอม	2
รูป 2.1 การตรวจจับลักษณะ Matrix.....	5
รูป 2.2 การตรวจจับแบบลักษณะ Vector.....	6
รูป 2.3 ตัวอย่างการทำ Thresholding.....	7
รูป 2.4 เอกสารรับรองผลการศึกษาศึกษา.....	8
รูป 2.5 ส่วนข้อมูลประวัติผู้เรียนบนเอกสารรับรองผลการศึกษาศึกษา	9
รูป 2.6 ส่วนข้อมูลประวัติการศึกษาศึกษาบนเอกสารรับรองผลการศึกษาศึกษา	9
รูป 3.1 ตัวอย่างค่าที่ได้จากการ Generate Code ข้อมูลบนเอกสารรับรองผลการศึกษาศึกษา.....	13
รูป 3.2 ข้อมูลประวัติผู้เรียนบนเอกสารรับรองผลการศึกษาศึกษาหลังเพิ่มรหัสความปลอดภัย	13
รูป 3.3 ภาพรวมการทำงานของโปรแกรม	14
รูป 3.4 Swimming lane แสดงขั้นตอนการทำงานของโปรแกรม	15
รูป 3.5 Flowchart ภาพรวมการทำงานของระบบ.....	16
รูป 3.6 Flowchart แสดงวิธีแยกค่า Security Code	17
รูป 3.7 Flowchart แสดงวิธีแยกค่าข้อมูลตัวบุคคล.....	18
รูป 3.8 Flowchart แสดงการเช็คความถูกต้องของชื่อและรหัสวิชา	18
รูป 3.9 Flowchart แสดงวิธีแยกค่ารหัสวิชาและเกรดที่ได้.....	19
รูป 3.10 Flowchart แสดงการเช็คความถูกต้องของคะแนน.....	19
รูป 3.11 Flowchart แสดงวิธีการหาชุดรหัสความปลอดภัย	20
รูป 3.12 โครงสร้างระบบ	20
รูป 3.13 ส่วนติดต่อผู้ใช้งาน	21
รูป 3.14 ฐานข้อมูลของระบบ.....	22
รูป 4.1 การทำงานหลักของระบบ	23
รูป 4.2 โครงสร้างระบบ	24
รูป 4.3 ผลการทดลองแสดงค่าความถูกต้อง ต่อ เวลาที่ใช้ในการประมวลผล โดยเฉลี่ยของรูปแต่ละ ความละเอียด	28
รูป 4.4 ผลการทดลองแสดงค่าความถูกต้องของรูปความละเอียด 360 dpi.....	31

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่ เช่น การค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

	หน้า
รูป 4.5 ผลการทดลองแสดงเวลาในการประมวลผลของรูปความละเอียด 360 dpi	32
รูป 5.6 ตัวอย่างรูปแบบในการแก้ไขโดยการแบบพิมพ์แก้ไข.....	36
รูป 5.7 การทำงานหลังจากเพิ่มขึ้นตอนการเรียนรู้การแก้ไขข้อมูลจากผู้ใช้	37
รูป 5.8 แผนภูมิสรุปข้อมูลผลการตรวจสอบเอกสารรับรองผลการศึกษา.....	38
รูป 5.9 ตัวอย่างการปรับแต่งเพื่อตรวจสอบเอกสารรับรองผลการศึกษาของมหาวิทยาลัยอื่นๆ.....	39
รูป 5.10 ส่วนการเพิ่มเอกสารรับรองการผลการศึกษาในรูปแบบใหม่.....	40



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

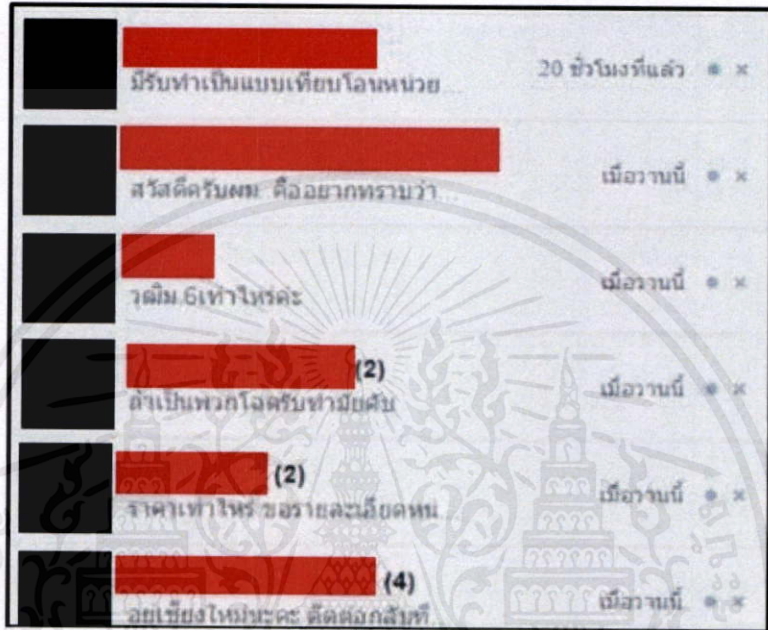
1.1 ความสำคัญและที่มาของโครงการ

ในปัจจุบันการสมัครงานของนักศึกษาหรือการที่บริษัทจะรับนักศึกษาเข้าทำงานจะมีหลักฐานการสมัครงานที่สำคัญคือ เอกสารรับรองผลการศึกษา(Transcript) เพราะข้อมูลบนเอกสารรับรองผลการศึกษามีข้อมูลที่เกี่ยวข้องกับประวัติการศึกษาของตัวนักศึกษาที่บริษัทต้องการทราบ เพื่อใช้ในการคัดเลือกนักศึกษาเข้ามาทำงานภายในบริษัทของตัวเอง ส่งผลให้ในปัจจุบันเกิดธุรกิจการปลอมแปลงเอกสารรับรองผลการศึกษา และแก้ไขข้อมูลบนเอกสารรับรองผลการศึกษาขึ้น โดยการปลอมแปลงจะมีทั้งการแก้ไขเกรดในบางวิชา การเพิ่มรายวิชา ฯลฯ แต่การปลอมแปลงที่พบเห็นมากที่สุดก็คือ การเปลี่ยนชื่อบนเอกสารรับรองผลการศึกษา เมื่อทำการสำรวจข้อมูลธุรกิจการปลอมแปลงเอกสาร พบว่าในปัจจุบันมีการรับปลอมแปลงเอกสารผลการศึกษาเป็นจำนวนมากและเปิดเผย รวมถึงใช้ Web Social Media เช่น Facebook ในการติดต่อรับจ้างงานปลอมแปลงดังรูปที่ 1.1 ที่แสดงเพจที่เปิดรับทำการปลอมแปลงผ่าน Web Social Media ของ Facebook



รูป 1.1 ตัวอย่างผู้ที่รับทำวุฒิการศึกษาปลอม

เมื่อลองติดต่อเพื่อใช้บริการ และสอบถามรายละเอียดต่างๆ จากผู้รับรบบลอมแปลงเอกสาร พบว่ามีการจ้างปลอมแปลงจริง ดังรูปที่ 1.2 และพบว่ามีจำนวนผู้ใช้งานเป็นจำนวนมากต่อวัน เพื่อความมั่นใจของหน่วยงานที่รับสมัครงาน จึงมีขั้นตอนที่ให้สถานศึกษาตรวจสอบและรับรองเอกสารทางการศึกษาก่อนนำมาใช้งานเสมอ



รูป 1.2 ตัวอย่างบุคคลที่สนใจทำวุฒิการศึกษาปลอม

สำหรับขั้นตอนการตรวจสอบและรับรองเอกสารของสถานศึกษาเริ่มตั้งแต่การส่งจดหมายขอใบรับรองจากสถานศึกษา จนถึงได้รับใบรับรองจากสถานศึกษาจะใช้ระยะเวลาประมาณ 1-2 สัปดาห์ จากปัญหาความล่าช้าดังกล่าว ผู้พัฒนาจึงสนใจที่จะพัฒนาระบบตรวจสอบเอกสารรับรองผลการศึกษา เพื่อตรวจสอบการปลอมแปลงเอกสารผลการศึกษา ซึ่งช่วยให้บริษัทสามารถตรวจสอบเอกสารทางการศึกษาได้โดยง่าย ช่วยลดระยะเวลาในการตรวจสอบเหลือประมาณ 20 นาทีต่อฉบับ และทำงานโดยอัตโนมัติสามารถลดภาระการทำงานของบุคลากรสำนักทะเบียนและประมวลผลได้

1.2 วัตถุประสงค์ของโครงการ

- 1) เพื่อพัฒนาระบบตรวจสอบภาพเอกสารรับรองผลการศึกษา
- 2) เพื่อช่วยลดระยะเวลาในการตรวจสอบเอกสารรับรองผลการศึกษาของหน่วยงานต่างๆ
- 3) เพื่ออำนวยความสะดวกในการตรวจสอบเอกสารรับรองผลการศึกษาหน่วยงานต่างๆ

1.3 ขอบเขตของโครงการ

- 1) สามารถตรวจสอบเอกสารรับรองผลการศึกษาของสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง โดยใช้ภาพของเอกสาร
- 2) ใช้ OCR Library ในการอ่านข้อมูลตัวอักษรบนเอกสารรับรองผลการศึกษาได้
- 3) สามารถนำการแฮช (Hash) มาใช้ในขั้นตอนที่การส่งข้อมูล
- 4) สรุปและนำเสนอผลงานเพื่อนำไปใช้งานหรือพัฒนาต่อยอดได้

1.4 วิธีการดำเนินงาน

- 1) ศึกษาทฤษฎีที่เกี่ยวข้อง ได้แก่ การทำงานของ OCR, OCR Library, C#, Database, ASP.net, Ajax, Json, MVC,
- 2) ทดสอบ OCR Library แต่ละผู้ผลิต
- 3) ลองเขียนโปรแกรมทดสอบเพื่อทดสอบความเป็นไปได้ของโครงการ
- 4) วางแผนการทำโครงการเพื่อให้เสร็จตามเวลาที่กำหนด
- 5) ออกแบบโครงสร้างของระบบ โดยจำลองถึงการใช้งานจริงของผู้ใช้งาน
- 6) พัฒนาตัวโปรแกรมที่ใช้ในการตรวจสอบ
- 7) ทำการทดลองใช้โปรแกรมเพื่อทดสอบผลลัพธ์
- 8) สรุปการทำโครงการ

1.5 ประโยชน์ที่คาดว่าจะได้รับ

- 1) สามารถนำไปใช้ตรวจสอบการปลอมแปลงเอกสารรับรองผลการศึกษาได้
- 2) อำนวยความสะดวกแก่ผู้ต้องการตรวจสอบเอกสารรับรองผลการศึกษา
- 3) ลดภาระการทำงานของสำนักทะเบียนและประมวลผลของสถาบัน
- 4) เป็นต้นแบบในการพัฒนาระบบเกี่ยวกับตรวจสอบเอกสารได้

1.6 ส่วนประกอบของปฏิญานิพนธ์

ปฏิญานิพนธ์ฉบับนี้ได้แบ่งเนื้อหาออกเป็น 5 บทด้วยกันคือ

- 1) บทที่ 1 บทนำ กล่าวถึงความสำคัญและที่มาของโครงการ วัตถุประสงค์ของโครงการ

เอกสารนี้เป็นเอกสารของโครงการ ขอบเขตของโครงการ วิธีการดำเนินการ ประโยชน์ที่คาดว่าจะได้รับ และส่วนประกอบ
ไม่ว่ากรณีใดๆ ทั้งสิ้น ของปฏิญานิพนธ์เปลี่ยนแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 2) บทที่ 2 ทฤษฎีที่เกี่ยวข้อง กล่าวถึงทฤษฎีพื้นฐานที่ใช้ในโครงการ

- 3) บทที่ 3 การออกแบบและพัฒนา กล่าวถึงรายละเอียดของโครงการนี้ ส่วนที่ได้ออกแบบและพัฒนาขึ้น การทำงานของระบบหรือชิ้นงานบรรยายโดยละเอียด
- 4) บทที่ 4 การทดลองและผลการทดลอง กล่าวถึงการเตรียมการทดลองทั้งการจัดเตรียมฮาร์ดแวร์ ซอฟต์แวร์ สภาวะแวดล้อมในการทำการทดลอง ข้อมูลทดสอบ การทำงานหรือการจำลองการทำงานของระบบ ผลการทดลอง ค่าสมรรถนะของระบบ การวัดประสิทธิภาพ ของระบบ การวิเคราะห์ผลการทดลองหรือผลการทำงาทั้งหมด
- 5) บทที่ 5 บทสรุป กล่าวถึงบทสรุปของโครงการ วิเคราะห์สิ่งที่ได้รับจากโครงการ ข้อจำกัด รวมถึงปัญหาอุปสรรคต่างๆ ของโครงการและข้อเสนอแนะสำหรับเป็นแนวทางในการพัฒนาต่อ



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

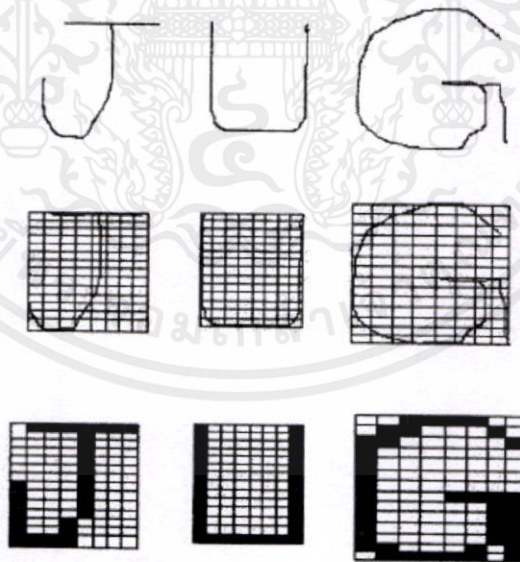
ทฤษฎีที่เกี่ยวข้อง

ระบบตรวจสอบผลการศึกษามีการใช้งาน OCR Library ในการอ่านตัวอักษรจากภาพเอกสาร รับรองผลการศึกษา และมีการใช้งานทฤษฎีแซช ในการเข้ารหัสเพื่อความปลอดภัยในขั้นตอนส่งข้อมูล ดังนั้นในส่วนของทฤษฎีที่เกี่ยวข้องจะอธิบายเกี่ยวกับ OCR วิธีการทำงาน ที่ทำให้สามารถอ่านภาพตัวอักษรและแปลงเป็นข้อมูลตัวอักษรได้ และอธิบายเกี่ยวกับรูปแบบเอกสารรับรองผลการศึกษาของสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง รวมถึงอธิบายเกี่ยวกับทฤษฎีแซช

2.1 Optical Character Recognition (OCR)

Optical Character Recognition คือการตรวจจับอักษรบนรูปภาพและแปลงตัวอักษรเหล่านั้นให้อยู่ในรูปแบบข้อมูลตัวหนังสือ โดยการกระบวนการดังกล่าวสามารถแบ่งได้เป็น 2 แบบ ดังนี้

2.1.1 แบบมองอักษรที่ตรวจจับในลักษณะเมตริกซ์



รูป 2.1 การตรวจจับลักษณะ Matrix

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้ทำไปใช้ประโยชน์ด้านการค้า โดยมองว่าอักษรที่ตรวจจับนั้นคือรูปภาพรูปหนึ่งทีประกอบไปด้วยพิกเซลหลายพิกเซลเรียงต่อกันในลักษณะรูปสี่เหลี่ยมที่แบ่งเป็นแถวเป็นหลัก เปรียบเสมือนเมตริกซ์หนึ่ง ให้แต่ละพิกเซลแทนข้อมูลในแต่ละตำแหน่งของเมตริกซ์จากนั้นจะนำเมตริกซ์ที่ได้ไปเปรียบเทียบกับ

ฐานข้อมูลอักษรที่เก็บข้อมูลแต่ละตัวอักษรในรูปแบบเมตริกซ์เช่นกัน หากค่าเมตริกซ์ที่เปรียบเทียบได้ใกล้เคียงกับอักษรตัวไหนมากที่สุด ก็จะแปลงภาพเป็นออกมาเป็นอักษรนั้น

2.1.2 แบบมองอักษรที่ตรวจจับในลักษณะ Vector



รูป 2.2 การตรวจจับแบบลักษณะ Vector

โดยจะมองอักษรบนภาพเกิดจากจุดและเส้นที่เรียงต่อกันจนกลายเป็นอักษร จากนั้นจะนำลักษณะการเรียงต่อกันของเส้น ไปเปรียบเทียบกับฐานข้อมูลอักษรที่เก็บลักษณะของเส้นของตัวอักษรต่างๆ ไว้ ลักษณะเส้นที่จับได้ใกล้เคียงกับลักษณะเส้นของอักษรตัวไหนในฐานข้อมูลมากที่สุด ก็จะแปลงภาพนั้นออกมาเป็นตัวอักษรดังกล่าว

2.2 การทำงานของ Tesseract OCR

Tesseract คือชื่อของ Library ที่ใช้งาน OCR และเปิดให้ใช้งานได้ฟรี รองรับระบบปฏิบัติการได้หลากหลาย สามารถอ่านภาพตัวอักษรและแปลงเป็นตัวอักษรได้หลายสามารถภาษา และถูกยอมรับว่าเป็นหนึ่งใน Library ที่มีความแม่นยำสูงในการใช้ OCR ในปัจจุบัน การทำงานของ Tesseract แบ่งออกเป็น 7 ขั้นตอน

1) ทำ Thresholding กับภาพที่รับเข้ามา

Thresholding คือขั้นตอนการทำภาพที่มีสีปกติ ให้เป็นรูป Binary Image คือมีแค่ ขาวกับดำ การทำ Thresholding นั้นไม่ยากนัก เพียงแค่ปรับค่าให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 2.3 ตัวอย่างการทำ Thresholding

- 2) ทำการพิจารณาภาพ Binary Image หาส่วนประกอบที่อยู่ติดกัน
- 3) ทำการค้นหาวรรทัดที่มีตัวอักษร และหาตัวอักษรในภาพ Binary Image
- 4) ทำการเรียนรู้คำศัพท์ เพื่อเพิ่มความแม่นยำในการอ่านตัวอักษร
- 5) รับค่าตัวอักษรที่อ่านได้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 รูปแบบเอกสารรับรองผลการศึกษาศึกษาของสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารรับรองผลการศึกษาศึกษาของสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังจะมีรูปแบบที่เหมือนกันตามรูปที่ 2.4

(Unofficial Transcript)

**Faculty of Engineering
Department of Computer Engineering**

Photo

Name: _____

Date of Birth: September 17, 1992 Student ID: _____

Date of Admission: 2011 Date of Graduation: N/A

Degree: Bachelor of Engineering

Major: Computer Engineering

COURSE TITLE	CREDIT	GRADE	COURSE TITLE	CREDIT	GRADE
1st Semester, Year 2011-2012					
0100001 ENGINEERING MATHEMATICS 1	3	B	0104203 COMPUTER NETWORKS	3	B+
0100009 ENGINEERING DRAWING	3	C	0104204 DATABASE SYSTEMS	3	A
0100009 ENGINEERING MECHANICS	3	C	0107205 ADVANCED DIGITAL SYSTEM DESIGN	3	A
0100003 GENERAL CHEMISTRY	3	B+	0107206 ADVANCED DIGITAL SYSTEM LABORATORY	1	B
0100004 GENERAL CHEMISTRY LABORATORY	1	B+	0100008 ADMINISTRATION OF COMPUTERS AND EDUCATIONAL TECHNOLOGY CENTER	3	B
0100001 GENERAL PHYSICS 1	3	B	0100009 COMMUNITY HEALTH	3	B
0100002 GENERAL PHYSICS LABORATORY 1	1	B+	2nd Semester, Year 2011-2012		
0100003 FOUNDATION ENGLISH 1	3	B+	0104107 THEORY OF COMPUTATION	3	C+
GP1 : 2.97 GP4 : 2.97			0104208 OPERATING SYSTEMS	3	A
3rd Semester, Year 2011-2012					
0100002 ENGINEERING MATHEMATICS 2	3	A	0104209 INFORMATION SYSTEMS ANALYSIS AND DESIGN	3	C
0100003 ENGINEERING MATERIALS 1	3	B	0104208 INFORMATION SYSTEMS ANALYSIS AND DESIGN LABORATORY	1	B
0100003 GENERAL PHYSICS 2	3	B	0104210 TCP/IP NETWORKS	3	B+
0100002 GENERAL PHYSICS LABORATORY 2	1	B+	0107202 CPU NETWORK LABORATORY	1	A
0100002 COMPUTER AND PROGRAMMING	3	B	0104209 INTRODUCTION TO LOCAL AND WIDE AREA NETWORKS	3	C+
0100003 FOUNDATION ENGLISH 2	3	A	0107202 LOCAL AND WIDE AREA NETWORKS LABORATORY	1	B+
0100003 FUNDAMENTAL LAWS FOR EVERYDAY LIFE	2	A	4th Semester, Year 2011-2012		
GP1 : 2.97 GP4 : 2.97			0100003 ENGINEERING MATHEMATICS 3	3	A
5th Semester, Year 2012-2013					
0104201 ADVANCED COMPUTER PROGRAMMING	3	B	0100004 INDUSTRIAL TRAINING	8	B
0107202 DIGITAL CIRCUIT AND LOGIC DESIGN	3	B+	6th Semester, Year 2012-2013		
0107202 DIGITAL CIRCUIT LABORATORY	1	B+	0104201 PROJECT	1	X
0107204 BASIC ELECTRONICS FOR COMPUTER ENGINEERING	3	B+	0104208 INFORMATION STORAGE AND RETRIEVAL	3	X
0107203 COMPUTER ENGINEERING LABORATORY	1	A	0107203 MULTIMEDIA SYSTEMS	3	X
0100002 INTRODUCED TO TRAIL LAWS	3	A	0107202 NETWORK PROGRAMMING	3	X
0100002 ENGLISH FOR PROFESSIONAL PURPOSES	3	A	0104201 REMAIN COMPUTER EXTRACTION	3	X
GP1 : 2.97 GP4 : 2.97			GP1 : 2.86 GP4 : 2.77		
7th Semester, Year 2012-2013					
0107204 DATA STRUCTURE AND ALGORITHM	3	C	Total number of credit course: 62		
0107207 DATA COMMUNICATION	3	A	Cumulative GPA: 2.81		
0107208 DATA COMMUNICATION LABORATORY	1	A	Dean's Approval		
0107209 COMPUTER ORGANIZATION AND ASSEMBLY LANGUAGE	3	A	Signature of Chair		
0107208 ASSEMBLY LANGUAGE LABORATORY	1	C	Chairship		
0107204 WEB TECHNOLOGY	3	B	The Associate Dean's Approval		
0100002 MATHEMATICS IN DAILY LIFE	3	B	Signature of Dean		
0100007 ENGLISH FOR BUSINESS	3	B	Signature of Dean		
GP1 : 2.55 GP4 : 2.55			GP1 : 2.86 GP4 : 2.77		
8th Semester, Year 2012-2013					
0100004 INDIVIDUAL STUDY	3	B+	Signature of Dean		
0100002 REMAIN AND COURSE	3	A	Signature of Dean		
GP1 : 2.74 GP4 : 2.74			GP1 : 2.86 GP4 : 2.77		
9th Semester, Year 2013-2014					
0107201 SOFTWARE ENGINEERING	3	B+	Signature of Dean		
0107202 COMPUTER ARCHITECTURE	3	B+	Signature of Dean		
GP1 : 2.74 GP4 : 2.74			GP1 : 2.86 GP4 : 2.77		

Date Issued: 01/05/2014 This document is unofficial transcript.

รูป 2.4 เอกสารรับรองผลการศึกษาศึกษาปกติ

ซึ่งในเอกสารรับรองผลการศึกษาศึกษาจะข้อมูลอยู่ 2 ส่วนคือข้อมูลประวัติส่วนตัว ดังรูปที่ 2.5 และ ข้อมูลผลการเรียนดังรูปที่ 2.6 ในส่วนของข้อมูลประวัติผู้เรียน จะเก็บข้อมูลดังนี้

1. ชื่อ - นามสกุล
2. วันเดือนปีเกิด
3. รหัสนักศึกษา
4. ปีที่เข้ามาเรียน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5. ปีที่จบ
6. คณะที่เรียน, สาขาที่เรียน

Faculty of Engineering
Department of Computer Engineering

Name Mr.Chanawee Chanpaengngoen

Date of Birth September 17, 1992

Student ID 54010275

Date of Admission 2011

Date of Graduation N/A

Degree Bachelor of Engineering

Major Computer Engineering

รูป 2.5 ส่วนข้อมูลประวัติผู้เรียนบนเอกสารรับรองผลการศึกษา

ส่วนต่อไปเป็นข้อมูลผลการเรียนของนักศึกษา ประกอบไปด้วยข้อมูลดังนี้

1. เทอมที่ลงทะเบียนเรียน
2. ชื่อ – รหัสวิชาที่ลงทะเบียน และหน่วยกิตของวิชานั้นๆ
3. เกรดรายวิชา
4. GPA / GPS
5. หน่วยกิตรวมทั้งหมด

COURSE TITLE	CREDIT	GRADE	COURSE TITLE	CREDIT	GRADE
1st Semester, Year, 2011-2012					
01006001 ENGINEERING MATHEMATICS 1	3	B	01076213 COMPUTER NETWORKS	3	B+
01006009 ENGINEERING DRAWING	3	C+	01076214 DATABASE SYSTEMS	3	A
01006010 ENGINEERING MECHANICS	3	C	01076215 ADVANCED DIGITAL SYSTEM DESIGN	3	A
05100193 GENERAL CHEMISTRY	3	B+	01076216 ADVANCED DIGITAL SYSTEM LABORATORY	1	B
05100194 GENERAL CHEMISTRY LABORATORY	1	B+	03260090 ADMINISTRATION OF INNOVATION AND EDUCATIONAL TECHNOLOGY CENTER	3	B
05300121 GENERAL PHYSICS 1	3	B	90303015 COMMUNITY HEALTH	3	B
05300122 GENERAL PHYSICS LABORATORY 1	1	B+	GPS : 2.47 GPA : 2.16		
90201001 FOUNDATION ENGLISH 1	3	B+	2nd Semester, Year, 2012-2014		
GPS : 2.97 GPA : 2.97					
2nd Semester, Year, 2011-2012					
01006002 ENGINEERING MATHEMATICS 2	3	B+	01076217 THEORY OF COMPUTATION	3	C+
01006011 ENGINEERING MATERIALS	3	B	01076218 OPERATING SYSTEMS	3	A
05300123 GENERAL PHYSICS 2	3	B	01076219 INFORMATION SYSTEMS ANALYSIS AND DESIGN	3	C
05300124 GENERAL PHYSICS LABORATORY 2	1	B+	01076220 INFORMATION SYSTEM ANALYSIS AND DESIGN LABORATORY	1	B
90102003 COMPUTERS AND PROGRAMMING	3	A	01076221 TCP/IP NETWORKS	3	B+
90201002 FOUNDATION ENGLISH 2	3	B	01076222 TCP/IP NETWORKS LABORATORY	1	A
90402005 FUNDAMENTAL LAWS FOR EVERYDAY LIFE	3	A	01076601 INTRODUCTION TO LOCAL AND WIDE AREA NETWORKS	3	C+
GPS : 3.42 GPA : 3.19					
1st Semester, Year, 2012-2013					
01006003 ENGINEERING MATHEMATICS 3	3	A	01076602 LOCAL AND WIDE AREA NETWORKS LABORATORY	1	B+
01076201 ADVANCED COMPUTER PROGRAMMING	3	B	GPS : 3.08 GPA : 3.11		
01076202 DIGITAL CIRCUIT AND LOGIC DESIGN	3	B+	3rd Semester, Year, 2013-2014		
01076203 DIGITAL CIRCUIT LABORATORY	1	A	01006004 INDUSTRIAL TRAINING	0	
01076204 BASIC ELECTRONICS FOR COMPUTER ENGINEERING	3	B+	GPS : GPA : 3.11		
01076205 COMPUTER ENGINEERING LABORATORY	1	A	01076301 PROJECT 1	1	X
90040002 INTRODUCTION TO THAI LAWS	3	A	01076526 INFORMATION STORAGE AND RETRIEVAL	3	X
90201024 ENGLISH FOR PROFESSIONAL PURPOSES	3	A	01076531 MULTIMEDIA SYSTEMS	3	X
GPS : 3.79 GPA : 3.16					
2nd Semester, Year, 2012-2013					
01076206 DATA STRUCTURE AND ALGORITHM	3	C	01076542 NETWORK PROGRAMMING	3	X
01076207 DATA COMMUNICATION	3	A	01076547 HUMAN COMPUTER INTERACTION	3	X
01076208 DATA COMMUNICATION LABORATORY	1	A	GPS : 0.00 GPA : 3.11		
01076209 COMPUTER ORGANIZATION AND ASSEMBLY LANGUAGE	3	A	Total number of credit earned: 125		
01076210 ASSEMBLY LANGUAGE LABORATORY	1	C	Cumulative GPA: 3.31		
01076541 WEB TECHNOLOGY	3	B	Transcript Closed		
90101002 MATHEMATICS IN DAILY LIFE	3	B	Checked by _____		
90201017 ENGLISH FOR BUSINESS	3	B	(Or XXXXXXXX XXXXXXXXXXXX)		
GPS : 3.15 GPA : 3.11					
3rd Semester, Year, 2012-2013					
03426034 INDIVIDUAL STUDY	3	B+	Total number of credit earned: 125		
90305003 HUMAN AND TOURISM	3	A			
GPS : 3.75 GPA : 3.14					
1st Semester, Year, 2013-2014					
01076211 SOFTWARE ENGINEERING	3	B+	Total number of credit earned: 125		
01076212 COMPUTER ARCHITECTURE	3	B+			
Continue next column					

รูป 2.6 ส่วนข้อมูลประวัติการศึกษาบนเอกสารรับรองผลการศึกษา

เอกสารนี้เป็นเอกสารที่ส่วนงานใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปขึ้นประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น หากมีข้อสงสัย กรุณาติดต่อแจ้งถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4 ทฤษฎี Hashing

การทำแฮช (Hashing) คือการแปลงข้อมูลแบบหนึ่ง เรียกว่าการทำแฮชข้อมูล เพื่อให้การเข้าถึงข้อมูลมีความปลอดภัยมากขึ้น ค่าที่ได้จากการแฮชจะเป็นตัวเลขและตัวอักษรที่ถูกสร้างขึ้นใหม่ โดยเลขที่สร้างออกมาใหม่จะมีขนาดที่สั้นกว่าเดิม และรูปแบบที่ได้จากการแฮชจะมีสูตรการสร้างอยู่ ขึ้นอยู่กับวิธีในการทำแฮช

ในปัจจุบันมีอัลกอริทึมหลายรูปแบบในการใช้งานฟังก์ชันแฮชขั้นตอนการตัดสินใจเลือกใช้งานอัลกอริทึมควรพิจารณาจากข้อมูลที่ต้องการนำไปแฮชและความยาวข้อมูลที่ต้องการรับกลับมา รวมถึงความปลอดภัยในการเข้ารหัสข้อมูล ในแต่ละอัลกอริทึมจะมีค่าต่างกัน ดังที่แสดงในตารางที่ 2.1

ตาราง 2.1 แสดงความแตกต่างของฟังก์ชันแฮช

Algorithm	Output size (bits)	Internal state size	Block size	Length size	Word size	Rounds
GOST	256	256	256	256	32	256
HAVAL	256/224/192/160/128	256	1,024	64	32	160/128/96
MD2	128	384	128	-	32	864
MD4	128	128	512	64	32	48
MD5	128	128	512	64	32	64
PANAMA	256	8,736	256	-	32	-
RadioGatún	Up to 608/1,216	58 words	3 words	-	1-64	-
RIPEMD	128	128	512	64	32	48
RIPEMD-128/256	128/256	128/256	512	64	32	64
RIPEMD-160	160	160	512	64	32	80

Algorithm	Output size (bits)	Internal state size	Block size	Length size	Word size	Rounds
RIPEMD-320	320	320	512	64	32	80
SHA-0	160	160	512	64	32	80
SHA-1	160	160	512	64	32	80
SHA-224, SHA-256	256/224	256	512	64	32	64
SHA-384, SHA-512, SHA-512/224, SHA-512/256	384/512/224/256	512	1,024	128	64	80
SHA-3	224/256/384/512[p 2]	1600	1600- 2*bits	-	64	24
SHA3-224	224	1600	1152	-	64	24
SHA3-256	256	1600	1088	-	64	24
SHA3-384	384	1600	832	-	64	24
SHA3-512	512	1600	576	-	64	24
Tiger(2)- 192/160/128	192/160/128	192	512	64	64	24
WHIRLPOOL	512	512	512	256	8	10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ผู้สืบค้นข้อมูลนี้ให้ออกไปเอง ผู้พัฒนาและผู้ใช้งานจึงเป็นฝ่ายรับผิดชอบเอกสารหรือข้อมูลที่ปรากฏไปใช้

ในการพัฒนาระบบตรวจสอบผู้พัฒนาได้เลือกใช้งานอัลกอริทึม SHA3-256 เนื่องจากใน

ปัจจุบัน SHA3 มีความปลอดภัยในการเข้ารหัสข้อมูลสูง และยังไม่มีการค้นพบวิธีการสร้าง

ข้อมูลให้มีค่าแฮชที่ซ้ำกันได้ เลข 256 จาก SHA3-256 หมายถึงจำนวนบิตข้อมูลที่ใส่เก็บค่า
ข้อมูลที่ได้จากการ แฮช โดย 1 ตัวอักษรจะใช้จำนวนบิตในการเก็บ 4 บิต โดยมีค่าที่เป็นไปได้
คือ 0-f ดังนั้นจำนวนตัวอักษรที่แฮชออกมาจากอัลกอริทึม SHA3-256 คือ 64 ตัวอักษร



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

การออกแบบและพัฒนา

ระบบตรวจสอบการปลอมแปลงเอกสารรับรองผลการศึกษาระบบที่ออกแบบไว้สำหรับให้ฝ่ายบุคคลของบริษัทเป็นผู้ใช้งาน โดยสามารถตรวจสอบเอกสารรับรองผลการศึกษาได้เอง ไม่จำเป็นต้องผ่านสำนักทะเบียนและประมวลผลของสถาบัน ส่งผลให้ใช้เวลาในการตรวจสอบเอกสารลดลง ผู้พัฒนาได้ออกแบบการตรวจสอบให้เป็น Web application เพื่อไว้เป็นทางเลือกหนึ่งในการตรวจสอบที่รวดเร็ว และสะดวกมากยิ่งขึ้น

สำหรับการจัดการเตรียมเอกสารรับรองผลการศึกษาเพื่อให้สามารถใช้โปรแกรมตรวจสอบได้ จำเป็นต้องเพิ่มข้อมูลแฮชขึ้นในตำแหน่งมุมซ้ายบนของเอกสารรับรองผลการศึกษา ค่าข้อมูลของชุดรหัสความปลอดภัยได้มาจากการนำข้อมูลผลการศึกษาไปแปลงค่า และเข้าฟังก์ชัน SHA3-256 เพื่อทำการแฮชและสร้างชุดข้อมูลขึ้น ค่า Code ที่ได้จะเป็นคิงรูปที่ 3.1 คือตัดจากค่าที่ได้ SHA3-256 ที่มี 64 ตัวอักษรเหลือ 12 ตัวอักษรเพื่อนำมาใช้งานและหลังจากนำไปใส่ในเอกสารรับรองผลการศึกษาดังรูปที่ 3.2 ภาพเอกสารรับรองผลการศึกษาดังนั้นจะสามารถใช้งานระบบตรวจสอบได้

21538d91bc71

รูป 3.1 ตัวอย่างค่าที่ได้จากการ Generate Code ข้อมูลบนเอกสารรับรองผลการศึกษา

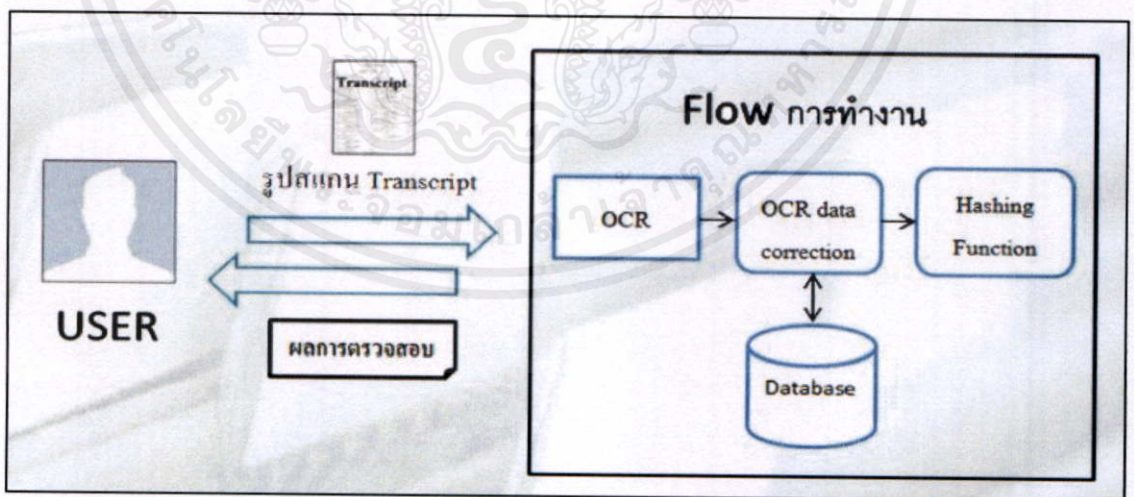
21538d91bc71	
(Unofficial Transcript)	
<i>Faculty of</i> Engineering <i>Department of</i> Computer Engineering	
<i>Name</i> Mr.Chanawee Chanpaengngoen	
<i>Date of Birth</i> September 17, 1992	<i>Student ID</i> 54010275
<i>Date of Admission</i> 2011	<i>Date of Graduation</i> N/A
<i>Degree</i> Bachelor of Engineering	
<i>Major</i> Computer Engineering	

เอกสารนี้เป็นเอกสารลับ
รูป 3.2 ข้อมูลประวัติผู้เรียนบนเอกสารรับรองผลการศึกษาดังนั้นจะสามารถใช้งานระบบตรวจสอบได้
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1 การทำงานของระบบ

ระบบตรวจสอบการปลอมแปลงเอกสารรับรองผลการศึกษาคือระบบที่ทำงานผ่าน Web application และทำงานบนเซิร์ฟเวอร์ โดยขั้นตอนการทำงานของระบบขั้นตอนแรกจะรองรับรูปเอกสารรับรองผลการศึกษาที่ผู้ใช้งานต้องการตรวจสอบส่งเข้ามาให้เซิร์ฟเวอร์จากนั้นเซิร์ฟเวอร์จะนำเอกสารรับรองผลการศึกษาที่ได้รับมานำไปเข้าขั้นตอนการตรวจสอบที่สามารถแสดงขั้นตอนการทำงานได้ดังรูป Flowchart ในส่วนที่จะอธิบายต่อไป

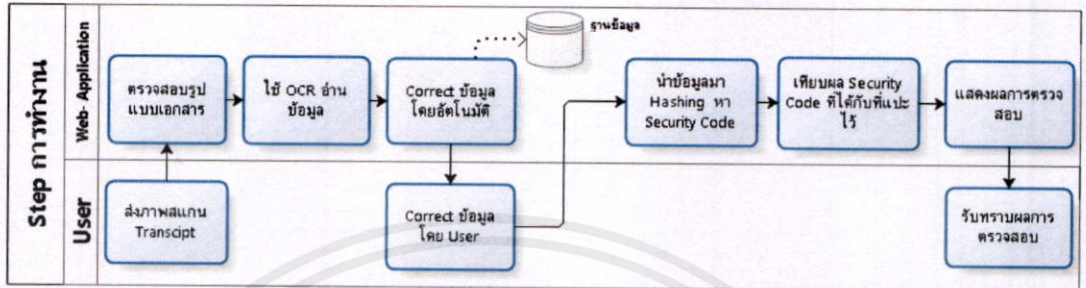
รูปที่ 3.3 จะแสดงถึงขั้นตอนการทำงานทั้งหมดของโปรแกรมตรวจสอบความถูกต้องของเอกสารรับรองผลการศึกษาเริ่มตั้งแต่ขั้นตอนแรกคือ รองรับภาพเอกสารรับรองผลการศึกษาจากผู้ใช้งาน โปรแกรมจากนั้นจะส่งไฟล์ภาพเอกสารรับรองผลการศึกษาเข้าไปทำงานในส่วนของการตรวจสอบการปลอมแปลง โดยแบ่งเป็นการทำงานเป็นลำดับขั้น ขั้นแรกจะใช้ OCR ในการอ่านข้อมูลบนเอกสารรับรองผลการศึกษาเพื่อให้ได้ผลลัพธ์เป็นตัวอักษร แต่การอ่านตัวอักษรโดยใช้ OCR จะมีความผิดพลาด จึงต้องมีขั้นตอนการแก้ไขข้อมูลที่ OCR อ่านได้ เพื่อช่วยให้ข้อมูลมีความถูกต้อง โดยขั้นตอนนี้จะมีการใช้ข้อมูลในฐานะข้อมูลมาช่วยในการทำการแก้ไขข้อมูลเพื่อความรวดเร็วขึ้น จากนั้นจะนำข้อมูลที่อ่านได้ไปหาค่าแฮชอีกกรอบหนึ่งเพื่อเพิ่มความปลอดภัยและเพื่อเป็นการตรวจสอบว่าข้อมูลเอกสารรับรองผลการศึกษานั้นมีความถูกต้องหรือไม่ จากนั้นจะส่งผลลัพธ์แจ้งผลการตรวจสอบให้กับผู้ใช้งาน โดยขั้นตอนทั้งหมดใช้เวลาประมาณ 20-40 วินาที



รูป 3.3 ภาพรวมการทำงานของโปรแกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับรายละเอียดการทำงานของแต่ละขั้นตอนหลักในรูปที่ 3.3 จะอธิบายโดยละเอียดในรูปที่ 3.4 ซึ่งอธิบายขั้นตอนการทำงานของโปรแกรมในรูปแบบ Swimming lane ที่มี 2 ส่วนคือ ส่วนของผู้ใช้งานและส่วนของโปรแกรมตรวจสอบ



รูป 3.4 Swimming lane แสดงขั้นตอนการทำงานของโปรแกรม

จากรูปที่ 3.4 ผู้ใช้งานมีการทำงาน 3 ขั้นตอนคือ

- 1) ส่งภาพเอกสารรับรองผลการศึกษาเข้าสู่ระบบตรวจสอบ
- 2) ทำการ Correction ข้อมูล
- 3) รับผลการตรวจสอบ

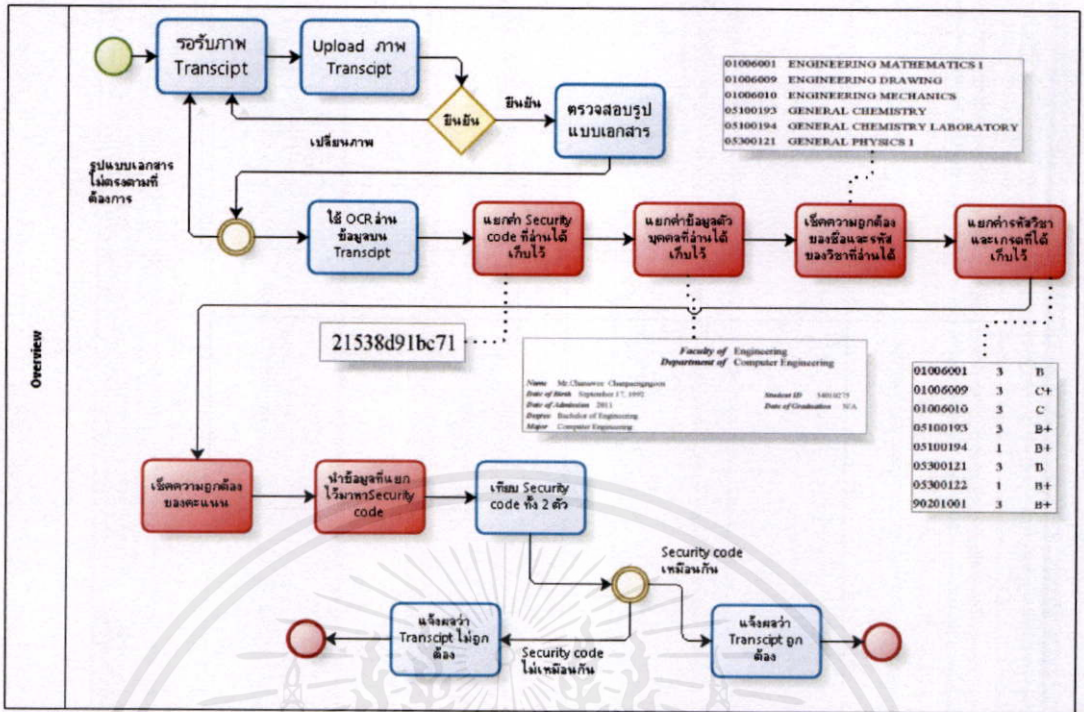
สาเหตุที่ต้องมีการทำ Correction ข้อมูลเนื่องจากการอ่านข้อมูลจากภาพเอกสารด้วย OCR Library ยังพบว่ามีผิดพลาดในการอ่านตัวอักษร โดยขั้นตอนการแก้ไขข้อมูลแบ่งเป็น 2 ขั้นตอน

- 1) แก้ไขโดยระบบ
- 2) แก้ไขโดยผู้ใช้งาน

ในการแก้ไขข้อมูลโดยระบบ เป็นการนำข้อมูลที่ใกล้เคียงที่สุดจากฐานข้อมูลมาทำการแก้ไข ซึ่งเป็นข้อมูลที่มีค่าต่างไม่เกิน 2 ตัวอักษร หากเกินกว่านั้นจะให้ผู้ใช้งานทำการแก้ไขต่อไป

ในการตรวจสอบความถูกต้องของเอกสาร ระบบจะสร้างชุดรหัสความปลอดภัยจากข้อมูลของเอกสารรับรองผลการศึกษาที่ประกอบด้วย ประวัติบุคคล ข้อมูลรหัสวิชาและผลการเรียน เพื่อนำมาเปรียบเทียบกับชุดรหัสที่อยู่บนเอกสาร หากมีค่าตรงกันจะระบุได้ว่าเอกสารดังกล่าวเป็นฉบับที่ไม่ถูกปลอมแปลง ดังรูปที่ 3.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 3.5 Flowchart ภาพรวมการทำงานของระบบ

- วงกลมสีเขียว แสดงจุดเริ่มการทำงานของโปรแกรม
- วงกลมสีแดงแสดงถึงจุดสิ้นสุดการทำงานของโปรแกรม
- กรอบสี่เหลี่ยมสีฟ้าแสดงถึง Process การทำงานของโปรแกรม
- กรอบสี่เหลี่ยมสีแดงแสดงถึง Process การทำงานของโปรแกรมที่มีอธิบายต่อเป็นใน Process ย่อยในหัวข้อที่ 3.2
- กรอบสี่เหลี่ยมสีเหลืองทำมุม 45 องศา แสดงถึง Gateway แสดงจุดที่มีทางเลือกการทำงาน โดยมีเงื่อนไขการทำงานมากกว่า 2 เส้นทาง ขึ้นอยู่กับผลลัพธ์ที่ได้จากความต้องการของผู้ใช้งาน
- วงกลมสีเหลืองแสดงจุดที่มีทางเลือกการทำงาน โดยมีเงื่อนไขการทำงานมากกว่า 2 เส้นทาง ขึ้นอยู่กับค่าที่ได้จาก Process ก่อนหน้านั้น

3.2 วิธีการตรวจสอบ

ระบบตรวจสอบเอกสารรับรองผลการศึกษาจะใช้ OCR Library ช่วยในการอ่านข้อมูลบน

ภาพเอกสารรับรองผลการศึกษา โดยนำข้อมูลดังกล่าวสร้างชุดรหัสความปลอดภัย สำหรับใช้ในการตรวจสอบการปลอมแปลงของเอกสาร ซึ่งมีรายละเอียดขั้นตอนการทำงานดังต่อไปนี้

- 1) ตรวจสอบรูปแบบของภาพเอกสารรับรองผลการศึกษา

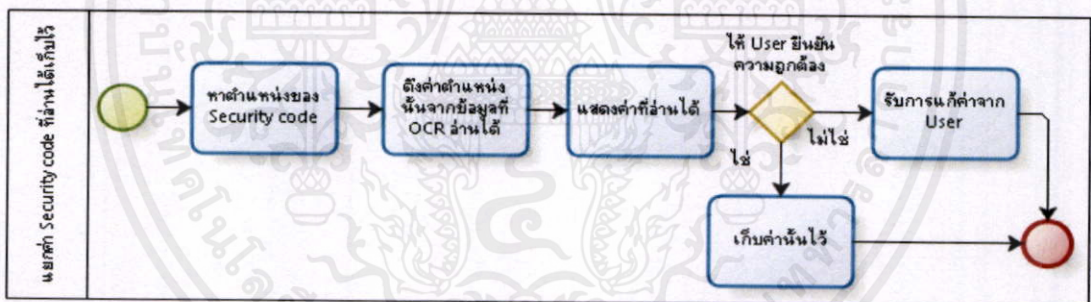
ในขั้นตอนนี้ ระบบจะทำการตรวจสอบรูปแบบของภาพเอกสารรับรองผลการศึกษา จาก ตำแหน่งและจำนวนของกล่องการตรวจจับข้อมูลที่ได้จากการทำ OCR ว่ามีตำแหน่งและจำนวน อยู่ในช่วงที่กำหนดไว้หรือไม่ หากไม่อยู่ในช่วงที่กำหนดจะหยุดและยกเลิกการทำงานขั้นถัดไป แล้วแจ้งเตือนให้ว่าเอกสารดังกล่าวมีรูปแบบที่ไม่ถูกต้อง

2) ใช้ OCR Library อ่านข้อมูลบนภาพเอกสารรับรองผลการศึกษา

หลังจากที่ระบบผ่านการตรวจสอบรูปแบบของเอกสารแล้ว จะทำการแปลงภาพเป็นอักษร โดยในการเก็บข้อมูลจะแบ่งการเก็บเป็นช่วงทุกครั้งที่พบระยะห่างระหว่างตัวอักษรมีมากกว่า ระยะห่างของเว้นวรรค แล้วไว้ในตัวแปร Array ตามลำดับ Index เพื่อใช้ในการจำแนกประเภท ข้อมูลต่อไป

3) ดึงข้อมูลชุดรหัสความปลอดภัย

ในขั้นตอนนี้จะทำการดึงข้อมูลชุดรหัสความปลอดภัย จากตัวแปร Array ที่ได้จาก ขั้นตอนที่ 2) โดยดึงข้อมูลจากตำแหน่ง Index แรกของตัวแปร เก็บไว้ในตัวแปรของ ชุดรหัสความปลอดภัย เนื่องจากชุดรหัสดังกล่าวอยู่บนซ้าย ซึ่งเป็นตำแหน่งแรกในการเริ่มทำ OCR คืออ่าน จากทางซ้ายไปขวา และจากด้านบนลงล่าง ดังการทำงานตามรูปที่ 3.6

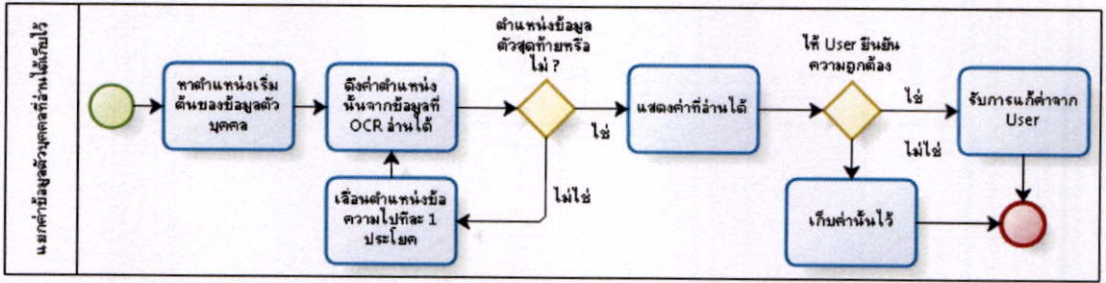


รูป 3.6 Flowchart แสดงวิธีแยกค่า Security Code

4) ดึงข้อมูลตัวบุคคล

ในขั้นตอนนี้จะทำการดึงข้อมูลตัวบุคคลจากตัวแปร Array ที่ได้จากขั้นตอนที่ 2) โดยอาศัย ตำแหน่งในแนวแกน Y ของรูปภาพที่มีค่าระหว่าง 5%-20% ของความสูงภาพเอกสาร ในการระบุขอบเขตของข้อมูลตัวบุคคลและการเรียงลำดับของข้อมูล โดยมีลำดับดังต่อไปนี้ ชื่อ-นามสกุล, วันเกิด, รหัสนักศึกษา, ปีที่เข้ามาศึกษา, ปีที่จบการศึกษา, คณะวิชาและภาควิชา ในการแยกชนิดของข้อมูลตัวบุคคล ดังการทำงานตามรูปที่ 3.8

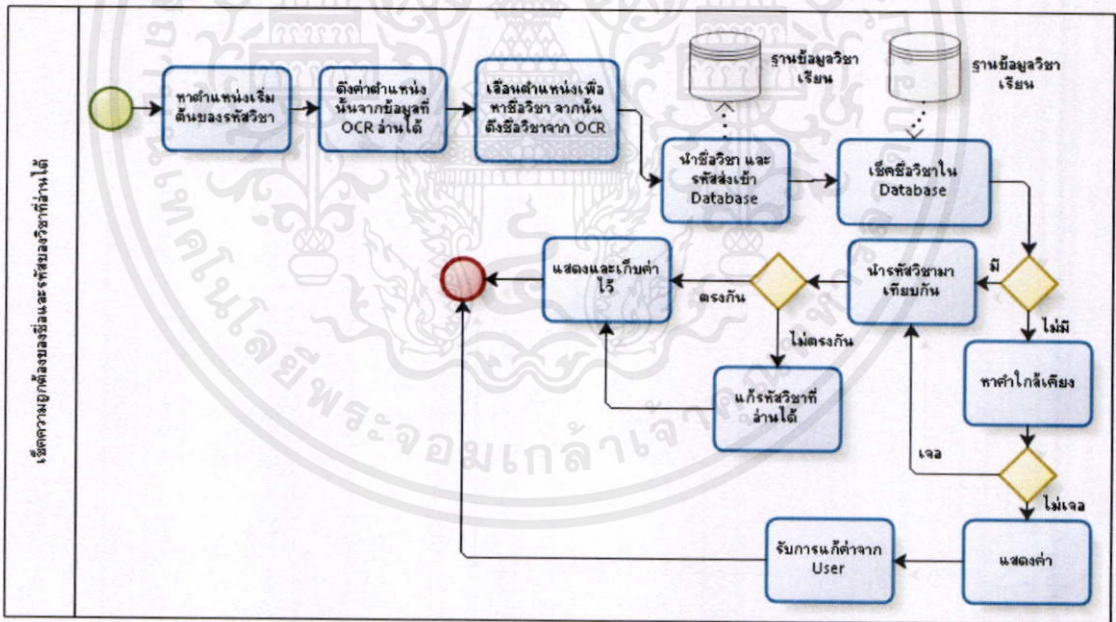
เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ซึ่งในพิธีการศึกษานี้เป็นอนุญาติให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 3.7 Flowchart แสดงวิธีแยกค่าข้อมูลตัวบุคคล

5) เช็คว่าความถูกต้องของชื่อวิชา และรหัสวิชา

ในขั้นตอนนี้จะทำการเช็คว่าความถูกต้องของรหัสและชื่อของแต่ละรายวิชา โดยส่งข้อมูลรหัสและชื่อวิชาที่ต้องการไปตรวจสอบกับฐานข้อมูลรายวิชาที่ประกอบด้วยรหัสและชื่อวิชา เพื่อหาว่ารหัสและชื่อวิชาสอดคล้องกับค่าในฐานข้อมูลหรือไม่ ถ้าขัดแย้งกันให้ทำการตรวจสอบอีกครั้งด้วยค่าที่ต่างกันไม่เกิน 2 อักขร หากผลยังคงขัดแย้งให้ระบุว่าข้อมูลวิชาดังกล่าวมีข้อผิดพลาดจะส่งให้ผู้ใช้งานทำการแก้ไขต่อไป ดังการทำงานตามรูปที่ 3.8

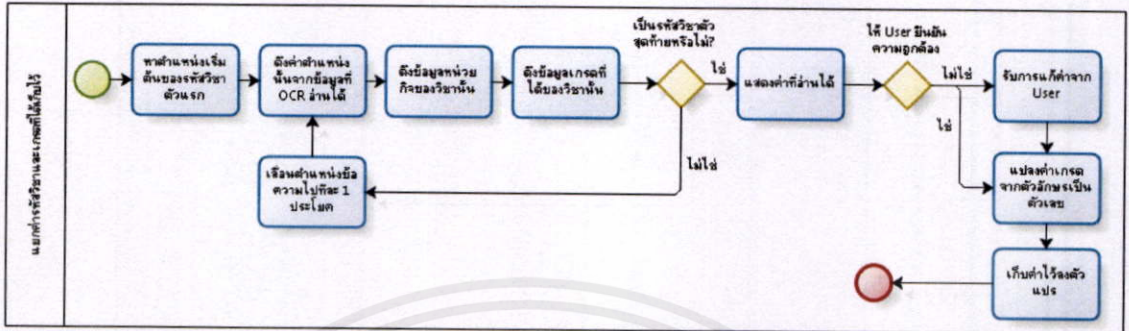


รูป 3.8 Flowchart แสดงการเช็คว่าความถูกต้องของชื่อและรหัสวิชา

6) ดึงคำรหัสวิชา และเกรดที่ได้เก็บไว้

เอกสารนี้เป็นเอกสารที่จัดทำขึ้นเพื่อใช้ในการเรียนการสอนเท่านั้น ไม่สามารถนำข้อมูลไปใช้โดยไม่ได้รับอนุญาต
 ในขั้นตอนนี้จะทำการดึงคำรหัสและเกรดรายวิชา ออกจากตัวแปร Array ที่ได้จากขั้นตอนที่ 2) โดยค้นหาค่าที่เป็นตัวเลขและมีจำนวน 8 หลัก ในการหารหัสวิชา จากนั้นให้ทำการเพิ่ม Index

จาก Index ที่พบรหัสวิชาไป 3 ตำแหน่งจะได้ Index ที่เก็บข้อมูลเกรด เนื่องจากข้อมูลรายวิชาจะมีการเรียงลำดับดังนี้ รหัสวิชา,ชื่อวิชา,หน่วยกิตและเกรด ดังการทำงานตามรูปที่ 3.9



รูป 3.9 Flowchart แสดงวิธีแยกค่ารหัสวิชาและเกรดที่ได้

7) ตรวจสอบความถูกต้องของคะแนนที่อ่านได้

ในขั้นตอนนี้จะทำการตรวจสอบในส่วนของข้อมูล GPS และ GPA โดยคำนวณจากข้อมูลเกรดและหน่วยกิตบนภาพเอกสารรับรองการศึกษา หากพบว่าค่าที่คำนวณได้ไม่ตรงกับค่าที่ปรากฏบนภาพเอกสาร จะระบุว่าข้อมูลนั้นผิดพลาด แล้วส่งต่อให้ผู้ใช้งานพิจารณาการแก้ไขต่อไป ดังการทำงานตามรูปที่ 3.10



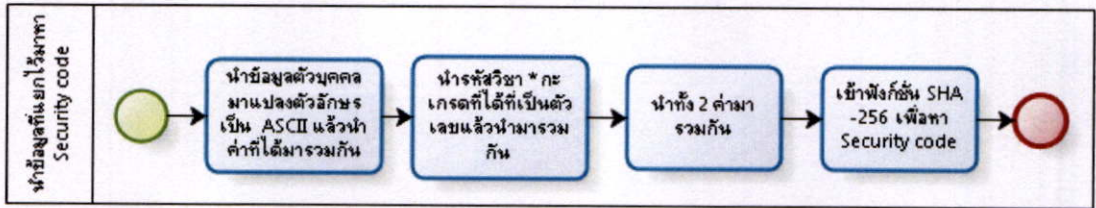
รูป 3.10 Flowchart แสดงการเช็คความถูกต้องของคะแนน

8) สร้างชุดรหัสความปลอดภัยจากข้อมูลภาพเอกสารรับรองผลการศึกษา

นำข้อมูลที่ได้จากขั้นตอนที่ 4) – 6) มาทำการ Summation ค่า ASCII ของตัวอักษรแต่ละตัวของแต่ละข้อมูล เพื่อหาผลรวมที่ได้จากข้อมูลทั้งหมด แล้วนำค่าที่ได้ทำการแฮชโดยใช้ฟังก์ชัน

เอกสารนี้เป็น SHA-256 จากนั้นเลือกข้อมูล 12 หลักแรกให้เป็นชุดรหัสความปลอดภัย ดังการทำงานตามรูปที่ 3.11

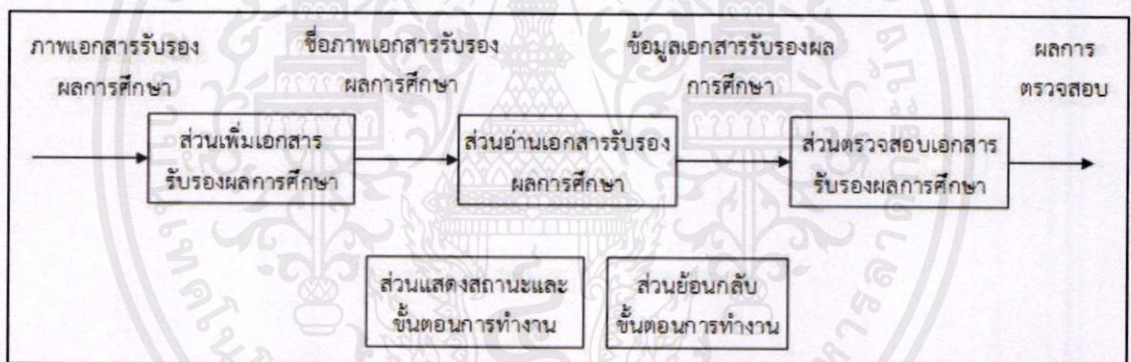
ไม่ว่ากรณีใด 3.11 ขึ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 3.11 Flowchart แสดงวิธีการหาชุดรหัสความปลอดภัย

- 9) นำชุดรหัสความปลอดภัยที่หาได้ มาเปรียบเทียบกับชุดรหัสความปลอดภัยที่เก็บไว้ตอนแรก
- หากค่าตรงกัน แสดงว่าเอกสารรับรองผลการศึกษาดังกล่าวไม่ผ่านการปลอมแปลง
 - หากค่าไม่ตรงกัน แสดงว่าเอกสารรับรองผลการศึกษาดังกล่าวเป็นเอกสารที่ผ่านการปลอมแปลง

3.3 โครงสร้างของระบบ

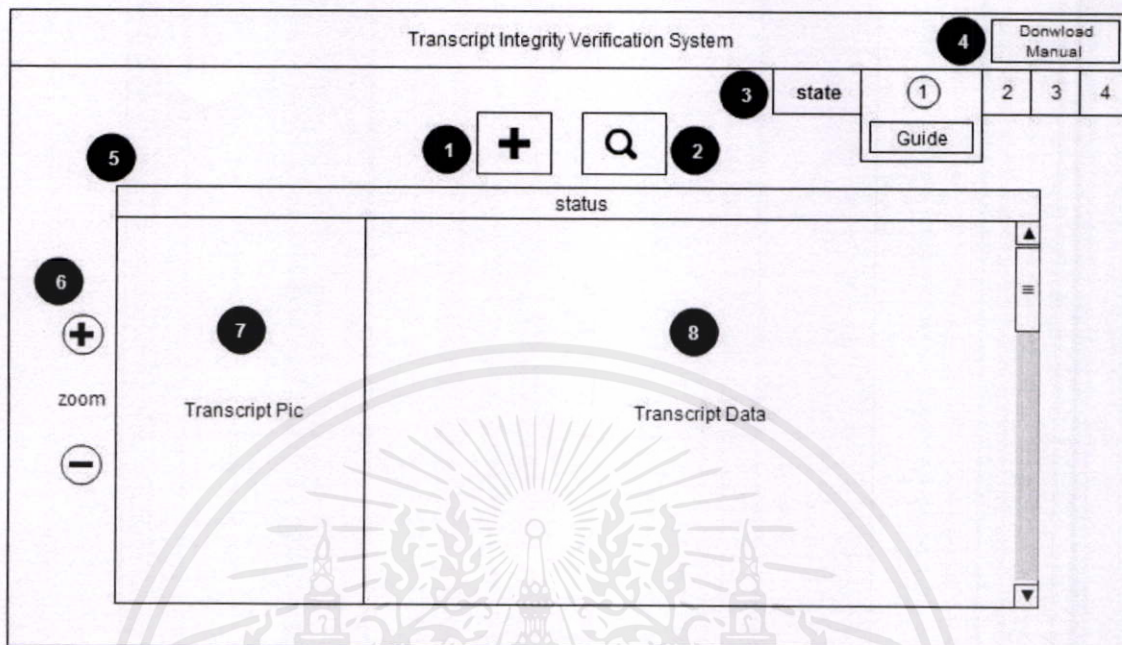


รูป 3.12 โครงสร้างระบบ

การทำงานของระบบแบ่งออกเป็น 5 ส่วน ได้แก่

- 1) ส่วนเพิ่มเอกสาร ทำหน้าที่ในการอัปโหลดภาพเอกสารรับรองผลการศึกษา
- 2) ส่วนอ่านภาพเอกสาร ทำหน้าที่แปลงภาพเอกสารรับรองผลการศึกษาเป็นข้อมูลตัวอักษร พร้อมทำการแก้ไขข้อมูลดังกล่าว
- 3) ส่วนตรวจสอบเอกสารรับรองผลการศึกษาทำหน้าที่สร้างแฮชจากข้อมูลเอกสารรับรองผล และตรวจสอบ Code ดังกล่าวเพื่อหาว่าข้อมูลเอกสารรับรองผลการศึกษานั้นถูกต้องหรือไม่
- 4) ส่วนแสดงสถานะและขั้นตอนการทำงาน
- 5) ส่วนย้อนกลับขั้นตอนการทำงานทำหน้าที่ย้อนกลับไปยังขั้นตอนการทำงานที่ผ่านมา

ส่วนติดต่อผู้ใช้งาน

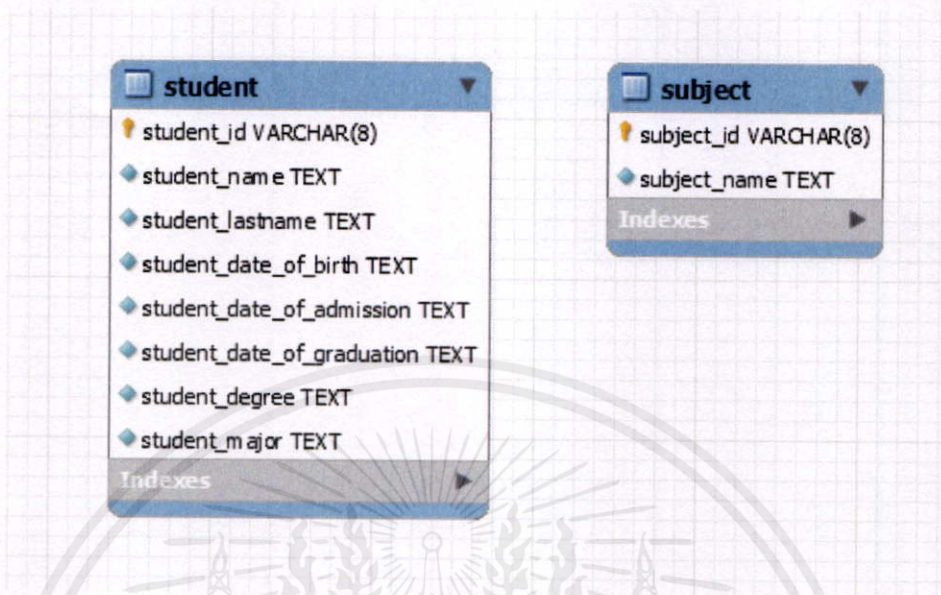


รูป 3.13 ส่วนติดต่อผู้ใช้งาน

- หมายเลข 1 ปุ่มเพิ่มรูปภาพเอกสารรับรองผลการศึกษา
- หมายเลข 2 ปุ่มตรวจสอบเอกสารรับรองผลการศึกษา
- หมายเลข 3 ส่วนแสดงขั้นตอนการทำงาน
- หมายเลข 4 ปุ่มดาวน์โหลดคู่มือการใช้งาน
- หมายเลข 5 ส่วนแสดงสถานะการทำงาน
- หมายเลข 6 ปุ่ม Zoom In/Out สำหรับรูปเอกสารรับรองผลการศึกษา
- หมายเลข 7 ส่วนแสดงรูปเอกสารรับรองผลการศึกษา
- หมายเลข 8 ส่วนแสดงและแก้ไขข้อมูลเอกสารรับรองผลการศึกษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ส่วนฐานข้อมูล



รูป 3.14 ฐานข้อมูลของระบบ

ฐานข้อมูลประกอบไปด้วย 2 ตาราง คือ

ตาราง student สำหรับเก็บข้อมูลส่วนบุคคลของนักศึกษา ประกอบด้วย

- student_id รหัสนักศึกษา
- student_name ชื่อ
- student_lastname นามสกุล
- student_date_of_birth วันเกิด
- student_date_of_admission ปีที่เข้าศึกษา
- student_date_of_graduation ปีที่จบการศึกษา
- student_degree ระดับปริญญาและคณะ
- student_major วิชาเอก

ตาราง subject สำหรับเก็บข้อมูลรายวิชาต่างๆ ประกอบด้วย

- subject_id รหัสวิชา
- subject_name ชื่อวิชา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

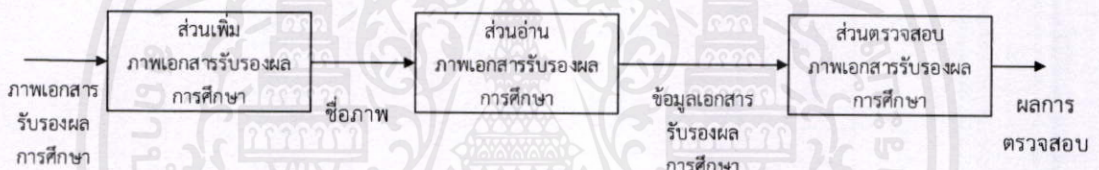
การทดลองและผลการทดลอง

ในบทนี้จะอธิบายถึงการทดลองเพื่อทดสอบประสิทธิภาพและขอบเขตการทำงานของระบบ โดยได้อธิบายถึงเครื่องมือและอุปกรณ์ที่ใช้ในการทดลอง สภาพแวดล้อมการทำงานของระบบระหว่างทำการทดลอง รวมถึงที่มาของแต่ละการทดลอง

ระบบตรวจสอบเอกสารรับรองผลการศึกษาระหว่างการทำงานหลักออกเป็น 3 ส่วนได้แก่

- 1) ส่วนเพิ่มเอกสารรับรองผลการศึกษา
- 2) ส่วนอ่านภาพเอกสารรับรองผลการศึกษา
- 3) ส่วนตรวจภาพส่วนเพิ่มเอกสารรับรองผลการศึกษา

ผังรูปที่ 4.1



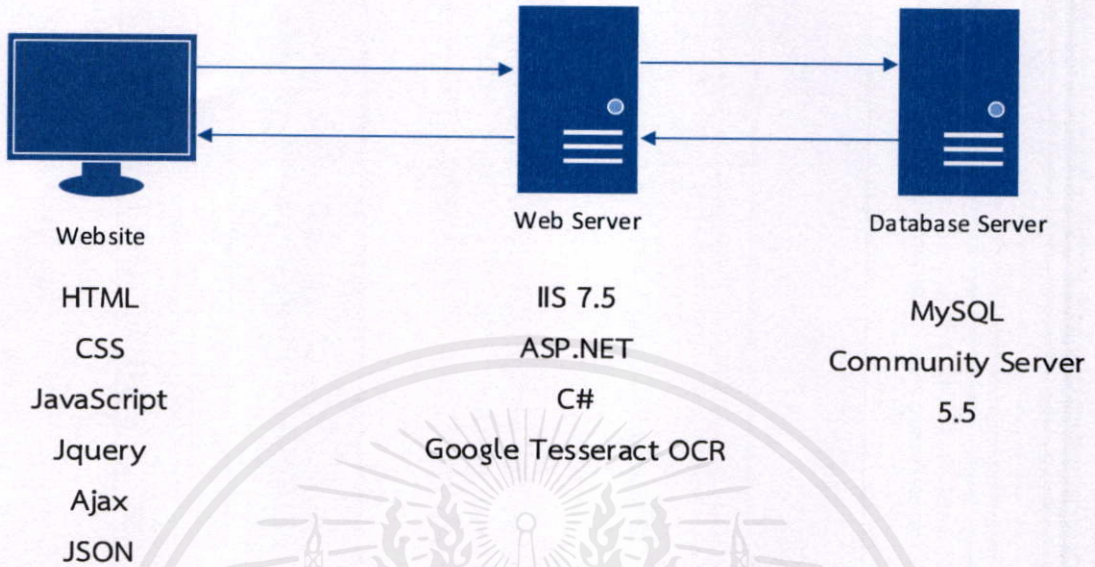
รูป 4.1 การทำงานหลักของระบบ

เพื่อทดสอบความถูกต้องและประสิทธิภาพในการทำงานจึงได้ทำการทดลองดังต่อไปนี้ โดยในแต่ละการทดลองจะใช้เครื่องมือและอุปกรณ์ดังต่อไปนี้เป็นหลัก ได้แก่

- 1) Notebook
 - CPU: Intel core i7-4710HQ 2.50 GHz
 - RAM: 4 GB
 - ระบบปฏิบัติการ: Windows 7 Professional 64bit
- 2) เครื่องสแกนจากเครื่องพิมพ์ EPSON รุ่น L210L
- 3) Internet Information Services (IIS) 7.5
- 4) MySQL Community Server 5.5
- 5) Chrome Web Browser 41

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยเครื่องมือและอุปกรณ์ดังกล่าวจะทำงานบนสภาพแวดล้อมการทำงาน ดังรูปต่อไปนี้



รูป 4.2 โครงสร้างระบบ

เนื่องจากไฟล์เอกสารรับรองผลการศึกษานำมาใช้งาน ได้มาจากการสแกนภาพที่สามารถกำหนดประเภทไฟล์ได้หลายประเภท เพื่อหาประเภทไฟล์ที่ระบบรองรับ จึงทำการทดลองดังต่อไปนี้

4.1 การทดลองนำเอกสารรับรองผลการศึกษาเข้าโปรแกรม OCR เพื่อหาประเภทไฟล์ที่รองรับ

4.1.1 จุดประสงค์การทดลอง

เพื่อหาประเภทไฟล์ที่ระบบรองรับ

4.1.2 เครื่องมือและอุปกรณ์ที่ใช้ในการทดลอง

ไฟล์เอกสารรับรองผลการศึกษาที่ได้จากการสแกน ประเภท JPEG, Bitmap, TIFF ,PDF

4.1.3 วิธีการทดลอง

- 1) นำไฟล์เอกสารรับรองผลการศึกษาเข้าสู่ระบบ
- 2) ตรวจสอบการทำงานของกระบวนการ OCR ว่าสามารถแปลงภาพเป็นอักษรได้หรือไม่
- 3) ทำซ้ำขั้นตอน 1) และ 2) โดยเปลี่ยนประเภทไฟล์ไปตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.4 ผลการทดลอง

หลังจากทำการทดลองไฟล์แต่ละประเภทแล้ว ผลการทดลองเป็นดังตาราง 4.1

ตาราง 4.1 ผลการทำกระบวนการ OCR ของไฟล์แต่ละประเภท

ประเภทไฟล์	ผลการทำกระบวนการ OCR
JPEG	สามารถแปลงภาพเป็นอักษรได้
Bitmap	สามารถแปลงภาพเป็นอักษรได้
TIFF	สามารถแปลงภาพเป็นอักษรได้
PDF	ไม่สามารถแปลงภาพเป็นอักษรได้

4.1.5 สรุปผลการทดลอง

จากผลการทดลองพบว่าประเภทไฟล์ที่ระบบรองรับได้แก่ JPEG, Bitmap, TIFF ส่วนไฟล์ประเภท PDF ระบบไม่ โดยไฟล์ประเภท JPEG เป็นประเภทที่เหมาะสม เนื่องจากใช้พื้นที่ในการเก็บข้อมูลน้อยที่สุด

เนื่องจากไฟล์ภาพเอกสารรับรองผลการศึกษาที่ใช้ในส่วนเพิ่มเอกสารรับรองผลการศึกษาได้มาจากการสแกนภาพ ที่สามารถกำหนดได้หลายความละเอียด เพื่อหาความละเอียดที่เหมาะสมที่สุดของไฟล์ภาพดังกล่าว จึงทำการทดลองดังต่อไปนี้

4.2 การทดลองนำเอกสารรับรองผลการศึกษาเข้าโปรแกรม OCR เพื่อหาค่า DPI ที่เหมาะสม

4.2.1 จุดประสงค์การทดลอง

เพื่อหาค่า DPI ที่เหมาะสมที่สุดของไฟล์ภาพเอกสารรับรองผลการศึกษาสำหรับนำเข้าระบบ

4.2.2 เครื่องมือและอุปกรณ์ที่ใช้ในการทดลอง

ไฟล์ภาพเอกสารรับรองผลการศึกษา ประเภท JPEG ที่ได้จากการสแกนความละเอียด 150,200,240,300,360,400,600,720,800,1200 dpi

4.2.3 วิธีการทดลอง

เอกสารนี้เป็นเอกสาร 1) นำเข้าไฟล์รูป จากความละเอียดสูงสุดไปต่ำสุด ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น 2) วัดความถูกต้องของข้อมูลและเวลาในการประมวลผล จำนวน 3 ครั้ง

โดยในการวัดความถูกต้องของข้อมูล ข้อมูลของเอกสารรับรองผลการศึกษา
แบ่งออกเป็น

- ส่วน ชื่อและนามสกุล , ส่วน รหัสนักศึกษา, ส่วน วันเกิด
- ส่วน วันเข้าศึกษา , ส่วน วันจบการศึกษา
- ส่วน คณะ , ส่วน ภาควิชา
- ส่วน รหัสวิชา และชื่อวิชา , ส่วน หน่วยกิต , ส่วน เกรด
- ส่วน GPS , ส่วน GPA

หลังจากนั้นนับส่วนที่ถูกต้องว่าคิดเป็นเปอร์เซ็นต์เท่าใดของทั้งหมด

3) คำนวณค่าเฉลี่ยของ ค่าความถูกต้องและเวลาที่ใช้ในการประมวลผล ของแต่ละความ
ละเอียด

4) คำนวณค่าค่าความถูกต้อง ต่อ เวลาที่ใช้ทั้งหมด ของแต่ละความละเอียด

4.2.4 ผลการทดลอง

หลังจากทำการทดลองเอกสารรับรองผลการศึกษาทุกความละเอียดแล้ว ผลการทดลองเป็น
ดังตาราง 4.2

ตาราง 4.2 ค่าความถูกต้องในการตรวจสอบและเวลาที่ใช้ในการประมวลผล

ความ ละเอียด (dpi)	การตรวจสอบ							
	ครั้งที่ 1		ครั้งที่ 2		ครั้งที่ 3		ค่าเฉลี่ย	
	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)
1200	-	-	-	-	-	-	-	-
800	-	-	-	-	-	-	-	-
720	-	-	-	-	-	-	-	-
600	-	-	-	-	-	-	-	-
400	98.95	23	98.95	23	98.95	23	98.95	23
360	98.95	22	98.95	22	98.95	22	98.95	22
300	95.79	21	95.79	21	95.79	21	95.79	21
240	93.16	21	93.16	20	93.16	21	93.16	20.67
200	93.16	23	93.16	23	93.16	23	93.16	23

ความ ละเอียด (dpi)	การตรวจสอบ							
	ครั้งที่ 1		ครั้งที่ 2		ครั้งที่ 3		ค่าเฉลี่ย	
	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)
150	80.53	22	80.53	21	80.53	21	80.53	21.33
96	7.90	20	7.90	20	7.90	20	7.90	20
72	0.53	6	0.53	6	0.53	6	0.53	6
50	0%	2	0%	2	0%	2	0%	2

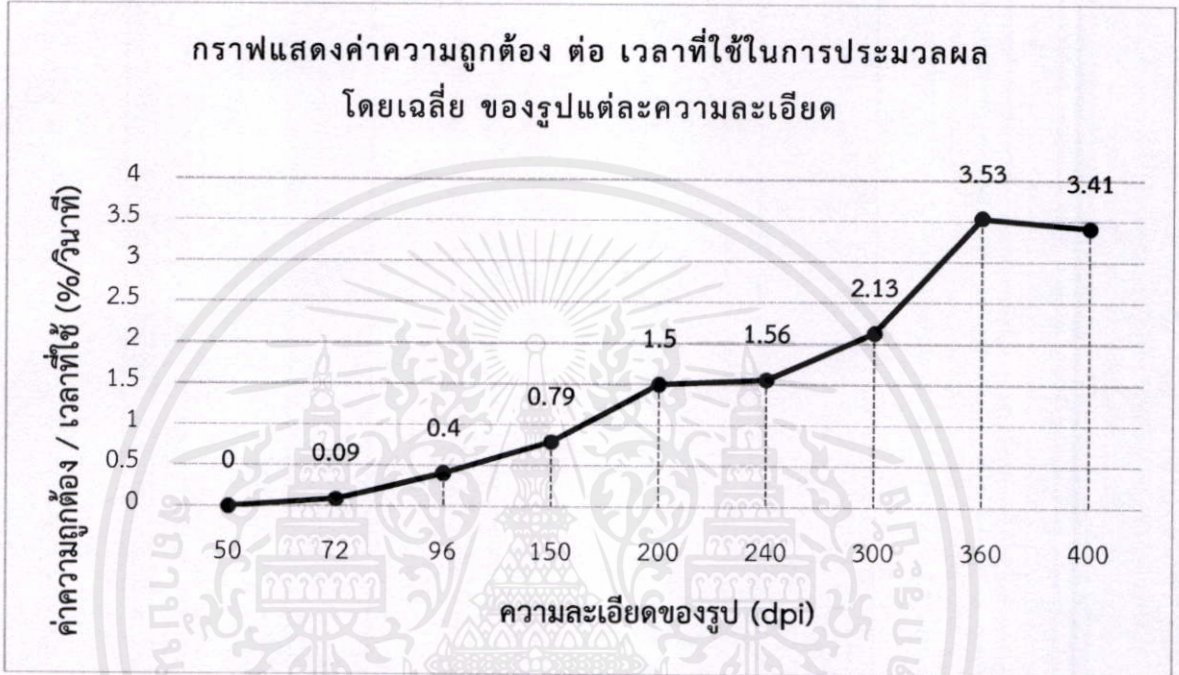
- หมายเหตุ ค่า “-” แทนค่าผลการทดลองที่ไม่สามารถบันทึกผลได้

ตาราง 4.3 ค่าความถูกต้อง ต่อ เวลาทั้งหมดที่ใช้ในการตรวจสอบ

ความ ละเอียด (dpi)	จำนวน ข้อมูลที่ผิด	เวลาที่ใช้ในการแก้ไข ข้อมูลที่ผิดโดยผู้ใช้ (วินาที)	เวลาที่ใช้ ทั้งหมด (วินาที)	ค่าความถูกต้อง ต่อ เวลาที่ใช้ ทั้งหมด (เปอร์เซ็นต์/วินาที)
1200	-	-	-	-
800	-	-	-	-
720	-	-	-	-
600	-	-	-	-
400	2	6	29	3.41
360	2	6	28	3.53
300	8	24	45	2.13
240	13	39	59.67	1.56
200	13	39	62	1.5
150	37	81	102.33	0.79
96	175	-	20	0.40
72	189	-	6	0.09
50	190	-	2	0.00

เอกสารนี้เป็นทรัพย์สินของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี หากมีการนำเอกสารนี้ไปใช้โดยไม่ได้รับอนุญาตให้ถือว่าผิดลิขสิทธิ์และต้องรับผิดชอบต่อเอกสารทุกครั้งที่มีการนำไปใช้

- เวลาที่ใช้ในการแก้ไขข้อมูลที่ผิดโดยผู้ใช้ = จำนวนข้อมูลที่ผิด คูณ 3 วินาที
- เวลาที่ใช้ทั้งหมด = เวลาที่ระบบใช้ในการประมวลผล+เวลาที่ใช้ในการแก้ไขข้อมูลที่ผิด โดยผู้ใช้
- หมายถึง ค่า “-” แทนค่าผลการทดลองที่ไม่สามารถบันทึกผลได้



รูป 4.3 ผลการทดลองแสดงค่าความถูกต้อง ต่อ เวลาที่ใช้ในการประมวลผลโดยเฉลี่ยของรูปแต่ละความละเอียด

4.2.5 สรุปผลการทดลอง

จากการทดลองพบว่าระบบรองรับความละเอียดของรูปในช่วง 50-400 dpi โดยเมื่อลดความละเอียดรูปลง

- ค่าความถูกต้องของข้อมูลมีแนวโน้มลดลง โดยที่ความละเอียด 96 dpi ลงไปข้อมูลของเอกสารรับรองผลการศึกษาก็จะเริ่มขาดหายไป ทำให้ไม่สามารถวัดเวลาที่ใช้ในการแก้ไขข้อมูลที่ผิดโดยผู้ใช้ได้ จนกระทั่งความละเอียด 50dpi เปอร์เซ็นต์ความถูกต้องจะเป็น 0

- ในส่วนของเวลาที่ใช้ในการประมวลผลมีแนวโน้มลดลงเช่น เมื่อเพิ่มความละเอียดรูปขึ้น

- ค่าความถูกต้องของข้อมูลมีแนวโน้มเพิ่มขึ้น จนกระทั่งความละเอียด360dpi ขึ้นไป ค่าความถูกต้องที่ได้เท่ากับ 99% ซึ่งเป็นค่าที่สูงสุด

- เวลาที่ใช้ในการประมวลผลมีแนวโน้มเพิ่มขึ้นเช่นกัน

เอกสารนี้เป็นเอกสารที่มอบให้สำหรับการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ในการค้า
 ไม่จําการถือใบอนุญาตอื่นใดที่ให้อัปโหลดข้อมูลขึ้นบนระบบและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หากพิจารณาค่าความถูกต้อง ต่อ เวลาที่ใช้ทั้งหมด พบว่าที่ความละเอียด 360dpi ได้ค่าที่ดีที่สุด ความละเอียดนี้จึงเหมาะสมที่สุด

เนื่องจากกระบวนการทำ OCR มีข้อผิดพลาดในการทำงาน จึงได้พัฒนาส่วนการปรับปรุงข้อมูล OCR โดยระบบ ซึ่งทำการปรับปรุงจากการค้นหาข้อมูลที่ใกล้เคียงในฐานข้อมูล เช่นเดียวกับที่ปรากฏในบท 3 หัวข้อ 3.1 เพื่อทดสอบประสิทธิภาพในการทำงานของส่วนดังกล่าวและประสิทธิภาพในการทำงานของระบบ ดังการทดลองต่อไปนี้

4.3 การทดสอบประสิทธิภาพ เมื่อใช้งานส่วนการปรับปรุงข้อมูล OCR โดยระบบ

4.3.1 จุดประสงค์การทดลอง

เพื่อหาค่าความถูกต้องและเวลาที่ใช้โดยเฉลี่ยในการตรวจสอบเอกสารรับรองผลการศึกษานี้ในกรณีที่มีการเปิดและปิดการใช้งานส่วนการปรับปรุงข้อมูล OCR โดยระบบ

4.3.2 เครื่องมือและอุปกรณ์ที่ใช้ในการทดลอง

ไฟล์ภาพเอกสารรับรองผลการศึกษาระเบท JPEG ที่ได้จากการสแกน ความละเอียด 360 dpi จากการสุ่มเลือกจากนักศึกษา สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง คณะวิศวกรรมศาสตร์ สาขาวิชา วิศวกรรมคอมพิวเตอร์ จำนวน 10 ฉบับ

4.3.3 วิธีการทดลอง

- 1) ปิดการใช้งานส่วนการปรับปรุงข้อมูล OCR โดยระบบ
- 2) นำไฟล์ภาพเอกสารรับรองผลการศึกษามาตรวจสอบ รูปละ 3 ครั้ง แล้วเก็บค่าความและเวลาที่ใช้ในการประมวลผล
- 3) เปิดการใช้งานส่วนการปรับปรุงข้อมูล OCR โดยระบบ จากนั้นทำซ้ำ ขั้นตอนที่ 1) และ 2) ตามลำดับ
- 4) คำนวณค่าเฉลี่ยของ ค่าความถูกต้องและเวลาที่ใช้ในการประมวลผลของทั้งสองกรณี

4.3.4 ผลการทดลอง

หลังจากปิดการใช้งานส่วนการปรับปรุงข้อมูล OCR โดยระบบ ผลการทดลองเป็นดังตารางต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 4.4 ค่าความถูกต้องในการตรวจสอบและเวลาที่ใช้ในการประมวลผล กรณีปิดการใช้งานส่วน
การปรับปรุงข้อมูล OCR โดยระบบ

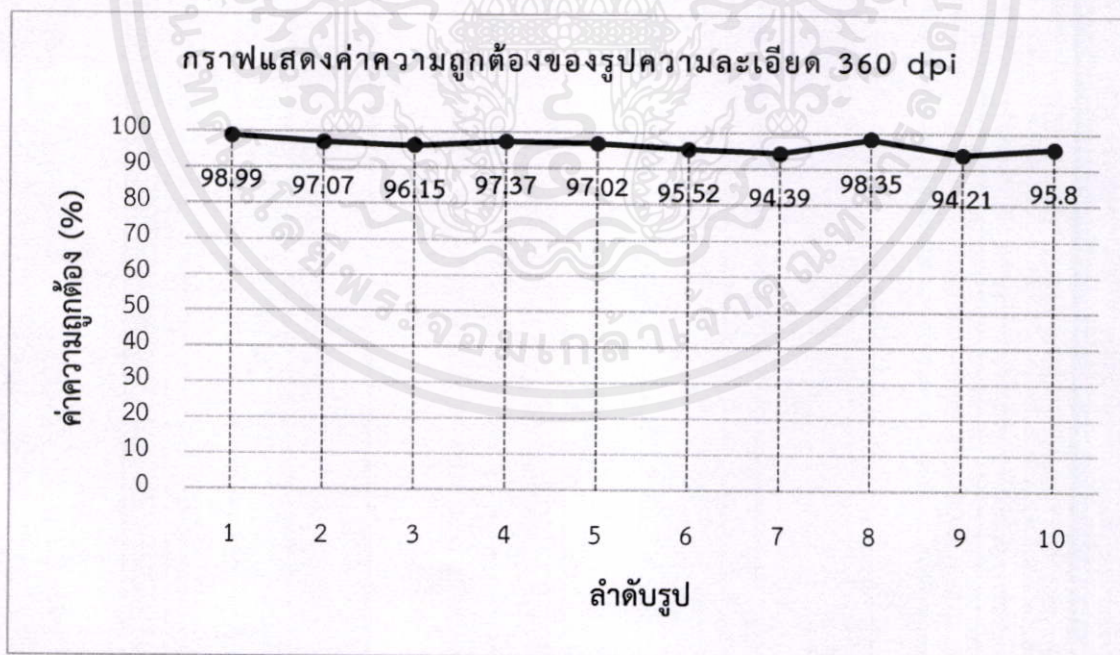
ลำดับ เอกสาร รับรอง ผล การศึกษา	การตรวจสอบ							
	ครั้งที่ 1		ครั้งที่ 2		ครั้งที่ 3		ค่าเฉลี่ย	
	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)
1	98.99	21	98.99	21	98.99	22	98.99	21.33
2	97.07	22	97.07	22	97.07	22	97.07	22
3	96.15	23	96.15	22	96.15	22	96.15	22.33
4	97.37	21	97.37	21	97.37	21	97.37	21
5	97.02	22	97.02	22	97.02	22	97.02	22
6	95.52	22	95.52	22	95.52	22	95.52	22
7	94.39	22	94.39	21	94.39	21	94.39	21.33
8	96.35	21	96.35	21	96.35	21	96.35	21
9	94.21	21	94.21	21	94.21	21	94.21	21
10	95.80	21	95.80	21	95.80	21	95.80	21
ค่าเฉลี่ยรวม							96.29	21.50

หลังจากเปิดใช้งานส่วนการปรับปรุงข้อมูล OCR โดยระบบ ผลการทดลองเป็นดังต่อไปนี้

ตาราง 4.5 ค่าความถูกต้องในการตรวจสอบและเวลาที่ใช้ในการประมวลผล กรณีที่เปิดการใช้งาน
ส่วนการปรับปรุงข้อมูล OCR โดยระบบ

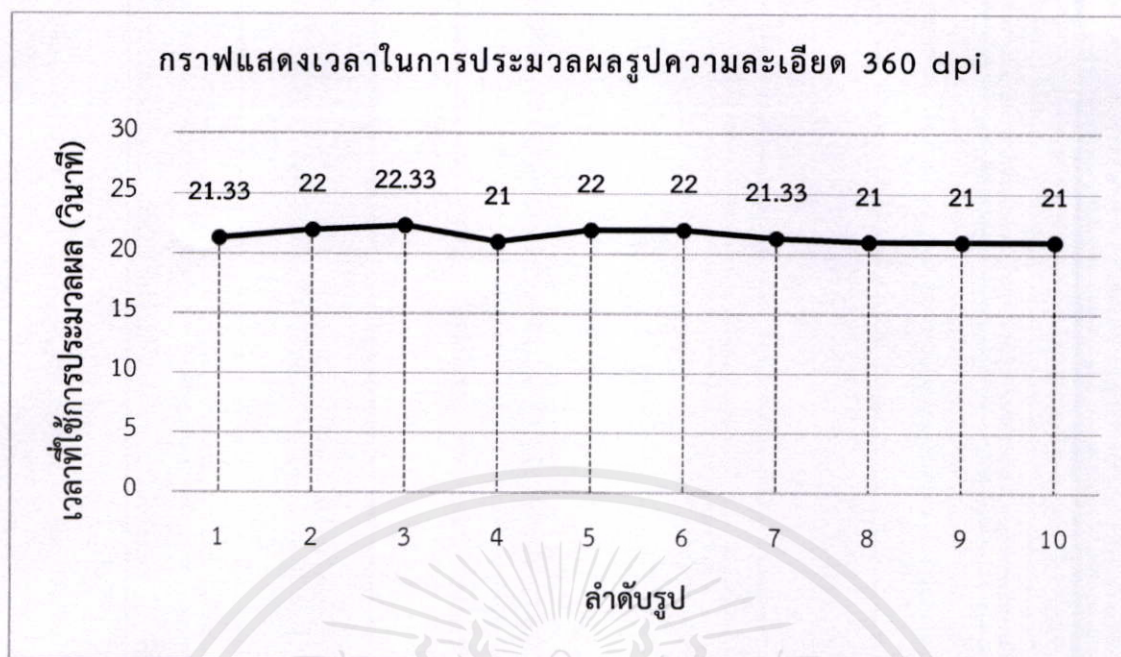
ลำดับ เอกสาร รับรอง ผล การศึกษา	การตรวจสอบ							
	ครั้งที่ 1		ครั้งที่ 2		ครั้งที่ 3		ค่าเฉลี่ย	
	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)
1	99.50	22	99.5	22	99.5	22	99.5	22
2	98.54	23	98.54	23	98.54	23	98.54	23

ลำดับ เอกสาร รับรอง ผล การศึกษา	การตรวจสอบ							
	ครั้งที่ 1		ครั้งที่ 2		ครั้งที่ 3		ค่าเฉลี่ย	
	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)	ความ ถูกต้อง (เปอร์เซ็นต์)	เวลาที่ ใช้ (วินาที)
3	98.08	23	98.08	23	98.08	23	98.08	23
4	97.90	21	97.90	21	97.90	21	97.90	21
5	98.51	23	98.51	22	98.51	22	98.51	22.33
6	96.52	23	96.52	22	96.52	22	96.52	22.33
7	96.94	22	96.94	21	96.94	21	96.94	21.33
8	96.88	22	96.88	21	96.88	21	96.88	21.33
9	96.84	21	96.84	21	96.84	21	96.84	21
10	98.42	21	98.42	21	98.42	21	98.42	21
ค่าเฉลี่ยรวม							97.81	21.83



รูป 4.4 ผลการทดลองแสดงค่าความถูกต้องของรูปความละเอียด 360 dpi

เอกสารนี้เป็นเอกสารสงวนไว้สำหรับการใช้งานเพื่อการศึกษาดูงาน มิใช่เพื่อเผยแพร่เป็นประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 4.5 ผลการทดลองแสดงเวลาในการประมวลผลของรูปความละเอียด 360 dpi

4.3.5 สรุปผลการทดลอง

จากการทดลองพบว่าเมื่อมีการเปิดฟังก์ชันการปรับปรุงข้อมูล OCR โดยระบบ ค่าความถูกต้องโดยเฉลี่ยจะเพิ่มขึ้น 1.52% คือ 97.81% ซึ่งต้องอาศัยการแก้ไขข้อมูลจากผู้ใช้งานและเวลาที่ใช้ในการประมวลผลโดยเฉลี่ยเพิ่มขึ้น 0.33 วินาที คือ 21.50 วินาที ซึ่งเป็นเวลาที่ยอมรับได้เมื่อเทียบกับวิธีการตรวจสอบเดิมที่ใช้เวลาประมาณ 2 สัปดาห์ เพื่อตรวจสอบความถูกต้องในแต่ละส่วนของการทำงาน จึงทำการทดลองดังต่อไปนี้

4.4 การทดลองความถูกต้องการทำงานของระบบ

4.4.1 จุดประสงค์การทดลอง

เพื่อตรวจสอบความถูกต้องในการทำงานของระบบ ในแต่ละส่วน ได้แก่

- 1) ส่วนนำเข้าภาพเอกสารรับรองผลการศึกษา
- 2) ส่วนอ่านเอกสารรับรองผลการศึกษา
- 3) ส่วนตรวจสอบเอกสารรับรองผลการศึกษา

4.4.2 เครื่องมือและอุปกรณ์ที่ใช้ในการทดลอง

ไฟล์ภาพเอกสารรับรองผลการศึกษา ประเภท JPEG ที่ได้จากการสแกน ความละเอียด 360 dpi จากการสุ่มเลือกจากนักศึกษา สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง คณะวิศวกรรมศาสตร์ สาขาวิชา วิศวกรรมคอมพิวเตอร์ จำนวน 10 ฉบับ

4.4.3 วิธีการทดลอง

- 1) นำเข้าภาพเอกสารรับรองผลการศึกษา
- 2) หากภาพเอกสารรับรองผลการศึกษาแสดงในส่วนแสดงภาพเอกสารรับรองผลการศึกษา ให้บันทึกผลว่า ผ่าน ถ้าไม่ ให้บันทึกผลว่า ไม่ผ่าน
- 3) หากขั้นตอนที่ 2) ผ่าน ให้ทำขั้นตอนต่อไป ถ้าหากไม่ผ่านให้หยุดทำการทดลอง และบันทึกผลส่วนการทำงานที่เหลือด้วยค่า “-”
- 4) กดปุ่มตรวจสอบเอกสารรับรองผลการศึกษา หากแสดงข้อมูลเอกสารรับรองผลการศึกษา ให้บันทึกผลว่า ผ่าน ถ้าไม่ ให้บันทึกว่าไม่ผ่าน
- 5) ทำซ้ำขั้นตอน 3) แต่เปลี่ยนขั้นตอนที่อ้างอิงเป็นขั้นตอนที่ 4)
- 6) ทำการกดปุ่ม Next เพื่อแสดงผลลัพธ์การตรวจสอบ ถ้าผลลัพธ์แสดงบนแถบสถานะให้บันทึกผลว่า ผ่าน ถ้าไม่ ให้บันทึกว่า ไม่ผ่าน
- 7) ทำซ้ำขั้นตอน 1)-6) โดยเปลี่ยนเป็นภาพเอกสารรับรองผลการศึกษาอื่นจนครบ 10 ภาพ

4.4.4 ผลการทดลอง

หลังจากทำการทดลองแต่ละภาพเอกสารรับรองผลการศึกษาทั้งหมดแล้ว ผลการทดลองเป็นดังต่อไปนี้

ตาราง 4.6 ความถูกต้องในการทำงานแต่ละส่วนของระบบ

ลำดับไฟล์ภาพเอกสารรับรองผลการศึกษา	ส่วนนำเข้าภาพเอกสารรับรองผลการศึกษา	ส่วนอ่านเอกสารรับรองผลการศึกษา	ส่วนตรวจสอบเอกสารรับรองผลการศึกษา
1	ผ่าน	ผ่าน	ผ่าน
2	ผ่าน	ผ่าน	ผ่าน
3	ผ่าน	ผ่าน	ผ่าน
4	ผ่าน	ผ่าน	ผ่าน
5	ผ่าน	ผ่าน	ผ่าน
6	ผ่าน	ผ่าน	ผ่าน
7	ผ่าน	ผ่าน	ผ่าน
8	ผ่าน	ผ่าน	ผ่าน
9	ผ่าน	ผ่าน	ผ่าน
10	ผ่าน	ผ่าน	ผ่าน

4.4.5 สรุปผลการทดลอง

จากผลการทดลองพบว่าระบบสามารถทำงานได้อย่างถูกต้องทุกส่วนการทำงานสำหรับทุกภาพเอกสารรับรองผลการศึกษา



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

บทสรุปและข้อเสนอแนะ

ในบทนี้จะเริ่มอธิบายจากส่วนของบทสรุปที่บอกถึงจุดประสงค์ของการพัฒนา หลักการทำงานของระบบ ประสิทธิภาพและขอบเขตของการทำงาน ในส่วนของปัญหาอุปสรรคและแนวทางการแก้ไขได้อธิบายถึงปัญหาที่พบในปัจจุบัน รวมทั้งแนวทางในการแก้ไข และในส่วนของแนวทางพัฒนาต่อได้อธิบายถึงแนวทางในการพัฒนาเพื่อเพิ่มความสามารถระบบ

5.1 บทสรุป

โครงการเรื่องระบบตรวจสอบการปลอมแปลงเอกสารรับรองผลการศึกษา เป็นเว็บไซต์สำหรับช่วยในการตรวจสอบการปลอมแปลงเอกสารรับรองผลการศึกษาที่เป็นภาษาอังกฤษของสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง โดยประยุกต์ใช้กระบวนการ OCR หรือการแปลงรูปภาพเป็นอักษรช่วยในการแปลงข้อมูลเอกสารรับรองผลการศึกษาที่อยู่ในรูปของไฟล์ภาพสแกน ให้อยู่ในรูปแบบตัวอักษร ในกรณีที่มีความผิดพลาดในการแปลงข้อมูลระบบจะมีการแก้ไขข้อมูลโดยอัตโนมัติ หากยังพบข้อมูลที่มีความผิดพลาดจะให้ผู้ใช้งานทำการแก้ไขในส่วนที่เหลือ จากนั้นจึงนำข้อมูลมาทำการเข้ารหัสเป็นชุดรหัสและใช้ชุดรหัสนี้ในการตรวจสอบความถูกต้องของข้อมูลกับชุดรหัสที่ถูกสร้างขึ้นมาจากฐานข้อมูลก่อนหน้านี้ รูปแบบของไฟล์ภาพสแกนที่นำมาใช้ในการตรวจสอบ คือไฟล์ภาพประเภท JPEG โดยที่ความละเอียด 360dpi เป็นความละเอียดที่เหมาะสมที่สุด ซึ่งให้ค่าความถูกต้องในการแปลงรูปภาพเป็นข้อความโดยเฉลี่ยที่ 97.81% โดยใช้เวลาในการประมวลผลโดยเฉลี่ยอยู่ที่ 21.83 วินาที ระบบนี้เหมาะสำหรับบุคคลที่หน้าที่ตรวจสอบเอกสารรับรองผลการศึกษาในสำนักทะเบียนของทางสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ซึ่งช่วยลดเวลาและความผิดพลาดในการตรวจสอบของการตรวจสอบเอกสารรับรองผลการศึกษาในปัจจุบัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 ปัญหาอุปสรรค และแนวทางการแก้ไข

5.2.1 ความละเอียดของไฟล์ภาพ

ปัญหา

ระบบตรวจสอบการปลอมแปลงเอกสารรับรองผลการศึกษาฯ รองรับไฟล์ภาพประเภท JPEG ความละเอียดตั้งแต่ 50-400 dpi โดยความละเอียดที่เหมาะสมที่สุดคือ 360dpi เนื่องจากไฟล์ภาพความละเอียดนี้มีขนาดเฉลี่ยอยู่ที่ 1.3 MB ซึ่งมีขนาดใหญ่ ทำให้เกิดปัญหาในการใช้พื้นที่หากมีจำนวนมาก

แนวทางแก้ไข

เปลี่ยนวิธีเก็บไฟล์ภาพจากวิธีการอัป โหลดไปเก็บบนเว็บเซิร์ฟเวอร์เป็นการเก็บไฟล์ภาพให้อยู่ในรูปแบบของ Base64 String แล้วใช้ String ชุดนั้นชั่วคราวขณะทำการตรวจสอบเอกสารรับรองผลการศึกษา เมื่อจบการทำงานระบบจะทำการลบค่าออก

5.2.2 การทำการแก้ไขโดยอัตโนมัติ

ปัญหา

ในส่วนการทำการแก้ไขโดยอัตโนมัติยังพบปัญหาในการแก้ไขเกรดแต่ละรายวิชา เช่น เกรด D+ หลังผ่านการทำกระบวนการ OCR จะได้ผลลัพธ์เป็น Db โดยไม่ได้รับการแก้ไข

แนวทางการแก้ไข

ระหว่างขั้นตอนการทำแก้ไขโดยอัตโนมัติให้ทำการตรวจ เกรด ที่ได้จากการทำกระบวนการ OCR หากพบว่าข้อมูลที่ได้มีข้อผิดพลาด และมีจำนวน 2 ตัวอักษร ให้ทำการแทนตัวอักษรสุดท้าย ด้วยอักษร “+”

5.3 แนวทางการพัฒนาต่อ

5.3.1 ส่วนการแก้ไขข้อมูล

ในการแก้ไขข้อมูลที่ได้จากกระบวนการทำ OCR โดยผู้ใช้งานแต่ละครั้ง ซึ่งเป็นข้อมูลที่ระบบไม่สามารถทำการแก้ไขได้จากวิธีการค้นหาข้อมูลที่ใกล้เคียงในฐานข้อมูล ส่วนใหญ่ข้อมูลที่แก้ไขจะเป็นข้อมูลเดิม เช่น ชื่อวิชา ดังรูปที่ 5.1

01006003 ENGINEERING MATHEMATICS 3

01005003 ENGINEERING MATHEMATICS 3

เอกสารนี้

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามนำข้อมูลไปเผยแพร่และต้องแจ้งปริมาณการนำข้อมูลไปใช้

รูป 5.1 ตัวอย่างรูปแบบในการแก้ไขโดยการแบบพิมพ์แก้ไข

จะมีการแก้ไขทุกครั้งที่ทำกรตรวจสอบเอกสารรับรองผลการศีกษา ด้วยเหตุนี้จึง
แก้ปัญหาโดยให้ระบบมีการเรียนรู้การแก้ไขจากผู้ใช้งาน หากการแก้ไขถูกต้องจะทำการเก็บข้อมูล
ไว้ในฐานข้อมูลที่มีการเก็บข้อมูลที่ผิดและข้อมูลที่ถูกต้องไว้ เช่นดังตาราง 5.1

ตาราง 5.1 รูปแบบตารางสำหรับเก็บข้อมูลการแก้ไขจากผู้ใช้

ข้อมูลที่ผิด	ข้อมูลที่ถูก
ENGINEERONGMATHIEMATICS3	ENGINEERONGMATHEMATICS3

จากรูปแบบการทำงาน ระบบจะทำการเช็คข้อมูลอยู่ 2 ครั้ง คือ

- 1) เช็คข้อมูลที่มีอยู่ในฐานข้อมูลหรือไม่
หากไม่พบจะทำการเช็คครั้งที่ 2)
- 2) เช็คข้อมูลใกล้เคียงกับข้อมูลที่มีอยู่ในฐานข้อมูลหรือไม่
โดยกระบวนการแก้ไขนี้จะทำการก่อนการเช็คข้อมูลครั้งที่ 1)



รูป 5.2 การทำงานหลังจากเพิ่มขั้นตอนการเรียนรู้การแก้ไขข้อมูลจากผู้ใช้

5.3.2 ส่วนการเก็บผลลัพธ์การทำงาน

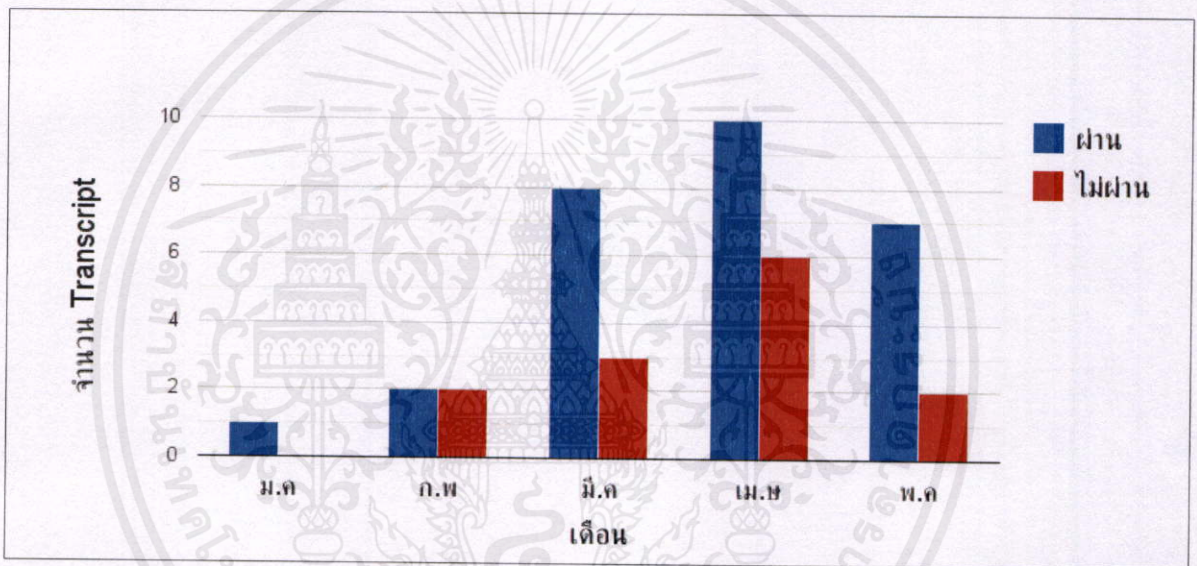
โดยทั่วไประบบจะทำการตรวจสอบเอกสารรับรองผลการศีกษานับต่อฉบับ โดยไม่มีการเก็บข้อมูลผลลัพธ์การทำงานเพื่อตรวจสอบข้อมูลย้อนหลัง จึงออกแบบส่วนการทำงานนี้ขึ้น โดยทุกครั้งที่มีการตรวจสอบข้อมูลจะมีการเก็บรวบรวมข้อมูลไว้ในฐานข้อมูล โดยผู้ใช้งานสามารถดูข้อมูลได้ ดังต่อตารางที่ 5.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 5.2 รูปแบบตารางสำหรับเก็บข้อมูลการใช้งานของผู้ใช้

วัน/ เวลา ที่ตรวจสอบ	ชื่อ	นามสกุล	ภาพเอกสารรับรอง ผลการศึกษา	ผลการตรวจสอบ
1/1/2019 9:00 AM	A	a	click	Pass
2/1/2019 13:00 AM	B	b	click	Not pass
3/1/2019 10:00 AM	C	c	click	Pass

นอกจากนี้สามารถแสดงข้อมูลสรุปรายวัน รายเดือน หรือรายปี เช่นดังรูปที่ 5.3



รูป 5.3 แผนภูมิสรุปข้อมูลผลการตรวจสอบเอกสารรับรองผลการศึกษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.3.3 ส่วนปรับแต่งเพื่อตรวจสอบเอกสารรับรองอื่นๆ

จากการทำงานในส่วนของการอ่านข้อมูลจากเอกสารรับรองผลการศึกษาที่มีการเก็บข้อมูลแต่ละส่วนเป็น Box เพื่อรองรับการตรวจสอบเอกสารรับรองอื่นจึงทำการเพิ่มส่วนกำหนดประเภทข้อมูลของ Box เพื่อใช้ในสร้างรูปแบบการอ่านข้อมูลใหม่และสามารถเก็บการตั้งค่าดังกล่าวเพื่อใช้ในครั้งต่อไป ดังรูปที่ 5.4

The screenshot shows a 'Create New Format' window with the following settings:

ประเภทข้อมูล	BOX ที่
ชื่อ นามสกุล	4
วันเกิด	6
รหัสนักศึกษา	8
ปีเข้าศึกษา	10

Format name: KMITL

save

รูป 5.4 ตัวอย่างการปรับแต่งเพื่อตรวจสอบเอกสารรับรองผลการศึกษาของมหาวิทยาลัยอื่นๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.3.4 ปรับแต่งส่วนเพิ่มเอกสารรับรองผลการศึกษา

หากระบบมีการปรับกระบวนการ OCR และ Auto Correction จนสามารถทำงานอัตโนมัติได้ ระบบสามารถปรับปรุงให้ตรวจสอบเอกสารรับรองผลการศึกษาได้หลายฉบับในคราวเดียวเพื่อประหยัดเวลาในการตรวจสอบ โดยมีการแสดงสถานะและความคืบหน้าการทำงานของแต่ละเอกสาร ดังรูปที่ 5.5

		+	START
C:\Pic\Transcript\Example.jpg	0%	X	
C:\Pic\Transcript\Example2.jpg	0%	X	

รูป 5.5 ส่วนการเพิ่มเอกสารรับรองการผลการศึกษาแบบใหม่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

Ray Smith,2007.“An Overview of the Tesseract OCR Engine” 629-633.In **Seventh International Conference on Document Analysis and Recognition**. Curitiba:IEEE Computer Society.

Ray Smith.2015. **Improving the quality of the output**. [Online]. Available:

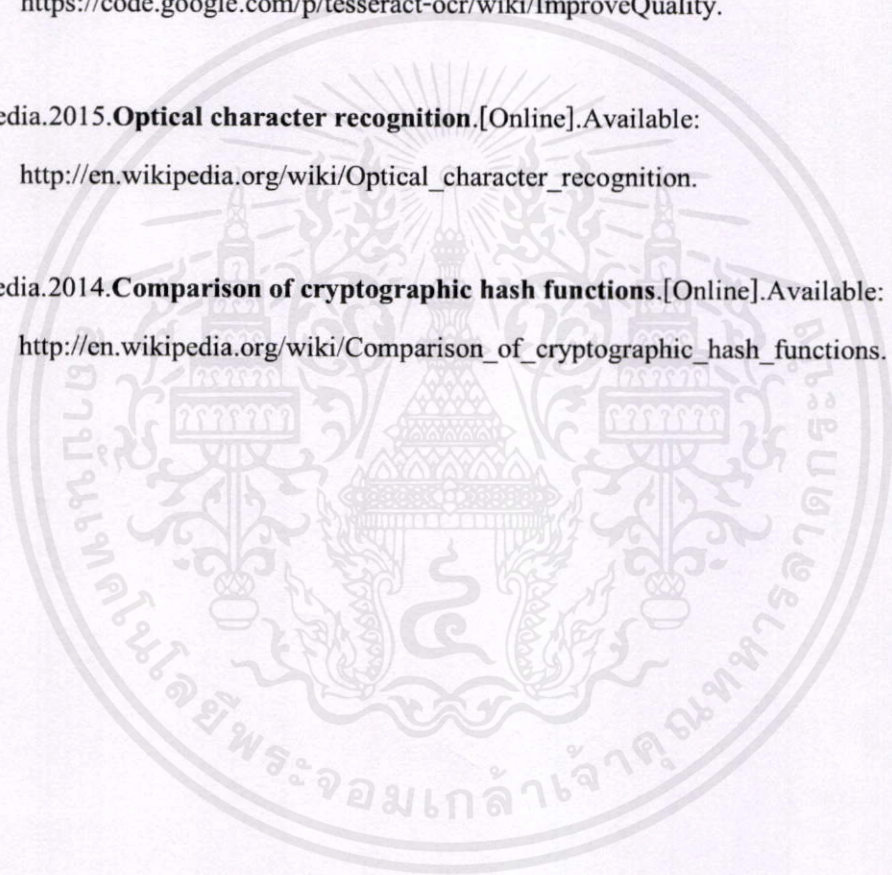
<https://code.google.com/p/tesseract-ocr/wiki/ImproveQuality>.

Wikipedia.2015.**Optical character recognition**. [Online]. Available:

http://en.wikipedia.org/wiki/Optical_character_recognition.

Wikipedia.2014.**Comparison of cryptographic hash functions**. [Online]. Available:

http://en.wikipedia.org/wiki/Comparison_of_cryptographic_hash_functions.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้