

การเปรียบเทียบตัวแบบการจำแนกกลุ่มการยุติการบำบัดรักษาผู้ติดยาเสพติด
แบบระบบสมัครใจในพื้นที่กรุงเทพมหานครโดยใช้เทคนิคการเรียนรู้ของเครื่อง

A COMPARATIVE OF CLASSIFICATION MODELS FOR PREMATURE
TERMINATION IN A VOLUNTARY DRUG REHABILITATION PROGRAM
IN BANGKOK USING MACHINE LEARNING TECHNIQUES



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์
ภาควิชาสถิติ คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2568

KMITL-2025-SC-M-050-028

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

A COMPARATIVE OF CLASSIFICATION MODELS FOR PREMATURE
TERMINATION IN A VOLUNTARY DRUG REHABILITATION PROGRAM
IN BANGKOK USING MACHINE LEARNING TECHNIQUES



A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE
DEGREE OF MASTER OF SCIENCE IN APPLIED STATISTICS
DEPARTMENT OF STATISTICS SCHOOL OF SCIENCE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2025

KMITL-2025-SC-M-050-028

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2025

SCHOOL OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	การเปรียบเทียบตัวแบบการจำแนกกลุ่มการยุติการบำบัดรักษาผู้ติดยาเสพติดแบบระบบสมัครใจในพื้นที่กรุงเทพมหานครโดยใช้เทคนิคการเรียนรู้ของเครื่อง
ชื่อนักศึกษา	นางสาวสุกฤษฎา เพชรชื่น
รหัสประจำตัว	65056091
ปริญญา	วิทยาศาสตรมหาบัณฑิต (สถิติประยุกต์)
ภาควิชา	สถิติ
พ.ศ.	2568
อาจารย์ที่ปรึกษาวิทยานิพนธ์	ผู้ช่วยศาสตราจารย์ ดร.ยุวดี กลุ่มวิเศษ

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของตัวแบบการจำแนก รวมถึงเปรียบเทียบวิธีการคัดเลือกตัวแปรอิสระสำหรับตัวแบบการจำแนกกลุ่มการยุติการบำบัดรักษาผู้ติดยาเสพติดแบบระบบสมัครใจในพื้นที่กรุงเทพมหานคร โดยใช้ข้อมูลผู้เข้ารับการบำบัดรักษาและฟื้นฟูผู้ติดยาเสพติดจากระบบข้อมูลการบำบัดรักษาและฟื้นฟูสมรรถภาพผู้ติดยาเสพติดของประเทศ (บสต.) จากกระทรวงสาธารณสุขในระบบสมัครใจระยะเวลา 6 ปีงบประมาณ ตั้งแต่ ตุลาคม พ.ศ. 2560 - กันยายน พ.ศ. 2566 จำนวน 3,241 คน ในการศึกษาครั้งนี้ผู้วิจัยใช้การเรียนรู้ของเครื่อง 3 เทคนิคในการจำแนก ได้แก่ การถดถอยโลจิสติก (Logistic Regression), เทคนิคป่าสุ่ม (Random Forest) และเทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) นำมาสร้างตัวแบบการจำแนกในรูปแบบที่แตกต่างกันแบ่งเป็นตัวแบบที่ใช้ตัวแปรอิสระครบทุกตัว, ตัวแบบที่ใช้ตัวแปรอิสระไม่ครบทุกตัวจากการคัดเลือกตัวแปร และตัวแบบที่ใช้ตัวแปรอิสระไม่ครบทุกตัวจากการคัดเลือกตัวแปร ร่วมกับการปรับค่าไฮเปอร์พารามิเตอร์ (Hyperparameter Tuning) และ ทำการวัดประสิทธิภาพตัวแบบโดยใช้เกณฑ์คะแนน F1(F1-score) หรือ ค่าเฉลี่ยฮาร์โมนิกของค่าความเที่ยง (Precision) และค่าความครบถ้วน (Recall) จากเมทริกซ์สับสน (Confusion Matrix) ผลการศึกษาพบว่าตัวแบบที่มีประสิทธิภาพที่ดีที่สุด คือ ตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ใช้ตัวแปรอิสระครบทุกตัว โดยมีค่าคะแนน F1-score = 98.85% และ Accuracy = 98.56% โดยปัจจัยที่ส่งผลกระทบต่อกรยุติการบำบัดรักษา 5 อันดับแรก ได้แก่ อาศัยอยู่กับใครในช่วง 30 วันที่ผ่านมา, ยาเสพติดที่ใช้ในปัจจุบัน, อาชีพ, ยาเสพติดที่ใช้ครั้งแรก และอายุ

คำสำคัญ : ตัวแบบการจำแนก, การปรับค่าไฮเปอร์พารามิเตอร์, การวัดประสิทธิภาพตัวแบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีกรนำไปใช้

Thesis Title	A Comparative of Classification Models for Premature Termination in a Voluntary Drug Rehabilitation Program in Bangkok Using Machine Learning Techniques
Student Name	Sukullaya Petchuen
Student ID	65056091
Degree	Master of Science (Applied Statistics)
Department	Statistics
Year	2025
Thesis Advisor	Asst. Prof. Dr. Yuwadee Klomwiset

Abstract

The objective of this research is to compare the performance of classification models and associated feature selection methods to classify the premature termination of voluntary drug treatment in Bangkok. Data was sourced from the National Narcotics Treatment and Rehabilitation Data System, under the Ministry of Public Health. The data covers six fiscal years, from October 2017 to September 2023 and includes 3,241 voluntary participants. In this work, we used three classification methods of machine learning including logistic regression, random forest, and support vector machine. These were used to create classification models in three forms including models with all independent variables, models with feature selection, and models with feature selection and hyperparameter tuning. In addition, we compared the performance of each model using the F1-score derived from the confusion matrix. The results revealed that the Support Vector Machine, utilizing all independent variables, outperformed the other models, achieving an F1-score of 98.85% and an accuracy of 98.56%. The five most significant predictors of premature termination from voluntary drug treatment programs were: who participants lived with (in the past 30 days), currently used drugs, occupation, first drug used and age.

Keywords : Classification Model, Hyperparameter Tuning, Evaluation Model

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลงได้ด้วยดี เนื่องจากได้รับความกรุณาอย่างสูงจากผู้ช่วยศาสตราจารย์ ดร.ยวดี กล่อมวิเศษ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่กรุณาให้คำแนะนำ ปรึกษา ตลอดจนปรับปรุงแก้ไขข้อบกพร่องต่าง ๆ ผู้วิจัยตระหนักถึงความตั้งใจจริงและความทุ่มเทของอาจารย์และขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ ที่นี้

ขอกราบขอบพระคุณ ดร. ชูเกียรติ ผุดพรมราช หัวหน้าแขนงวิชาวิทยาการข้อมูลและการวิเคราะห์สถิติเชิงธุรกิจ มหาวิทยาลัยราชภัฏสวนสุนันทา ที่กรุณาเป็นประธานกรรมการ และผู้ช่วยศาสตราจารย์ ดร. กนกวรรณ ลีโรจนประภา อาจารย์ประจำ ภาควิชาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ที่กรุณาเป็นกรรมการสอบและให้คำชี้แนะ ทำให้วิทยานิพนธ์ฉบับนี้สมบูรณ์ยิ่งขึ้น

ขอขอบคุณ สำนักงานป้องกันและบำบัดการติดยาเสพติด ที่ให้ความอนุเคราะห์ข้อมูลการทำวิทยานิพนธ์ในครั้งนี้ รวมทั้งขอขอบคุณหัวหน้า พี่ และเพื่อน กลุ่มแผนงานและระบบข้อมูลยาเสพติด สำนักงานป้องกันและบำบัดการติดยาเสพติด ที่ให้การสนับสนุน ให้ความช่วยเหลือ พร้อมทั้งให้คำแนะนำเกี่ยวกับข้อมูลในเชิงลึก

ขอขอบคุณ นางสาวมินตรา ชิมโสม และนายจิรเมธ ผาชลา นักศึกษาปริญญาโท สาขาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ที่ให้กำลังใจและช่วยเหลือเกื้อกูลกันมาโดยตลอด

ขอขอบคุณครอบครัว ที่คอยให้กำลังใจและให้การสนับสนุนในการเรียนต่อปริญญาโท และนางสาวณัฏฐา เชื้อสุข ที่คอยเป็นกำลังใจ เป็นแรงผลักดันและช่วยเหลือมาโดยตลอด

สุกุลยา เพชรชื่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ก
บทคัดย่อภาษาอังกฤษ.....	ข
กิตติกรรมประกาศ.....	ค
สารบัญ.....	ง
สารบัญตาราง	ช
สารบัญรูป	ญ
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของงานวิจัย	3
1.3 ขอบเขตของงานวิจัย.....	3
1.4 เกณฑ์การตัดสินใจ.....	3
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	4
1.6 นิยามศัพท์	4
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	5
2.1 ทฤษฎีที่เกี่ยวข้อง	5
2.1.1 สถิติเชิงพรรณนา.....	5
2.1.2 การสุ่มตัวอย่าง.....	5
2.1.2.1 การสุ่มตัวอย่างแบบแบ่งชั้นภูมิ (Stratified Sampling).....	5
2.1.3 สถิติเชิงอนุมาน.....	6
2.1.4 การวิเคราะห์ความสัมพันธ์ของตัวแปรเชิงคุณภาพ 2 ตัวแปร.....	6
2.1.4.1 เงื่อนไขของการใช้สถิติทดสอบเพียร์สัน-ไคสแควร์.....	7
2.1.5 การหาความสัมพันธ์.....	7
2.1.5.1 ค่าสหสัมพันธ์ Cramer's V.....	7
2.2 การเรียนรู้ของเครื่อง (Machine Learning).....	7
2.2.1 การเรียนรู้แบบมีผู้สอน (Supervised Learning).....	7
2.2.1.1 การจำแนกประเภท (Classification).....	7
2.2.1.2 การถดถอยเชิงเส้น (Linear Regression).....	7
2.2.2 การเรียนรู้โดยไม่มีผู้สอน (Unsupervised Learning)	8
2.2.3 การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning)	8
2.3 วิธีการจำแนก.....	8
2.3.1 วิธีการถดถอยโลจิสติก	8
2.3.1.1 การถดถอยโลจิสติกแบบทวิ.....	8
2.3.2 เทคนิคต้นไม้ตัดสินใจ.....	9

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
2.3.2.1 โหนด (Node).....	9
2.3.2.2 กิ่ง (Branch).....	9
2.3.2.3 ใบ (Leaf)	9
2.3.3 เทคนิคป่าสุ่ม	10
2.3.4 เทคนิคซัพพอร์ตเวกเตอร์แมชชีน	11
2.4 การคัดเลือกตัวแปร	11
2.4.1 Enter method.....	11
2.4.2 Forward Selection.....	11
2.4.3 Backward Elimination	11
2.4.4 Stepwise Selection	11
2.5 การวัดประสิทธิภาพ	12
2.5.1 การวัดประสิทธิภาพตัวแบบด้วยวิธีการตรวจสอบไขว้แบบ K รอบ	12
2.5.2 การวัดประสิทธิภาพตัวแบบด้วยเมทริกซ์สับสน (Confusion Matrix)	12
2.6 งานวิจัยที่เกี่ยวข้อง.....	13
บทที่ 3 วิธีการดำเนินงานวิจัย	16
3.1 โครงสร้างของขั้นตอนและวิธีการดำเนินงานวิจัย	16
3.2 การเก็บรวบรวมข้อมูล.....	18
3.3 การเตรียมข้อมูล.....	18
3.4 การวิเคราะห์ตัวแปร.....	21
3.4.1 วิเคราะห์โดยใช้สถิติเชิงพรรณนา (Descriptive Statistics).....	21
3.4.2 วิเคราะห์ผลโดยใช้สถิติเชิงอนุมาน (Inferential Statistics).....	21
3.4.2.1 ตรวจสอบความสัมพันธ์ระหว่างตัวแปรอิสระทุกตัวกับตัวแปรตาม..	21
3.4.2.2 ตรวจสอบความสัมพันธ์ระหว่างตัวแปรอิสระด้วยตนเอง	22
3.4.3 วิเคราะห์โดยผู้เชี่ยวชาญด้านการบำบัดรักษาผู้ติดยาเสพติด	22
3.4.4 วิเคราะห์โดยใช้การเรียนรู้ของเครื่อง (Machine Learning).....	22
3.5 การสร้างตัวแบบ	23
3.5.1 การสุ่มตัวอย่างและการแบ่งข้อมูล.....	23
3.5.2 การสร้างตัวแบบ	23
3.5.2.1 สร้างตัวแบบโดยใช้ตัวแปรอิสระครบทุกตัว (Full Model).....	23
3.5.2.2 สร้างตัวแบบโดยใช้ตัวแปรอิสระไม่ครบทุกตัวจากการคัดเลือกตัวแปร...	23
3.5.2.3 สร้างตัวแบบโดยใช้ตัวแปรอิสระไม่ครบทุกตัวจากการคัดเลือก ตัวแปรร่วมกับการปรับค่าไฮเปอร์พารามิเตอร์	23
3.5.3 การประเมินประสิทธิภาพด้วยวิธีการตรวจสอบไขว้แบบ 10 รอบ	24
3.6 การประเมินประสิทธิภาพตัวแบบ.....	24

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
บทที่ 4 ผลการวิจัย	25
4.1 ผลการวิเคราะห์ตัวแปร	25
4.1.1 วิเคราะห์โดยใช้สถิติเชิงพรรณนา (Descriptive Statistics).....	25
4.1.1.1 การวิเคราะห์ข้อมูลแบบตารางทางเดียว	25
4.1.2 วิเคราะห์ผลโดยใช้สถิติเชิงอนุมาน (Inferential Statistics)	32
4.1.2.1 ผลการตรวจสอบความสัมพันธ์ระหว่างตัวแปรอิสระทุกตัวกับตัวแปรตาม	32
4.1.2.2 ผลการตรวจสอบความสัมพันธ์ระหว่างตัวแปรอิสระด้วยกันเอง ...	43
4.1.3 วิเคราะห์โดยใช้การเรียนรู้ของเครื่อง.....	44
4.1.3.1 ผลการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระที่ผ่าน	
การคัดเลือกกับตัวแปรตามโดยใช้การถดถอยโลจิสติกอย่างง่าย	
(Univariate Logistic Regression).....	44
4.1.3.2 การหาค่า Feature Importance.....	47
4.2 ผลการสร้างตัวแบบ.....	50
4.2.1 ผลการสร้างตัวแบบโดยใช้ตัวแปรอิสระครบทุกตัว	50
4.2.1.1 ผลการสร้างตัวแบบการถดถอยโลจิสติก.....	50
4.2.1.2 ผลการสร้างตัวแบบป่าสุ่ม	53
4.2.1.3 ผลการสร้างตัวแบบซัพพอร์ตเวกเตอร์แมชชีน	53
4.2.2 ผลการสร้างตัวแบบโดยใช้ตัวแปรอิสระไม่ครบทุกตัวจากการคัดเลือกตัวแปร .	54
4.2.2.1 ผลการสร้างตัวแบบการถดถอยโลจิสติก.....	54
4.2.2.2 ผลการสร้างตัวแบบป่าสุ่ม	56
4.2.2.3 ผลการสร้างตัวแบบซัพพอร์ตเวกเตอร์แมชชีน	57
4.2.3 ผลสร้างตัวแบบโดยใช้ตัวแปรอิสระไม่ครบทุกตัวจากการคัดเลือก	
ตัวแปรร่วมกับการปรับค่าไฮเปอร์พารามิเตอร์.....	59
4.3 ผลการประเมินประสิทธิภาพตัวแบบ	61
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ	65
5.1 สรุปผลการวิจัย	65
5.2 ข้อเสนอแนะ	65
บรรณานุกรม	66
ภาคผนวก.....	68
ภาคผนวก ก	69
ภาคผนวก ข	70
ภาคผนวก ค	111
ประวัติผู้เขียน.....	112

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
2.1 เมทริกซ์สับสน (Confusion Matrix).....	12
3.1 รายละเอียดของข้อมูลตัวแปร.....	18
3.2 พารามิเตอร์ของตัวแบบป่าสุ่มและตัวแบบซัพพอร์ตเวกเตอร์แมชชีน.....	24
4.1 จำนวนและร้อยละของผลการบำบัดรักษาผู้ติดยาเสพติดแบบระบบสมัครใจ	25
4.2 จำนวนและร้อยละของเพศผู้ติดยาเสพติดแบบระบบสมัครใจ	26
4.3 จำนวนและร้อยละของอายุผู้ติดยาเสพติดแบบระบบสมัครใจ.....	26
4.4 จำนวนและร้อยละของสถานภาพสมรสผู้ติดยาเสพติดแบบระบบสมัครใจ	26
4.5 จำนวนและร้อยละของการศึกษาผู้ติดยาเสพติดแบบระบบสมัครใจ	27
4.6 จำนวนและร้อยละของอาชีพผู้ติดยาเสพติดแบบระบบสมัครใจ	27
4.7 จำนวนและร้อยละของรายได้ต่อเดือนผู้ติดยาเสพติดแบบระบบสมัครใจ	28
4.8 จำนวนและร้อยละของการอาศัยอยู่กับใคร (ในช่วง 30 วันที่ผ่านมา) ผู้ติดยาเสพติดแบบระบบสมัครใจ	28
4.9 จำนวนและร้อยละของความสัมพันธ์ในครอบครัวผู้ติดยาเสพติดแบบระบบสมัครใจ	29
4.10 จำนวนและร้อยละของอายุที่ใช้ยาเสพติดครั้งแรกผู้ติดยาเสพติดแบบระบบสมัครใจ	29
4.11 จำนวนและร้อยละของยาเสพติดที่ใช้ครั้งแรกผู้ติดยาเสพติดแบบระบบสมัครใจ	30
4.12 จำนวนและร้อยละของสาเหตุการใช้ยาเสพติดครั้งแรกผู้ติดยาเสพติดแบบระบบสมัครใจ	30
4.13 จำนวนและร้อยละของยาเสพติดที่ใช้ในปัจจุบันผู้ติดยาเสพติดแบบระบบสมัครใจ	31
4.14 จำนวนและร้อยละของระยะเวลาที่ใช้ยาเสพติดผู้ติดยาเสพติดแบบระบบสมัครใจ.....	31
4.15 จำนวนและร้อยละของเคยรับการบำบัดรักษาหรือไม่ผู้ติดยาเสพติดแบบระบบสมัครใจ	32
4.16 จำนวนและร้อยละของสาเหตุที่เข้ารับการบำบัดรักษาผู้ติดยาเสพติดแบบระบบสมัครใจ	32
4.17 ค่าสถิติต่าง ๆ ของเพศผู้ติดยาเสพติดแบบระบบสมัครใจ	33
4.18 ค่าสถิติต่าง ๆ ของอายุผู้ติดยาเสพติดแบบระบบสมัครใจ	33
4.19 ค่าสถิติต่าง ๆ ของสถานภาพสมรสผู้ติดยาเสพติดแบบระบบสมัครใจ.....	34
4.20 ค่าสถิติต่าง ๆ ของการศึกษาของผู้ติดยาเสพติดแบบระบบสมัครใจ	34
4.21 ค่าสถิติต่าง ๆ ของอาชีพผู้ติดยาเสพติดแบบระบบสมัครใจ.....	35
4.22 ค่าสถิติต่าง ๆ ของรายได้ต่อเดือนของผู้ติดยาเสพติดแบบระบบสมัครใจ	36
4.23 ค่าสถิติต่าง ๆ ของการอาศัยอยู่กับใคร (ในช่วง 30 วันที่ผ่านมา) ผู้ติดยาเสพติดแบบระบบสมัครใจ	36
4.24 ค่าสถิติต่าง ๆ ของความสัมพันธ์ในครอบครัวของผู้ติดยาเสพติดแบบระบบสมัครใจ	37
4.25 ค่าสถิติต่าง ๆ ของอายุที่ใช้ยาเสพติดครั้งแรกของผู้ติดยาเสพติดแบบระบบสมัครใจ	38
4.26 ค่าสถิติต่าง ๆ ของยาเสพติดที่ใช้ครั้งแรกของผู้ติดยาเสพติดแบบระบบสมัครใจ	39
4.27 ค่าสถิติต่าง ๆ ของสาเหตุการใช้ยาเสพติดครั้งแรกของผู้ติดยาเสพติด แบบระบบสมัครใจ	39
4.28 ค่าสถิติต่าง ๆ ของยาเสพติดที่ใช้ในปัจจุบันของผู้ติดยาเสพติดแบบระบบสมัครใจ	40

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

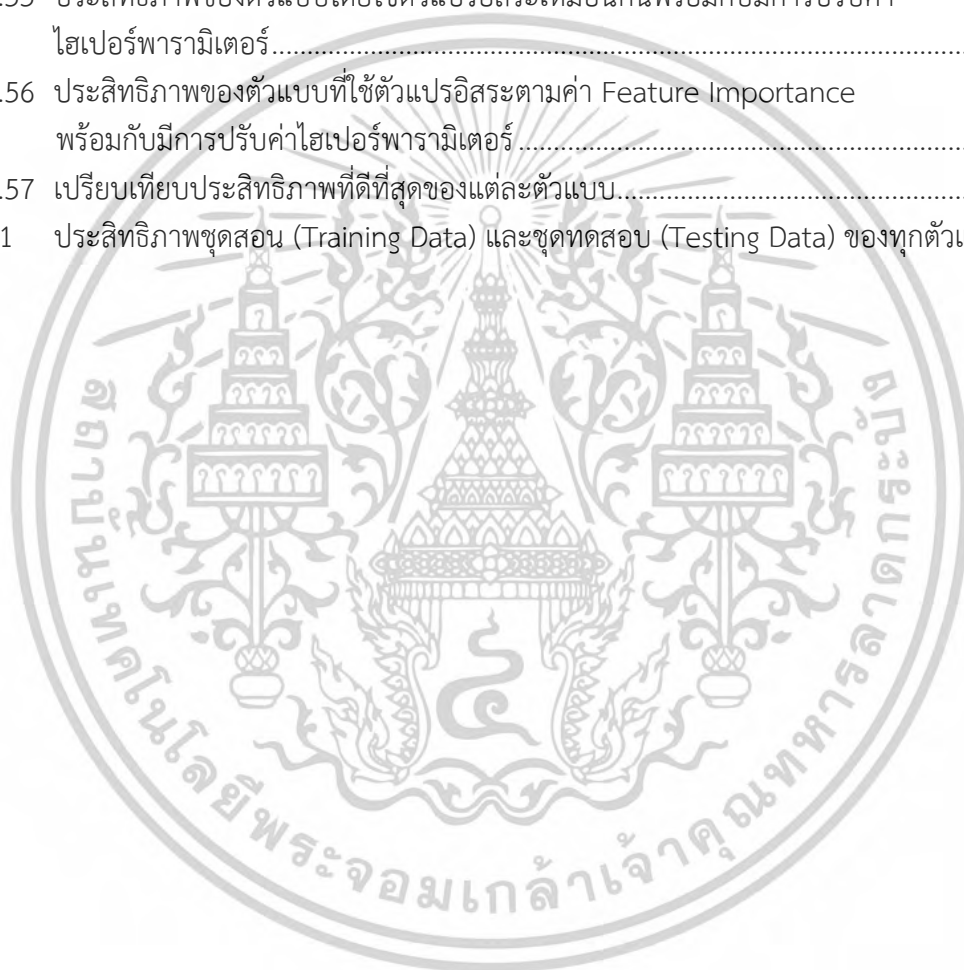
สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.29 ค่าสถิติต่าง ๆ ของระยะเวลาที่ใช้ยาเสพติดของผู้ติดยาเสพติดแบบระบบสมัครใจ	41
4.30 ค่าสถิติต่าง ๆ ของการเคยรับการบำบัดรักษาของผู้ติดยาเสพติดแบบระบบสมัครใจ	41
4.31 ค่าสถิติต่าง ๆ ของสาเหตุที่เข้ารับการบำบัดรักษาของผู้ติดยาเสพติด แบบระบบสมัครใจ.....	42
4.32 ผลการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระด้วยตนเอง	43
4.33 ผลการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระที่ผ่านการคัดเลือกกับตัวแปรตาม โดยใช้การถดถอยโลจิสติกอย่างง่าย (Univariate Logistic Regression).....	44
4.34 ค่า Odds Ratio ตัวแปรอิสระต่อการไม่ครบกำหนดการรักษา	50
4.35 ผลการทำนายด้วยตัวแบบการถดถอยโลจิสติก.....	53
4.36 ผลการทำนายด้วยตัวแบบป่าสุ่ม.....	53
4.37 ผลการทำนายด้วยตัวแบบซัพพอร์ตเวกเตอร์แมชชีน	54
4.38 ผลการคัดเลือกตัวแปรของตัวแบบการถดถอยโลจิสติก	54
4.39 ผลการจำแนกด้วยตัวแบบการถดถอยโลจิสติกที่ทำการคัดเลือกตัวแปร ด้วยวิธี stepwise selection.....	55
4.40 ผลการจำแนกด้วยตัวแบบการถดถอยโลจิสติกที่ทำการคัดเลือกตัวแปรด้วย ค่า Feature Importance.....	55
4.41 ผลการคัดเลือกตัวแปรของตัวแบบป่าสุ่ม.....	56
4.42 ผลการทำนายด้วยตัวแบบป่าสุ่มที่ทำการคัดเลือกตัวแปรโดยใช้ตัวแปรเหมือนกับตัว แบบการถดถอยโลจิสติก	57
4.43 ผลการทำนายด้วยตัวแบบป่าสุ่มที่ทำการคัดเลือกตัวแปรด้วยค่า Feature Importance.	57
4.44 ผลการคัดเลือกตัวแปรของตัวแบบซัพพอร์ตเวกเตอร์แมชชีน	58
4.45 ผลการทำนายด้วยตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ทำการคัดเลือกตัวแปร โดยใช้ตัวแปรเหมือนกับตัวแบบการถดถอยโลจิสติก	58
4.46 ผลการทำนายด้วยตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ทำการคัดเลือกตัวแปรด้วย ค่า Feature Importance.....	59
4.47 ค่าพารามิเตอร์หลังจากมีการปรับค่าไฮเปอร์พารามิเตอร์กับตัวแบบ.....	59
4.48 ผลการทำนายด้วยตัวแบบป่าสุ่มที่ทำการคัดเลือกตัวแปรโดยใช้ตัวแปรเหมือนกับ ตัวแบบการถดถอยโลจิสติกพร้อมกับทำการปรับค่าไฮเปอร์พารามิเตอร์	60
4.49 ผลการทำนายด้วยตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ทำการคัดเลือกตัวแปร โดยตัวแปรใช้เหมือนกับตัวแบบการถดถอยโลจิสติกพร้อมกับทำการปรับค่าไฮเปอร์ พารามิเตอร์.....	60
4.50 ผลการทำนายด้วยตัวแบบป่าสุ่มที่ทำการคัดเลือกตัวแปรด้วยค่า Feature Importance พร้อมกับทำการปรับค่าไฮเปอร์พารามิเตอร์.....	60

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.51 ผลการทำนายด้วยตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ทำการคัดเลือกตัวแปร ด้วยค่า Feature Importance พร้อมกับทำการปรับค่าไฮเปอร์พารามิเตอร์.....	61
4.52 ประสิทธิภาพของตัวแบบที่ใช้ตัวแปรอิสระครบทุกตัว.....	61
4.53 ประสิทธิภาพของตัวแบบที่ใช้ตัวแปรอิสระเหมือนกัน.....	62
4.54 ประสิทธิภาพของตัวแบบที่ใช้ตัวแปรอิสระตามค่า Feature Importance.....	62
4.55 ประสิทธิภาพของตัวแบบโดยใช้ตัวแปรอิสระเหมือนกันพร้อมกับการปรับค่า ไฮเปอร์พารามิเตอร์.....	62
4.56 ประสิทธิภาพของตัวแบบที่ใช้ตัวแปรอิสระตามค่า Feature Importance พร้อมกับการปรับค่าไฮเปอร์พารามิเตอร์.....	63
4.57 เปรียบเทียบประสิทธิภาพที่ดีที่สุดของแต่ละตัวแบบ.....	63
ก1 ประสิทธิภาพชุดสอน (Training Data) และชุดทดสอบ (Testing Data) ของทุกตัวแบบ.	69



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่	หน้า
2.1 การเลือกตัวอย่างแบบแบ่งชั้น	6
2.2 เทคนิคต้นไม้ตัดสินใจ (Decision tree).....	10
2.3 เทคนิคป่าสุ่ม (Random Forest).....	10
2.4 วิธีการแบบ 10-Fold Cross Validation	12
3.1 ขั้นตอนและวิธีการดำเนินงานวิจัย	17
4.1 คะแนนความสำคัญของตัวแปรของตัวแบบการถดถอยโลจิสติก	48
4.2 คะแนนความสำคัญของตัวแปรของตัวแบบป่าสุ่ม	48
4.3 คะแนนความสำคัญของตัวแปรของตัวแบบซัพพอร์ตเวกเตอร์แมชชีน	49



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

สถานการณ์ยาเสพติดในพื้นที่กรุงเทพมหานคร พบว่าเกือบตลอดทั้งปีงบประมาณ พ.ศ. 2566 ยังคงเผชิญกับปัญหายาเสพติดอย่างต่อเนื่อง ซึ่งมีการลักลอบนำเข้ามาในพื้นที่เป็นจำนวนมาก และปรากฏข้อมูลบ่งชี้ว่าพื้นที่กรุงเทพมหานครเป็นแหล่งพักพิงยาเสพติด ก่อนที่จะกระจายยาเสพติดไปสู่กลุ่มตลาดผู้เสพในพื้นที่ รวมถึงพักรอเพื่อลำเลียงไปยังภาคใต้และพื้นที่อื่น ๆ ทั้งในและนอกประเทศ และจากระบบข้อมูลการบำบัดรักษาและฟื้นฟูผู้ติดยาเสพติดของประเทศ (บสต.) พบว่ามีผู้เข้ารับการบำบัดฟื้นฟูเพียง 200,000 กว่าราย ยังมีผู้เสพยาเสพติดจำนวนมากที่ไม่เข้าระบบการบำบัดฟื้นฟูซึ่งปัญหาหนึ่งของการไม่เข้ารับการบำบัดฟื้นฟู คือ ผู้เสพยา ผู้ป่วย ไม่กล้าเปิดเผยตัวเพื่อเข้าสู่กระบวนการบำบัดฟื้นฟูเพราะเกรงว่าจะถูกจับกุม ต้องโทษ ทำให้ปัญหายาเสพติดในพื้นที่กรุงเทพมหานครยังคงทวีความรุนแรงมากขึ้นทุกวันและส่งผลกระทบต่อหลาย ๆ ด้าน ทั้งด้านสุขภาพกายและสุขภาพจิตของผู้เสพยา ส่งผลกระทบต่อสังคมและคนรอบข้าง รวมทั้งในด้านการพัฒนาประเทศ ผลกระทบเหล่านี้นำไปสู่การเกิดนโยบายในการแก้ไขปัญหายาเสพติดที่มุ่งเน้นการดูแลสุขภาพผู้ป่วยไปพร้อมกับการแก้ไขปัญหาสังคมและยังคงมีการป้องกันและปราบปรามยาเสพติดอย่างเข้มงวด

การดำเนินงานการบำบัดรักษาและฟื้นฟูสมรรถภาพผู้ติดยาเสพติดตามประมวลกฎหมายยาเสพติด มุ่งเน้นการดูแลสุขภาพผู้ใช้ ผู้เสพยา ผู้ติดยาเสพติด ด้วยกลไกสาธารณสุขแทนการดำเนินคดีทางอาญา โดยเน้นการแก้ไขปัญหาด้านสุขภาพ พฤติกรรม และสังคม ควบคู่ไปกับการบำบัดรักษา รวมถึงการคัดกรอง การประเมินความรุนแรง การบำบัดด้วยยา การฟื้นฟูสมรรถภาพ การลดอันตรายจากยาเสพติดและการติดตามหลังการบำบัดรักษา ซึ่งใช้ระยะเวลาต่อเนื่องและยาวนาน เพื่อที่จะลด ละ เลิกยาเสพติด หากผู้ป่วยเข้ารับการบำบัดรักษาครบตามเกณฑ์ ย่อมทำให้มีสุขภาพกายและสุขภาพจิตที่ดีขึ้น สามารถดำรงชีวิตในสังคมได้อย่างปกติสุข หากผู้ป่วยยาเสพติดเข้ารับการบำบัดรักษาไม่ครบตามเกณฑ์การรักษาที่กำหนดไว้ ย่อมส่งผลกระทบต่อสุขภาพกายและสุขภาพจิตของผู้ป่วย ผู้ป่วยมีโอกาสกลับมาเสพยาเสพติดได้อีกครั้ง อาจก่อให้เกิดปัญหาทางสังคมเพิ่มมากขึ้น เช่น การทะเลาะวิวาท การก่ออาชญากรรมต่าง ๆ ผู้ป่วยที่เข้ารับการบำบัดรักษา ย่อมมีลักษณะหรือปัจจัยที่แตกต่างกัน ซึ่งอาจส่งผลกระทบต่อพฤติกรรมการบำบัดรักษาของผู้ป่วยนั้น ๆ ด้วย เช่น เพศ สถานภาพสมรส ยาเสพติดหลักที่ใช้ระยะเวลาในการใช้ยาเสพติด เป็นต้น

งานวิจัยในช่วงสิบปีที่ผ่านมา มีการศึกษาเกี่ยวกับปัจจัยที่ส่งผลต่อการบำบัดรักษาผู้ป่วยยาเสพติดมากมาย ผู้วิจัยขอยกมาโดยสังเขป ดังนี้

ชลอวัฒน์ อินปา และพิศิษฐ์ ศรีประเสริฐ (2564) ได้ศึกษาปัจจัยที่ส่งผลต่อการเข้ารับการบำบัดรักษาครบตามเกณฑ์ของผู้ป่วยยาเสพติดในจังหวัดตาก โดยใช้สถิติเชิงพรรณนาและสถิติเชิงอนุมาน ได้แก่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน สถิติ Chi-square test และการวิเคราะห์ถดถอยพหุโลจิสติกส์ ในการวิเคราะห์ข้อมูลผู้ป่วยยาเสพติดในจังหวัดตากที่เข้ารับการบำบัดรักษาปีงบประมาณ 2563 จำนวน 1,293 ราย จากระบบข้อมูลการบำบัดรักษาและฟื้นฟูผู้ติดยาเสพติดของประเทศ (บสต.) กระทรวงสาธารณสุข ผลการศึกษาพบว่า ปัจจัยที่ส่งผลต่อการเข้ารับการบำบัดรักษา

ครบตามเกณฑ์ของผู้ป่วยยาเสพติด ได้แก่ (1) ปัจจัยส่วนบุคคล คือ สถานภาพสมรส และยาเสพติด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่โดยไม่ได้รับอนุญาต

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลักที่ใช้เสพ (2) ปัจจัยด้านการรักษา คือ ความสมัครใจ และวิธีการรักษาด้วยเมทาโดน (3) ปัจจัยด้านการสนับสนุนทางสังคม คือ ความสัมพันธ์ในครอบครัวที่ยอมรับและช่วยเหลือ

สมพร สุวรรณมาโจ และคณะ (2562) ได้ศึกษาปัจจัยทำนายในการคงอยู่ในระบบระยะเวลา 1 ปีของผู้บำบัดด้วยเมทาโดนระยะยาวในคลินิกเมทาโดน ณ สถาบันบำบัดรักษาและฟื้นฟูผู้ติดยาเสพติดแห่งชาติบรมราชชนนี โดยใช้สถิติเชิงพรรณนาในรูปแบบของความถี่ ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน วิเคราะห์หาตัวแปรที่มีความสัมพันธ์ของผู้ที่ได้รับยาเมทาโดนด้วย Chi-square และวิเคราะห์ปัจจัยทำนายด้วยการวิเคราะห์การถดถอยโลจิสติกส์ วิเคราะห์ข้อมูลที่เก็บจากเวชระเบียนของผู้ป่วยเสพติดสารกลุ่มโอปิออยด์ที่รับการบำบัดด้วยเมทาโดนระยะยาวและแบบรายงานการบำบัดรักษาผู้ติดยาเสพติด (บสต.) ระหว่างปีงบประมาณ 2558 จนถึง 2561 จำนวน 173 ราย ผลการศึกษาพบว่า ปัจจัยทำนายในการคงอยู่ในระบบระยะเวลา 1 ปีของผู้บำบัดด้วยเมทาโดนระยะยาวในคลินิกเมทาโดน ณ สถาบันบำบัดรักษาและฟื้นฟูผู้ติดยาเสพติดแห่งชาติบรมราชชนนี คือ ผู้ป่วยมีการประกอบอาชีพ และมีรายได้ต่อเดือนมากกว่า 10,000 บาท

จรรยาตรี โคมพุดชา และชัยภัทร อีระชาญไชย (2558) ได้ศึกษาปัจจัยที่มีความสัมพันธ์กับการบำบัดครบกำหนดของผู้ใช้สารเสพติดในระบบสมัครใจ จังหวัดนครราชสีมา ปีงบประมาณ 2555-2556 โดยวิเคราะห์และนำเสนอข้อมูลด้วยสถิติเชิงพรรณนา เป็นจำนวน ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน เปรียบเทียบความแตกต่างของสัดส่วนของกลุ่มตัวอย่าง 2 กลุ่ม คือ กลุ่มบำบัดครบกำหนด และกลุ่มบำบัดไม่ครบกำหนด ด้วยสถิติ Chi-square test และ Fisher's exact test เปรียบเทียบความแตกต่างของค่าเฉลี่ยหรือมัธยฐานของตัวอย่าง 2 กลุ่ม ด้วยสถิติ Kruskal Wallis Test โดยกำหนดค่านัยสำคัญทางสถิติที่ 0.05 โดยใช้ข้อมูลการบำบัดผู้ใช้สารเสพติดในระบบสมัครใจ ในสถานบริการสาธารณสุขของรัฐ จังหวัดนครราชสีมา ปีงบประมาณ 2555-2556 จากกระบบรายงานระบบติดตามและเฝ้าระวังปัญหาสารเสพติดของกระทรวงสาธารณสุข (บสต.) จำนวน 8,784 ราย ผลการศึกษาพบว่า ผู้เข้ารับการบำบัดได้รับการบำบัดครบกำหนด 8,187 ราย (ร้อยละ 93.2) ปัจจัยที่มีความสัมพันธ์กับการบำบัดครบกำหนดของผู้รับการบำบัด ได้แก่ ที่อยู่อาศัย วิธีการใช้สารเสพติด สาเหตุสำคัญในการใช้สารเสพติด เหตุผลในการเข้ารับการบำบัด รูปแบบวิธีการบำบัด และจำนวนวันบำบัดรักษา

เสาวนีย์ พันธุ์พัฒนกุล และณัฐณิย์ มีมนต์ (2558) ได้ศึกษาปัจจัยที่สัมพันธ์กับการบรรลุผลกระบวนการบำบัดรักษาฟื้นฟูสมรรถภาพและติดตามของผู้รับการบำบัดยาเสพติด จังหวัดพะเยา ทำการวิเคราะห์และแสดงผลข้อมูลสถิติเชิงพรรณนาด้วยการกระจายความถี่แบบร้อยละและทดสอบความสัมพันธ์ระหว่างปัจจัยต่างๆกับการบรรลุผลกระบวนการบำบัดรักษาและติดตามผู้เสพยาเสพติดด้วยสถิติ Chi-square หรือ Fisher' Exact Test กำหนดค่านัยสำคัญทางสถิติที่ 0.05 โดยใช้ข้อมูลผู้เสพยาเสพติดที่เข้ารับการบำบัดรักษาและฟื้นฟูสมรรถภาพในสถานบริการสาธารณสุขทุกแห่งที่ตั้งอยู่ในเขตจังหวัดพะเยา ระหว่างปีงบประมาณ 2551-2556 จากกระบบติดตามและเฝ้าระวังปัญหาสารเสพติดของกระทรวงสาธารณสุข (บสต.) จำนวน 8,408 ราย ผลการศึกษาพบว่า ปัจจัยที่สัมพันธ์กับการบรรลุผลกระบวนการบำบัดรักษาฟื้นฟูสมรรถภาพและติดตามผู้ติดยาเสพติดที่สำคัญ ได้แก่ เพศ อายุ สถานภาพสมรส ระดับการศึกษา เหตุผลที่เข้ารับการบำบัด และสถานการณ์ยาเสพติด

จากที่กล่าวมาข้างต้น พบว่าที่ผ่านมายังไม่มีการศึกษาเกี่ยวกับการยุติการรักษาของผู้ติดยาเสพติด และใช้การถดถอยโลจิสติกส์ในการวิเคราะห์มากนัก ผู้วิจัยจึงมีวัตถุประสงค์ที่จะทำการศึกษานี้เกี่ยวกับการยุติการรักษาของผู้ติดยาเสพติด ปัจจัยที่ส่งผลต่อการยุติการรักษาของผู้ติดยาเสพติด และเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ต่อยอดการวิเคราะห์ทางสถิติ จึงได้นำข้อมูลผู้เข้ารับการบำบัดรักษาและฟื้นฟูผู้ติดยาเสพติด ระบบสมัครใจในพื้นที่กรุงเทพมหานคร จากระบบข้อมูลการบำบัดรักษาและฟื้นฟูสมรรถภาพผู้ติดยาเสพติดของประเทศ (บสต.) จากกระทรวงสาธารณสุข มาสร้างและเปรียบเทียบประสิทธิภาพการจำแนกกลุ่มข้อมูลการยุติการบำบัดรักษาผู้ติดยาเสพติด รวมถึงระบุปัจจัยที่มีผลต่อการยุติการรักษาของผู้ติดยาเสพติด และทำนายลักษณะผู้เข้ารับการบำบัดรักษา ด้วยการใช้วิธีการจำแนกข้อมูล 3 วิธี คือ วิธีการถดถอยโลจิสติก (Logistic Regression) วิธีป่าสุ่ม (Random Forest) และวิธีซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) เพื่อวางแผนในการพัฒนากระบวนการบำบัดรักษาผู้ติดยาเสพติดให้มีประสิทธิภาพมากขึ้นและลดโอกาสในการเกิดการยุติการบำบัดรักษา

1.2 วัตถุประสงค์ของงานวิจัย

- 1) เพื่อศึกษาวิธีการเปรียบเทียบและสร้างตัวแบบการจำแนกกลุ่มการยุติการบำบัดรักษาผู้ติดยาเสพติด โดยใช้การเรียนรู้ของเครื่อง 3 วิธี คือ วิธีการถดถอยโลจิสติก (Logistic Regression) วิธีป่าสุ่ม (Random Forest) และวิธีซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine)
- 2) เพื่อเปรียบเทียบวิธีการคัดเลือกตัวแปรอิสระสำหรับตัวแบบการจำแนกกลุ่มการยุติการบำบัดรักษาผู้ติดยาเสพติด

1.3 ขอบเขตของงานวิจัย

- 1) ข้อมูลที่ใช้ในการศึกษาเป็นข้อมูลการบำบัดรักษาของผู้ติดยาเสพติดระบบสมัครใจในพื้นที่กรุงเทพมหานคร จากระบบข้อมูลการบำบัดรักษาและฟื้นฟูผู้ติดยาเสพติดของประเทศ (บสต.) กระทรวงสาธารณสุข เป็นระยะเวลา 6 ปีงบประมาณ ตั้งแต่ ตุลาคม พ.ศ. 2560 - กันยายน พ.ศ. 2566 โดยเลือกศึกษาเฉพาะข้อมูลการบำบัดรักษาครั้งล่าสุด (ในกรณีที่ผู้ป่วยเข้ามาบำบัดหลายครั้งในช่วง 6 ปีงบประมาณ ที่มีรหัสการเข้ารับบำบัดหลายรหัส) ที่มีการระบุว่าบำบัดครบกำหนด และบำบัดไม่ครบกำหนด มีรูปแบบการบำบัดรักษาแบบผู้ป่วยนอกในสถานพยาบาล และมีข้อมูลในระบบการบำบัดรักษาและฟื้นฟูผู้ติดยาเสพติดครบถ้วน จำนวน 3,241 คน มีจำนวนตัวแปรอิสระ 15 ตัวแปร
- 2) ศึกษาการจำแนกข้อมูลด้วยวิธีการถดถอยโลจิสติก วิธีป่าสุ่ม และวิธีซัพพอร์ตเวกเตอร์แมชชีน และการเพิ่มประสิทธิภาพตัวแบบด้วยการคัดเลือกตัวแปรแบบแรปเปอร์และการปรับค่าไฮเปอร์พารามิเตอร์ พร้อมทั้งคำนวณค่าทางสถิติในการจัดลำดับตัวแปร เช่น ค่า p value และค่า feature importance
- 3) เปรียบเทียบประสิทธิภาพตัวแบบ โดยใช้เกณฑ์การตัดสินใจ คือ เกณฑ์คะแนน F1 หรือ ค่าเฉลี่ยฮาร์โมนิกของค่าความเที่ยง (Precision) และ ค่าความครบถ้วน (Recall) จากตารางสรุปผลลัพธ์การทำนาย (Confusion Matrix) เพื่อหาตัวแบบที่ดีที่สุด
- 4) วิเคราะห์ข้อมูลโดยใช้โปรแกรมอาร์ R เวอร์ชัน 4.3.2

1.4 เกณฑ์การตัดสินใจ

เกณฑ์คะแนน F1 หรือ ค่าเฉลี่ยฮาร์โมนิกของค่าความเที่ยง (Precision) และค่าความครบถ้วน (Recall) จากตารางสรุปผลลัพธ์การทำนาย (Confusion Matrix) โดยตัวแบบที่ให้คะแนน F1 มากที่สุดจะเป็นตัวแบบที่ดีที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.5 ประโยชน์ที่คาดว่าจะได้รับ

- 1) ได้ตัวแบบที่เหมาะสมสำหรับการจำแนกกลุ่มการยุติการบำบัดรักษาผู้ติดยาเสพติด
- 2) ทราบวิธีการคัดเลือกตัวแปรอิสระสำหรับตัวแบบการจำแนกกลุ่มการยุติการบำบัดรักษาผู้ติดยาเสพติด
- 3) ทราบประสิทธิภาพการจำแนกกลุ่มการยุติการบำบัดรักษาผู้ติดยาเสพติดของตัวแบบด้วยวิธีการถดถอยโลจิสติก, วิธีป่าสุ่ม และ วิธีซัพพอร์ตเวกเตอร์แมชชีน
- 4) สามารถนำผลการครบกำหนดตามเกณฑ์การบำบัดรักษาที่ได้จากการพยากรณ์มาใช้เป็นแนวทางในการวางแผนโปรแกรมการบำบัดรักษาสำหรับผู้ติดยาเสพติดนั้น ๆ เพื่อป้องกันและลดโอกาสในการเกิดการยุติการบำบัดรักษา

1.6 นิยามศัพท์

- 1) กลุ่มบำบัดครบกำหนด หมายถึง ผู้ติดยาเสพติดที่เข้ารับการบำบัดในกระบวนการบำบัดรักษาครบระยะเวลาตามรูปแบบการบำบัดรักษาของสถานบริการสุขภาพ สังกัดกระทรวงสาธารณสุข
- 2) กลุ่มบำบัดไม่ครบกำหนด หมายถึง ผู้ติดยาเสพติดที่เข้ารับการบำบัดในกระบวนการบำบัดรักษาไม่ครบระยะเวลาตามรูปแบบการบำบัดรักษาของสถานบริการสุขภาพ สังกัดกระทรวงสาธารณสุข หรือเรียกว่า กลุ่มยุติการบำบัดการรักษา
- 3) บำบัดรักษา หมายถึง การบำบัดรักษาผู้ติดยาเสพติดซึ่งรวมถึงการคัดกรอง การประเมินความรุนแรง การบำบัดด้วยยา การฟื้นฟูสมรรถภาพ การลดอันตรายจากยาเสพติดและการติดตามหลังการบำบัดรักษา
- 4) ระบบสมัครใจ หมายถึง ระบบการบำบัดโดยผู้ป่วยเข้ารับการบำบัดรักษาโดยสมัครใจด้วยตนเองหรือผู้ปกครอง ญาติ, ผู้นำชุมชน
- 5) ผู้ป่วยนอก หมายถึง การบำบัดรักษาผู้ป่วยแบบไป-กลับไม่พักค้างในสถานพยาบาล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

งานวิจัยนี้เป็นการศึกษาปัจจัยที่ส่งผลต่อกลุ่มการยุติการบำบัดรักษาผู้ติดยาเสพติดแบบระบบสมัครใจในพื้นที่กรุงเทพมหานครโดยประยุกต์ใช้การเรียนรู้ของเครื่อง ซึ่งในบทนี้จะกล่าวถึงทฤษฎีที่ใช้ในการศึกษาและงานวิจัยที่เกี่ยวข้องดังนี้

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 สถิติเชิงพรรณนา

2.1.2 การสุ่มตัวอย่าง

2.1.3 สถิติเชิงอนุมาน

2.1.4 การวิเคราะห์ความสัมพันธ์ของตัวแปรเชิงคุณภาพ 2 ตัวแปร

2.1.5 การหาความสัมพันธ์

2.2 การเรียนรู้ของเครื่อง (Machine Learning)

2.2.1 การเรียนรู้แบบมีผู้สอน (Supervised Learning)

2.2.2 การเรียนรู้โดยไม่มีผู้สอน (Unsupervised Learning)

2.2.3 การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning)

2.3 วิธีการจำแนก

2.3.1 วิธีการถดถอยโลจิสติก

2.3.2 เทคนิคต้นไม้ตัดสินใจ

2.3.3 เทคนิคป่าสุ่ม

2.3.4 เทคนิคซัพพอร์ตเวกเตอร์แมชชีน

2.4 การคัดเลือกตัวแปร

2.5 การวัดประสิทธิภาพ

2.6 งานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 สถิติเชิงพรรณนา

เป็นสถิติที่ใช้ในการบรรยายลักษณะที่สำคัญของกลุ่มข้อมูล โดยที่เก็บรวบรวมข้อมูลมาจากตัวอย่างหรือประชากร สถิติเชิงพรรณนา ได้แก่ จำนวน ร้อยละ ค่ากลาง (ค่าเฉลี่ย ค่ามัธยฐาน ค่าฐานนิยม) ค่าการกระจาย (ค่าแปรปรวน ค่าเบี่ยงเบนมาตรฐาน ค่าเปอร์เซ็นต์ไทล์ สัมประสิทธิ์ความผันแปร ฯลฯ) กราฟต่าง ๆ (กัลยา วานิชปัญษา, 2567)

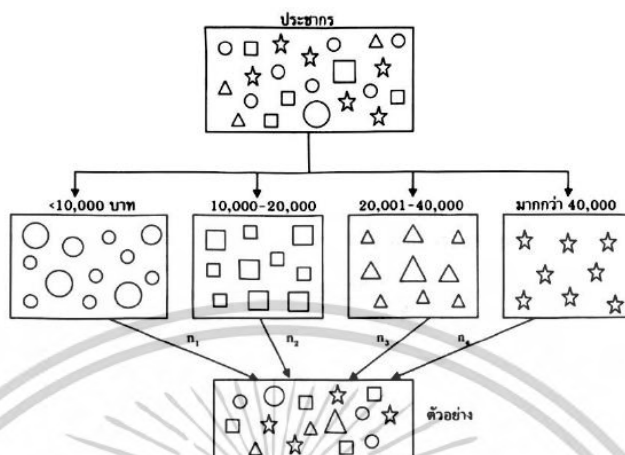
2.1.2 การสุ่มตัวอย่าง

2.1.2.1 การสุ่มตัวอย่างแบบแบ่งชั้นภูมิ (Stratified Sampling)

เป็นการสุ่มตัวอย่างจากประชากรที่มีจำนวนมากและมีความแตกต่างกัน โดยจะทำการแบ่งประชากรออกเป็นกลุ่มย่อยหลาย ๆ กลุ่ม เรียกว่า ชั้นภูมิ (Stratum) มีหลักการในการแบ่งคือ ต้องการให้หน่วยที่อยู่ในชั้นภูมิเดียวกันมีความคล้ายกันหรือเหมือนกันในเรื่องที่สนใจศึกษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ส่วนหน่วยที่อยู่ต่างชั้นภูมิกันมีความแตกต่างกันในเรื่องที่สนใจศึกษา การสุ่มตัวอย่างแบบนี้จะทำให้ได้ข้อมูลที่มีความครบถ้วนและครอบคลุม (กัลยา วานิชบัญชา, 2564) ดังแสดงในรูปที่ 2.1



รูปที่ 2.1 การเลือกตัวอย่างแบบแบ่งชั้น (กัลยา วานิชบัญชา, 2567)

2.1.3 สถิติเชิงอนุมาน

เป็นสถิติที่ใช้ศึกษาข้อมูลจากกลุ่มตัวอย่างแล้วนำผลสรุปที่ได้จากกลุ่มตัวอย่างมาสรุปอ้างอิงลักษณะประชากรแบ่งเป็น การประมาณค่า (Estimation) และการทดสอบสมมติฐาน (Hypothesis Testing) (กัลยา วานิชบัญชา, 2567)

2.1.4 การวิเคราะห์ความสัมพันธ์ของตัวแปรเชิงคุณภาพ 2 ตัวแปร

ในการศึกษาความสัมพันธ์ของตัวแปรเชิงคุณภาพ 2 ตัวที่เป็นสเกลนามบัญญัติหรือสเกลอันดับ (กัลยา วานิชบัญชา, 2567) เช่น เพศกับอาชีพ จะใช้สถิติเพียร์สันไคสแควร์ในการศึกษา โดยเขียนสมมติฐานเพื่อการทดสอบ ดังนี้

H_0 : ตัวแปรทั้ง 2 ตัวเป็นอิสระกัน

H_1 : ตัวแปรทั้ง 2 ตัวไม่เป็นอิสระกัน

สถิติที่ใช้ทดสอบ : เพียร์สันไคสแควร์ (Pearson Chi-Square)

โดยที่ Pearson Chi-Square
$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (2.1)$$

O_{ij} = จำนวนหรือความถี่ของข้อมูลที่มีลักษณะที่ i ของตัวแปรที่หนึ่งและมีลักษณะที่ j ของตัวแปรที่สอง

r = จำนวนลักษณะของตัวแปรที่หนึ่ง

c = จำนวนลักษณะของตัวแปรที่สอง

E_{ij} = ความถี่หรือจำนวนที่คาดไว้ของลักษณะที่ i ของตัวแปรที่หนึ่งและมีลักษณะที่ j ของตัวแปรที่สอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.4.1 เงื่อนไขของการใช้สถิติทดสอบเพียร์สัน-ไคสแควร์

1. ความถี่ หรือจำนวนที่คาดหวัง E_{ij} จะต้องไม่ต่ำกว่า 5 หรือ $E_{ij} \geq 5$; $i = 1, 2, \dots, r$, $j = 1, 2, \dots, c$ กรณีที่มีค่า $E_{ij} < 5$ จะสามารถมี $E_{ij} < 5$ ได้ไม่เกิน 20% ของจำนวน cell ทั้งหมด และไม่มี cell ใดที่มีค่า $E_{ij} < 1$

2. กรณีที่มีตาราง ขนาด 2×2 ($r=2$, $c=2$) สถิติทดสอบเพียร์สัน-ไคสแควร์ จะเป็น

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(|O_{ij} - E_{ij}| - 0.5)^2}{E_{ij}} \quad (2.2)$$

เมื่อทำการทดสอบแล้วพบว่าตัวแปรทั้ง 2 ตัวไม่เป็นอิสระกัน กล่าวคือตัวแปรทั้ง 2 ตัวมีความสัมพันธ์กัน จะทำการหาความสัมพันธ์ของตัวแปรอีกครั้ง โดยใช้ค่าสหสัมพันธ์คราเมอร์ เพื่อหาว่าตัวแปรทั้ง 2 ตัวมีความสัมพันธ์กันมากน้อยเพียงใด

2.1.5 การหาความสัมพันธ์

เป็นการหาว่าตัวแปร 2 ตัวมีความสัมพันธ์กันมากน้อยเพียงใด (กัลยา วานิชบัญชา, 2567)

2.1.5.1 ค่าสหสัมพันธ์ Cramer's V

$$V = \sqrt{\frac{\chi^2}{n \cdot \min\{(r-1), (c-1)\}}} \quad \begin{matrix} \text{โดยที่ } 0 < V < 1 \\ 0 < C < 1 \end{matrix} \quad (2.3)$$

- ถ้า V เข้าใกล้ 1 แสดงว่าตัวแปรเชิงกลุ่มทั้ง 2 ตัวมีความสัมพันธ์กันมาก
- ถ้า V เข้าใกล้ 0 แสดงว่าตัวแปรเชิงกลุ่มทั้ง 2 ตัวไม่มีความสัมพันธ์กันหรือความสัมพันธ์กันน้อยมาก

2.2 การเรียนรู้ของเครื่อง (Machine Learning)

เป็นการที่ทำให้อัลกอริทึมสามารถเรียนรู้สิ่งต่าง ๆ และพัฒนาการทำงานให้ดีขึ้นได้ด้วยตนเอง จากชุดข้อมูล มีหลักการการทำงานคือทำการบ่อนชุดข้อมูล เพื่อให้อัลกอริทึมเรียนรู้และสามารถทำนายผลลัพธ์ได้ และเพื่อให้ได้ผลลัพธ์ที่แม่นยำมากขึ้นจึงมีการบ่อนชุดข้อมูลใหม่ ๆ และการฝึกฝนอย่างสม่ำเสมอ การทำงานของการเรียนรู้ของเครื่อง มี 3 รูปแบบ (cyberelite, 2022)

2.2.1 การเรียนรู้แบบมีผู้สอน (Supervised Learning)

เป็นการบ่อนชุดข้อมูลเพื่อให้อัลกอริทึมเกิดการเรียนรู้ โดยจะเรียกชุดข้อมูลนั้นว่าข้อมูลชุดสอน (Training Data) จากนั้นจึงนำข้อมูลชุดทดสอบ (Testing Data) มาทดสอบ ให้อัลกอริทึมทำการประมวลผลก่อนจะนำไปใช้จริง การเรียนรู้แบบมีผู้สอนแบ่งออกเป็น 2 ประเภท คือ

2.2.1.1 การจำแนกประเภท (Classification) ใช้กับตัวแปรตามที่เป็นข้อมูลเชิงคุณภาพ

2.2.1.2 การถดถอยเชิงเส้น (Linear Regression) ใช้กับตัวแปรตามที่เป็นข้อมูล

เชิงปริมาณ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.2 การเรียนรู้โดยไม่มีผู้สอน (Unsupervised Learning)

เป็นการเรียนรู้ที่ให้อัลกอริทึมนั้นสามารถเรียนรู้ได้ด้วยตนเอง โดยไม่ต้องมีค่าเป้าหมายของแต่ละข้อมูล ซึ่งวิธีการคือใส่ข้อมูลต่าง ๆ และกำหนดสิ่งที่ต้องการจากข้อมูลเหล่านั้น ทำให้อัลกอริทึมวิเคราะห์จากการจำแนกและสร้างแบบแผนจากข้อมูลที่ได้รับมา เครื่องจักรนี้จะไม่สามารถเอาไปประมวล/จัดหมวดหมู่ (Classification) ได้แต่จะใช้วิธีการแบ่งกลุ่มแทน (Clustering)

2.2.3 การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning)

เป็นวิธีการเรียนรู้แบบหนึ่งที่ใช้การเรียนรู้เกิดมาจากการปฏิสัมพันธ์ (interaction) ระหว่างผู้เรียนรู้ (agent) กับสิ่งแวดล้อม (environment) ที่มีการเรียนรู้สิ่งต่าง ๆ จากผู้เรียนรู้ ภายใต้การเลือกกระทำสิ่งต่าง ๆ ให้ได้ผลลัพธ์ที่มากที่สุด

2.3 วิธีการจำแนก

2.3.1 วิธีการถดถอยโลจิสติก

เป็นวิธีในการศึกษาความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระว่ามีตัวแปรอิสระใดบ้างที่ส่งผลกับตัวแปรตาม ซึ่งวิธีการนี้มีหลักการคล้ายกับวิธีการถดถอยเชิงเส้น โดยมีวัตถุประสงค์เพื่อใช้ตัวแปรอิสระเหล่านี้ทำนายโอกาสที่จะเกิดเหตุการณ์ที่สนใจจากสมการถดถอยโลจิสติกที่เหมาะสม (กัลยา วานิชปัญญา, 2564) ซึ่งวิธีการถดถอยโลจิสติกนี้จะนำมาใช้ในกรณีที่ตัวแปรตามเป็นตัวแปรเชิงคุณภาพทั้งกรณีตัวแปรทวิภาค (Dichotomous Variable) และตัวแปรพหุวิภาค (Polytomous Variable)

2.3.1.1 การถดถอยโลจิสติกแบบทวิ (Binary Logistic Regression) ใช้เมื่อตัวแปรตามเป็นตัวแปรเชิงคุณภาพที่มีค่าได้เพียง 2 ค่า เช่น การยุติการบำบัดรักษาของผู้ติดยาเสพติด และการรับการบำบัดรักษาของผู้ติดยาเสพติดจนครบโปรแกรม

โดยปกติแล้วสมการถดถอยอย่างง่ายที่แสดงความสัมพันธ์ระหว่าง x กับ y จะอยู่ในรูปเชิงเส้น ดังนี้

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon \quad (2.4)$$

หรือ
$$E(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (2.5)$$

โดยที่ $-\infty < E(Y) < \infty$

แต่สำหรับการถดถอยโลจิสติก ความสัมพันธ์ระหว่าง x กับ y จะไม่ได้อยู่ในรูปเชิงเส้น แต่จะอยู่ในรูป

$$E(Y) = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}} \quad (2.6)$$

โดยที่ $E(Y)$ คือความน่าจะเป็นของเหตุการณ์ที่สนใจ และ $0 \leq E(Y) \leq 1$ เรียกว่า Logistic Function

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แต่ก็สามารถเขียนในรูปเชิงเส้นได้ ดังนี้

$$\text{ให้ Odds} = \frac{P(\text{เกิดเหตุการณ์})}{P(\text{ไม่เกิดเหตุการณ์})} \quad (2.7)$$

Odds หรือ Odds Ratio เป็นค่าที่แสดงถึงโอกาสที่จะเกิดเหตุการณ์เป็นกี่เท่าของโอกาสที่จะไม่เกิดเหตุการณ์ เช่น odds ratio = 3.5 แสดงว่าโอกาสที่จะเกิดเหตุการณ์เป็น 3.5 เท่า

$$\text{Odds} = \frac{E(Y)}{1 - E(Y)} \quad (2.8)$$

$$= \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}} \quad (2.9)$$

$$= \frac{1}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}} \quad (2.10)$$

$$\log(\text{odds}) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p \quad (2.11)$$

เมื่อ x_1, x_2, \dots, x_p คือ ตัวแปรอิสระ และ $\beta_0, \beta_1, \dots, \beta_p$ คือ สัมประสิทธิ์การถดถอย
ในการศึกษาตัวแบบการถดถอยโลจิสติก มีการเขียนสมมติฐานเพื่อการทดสอบ ดังนี้
 $H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$
 $H_1 : \text{มี } \beta_i \neq 0 \text{ อย่างน้อย 1 ค่า ; } i=1, 2, \dots, p$

2.3.2 เทคนิคต้นไม้ตัดสินใจ

เป็นเทคนิคที่ให้ผลลัพธ์ในลักษณะของโครงสร้างต้นไม้ซึ่งจะจำแนกข้อมูลต่าง ๆ โดยใช้คุณลักษณะ (Attribute) ของข้อมูลในการจำแนก (สุรวีชร ศรีเปารยะ, 2560) ต้นไม้ตัดสินใจประกอบด้วยส่วนหลัก ๆ 3 ส่วน ดังนี้

2.3.2.1 โหนด (Node)

คุณลักษณะต่าง ๆ ที่ใช้ในการทดสอบเพื่อจำแนกข้อมูล โดยโหนดที่อยู่สูงสุด เรียกว่า โหนดราก (root node)

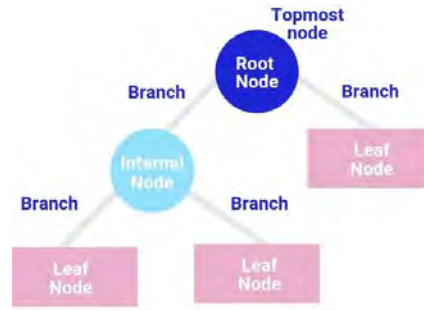
2.3.2.2 กิ่ง (Branch)

เป็นกิ่งของต้นไม้ที่แสดงถึงค่าที่เป็นไปได้ของคุณลักษณะที่ถูกทดสอบ

2.3.2.3 ใบ (Leaf)

กลุ่มของข้อมูลหรือผลลัพธ์ที่ได้จากการจำแนก เป็นสิ่งที่อยู่ล่างสุดของต้นไม้ตัดสินใจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.2 เทคนิคต้นไม้ตัดสินใจ (Decision tree) (Akash N H, 2020)

ต้นไม้ตัดสินใจมีค่าวัดความสามารถของการจัดกลุ่มของแต่ละคุณลักษณะ (ภรณ์ยา ปาลวิสุทธิ, 2559) ดังต่อไปนี้

Gini Index ค่าที่บ่งบอกว่าคุณลักษณะหรือปัจจัยใดควรนำมาใช้เป็นคุณลักษณะในการแบ่งกลุ่ม

$$Gini(t_i) = 1 - \sum_{i=1}^N [p(t_i)]^2 \tag{2.12}$$

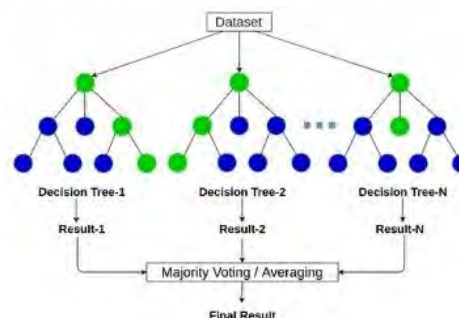
Entropy ค่าคาดคะเนของข้อมูลเป็นค่าที่แยกโดยใช้ลักษณะประจำของอัลกอริทึม

$$Entropy(t_i) = 1 - \sum_{i=1}^N [p(t_i)] \log p(t_i) \tag{2.13}$$

โดยที่ t_i คือ คุณลักษณะที่นำมาวัดค่า Entropy
 $P(t_i)$ คือ สัดส่วนของจำนวนสมาชิกกลุ่ม i กับจำนวนสมาชิกทั้งหมดของกลุ่มตัวอย่าง

2.3.3 เทคนิคป่าสุ่ม

เป็นเทคนิคที่ถูกพัฒนามาจากเทคนิคต้นไม้ตัดสินใจซึ่งจะทำการสร้างต้นไม้ตัดสินใจหลาย ๆ ต้น ที่มีคุณลักษณะและข้อมูลที่แตกต่างกัน โดยใช้ชุดข้อมูลใหม่ที่เกิดจากการสุ่มใหม่ (Resampling) จากข้อมูลทั้งหมดตามจำนวนที่ต้องการ ซึ่งผลลัพธ์ของเทคนิคป่าสุ่มสำหรับปัญหาการจำแนกประเภท จะได้จากการนำผลลัพธ์ของต้นไม้ตัดสินใจแต่ละต้นมาทำการโหวต (Majority Vote) และผลลัพธ์ที่มีผลโหวตมากที่สุดจะเป็นผลลัพธ์การจำแนกของเทคนิคป่าสุ่ม



รูปที่ 2.3 เทคนิคป่าสุ่ม (Random Forest) (Abhishek Jain, 2024)
 เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้เพื่อการศึกษาเท่านั้น เมื่อผู้เผยแพร่ข้อมูลไปใช้โดยไม่ผ่านการคัดค้าน
 ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.4 เทคนิคซ์พอร์ตเวกเตอร์แมชชีน

แนวคิดของซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM) คือ การค้นหาไฮเปอร์เพลน (Hyperplane) ที่สามารถแยกกลุ่มข้อมูล $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$ ออกจากกันได้ อย่างเหมาะสมที่สุด โดยคำนึงถึงการเพิ่มระยะห่าง (Margin) ให้มากที่สุด เทคนิคนี้สามารถใช้ได้กับข้อมูลที่เป็นเชิงเส้น (linear) และข้อมูลที่ไม่เป็นเชิงเส้น (non-linear) โดยข้อมูลที่ไม่เป็นเชิงเส้นจะใช้เคอร์เนลทริก (kernel trick) ในการเปลี่ยนข้อมูลไปยังปริภูมิใหม่ที่ทำให้ข้อมูลนั้นเป็นข้อมูลที่เป็นเชิงเส้น สำหรับกรณีพื้นฐานไฮเปอร์เพลนสามารถกำหนดได้ในรูปของ (W, b) ซึ่งสามารถเขียนเป็นปัญหาการหาค่าเหมาะที่สุด (Optimization Problem) (Jakkula Vikramaditya, 2006) ได้ดังนี้

$$\text{minimize}_{w,b} \frac{1}{2} \|w\|^2 \quad (2.14)$$

$$\text{ภายใต้เงื่อนไข} \quad y_i (\langle w, x_i \rangle + b) \geq 1, \quad i = 1, 2, \dots, m \quad (2.15)$$

โดยที่ W คือเวกเตอร์น้ำหนัก (Weight vector) และ b คือความคลาดเคลื่อน

ทั้งนี้ระยะห่างไฮเปอร์เพลนและจุดข้อมูลที่ใกล้ที่สุดสามารถคำนวณได้จาก

$$\gamma = \frac{1}{\|w\|^2} \quad (2.16)$$

2.4 การคัดเลือกตัวแปร

เป็นการเลือกตัวแปรอิสระที่มีอิทธิพลต่อตัวแปรตามเข้าสู่สมการถดถอยเพื่อเพิ่มความถูกต้องในการพยากรณ์ของตัวแบบ มี 4 เทคนิค (กัลยา วานิชบัญชา, 2564) ดังนี้

2.4.1 Enter method เป็นการนำตัวแปรอิสระทุกตัวเข้าสู่สมการถดถอย โดยผู้วิจัยต้องตัดสินใจว่าจะเลือกตัวแปรอิสระใดบ้าง ที่ควรจะเข้าไปในสมการถดถอย โดยพิจารณาจากค่าความสัมพันธ์ของตัวแปรอิสระ หรือ ค่าสถิติทดสอบ หรือ ค่านัยสำคัญทางสถิติ (Significance)

2.4.2 Forward Selection เป็นการคัดเลือกตัวแปรอิสระเข้าสู่สมการถดถอยครั้งละ 1 ตัว โดยจะทำการเลือกหลาย ๆ ครั้ง ในแต่ละครั้งจะเลือกตัวแปรอิสระเข้าสู่สมการเพียง 1 ตัว โดยเลือกตัวแปรอิสระตัวที่ทำให้ความสัมพันธ์กับตัวแปรตาม Y เพิ่มขึ้นมากที่สุด และจะหยุดนำตัวแปรอิสระเข้าสู่สมการ เมื่อนำเข้ามาเพิ่มแล้วไม่ได้ทำให้ความสัมพันธ์กับตัวแปรตามเพิ่มมากขึ้น

2.4.3 Backward Elimination เป็นวิธีการเลือกตัวแปรอย่างตรงข้ามกับวิธี Forward Selection โดยจะทำการตัดตัวแปรอิสระที่ไม่มีความสัมพันธ์กับ Y ออกจากสมการความถดถอยครั้งละ 1 ตัว ทำไปเรื่อย ๆ จนไม่สามารถตัดตัวแปรอิสระใดออกจากสมการได้จึงหยุด

2.4.4 Stepwise Selection เป็นวิธีการเลือกตัวแปรอิสระเข้าสู่สมการถดถอย วิธีการนี้ นิยมใช้มากวิธีนี้เป็นการผสมผสานระหว่าง Forward Selection และ Backward Elimination โดยจะทำการเลือกตัวแปรอิสระตัวแรกที่มีความสัมพันธ์กับตัวแปรตามมากที่สุดเข้าสู่สมการ จากนั้นเลือกตัวแปรอิสระตัวที่สองที่มีความสัมพันธ์กับตัวแปรตามเข้าสู่สมการแล้วใช้วิธีการของ Forward Selection พิจารณาตัวแปรอิสระ ในขณะที่เดียวกันจะพิจารณาตัวแปรอิสระตัวแรกที่อยู่ในสมการ

ว่าสมควรจะถูกเลือกออกจากสมการหรือไม่ ตามวิธีการของ Backward Elimination เพื่อป้องกัน

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้เผยแพร่โดยระบบออนไลน์การศึกษ

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

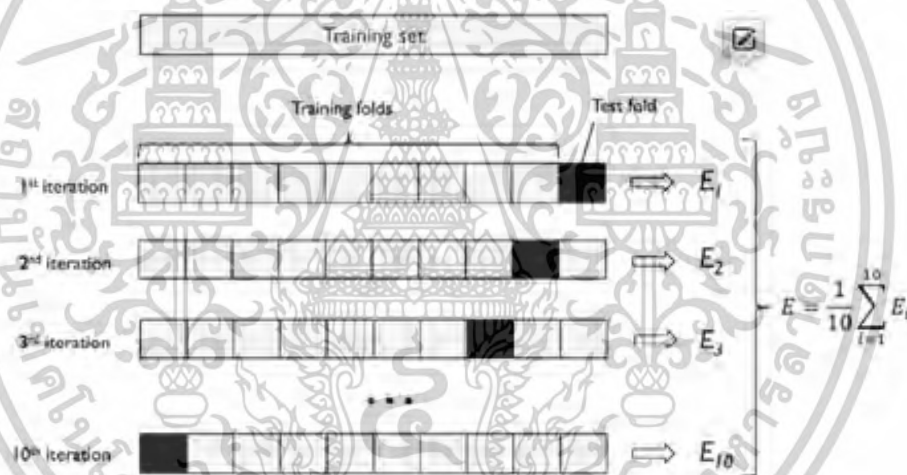
ไม่ให้นำตัวแปรอิสระเดิมเข้าและออกจากสมการ ทำจนกระทั่งไม่สามารถเลือกตัวแปรอิสระใดเข้าและไม่สามารถตัดตัวแปรอิสระใดออกจากสมการได้อีก

2.5 การวัดประสิทธิภาพ

การวัดประสิทธิภาพตัวแบบการจำแนกเป็นขั้นตอนที่มีความสำคัญ เพื่อให้ทราบถึงประสิทธิภาพของตัวแบบที่ถูกสร้างขึ้นมา รวมทั้งสามารถเปรียบเทียบแบบจำลองเพื่อเลือกตัวแบบที่ดีที่สุดในการนำไปใช้ประมวลผลข้อมูล โดยมีวิธีการวัดประสิทธิภาพของตัวแบบ ดังนี้

2.5.1. การวัดประสิทธิภาพตัวแบบด้วยวิธีการตรวจสอบไขว้แบบ K รอบ

เป็นการทดสอบประสิทธิภาพของตัวแบบหลังจากทำการสร้างตัวแบบเพื่อป้องกันการเกิด Overfitting โดยมีหลักการทำงานคือนำข้อมูลชุดสอนที่ใช้สร้างตัวแบบมาทำการแบ่งข้อมูลออกเป็น K ส่วนเท่า ๆ กัน เพื่อใช้เป็นข้อมูลชุดสอนและข้อมูลชุดทดสอบ เช่น 10 - Fold Cross Validation โดยข้อมูลทั้งหมดจะถูกแบ่งเป็น 10 ส่วนเท่า ๆ กัน ใช้ 9 ส่วน เป็นชุดสอน และ 1 ส่วน เป็นชุดทดสอบ ซึ่งจะทำเช่นนี้ไปเรื่อย ๆ จนครบทั้ง 10 ส่วน จากนั้นนำผลลัพธ์ที่ได้ในแต่ละรอบมาทำการหาผลโหวตเพื่อใช้ทดสอบประสิทธิภาพของตัวแบบ



รูปที่ 2.4 วิธีการแบบ 10-Fold Cross Validation (Anubha Singh, 2020)

2.5.2 การวัดประสิทธิภาพตัวแบบด้วยเมทริกซ์สับสน (Confusion Matrix)

เป็นการใช้ค่าที่ได้จากเมทริกซ์สับสน (Confusion Matrix) ซึ่งเป็นตารางสรุปจำนวนข้อมูลที่ได้จากการจำแนกโดยตัวแบบการจำแนก มีทั้งข้อมูลที่จำแนกได้อย่างถูกต้องและไม่ถูกต้อง (ศรีณย์พร กาญจนสุนทร, 2566)

ตารางที่ 2.1 เมทริกซ์สับสน (Confusion Matrix)

Predicted	Actual	
	Positive (1)	Negative (0)
Positive (1)	TP	FP
Negative (0)	FN	TN

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับวิธีในการวัดประสิทธิภาพของตัวแบบการจำแนก มีดังต่อไปนี้

- ค่าความถูกต้อง (Accuracy) เป็นการวัดประสิทธิภาพของตัวแบบโดยรวม เป็นวิธีที่นิยมใช้มากที่สุดสำหรับการประเมินประสิทธิภาพการจำแนกประเภท มีการคำนวณดังนี้

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (2.17)$$

- ค่าความเที่ยง (Precision) หรือค่าทำนายผลบวก (Positive Predictive Value: PPV) คือ อัตราส่วนระหว่างผลลัพธ์การทำนายและผลลัพธ์จริงเป็นจริง (TP) ต่อจำนวนตัวอย่างที่ผลลัพธ์ทำนายเป็นจริงทั้งหมด มีการคำนวณดังนี้

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2.18)$$

- ค่าความครบถ้วน (Recall) หรืออัตราผลบวกจริง (True Positive Rate: TPR) คือจำนวนตัวอย่างที่เป็นบวกที่ทำนายได้ถูกต้องเทียบกับจำนวนตัวอย่างที่เป็นบวกทั้งหมด มีการคำนวณดังนี้

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2.19)$$

- คะแนน F1 (F1-score) คือค่าเฉลี่ยฮาร์โมนิก (Harmonic mean) ของ Precision และ Recall มีการคำนวณดังนี้

$$\text{F1 - Score} = 2 \times \left(\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \right) \quad (2.20)$$

โดยที่ True Positive (TP) คือ จำนวนข้อมูลที่ทำนายได้ตรงกับสิ่งที่เกิดขึ้น กรณีทำนายว่าเกิดเหตุการณ์ที่สนใจและเหตุการณ์ที่สนใจเกิดขึ้นจริง

True Negative (TN) คือ จำนวนข้อมูลที่ทำนายได้ตรงกับสิ่งที่เกิดขึ้น กรณีทำนายว่าไม่เกิดเหตุการณ์ที่สนใจและเหตุการณ์ที่สนใจไม่เกิดขึ้นจริง

False Positive (FP) คือ จำนวนข้อมูลที่ทำนายได้ไม่ตรงกับสิ่งที่เกิดขึ้น กรณีทำนายว่าจะเกิดเหตุการณ์ที่สนใจแต่ความเป็นจริงคือไม่เกิดเหตุการณ์ที่สนใจ

False Negative (FN) คือ จำนวนข้อมูลที่ทำนายได้ไม่ตรงกับสิ่งที่เกิดขึ้น กรณีทำนายว่าจะไม่เกิดเหตุการณ์ที่สนใจแต่ความเป็นจริงคือเกิดเหตุการณ์ที่สนใจ

2.6 งานวิจัยที่เกี่ยวข้อง

ชลอวัฒน์ อินปา และพิศิษฐ์ ศรีประเสริฐ (2564) ได้ศึกษาปัจจัยที่ส่งผลต่อการเข้ารับ การบำบัดรักษาครบตามเกณฑ์ของผู้ป่วยยาเสพติดในจังหวัดตาก โดยใช้สถิติเชิงพรรณนาและสถิติ เชิงอนุมาน ได้แก่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน สถิติ Chi-square test และการวิเคราะห์ ฤทธอยพหุโลจิสติกส์ ในการวิเคราะห์ข้อมูลผู้ป่วยยาเสพติดในจังหวัดตากที่เข้ารับการรักษา เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่โดยไม่ ได้รับความเห็นชอบจากเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปีงบประมาณ 2563 จำนวน 1,293 ราย จากระบบข้อมูลการบำบัดรักษาและฟื้นฟูผู้ติดยาเสพติดของ ประเทศ (บสต.) กระทรวงสาธารณสุข ผลการศึกษาพบว่า ปัจจัยที่ส่งผลต่อการเข้ารับการรักษา ครอบคลุมเกณฑ์ของผู้ป่วยยาเสพติด ได้แก่ (1) ปัจจัยส่วนบุคคล คือ สถานภาพสมรส และยาเสพติดหลักที่ใช้เสพ (2) ปัจจัยด้านการรักษา คือ ความสมัครใจ และวิธีการรักษาด้วยเมทาโดน (3) ปัจจัยด้านการสนับสนุนทางสังคม คือ ความสัมพันธ์ในครอบครัวที่ยอมรับและช่วยเหลือ

สมพร สุวรรณมาโจ และคณะ (2562) ได้ศึกษาปัจจัยทำนายในการคงอยู่ในระบบระยะเวลา 1 ปี ของผู้บำบัดด้วยเมทาโดนระยะยาวในคลินิกเมทาโดน ณ สถาบันบำบัดรักษาและฟื้นฟูผู้ติดยาเสพติดแห่งชาติบรมราชชนนี โดยใช้สถิติเชิงพรรณนาในรูปแบบของความถี่ ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน วิเคราะห์หาตัวแปรที่มีความสัมพันธ์ของผู้ที่ได้รับยาเมทาโดนด้วย Chi-square และวิเคราะห์ปัจจัยทำนายด้วยการวิเคราะห์การถดถอยโลจิสติกส์ วิเคราะห์ข้อมูลที่เก็บจากเวชระเบียนของผู้ป่วยเสพยาเสพติดสารกลุ่มโอปิออยด์ที่รับการรักษาด้วยเมทาโดนระยะยาวและแบบรายงานการบำบัดรักษาผู้ติดยาเสพติด (บสต.) ระหว่างปีงบประมาณ 2558 จนถึง 2561 จำนวน 173 ราย ผลการศึกษาพบว่า ปัจจัยทำนายในการคงอยู่ในระบบระยะเวลา 1 ปี ของผู้บำบัดด้วยเมทาโดนระยะยาวในคลินิกเมทาโดน ณ สถาบันบำบัดรักษาและฟื้นฟูผู้ติดยาเสพติดแห่งชาติบรมราชชนนี คือ ผู้ป่วยมีการประกอบอาชีพ และมีรายได้ต่อเดือนมากกว่า 10,000 บาท

จรรยาตรี โคมพุดชา และชัยภัทร อีร์ชาญไชย (2558) ได้ศึกษาปัจจัยที่มีความสัมพันธ์กับการบำบัดครบกำหนดของผู้ใช้สารเสพติดในระบบสมัครใจ จังหวัดนครราชสีมา ปีงบประมาณ 2555-2556 โดยวิเคราะห์และนำเสนอข้อมูลด้วยสถิติเชิงพรรณนา เป็นจำนวน ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน เปรียบเทียบความแตกต่างของสัดส่วนของกลุ่มตัวอย่าง 2 กลุ่ม คือ กลุ่มบำบัดครบกำหนด และกลุ่มบำบัดไม่ครบกำหนด ด้วยสถิติ Chi-square test และ Fisher's exact test เปรียบเทียบความแตกต่างของค่าเฉลี่ยหรือมัธยฐานของตัวอย่าง 2 กลุ่ม ด้วยสถิติ Kruskal Wallis Test โดยกำหนดค่านัยสำคัญทางสถิติที่ 0.05 โดยใช้ข้อมูลการบำบัดผู้ใช้สารเสพติดในระบบสมัครใจ ในสถานบริการสาธารณสุขของรัฐ จังหวัดนครราชสีมา ปีงบประมาณ 2555-2556 จากกระบบรายงานระบบติดตามและเฝ้าระวังปัญหาสารเสพติดของกระทรวงสาธารณสุข (บสต.) จำนวน 8,784 ราย ผลการศึกษาพบว่า ผู้เข้ารับการรักษาได้รับการบำบัดครบกำหนด 8,187 ราย (ร้อยละ 93.2) ปัจจัยที่มีความสัมพันธ์กับการบำบัดครบกำหนดของผู้รับการรักษา ได้แก่ ที่อยู่อาศัย วิธีการใช้สารเสพติด สาเหตุสำคัญในการใช้สารเสพติด เหตุผลในการเข้ารับการรักษา รูปแบบวิธีการบำบัด และจำนวนวันบำบัดรักษา

เสาวนีย์ พันธุ์พัฒนกุล และณัฐณีย์ มีมนต์ (2558) ได้ศึกษาปัจจัยที่สัมพันธ์กับการบรรลุผล กระบวนการบำบัดรักษาฟื้นฟูสมรรถภาพและติดตามของผู้รับการรักษาติดยาเสพติด จังหวัดพะเยา ทำการวิเคราะห์และแสดงผลข้อมูลสถิติเชิงพรรณนาด้วยการกระจายความถี่แบบร้อยละและทดสอบความสัมพันธ์ระหว่างปัจจัยต่างๆกับการบรรลุผลกระบวนการบำบัดรักษาและติดตามผู้เสพยาเสพติด ด้วยสถิติ Chi-square หรือ Fisher' Exact Test กำหนดค่านัยสำคัญทางสถิติที่ 0.05 โดยใช้ข้อมูลผู้เสพยาเสพติดที่เข้ารับการรักษาและฟื้นฟูสมรรถภาพในสถานบริการสาธารณสุขทุกแห่งที่ตั้งอยู่ในเขตจังหวัดพะเยา ระหว่างปีงบประมาณ 2551-2556 จากกระบบติดตามและเฝ้าระวังปัญหาสารเสพติดของกระทรวงสาธารณสุข (บสต.) จำนวน 8,408 ราย ผลการศึกษาพบว่า ปัจจัยที่สัมพันธ์กับการบรรลุผลกระบวนการบำบัดรักษาฟื้นฟูสมรรถภาพและติดตามผู้ติดยาเสพติดที่สำคัญ ได้แก่

เพศ อายุ สถานภาพสมรส ระดับการศึกษา เหตุผลที่เข้ารับการรักษา และสถานการณ์ยาเสพติด เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อังกูร ภัทรารกร และ นันทา ชัยพิชิตพันธ์ (2558) ได้ศึกษาการเปรียบเทียบผลการบำบัดฟื้นฟูสมรรถภาพแบบเข้มข้นทางสายใหม่ (FAST Model) ระหว่างผู้ป่วยเสพติดระบบสมัครใจกับระบบบังคับบำบัด:กรณีศึกษาสถาบันธัญญารักษ์ ทำการวิเคราะห์ข้อมูลส่วนบุคคลและนำเสนอเป็นจำนวน ร้อยละ ค่าส่วนเบี่ยงเบนมาตรฐาน ส่วนความสัมพันธ์ระหว่างปัจจัยต่างๆกับการเลิกใช้ยาเสพติด ใช้สถิติวิเคราะห์ คือ chi-square, Fisher's exact test และ Odds ratio โดยใช้ข้อมูลย้อนหลังจากเวชระเบียนผู้ป่วยใน เป็นผู้ป่วยโรคติดยาเสพติดทั้งเพศชายและหญิงที่เข้ารับการบำบัดในระบบสมัครใจและบังคับบำบัดตามรูปแบบ FAST Model ในสถาบันธัญญารักษ์ ตั้งแต่วันที่ 1 ตุลาคม 2554 - 30 กันยายน 2555 จำนวน 368 ราย ผลการศึกษาพบว่า ผู้ป่วยที่เข้ารับการบำบัดรักษาในระบบสมัครใจมีโอกาสเลิกใช้ยาเสพติดมากกว่าผู้ป่วยที่เข้ารับการบำบัดในระบบบังคับบำบัด 1.56 เท่า

กนกลักษณ์ ศิริรุ่งวัฒนากุล (2562) ได้ศึกษาประสิทธิผลและปัจจัยที่มีผลต่อการบำบัดรักษาผู้ติดยาและสารเสพติดโดยใช้รูปแบบการบำบัดความคิดและพฤติกรรมแบบประยุกต์ของโรงพยาบาลสวรรค์ประชารักษ์ ทำการวิเคราะห์ข้อมูลด้วยสถิติเชิงพรรณนา และวิเคราะห์ปัจจัยที่สัมพันธ์กับประสิทธิผลของการบำบัดด้วยสถิติ Chi-square หรือ Fisher's Exact Test โดยกำหนดระดับนัยสำคัญทางสถิติที่ P น้อยกว่า 0.05 โดยใช้ข้อมูลผู้ป่วยยาเสพติดที่มารับการบำบัดรักษาครั้งแรกที่แผนกผู้ป่วยนอกของกลุ่มงานจิตเวชและยาเสพติดโรงพยาบาลสวรรค์ประชารักษ์ในช่วงระยะเวลาตั้งแต่ 1 กรกฎาคม 2555 ถึง 30 มกราคม 2561 และบำบัดครบโปรแกรม จากฐานข้อมูลระบบรายงาน ระบบติดตาม และการเฝ้าระวังปัญหาเสพติดของกระทรวงสาธารณสุข (บสต.) ผลการศึกษาพบว่า ผู้ป่วยหยุดเสพที่ 3 เดือน หลังครบการบำบัดร้อยละ 76.9 ปัจจัยที่สัมพันธ์กับประสิทธิผลการบำบัดรักษา คือ การบำบัดรักษาแบบสมัครใจ ($P=0.038$)

อารี สุภาวงศ์ (2559) ได้ศึกษาประสิทธิผลของการบำบัดรักษาผู้ป่วยยาเสพติดตามรูปแบบจิตสังคมบำบัดแบบประยุกต์ โรงพยาบาลทุ่งสง ทำการวิเคราะห์ข้อมูลด้วยสถิติเชิงพรรณนา และวิเคราะห์ปัจจัยที่สัมพันธ์กับประสิทธิผลของการบำบัดด้วยสถิติ Chi-square โดยใช้ข้อมูลผู้ป่วยยาเสพติดที่มารับการบำบัดรักษาที่โรงพยาบาลทุ่งสง ตั้งแต่ 1 ตุลาคม 2556 ถึง 30 กันยายน 2557 จำนวน 279 คน จากแบบเก็บรวบรวมข้อมูลที่ผู้วิจัยสร้างขึ้น และแบบรายงานการบำบัดรักษาผู้ใช้สารเสพติด ส่วนที่ 3 (บสต.3) ของผู้ป่วยเฉพาะราย ผลการศึกษาพบว่า ผู้ป่วยยาเสพติดร้อยละ 50.5 มีประสิทธิผลการบำบัด เมื่อทดสอบความสัมพันธ์พบว่า สถานภาพสมรสมีความสัมพันธ์กับประสิทธิผลการบำบัดอย่างมีนัยสำคัญทางสถิติ ($p=0.001$)

K. M. Rashedul et al. (2021) ได้ศึกษาการเปรียบเทียบการเรียนรู้ของเครื่องเพื่อที่จะทำนายผู้ติดยาเสพติดในบังคลาเทศและปัจจัยที่สัมพันธ์กับการติดยาเสพติด ทำการวิเคราะห์ข้อมูลด้วยสถิติ Chi-square และ PCA (Principal Component Analysis) พร้อมทั้งเปรียบเทียบประสิทธิภาพการเรียนรู้ของเครื่องทั้ง 7 อัลกอริทึม ได้แก่ การถดถอยโลจิสติก, ต้นไม้ตัดสินใจ, ป่าสุ่ม, ซัพพอร์ตเวกเตอร์แมชชีน, Naive Bayes (Bernoulli), Naive Bayes (Gaussian) และ XGBoost โดยใช้ข้อมูลผู้ติดยาเสพติดจากสถานบำบัดและผู้ไม่ติดยาเสพติดจากมหาลัยต่างๆในเมืองธากา ประเทศบังคลาเทศ เป็นเพศชายที่มีอายุ 17 ถึง 45 ปี จำนวน 188 คน ผลการศึกษาพบว่า อัลกอริทึมป่าสุ่มมีค่าความถูกต้องมากที่สุด 97.3484% และปัจจัยที่ส่งผลต่อการติดยาเสพติด คือ อายุ สถานะครอบครัว จำนวนสมาชิกในครอบครัว และการอยู่ข้างนอกในตอนกลางคืน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

วิธีการดำเนินงานวิจัย

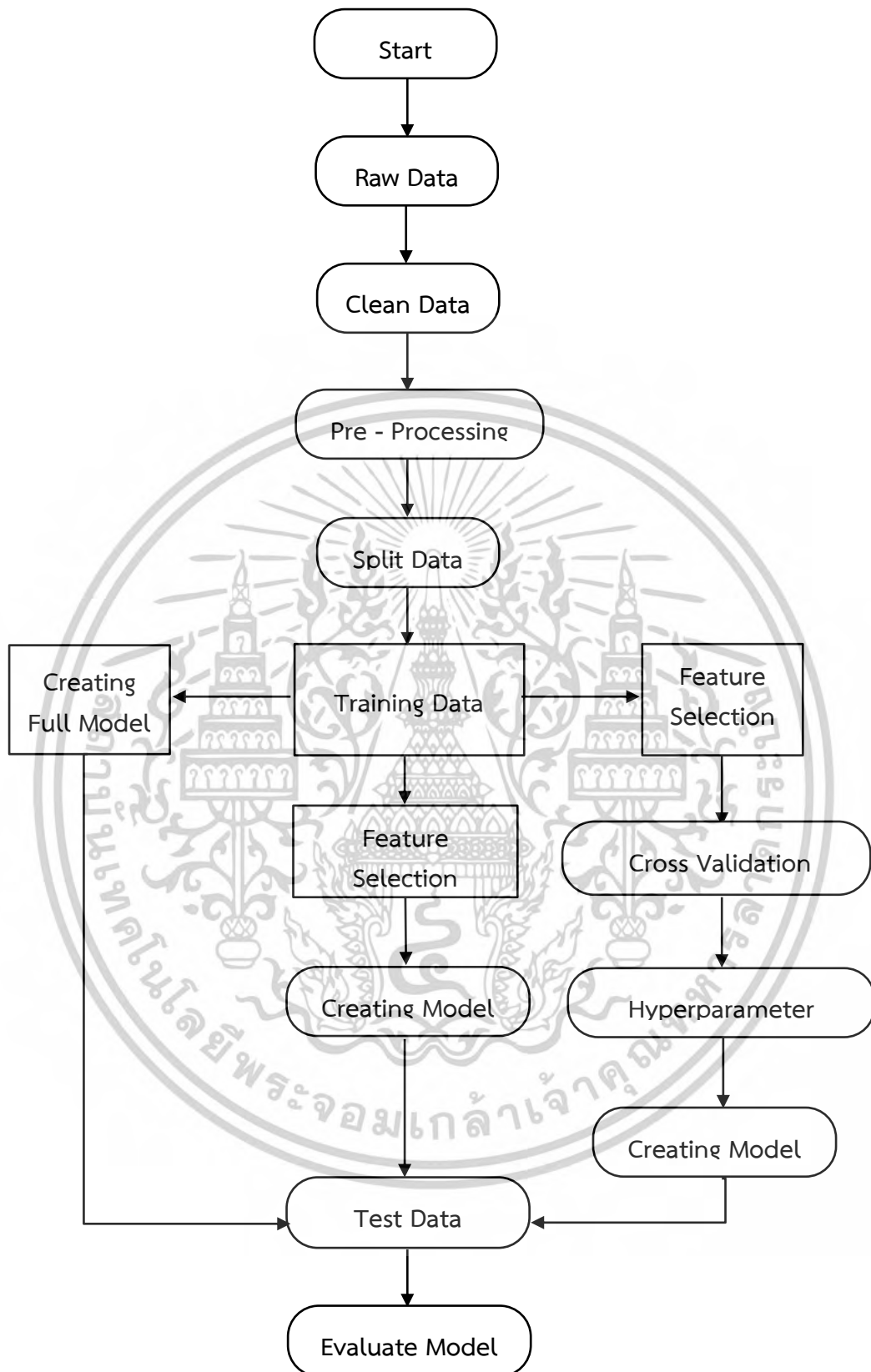
ในบทนี้จะเป็นการกล่าวถึงขั้นตอนการดำเนินการวิจัย ซึ่งขั้นตอนการดำเนินงานประกอบด้วย

- 3.1 โครงสร้างของขั้นตอนและวิธีการดำเนินงานวิจัย
- 3.2 การเก็บรวบรวมข้อมูล
- 3.3 การเตรียมข้อมูล
- 3.4 การวิเคราะห์ตัวแปร
- 3.5 การสร้างตัวแบบ
- 3.6 การประเมินประสิทธิภาพตัวแบบ

3.1 โครงสร้างของขั้นตอนและวิธีการดำเนินงานวิจัย

งานวิจัยนี้เริ่มต้นจากการนำข้อมูลผู้เข้ารับการรักษาของผู้ติดยาเสพติดระบบสมัครใจในพื้นที่กรุงเทพมหานครที่ได้จากการเก็บรวบรวมข้อมูลมาทำความสะอาดข้อมูล (Clean Data) และเตรียมข้อมูล (Pre-Processing) เพื่อให้ข้อมูลพร้อมที่จะนำไปวิเคราะห์ในโปรแกรม R โดยเริ่มจากการวิเคราะห์ตัวแปรอิสระก่อน จากนั้นจึงทำการสร้างตัวแบบโดยเริ่มจากการสุ่มตัวอย่างแบบชั้นภูมิ (Stratified Sampling) และแบ่งข้อมูล (Spilt Data) จำนวน 3,241 ข้อมูล ออกเป็น 2 ส่วน คิดเป็นสัดส่วน 70:30 โดยข้อมูลชุดสอน (Training Set) จำนวน 2,268 ข้อมูล ใช้สำหรับสร้างตัวแบบ และข้อมูลชุดทดสอบ (Test Set) จำนวน 973 ข้อมูล ซึ่งใช้สำหรับเปรียบเทียบประสิทธิภาพของตัวแบบ

เมื่อแบ่งข้อมูลเรียบร้อยแล้วก็นำข้อมูลชุดสอนมาทำการสร้างตัวแบบโดยใช้การเรียนรู้ของเครื่อง 3 เทคนิค ได้แก่ ตัวแบบการถดถอยโลจิสติก ตัวแบบป่าสุ่ม และตัวแบบซัพพอร์ตเวกเตอร์แมชชีน ในรูปแบบที่แตกต่างกัน แบ่งเป็นตัวแบบที่ใช้ตัวแปรอิสระครบทุกตัว, ตัวแบบที่ใช้ตัวแปรอิสระไม่ครบทุกตัวจากการคัดเลือกตัวแปร และตัวแบบที่ใช้ตัวแปรอิสระไม่ครบทุกตัวจากการคัดเลือกตัวแปรร่วมกับการปรับค่าไฮเปอร์พารามิเตอร์แบบ Grid Search พร้อมทั้งนำทุกตัวแบบมาทำการตรวจสอบไขว้แบบ 10 รอบ เพื่อประเมินประสิทธิภาพ จากนั้นจึงนำตัวแบบแต่ละตัวแบบมาทำการประเมินประสิทธิภาพอีกครั้งด้วยเมทริกซ์สับสน (Confusion Matrix) โดยใช้ข้อมูลชุดทดสอบ รวมทั้งเปรียบเทียบประสิทธิภาพของแต่ละตัวแบบโดยวัดจากค่าคะแนน F1 ดังรูปที่ 3.1



รูปที่ 3.1 ขั้นตอนและวิธีการดำเนินงานวิจัย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 การเก็บรวบรวมข้อมูล

งานวิจัยนี้ใช้ข้อมูลทุติยภูมิของผู้เข้ารับการรักษาของผู้ติดยาเสพติดระบบสมัครใจในพื้นที่กรุงเทพมหานคร จากระบบข้อมูลการบำบัดรักษาและฟื้นฟูผู้ติดยาเสพติดของประเทศ (บสต.) กระทรวงสาธารณสุข เป็นระยะเวลา 6 ปีงบประมาณ ตั้งแต่ ตุลาคม พ.ศ. 2560 - กันยายน พ.ศ. 2566 โดยเลือกศึกษาเฉพาะข้อมูลการบำบัดรักษาครั้งสุดท้าย (ในกรณีที่ผู้ป่วยเข้ามาบำบัดหลายครั้งในช่วง 6 ปีงบประมาณ ที่มีรหัสการเข้ารับบำบัดหลายรหัส) ที่มีการระบุว่าบำบัดครบกำหนด และบำบัดไม่ครบกำหนด มีรูปแบบการบำบัดรักษาแบบผู้ป่วยนอกในสถานพยาบาล และมีข้อมูลในระบบการบำบัดรักษาและฟื้นฟูผู้ติดยาเสพติดครบถ้วน จำนวน 5,190 คน

3.3 การเตรียมข้อมูล

ทำการตรวจสอบรายละเอียดความครบถ้วนของข้อมูลผู้เข้ารับการรักษาของผู้ติดยาเสพติดระบบสมัครใจในพื้นที่กรุงเทพมหานคร จากระบบข้อมูลการบำบัดรักษาและฟื้นฟูผู้ติดยาเสพติดของประเทศ (บสต.) กระทรวงสาธารณสุข จำนวน 5,190 คน จากนั้นทำความสะอาดข้อมูล (Data Cleansing) โดยตัดข้อมูลที่ไม่ถูกต้อง และไม่สมบูรณ์ออกจากชุดข้อมูล เหลือข้อมูลที่นำมาใช้ในการวิเคราะห์ จำนวน 3,241 คน จากนั้นจัดการข้อมูล (Data Processing) โดยเปลี่ยนข้อมูลให้อยู่ในรูปแบบที่ต้องการ ซึ่งข้อมูลที่ได้ประกอบไปด้วยข้อมูลตัวแปรทั้งหมด 16 ตัวแปร ดังตารางที่ 3.1

ตารางที่ 3.1 รายละเอียดของข้อมูลตัวแปร

ลำดับ	ข้อมูลตัวแปร (Variables)	คำอธิบายข้อมูลตัวแปร (Description)
1	Gender (เพศ)	Male คือ ชาย (gender01) Female คือ หญิง (gender02)
2	Age (อายุ)	10 – 19 คือ อายุ 10 ถึง 19 ปี (age01) 20 – 29 คือ อายุ 20 ถึง 29 ปี (age02) 30 – 39 คือ อายุ 30 ถึง 39 ปี (age03) 40 – 49 คือ อายุ 40 ถึง 49 ปี (age04) 50 years or more คือ อายุตั้งแต่ 50 ปีขึ้นไป (age05)
3	Marital_status (สถานภาพสมรส)	Divorce/Widowed/Separated คือ หย่า/หม้าย/ แยกกันอยู่ (marital01) Married คือ สมรส (marital02) Single คือ โสด (marital03)
4	Education (การศึกษา)	Primary คือ ประถมศึกษา (edu01) Junior_High_School คือ มัธยมศึกษาตอนต้น (edu02) Senior_High_School คือ มัธยมศึกษาตอนปลาย (edu03) Certificate คือ อนุปริญญา หรือ ปวช.,ปวส. (edu04) Bachelor_Degrees_upper คือ ตั้งแต่ปริญญาตรีขึ้นไป (edu05) Uneducation คือ ไม่มีการศึกษา (edu06)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.1 (ต่อ) รายละเอียดของข้อมูลตัวแปร

ลำดับ	ข้อมูลตัวแปร (Variables)	คำอธิบายข้อมูลตัวแปร (Description)
5	Occupation (อาชีพ)	Civil_servant/State_enterprise คือ ข้าราชการ/รัฐวิสาหกิจ (occupa01) Merchant คือ พ่อค้า/แม่ค้า (occupa02) Office_and_factory_worker คือ พนักงานออฟฟิศ/ พนักงานโรงงาน (occupa03) Student คือ นักเรียน (occupa04) Work_for_hire/labor/communication/agriculturist คือ รับจ้าง/ผู้ใช้แรงงาน/พนักงานคมนาคม/เกษตรกร (occupa05) Unemployed คือ ไม่มีอาชีพ (occupa06) Others คือ อาชีพอื่น ๆ นอกเหนือจากที่กล่าวมา (occupa07)
6	Income_per_month (รายได้ต่อเดือน : บาท)	less_than_or_equal_to_5,000 คือ น้อยกว่า หรือ เท่ากับ 5,000 บาท (income01) 5,001-10,000 คือ 5,001 ถึง 10,000 บาท (income02) 10,001-15,000 คือ 10,001 ถึง 15,000 บาท (income03) more_than_15,001 คือ มากกว่า หรือเท่ากับ 15,001 บาท (income04) no_income คือ ไม่มีรายได้ (income05)
7	live with whom in 30 days ago (อาศัยอยู่กับใครในช่วง 30 วัน ที่ผ่านมา)	Alone คือ อาศัยอยู่คนเดียว (live01) father mother and others คือ อยู่กับพ่อ แม่ และคนอื่น ๆ (live02) father or mother and others คือ อยู่กับพ่อหรือแม่ และคนอื่น ๆ (live03) father_and_mother คือ อยู่กับพ่อและแม่ (live04) father_or_mother คือ อยู่กับพ่อหรือแม่ (live05) friend คือ อยู่กับเพื่อน (live06) relative คือ อยู่กับญาติ (ที่ไม่ใช่พ่อแม่) (live07) spouse_and_others คือ อยู่กับคู่สมรสและคนอื่น ๆ (live08) spouse_or_child คือ อยู่กับคู่สมรสหรือบุตร (live09)
	Family relation (ความสัมพันธ์ในครอบครัว)	divorce/separated คือ หย่าร้าง/แยกกันอยู่ (fam01) father_and_mother_died คือ พ่อและแม่เสียชีวิต (fam02) father_or_mother_died คือ พ่อหรือแม่เสียชีวิต (fam03) stay together smoothly คือ อยู่ร่วมกันอย่างราบรื่น (fam04) stay together not smoothly คือ อยู่ร่วมกันอย่าง ไม่ราบรื่น (fam05)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้ใช้ประโยชน์ด้านการศึกษา

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.1 (ต่อ) รายละเอียดของข้อมูลตัวแปร

ลำดับ	ข้อมูลตัวแปร (Variables)	คำอธิบายข้อมูลตัวแปร (Description)
9	age_first_time_used_drug (อายุที่ใช้ยาเสพติดครั้งแรก)	10 – 19 คือ อายุ 10 ถึง 19 ปี (agefirst01) 20 – 29 คือ อายุ 20 ถึง 29 ปี (agefirst02) 30 – 39 คือ อายุ 30 ถึง 39 ปี (agefirst03) 40 – 49 คือ อายุ 40 ถึง 49 ปี (agefirst04) 50 years or more คือ อายุตั้งแต่ 50 ปีขึ้นไป (agefirst05)
10	drug_first_time (ยาเสพติดที่ใช้ครั้งแรก)	Cannabis คือ กัญชา (drugfirst01) Heroin คือ เฮโรอีน (drugfirst02) Ice คือ ยาไอซ์ (drugfirst03) Ketamine คือ ยาเค (drugfirst04) Kratom คือ กระต่อม (drugfirst05) Others คือ ยาอื่น ๆ นอกเหนือจากที่กล่าวมา (drugfirst06) Amphetamine คือ ยาบ้า หรือ แอมเฟตามีน (drugfirst07)
11	cause_use_drugs(first_time) (สาเหตุการใช้ยาเสพติดครั้งแรก)	be_cured คือ ทำให้หายป่วย (causeuse01) friend_invited คือ เพื่อนชักชวน (causeuse02) funny คือ เพื่อความสนุก (causeuse03) help_in_work คือ ช่วยในการทำงาน (causeuse04) uncomfortable คือ ไม่สบายใจ (causeuse05) want_to_try คือ อยากลอง (causeuse06)
12	drugs(present) (ยาเสพติดที่ใช้ในปัจจุบัน)	Cannabis คือ กัญชา (drugsresent01) Heroin คือ เฮโรอีน (drugsresent02) Ice คือ ยาไอซ์ (drugsresent03) Ketamine คือ ยาเค (drugsresent04) Kratom คือ กระต่อม (drugsresent05) Others คือ ยาอื่น ๆ นอกเหนือจากที่กล่าวมา (drugsresent06) Amphetamine คือ ยาบ้า หรือ แอมเฟตามีน (drugsresent07)
13	time_use_year (ระยะเวลาที่ใช้ยาเสพติด)	less_than_or_equal_to_1 คือ น้อยกว่า หรือ เท่ากับ 1 ปี (timeuse01) >1-5 คือ มากกว่า 1 ปี ถึง 5 ปี (timeuse02) >5-10 คือ มากกว่า 5 ปี ถึง 10 ปี (timeuse03) >10-15 คือ มากกว่า 10 ปี ถึง 15 ปี (timeuse04) more than 15 years คือ มากกว่า 15 ปี (timeuse05)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้เพื่อการศึกษาเท่านั้น เมื่อผู้เผยแพร่เห็นเป็นประโยชน์ต่อการนำเอกสารนี้ไปใช้

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.1 (ต่อ) รายละเอียดของข้อมูลตัวแปร

ลำดับ	ข้อมูลตัวแปร (Variables)	คำอธิบายข้อมูลตัวแปร (Description)
14	ever_received_treatment (เคยรับการบำบัดรักษาหรือไม่)	Yes คือ เคย (evertreatment01) No คือ ไม่เคย (evertreatment02)
15	cause_receive_treatment (สาเหตุที่เข้ารับการบำบัดรักษา)	don't_have_money คือ ไม่มีเงิน (causetreatment01) family_force/asked คือ ครอบครัวบังคับหรือขอร้อง (causetreatment02) mental_health_problem คือ มีปัญหาด้านสุขภาพจิต (causetreatment03) physical_health_problem คือ มีปัญหาด้านสุขภาพกาย (causetreatment04) company_conditions คือ เงื่อนไขของสถานประกอบการ (causetreatment05) school_conditions คือ เงื่อนไขของสถานศึกษา (causetreatment06) want_to_stop คือ อยากเลิก (causetreatment07) wasted คือ สิ้นเปลือง (causetreatment08)
16	Evaluation_treatment (ผลการบำบัดรักษา)	Complete คือ ครบกำหนด Uncomplete คือ ไม่ครบกำหนด

3.4 การวิเคราะห์ตัวแปร

ผู้วิจัยเริ่มต้นโดยการใช้สถิติในการวิเคราะห์ตัวแปร เริ่มจากใช้สถิติเชิงพรรณนา (Descriptive Statistics) และสถิติเชิงอนุมาน (Inferential Statistics) จากนั้นจึงนำผลการวิเคราะห์ตัวแปรที่ได้จากการใช้สถิติไปปรึกษาและแลกเปลี่ยนความคิดเห็นกับผู้เชี่ยวชาญด้านการบำบัดรักษา ผู้ติดยาเสพติด จากนั้นทำการวิเคราะห์ตัวแปรอีกครั้งในส่วนของการเรียนรู้ของเครื่อง (Machine Learning)

3.4.1 วิเคราะห์โดยใช้สถิติเชิงพรรณนา (Descriptive Statistics)

ผู้วิจัยใช้สถิติเชิงพรรณนาวิเคราะห์และนำเสนอข้อมูล โดยใช้ตารางทางเดียวในการแสดงจำนวน และร้อยละของตัวแปรแต่ละตัวที่ใช้ในงานวิจัย และใช้ตารางสองทางในการแสดงจำนวน ร้อยละ และค่าสถิติต่าง ๆ ระหว่างตัวแปรอิสระแต่ละตัวกับตัวแปรตาม

3.4.2 วิเคราะห์โดยใช้สถิติเชิงอนุมาน (Inferential Statistics)

ผู้วิจัยใช้สถิติเชิงอนุมานในการวิเคราะห์ตัวแปรโดยทำการทดสอบสมมติฐานตรวจสอบความสัมพันธ์ระหว่างตัวแปร ดังนี้

3.4.2.1 ตรวจสอบความสัมพันธ์ระหว่างตัวแปรอิสระทุกตัวกับตัวแปรตาม

ผู้วิจัยทำการตรวจสอบความสัมพันธ์ระหว่างตัวแปรอิสระทุกตัว ได้แก่ เพศ, อายุ, สถานภาพสมรส, การศึกษา, อาชีพ, รายได้ต่อเดือน, อาศัยอยู่กับใครในช่วง 30 วันที่ผ่านมา, ความสัมพันธ์ในครอบครัว, อายุที่เข้ายาเสพติดครั้งแรก, ยาเสพติดที่ใช้ครั้งแรก, สาเหตุการเข้ายาเสพติด เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับบริการเชิงงานเพื่อการศึกษาเท่านั้น เมื่อนำมาใช้บนเว็บไซต์ของหน่วยงานราชการ ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ครั้งแรก, ยาเสพติดที่ใช้ในปัจจุบัน, ระยะเวลาที่ใช้ยาเสพติด, เคยรับการบำบัดรักษาหรือไม่, สาเหตุที่เข้ารับการบำบัดรักษา กับตัวแปรตาม นั่นคือ ผลการบำบัดรักษา ว่ามีความสัมพันธ์กันหรือไม่ โดยทำการทดสอบสมมติฐานด้วยสถิติทดสอบไคสแควร์ (Chi-square Test) ที่ระดับความเชื่อมั่น 95%

สมมติฐานทางสถิติ

H_0 : ตัวแปรอิสระตัวที่ i ไม่มีความสัมพันธ์กับผลการบำบัดรักษา

H_1 : ตัวแปรอิสระตัวที่ i มีความสัมพันธ์กับผลการบำบัดรักษา

หากทำการทดสอบสมมติฐานเรียบร้อยแล้วแล้วพบว่ามีตัวแปรอิสระใดที่ไม่มีความสัมพันธ์กับตัวแปรตามจะทำการตัดตัวแปรอิสระนั้นออก

3.4.2.2 ตรวจสอบความสัมพันธ์ระหว่างตัวแปรอิสระด้วยตนเอง

หลังจากตรวจสอบความสัมพันธ์ระหว่างตัวแปรอิสระทุกตัวกับตัวแปรตามเรียบร้อยแล้ว ผู้วิจัยทำการตรวจสอบความสัมพันธ์ระหว่างตัวแปรอิสระด้วยตนเองว่ามีความสัมพันธ์กันหรือไม่ โดยการทดสอบสมมติฐานด้วยสถิติทดสอบไคสแควร์ (Chi-square Test) ที่ระดับความเชื่อมั่น 95%

สมมติฐานทางสถิติ

H_0 : ตัวแปรอิสระตัวที่ i ไม่มีความสัมพันธ์กับตัวแปรอิสระตัวที่ j

H_1 : ตัวแปรอิสระตัวที่ i มีความสัมพันธ์กับตัวแปรอิสระตัวที่ j

หากทำการทดสอบแล้วพบว่าตัวแปรอิสระทั้ง 2 ตัว มีความสัมพันธ์กันจะทำการหาค่าสหสัมพันธ์คราเมอร์ (Cramer's V Correlation) เพื่อตรวจสอบว่าตัวแปรอิสระทั้ง 2 ตัว มีความสัมพันธ์กันมากน้อยเพียงใด

3.4.3 วิเคราะห์โดยผู้เชี่ยวชาญด้านการบำบัดรักษาผู้ติดยาเสพติด

หลังจากทำการวิเคราะห์ตัวแปรโดยใช้สถิติเชิงอนุมานเรียบร้อยแล้ว ผู้วิจัยทำการวิเคราะห์ตัวแปรต่อโดยนำตัวแปรอิสระที่ผ่านการคัดเลือกเหล่านี้ไปปรึกษาและแลกเปลี่ยนความคิดเห็นกับผู้เชี่ยวชาญด้านการบำบัดรักษาผู้ติดยาเสพติดเพื่อให้แน่ใจว่าตัวแปรเหล่านั้นมีความเกี่ยวข้องและมีนัยสำคัญเชิงวิชาการและในทางปฏิบัติ

3.4.4 วิเคราะห์โดยใช้การเรียนรู้ของเครื่อง (Machine Learning)

ผู้วิจัยวิเคราะห์ตัวแปรโดยหาความสัมพันธ์ระหว่างตัวแปรอิสระที่ผ่านการคัดเลือกกับตัวแปรตามโดยใช้การถดถอยโลจิสติกอย่างง่าย (Univariate Logistic Regression) และการหาค่า Feature Importance

3.5 การสร้างตัวแบบ

สำหรับการสร้างตัวแบบการจำแนกในงานวิจัยนี้ จะทำการสร้างตัวแบบโดยใช้การเรียนรู้ของเครื่อง 3 เทคนิค ได้แก่ เทคนิคการถดถอยโลจิสติก, เทคนิคป่าสุ่ม และเทคนิคซัพพอร์ตเวกเตอร์แมชชีน ซึ่งจะทำการสร้างตัวแบบการจำแนกจากทั้ง 3 เทคนิค ในรูปแบบที่แตกต่างกัน แบ่งเป็นตัวแบบที่ใช้ตัวแปรอิสระครบทุกตัว, ตัวแบบที่ใช้ตัวแปรอิสระไม่ครบทุกตัวจากการคัดเลือกตัวแปร และตัวแบบที่ใช้ตัวแปรอิสระไม่ครบทุกตัวจากการคัดเลือกตัวแปรร่วมกับการปรับค่าไฮเปอร์พารามิเตอร์ โดยในการสร้างตัวแบบการจำแนกมีขั้นตอนต่าง ๆ ดังนี้

3.5.1 การสุ่มตัวอย่างและการแบ่งชุดข้อมูล

ก่อนที่จะทำการสร้างตัวแบบ ผู้วิจัยทำการสุ่มตัวอย่างแบบชั้นภูมิ (Stratified Sampling) ในขั้นตอนการแบ่งข้อมูล (Spilt Data) ด้วย โดยทำการแบ่งข้อมูลจำนวน 3,241 ข้อมูล ออกเป็น 2 ส่วน เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้หาไปใช้ประโยชน์อื่นใดโดยไม่ได้รับอนุญาตจากเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คิดเป็นสัดส่วน 70:30 โดยข้อมูลชุดสอน (Training Set) จำนวน 2,268 ข้อมูล และข้อมูลชุดทดสอบ (Test Set) จำนวน 973 ข้อมูล

3.5.2 การสร้างตัวแบบ

3.5.2.1 สร้างตัวแบบโดยใช้ตัวแปรอิสระครบทุกตัว (Full Model)

หลังจากแบ่งข้อมูลเรียบร้อยแล้ว ผู้วิจัยเริ่มทำการสร้างตัวแบบจากทั้ง 3 เทคนิค ได้แก่ เทคนิคการถดถอยโลจิสติก, เทคนิคป่าสุ่ม และเทคนิคซัพพอร์ตเวกเตอร์แมชชีน โดยใช้ข้อมูลชุดสอน และใช้ตัวแปรอิสระทุกตัวที่ผ่านการวิเคราะห์ตัวแปรในขั้นตอนที่ 3.4

3.5.2.2 สร้างตัวแบบโดยใช้ตัวแปรอิสระไม่ครบทุกตัวจากการคัดเลือกตัวแปร

หลังจากสร้างตัวแบบโดยใช้ตัวแปรอิสระครบทุกตัวเรียบร้อยแล้ว ตัวแบบในรูปแบบถัดไปที่ผู้วิจัยสร้างคือตัวแบบที่ใช้ตัวแปรอิสระไม่ครบทุกตัวจากการคัดเลือกตัวแปร ซึ่งผู้วิจัยเริ่มทำการคัดเลือกตัวแปรในตัวแบบการถดถอยโลจิสติกก่อน จากนั้นนำผลที่ได้จากการคัดเลือก ได้แก่ จำนวนตัวแปรอิสระ และตัวแปรอิสระที่ใช้ในตัวแบบการถดถอยโลจิสติก มาใช้ในการคัดเลือกตัวแปร

สำหรับตัวแบบการถดถอยโลจิสติก ผู้วิจัยคัดเลือกตัวแปรด้วยวิธี Stepwise Selection มีขั้นตอนดังต่อไปนี้

ขั้นที่ 1 ใช้คำสั่ง step ในโปรแกรม R เพื่อทำการคัดเลือกตัวแปรอิสระที่เหมาะสมที่สุดตามหลักการของ Stepwise Selection

ขั้นที่ 2 ตรวจสอบปัญหา Multicollinearity โดยใช้ค่า VIF สำหรับตัวแปรอิสระในตัวแบบที่ได้จากคำสั่งในขั้นตอนแรก

ขั้นที่ 3 พิจารณาค่า VIF ของตัวแปรอิสระทุกตัว หากมีค่า VIF น้อยกว่า 10 แสดงว่าตัวแบบนี้ไม่มีปัญหา Multicollinearity และสามารถใช้ในการจำแนกได้อย่างเหมาะสม

เมื่อได้จำนวนและตัวแปรอิสระเรียบร้อยแล้ว จะทำการคัดเลือกตัวแปรอีกครั้งในแต่ละตัวแบบนั้นคือตัวแบบการถดถอยโลจิสติก ตัวแบบป่าสุ่ม และตัวแบบซัพพอร์ตเวกเตอร์แมชชีน แบ่งเป็น 2 วิธี ดังนี้

วิธีที่ 1 วิธีการคัดเลือกตัวแปรโดยใช้ชุดตัวแปรเหมือนกัน

เป็นการคัดเลือกตัวแปรอิสระโดยใช้ชุดตัวแปรเดียวกันกับที่ได้จากการคัดเลือกตัวแปรในตัวแบบการถดถอยโลจิสติก

วิธีที่ 2 วิธีการคัดเลือกตัวแปรด้วยค่า Feature Importance

เป็นการคัดเลือกตัวแปรอิสระที่มีค่า Feature Importance มากที่สุด ไล่ตามลำดับจนได้จำนวนตัวแปรอิสระเท่ากับจำนวนตัวแปรที่กำหนด รูปแบบแรกประกอบไปด้วย ขั้นตอนต่อไปนี้

ขั้นที่ 1 กำหนดจำนวนตัวแปรอิสระที่จะใช้ ให้เท่ากับจำนวนตัวแปรอิสระที่ได้จากการคัดเลือกในตัวแบบการถดถอยโลจิสติก

ขั้นที่ 2 ทำการคัดเลือกตัวแปรอิสระที่มีค่า Feature Importance มากที่สุด ไล่ตามลำดับจากมากไปหาน้อย จนครบตามจำนวนตัวแปรที่กำหนดไว้ในขั้นที่ 1

3.5.2.3 สร้างตัวแบบโดยใช้ตัวแปรอิสระไม่ครบทุกตัวจากการคัดเลือกตัวแปร ร่วมกับการปรับค่าไฮเปอร์พารามิเตอร์

หลังจากสร้างตัวแบบโดยใช้ตัวแปรอิสระไม่ครบทุกตัวจากการคัดเลือกตัวแปรเรียบร้อยแล้ว ผู้วิจัยสร้างตัวแบบอีกครั้งโดยนำตัวแบบในขั้นตอนที่ 3.5.2.2 ทุกตัวแบบมาทำการปรับค่าไฮเปอร์พารามิเตอร์ โดยผู้วิจัยทำการปรับค่าไฮเปอร์พารามิเตอร์แบบ Grid Search ซึ่งเป็นกระบวนการค้นหาเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่าไฮเปอร์พารามิเตอร์จากค่าที่ผู้วิจัยกำหนดไว้ และเลือกค่าที่ทำให้ตัวแบบมีประสิทธิภาพสูงที่สุด ซึ่งมีรายละเอียดดังนี้

3.5.2.3.1 ตัวแบบการถดถอยโลจิสติก

ทำการปรับค่าไฮเปอร์พารามิเตอร์แบบ Grid Search ร่วมกับ 10-Fold Cross Validation

3.5.2.3.2 ตัวแบบป่าสุ่ม

ทำการปรับค่าไฮเปอร์พารามิเตอร์ ทั้งหมด 2 พารามิเตอร์ ดังนี้

1. mtry คือ จำนวนตัวแปรอิสระที่ใช้ในการจำแนกในแต่ละโหนด โดยค่าที่กำหนด คือ 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15
2. trees คือ จำนวนต้นไม้ที่ใช้ในตัวแบบ โดยค่าที่กำหนด คือ 250, 500, 750, 1000, 1250, 1500

3.5.2.3.3 ตัวแบบซัพพอร์ตเวกเตอร์แมชชีน

ทำการปรับค่าไฮเปอร์พารามิเตอร์ ทั้งหมด 2 ตัว ดังนี้

1. C คือ ขนาดของการทำ Regularization โดยค่าที่กำหนด คือ 0.01, 0.1, 1, 10, 100, 1000
2. Rbf_sigma คือ สัมประสิทธิ์ของ kernel โดยค่าที่กำหนด คือ 0.01, 0.1, 1, 10, 100

ตารางที่ 3.2 พารามิเตอร์ของตัวแบบป่าสุ่มและตัวแบบซัพพอร์ตเวกเตอร์แมชชีน

ตัวแบบ	พารามิเตอร์
Random Forest	mtry: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15] trees: [250, 500, 750, 1000, 1250, 1500]
Support Vector Machine	C: [0.01, 0.1, 1, 10, 100, 1000] Rbf_sigma: [0.01, 0.1, 1, 10, 100]

3.5.3 การประเมินประสิทธิภาพด้วยวิธีการตรวจสอบไขว้แบบ 10 รอบ (10-Fold Cross Validation)

หลังจากทำการสร้างตัวแบบแต่ละตัวแบบเรียบร้อยแล้ว ผู้วิจัยทำการประเมินประสิทธิภาพแต่ละตัวแบบด้วยวิธีการตรวจสอบไขว้แบบ 10 รอบ โดยนำข้อมูลชุดสอนที่ใช้ในการสร้างตัวแบบมาทำการแบ่งข้อมูลออกเป็น 10 ส่วนเท่า ๆ กัน ใช้ 9 ส่วน เป็นชุดสอน (Training Data) และ 1 ส่วน เป็นชุดทดสอบ (Test Data) ซึ่งจะทำเช่นนี้ไปเรื่อย ๆ จนครบทั้ง 10 ส่วน จากนั้นนำผลลัพธ์ที่ได้ในแต่ละรอบมาทำการหาค่าเฉลี่ย เพื่อใช้ทดสอบประสิทธิภาพของตัวแบบ

3.6 การประเมินประสิทธิภาพตัวแบบ

ผู้วิจัยใช้การประเมินประสิทธิภาพด้วยเมทริกซ์สับสน (Confusion Matrix) โดยให้ตัวแบบทำการจำแนกข้อมูลชุดทดสอบและนำผลลัพธ์การจำแนกใส่ในเมทริกซ์สับสน จากนั้นคำนวณคะแนน F1 จากเมทริกซ์สับสน ซึ่งในการประเมินและเปรียบเทียบประสิทธิภาพตัวแบบ โดยตัวแบบใดให้คะแนน F1 มาก มีความหมายว่า มีประสิทธิภาพในการจำแนกประเภทสูง

บทที่ 4

ผลการวิจัย

ในบทนี้จะกล่าวถึงผลการวิเคราะห์ของข้อมูลการเปรียบเทียบตัวแบบการจำแนกและปัจจัยที่ส่งผลต่อกลุ่มการยุติการบำบัดรักษาผู้ติดยาเสพติดแบบระบบสมัครใจในพื้นที่กรุงเทพมหานคร ซึ่งผู้วิจัยใช้ข้อมูลชุดสอนทั้งหมด 2,268 ข้อมูล ในการสร้างตัวแบบได้ 15 ตัวแบบที่แตกต่างกัน และใช้ข้อมูลชุดทดสอบทั้งหมด 973 ข้อมูล ในการประเมินประสิทธิภาพของตัวแบบด้วยคะแนน F1 ได้ผลลัพธ์ ดังนี้

- 4.1 ผลการวิเคราะห์ตัวแปร
- 4.2 ผลการสร้างตัวแบบ
- 4.3 ผลการประเมินประสิทธิภาพตัวแบบ

4.1 ผลการวิเคราะห์ตัวแปร

ผู้วิจัยใช้สถิติเชิงพรรณนา (Descriptive Statistics) ในการวิเคราะห์และนำเสนอข้อมูลในรูปแบบของความถี่และร้อยละ และใช้สถิติเชิงอนุมาน (Inferential Statistics) ในการทดสอบสมมติฐาน หาค่าความสัมพันธ์ระหว่างตัวแปรโดยใช้ค่าสถิติ Chi-Square, Cramér's V correlation และ Odds Ratio

4.1.1 วิเคราะห์โดยใช้สถิติเชิงพรรณนา (Descriptive Statistics)

4.1.1.1 การวิเคราะห์ข้อมูลแบบตารางทางเดียว

ผู้วิจัยใช้ตารางทางเดียวในการแสดงจำนวนและร้อยละของตัวแปรแต่ละตัวที่ใช้ในงานวิจัย ดังตารางที่ 4.1 – 4.16

ตารางที่ 4.1 จำนวนและร้อยละของผลการบำบัดรักษาผู้ติดยาเสพติดแบบระบบสมัครใจ

ผลการบำบัดรักษา	จำนวน	ร้อยละ
ไม่ครบกำหนด	2,008	61.96
ครบกำหนด	1,233	38.04
รวม	3,241	100.00

จากตารางที่ 4.1 พบว่า ผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจที่มีผลการบำบัดรักษาไม่ครบกำหนด มีจำนวน 2,008 คน คิดเป็นร้อยละ 61.96 และผู้ที่มีผลการบำบัดรักษาครบกำหนด มีจำนวน 1,233 คน คิดเป็นร้อยละ 38.04

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.2 จำนวนและร้อยละของเพศผู้ติดยาเสพติดแบบระบบสมัครใจ

เพศ	จำนวน	ร้อยละ
ชาย	2,880	88.86
หญิง	361	11.14
รวม	3,241	100.00

จากตารางที่ 4.2 พบว่า ผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจส่วนใหญ่เป็นเพศชาย มีจำนวน 2,880 คน คิดเป็นร้อยละ 88.86 และเป็นเพศหญิง จำนวน 361 คน คิดเป็นร้อยละ 11.14

ตารางที่ 4.3 จำนวนและร้อยละของอายุผู้ติดยาเสพติดแบบระบบสมัครใจ

อายุ	จำนวน	ร้อยละ
10 - 19 ปี	372	11.48
20 - 29 ปี	1,718	53.01
30 - 39 ปี	684	21.10
40 - 49 ปี	374	11.54
ตั้งแต่ 50 ปีขึ้นไป	93	2.87
รวม	3,241	100.00

จากตารางที่ 4.3 พบว่า ประมาณครึ่งหนึ่งของผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจ อยู่ในช่วงอายุ 20 – 29 ปี มีจำนวน 1,718 คน คิดเป็นร้อยละ 53.01 รองลงมาคืออยู่ในช่วงอายุ 30 – 39 ปี มีจำนวน 684 คน คิดเป็นร้อยละ 21.10 ส่วนผู้ที่อยู่ในช่วงอายุ 40 – 49 ปี มีจำนวน 374 คน คิดเป็นร้อยละ 11.54 มีจำนวนมากกว่าผู้ที่อยู่ในช่วงอายุ 10 – 19 ปี ที่มีจำนวน 372 คน คิดเป็นร้อยละ 11.48 เพียง 2 คน และผู้ที่อยู่ในช่วงอายุตั้งแต่ 50 ปีขึ้นไป มีจำนวนน้อยที่สุดเพียง 93 คน คิดเป็นร้อยละ 2.87

ตารางที่ 4.4 จำนวนและร้อยละของสถานภาพสมรสผู้ติดยาเสพติดแบบระบบสมัครใจ

สถานภาพสมรส	จำนวน	ร้อยละ
หย่า/หม้าย/แยกกันอยู่	295	9.11
สมรส	1,070	33.01
โสด	1,876	57.88
รวม	3,241	100.00

จากตารางที่ 4.4 พบว่า ผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจ จำนวน 1,876 คน คิดเป็นร้อยละ 57.88 มีสถานะโสด รองลงมาคือ สมรส มีจำนวน 1,070 คน คิดเป็นร้อยละ 33.01 และผู้ที่มีสถานะหย่า/หม้าย/แยกกันอยู่ มีจำนวนน้อยที่สุด คือ 295 คน คิดเป็นร้อยละ 9.11

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.5 จำนวนและร้อยละของการศึกษาผู้ติดยาเสพติดแบบระบบสมัครใจ

การศึกษา	จำนวน	ร้อยละ
ประถมศึกษา	861	26.57
มัธยมศึกษาตอนต้น	1,152	35.54
มัธยมศึกษาตอนปลาย	592	18.27
อนุปริญญา / ปวช., ปวส.	379	11.69
ตั้งแต่ปริญญาตรีขึ้นไป	213	6.57
ไม่มีการศึกษา	44	1.36
รวม	3,241	100.00

จากตารางที่ 4.5 พบว่า ผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจ ส่วนใหญ่มีการศึกษาอยู่ในระดับมัธยมศึกษาตอนต้น จำนวน 1,152 คน คิดเป็นร้อยละ 35.54 รองลงมาคืออยู่ในระดับประถมศึกษา มีจำนวน 861 คน คิดเป็นร้อยละ 26.57 ส่วนผู้ที่มีการศึกษาอยู่ในระดับมัธยมศึกษาตอนปลาย มีจำนวน 592 คน คิดเป็นร้อยละ 18.27 ผู้ที่มีการศึกษาอยู่ในระดับอนุปริญญา / ปวช., ปวส. มีจำนวนใกล้เคียงกับผู้ที่มีการศึกษาอยู่ในระดับตั้งแต่ปริญญาตรีขึ้นไป นั่นคือ มีจำนวน 379 คน และ 213 คน คิดเป็นร้อยละ 11.69 และ 6.57 ตามลำดับ และผู้ที่ไม่มีการศึกษา มีจำนวนน้อยที่สุดเพียง 44 คน คิดเป็นร้อยละ 1.36

ตารางที่ 4.6 จำนวนและร้อยละของอาชีพผู้ติดยาเสพติดแบบระบบสมัครใจ

อาชีพ	จำนวน	ร้อยละ
ข้าราชการ/รัฐวิสาหกิจ	64	1.97
พ่อค้า/แม่ค้า	253	7.81
พนักงานออฟฟิศ/พนักงาน โรงงาน	280	8.64
นักเรียน	144	4.44
รับจ้าง/ผู้ใช้แรงงาน/พนักงาน คมนาคม/เกษตรกร	1,459	45.02
ไม่มีอาชีพ	947	29.22
อาชีพอื่น ๆ นอกเหนือจาก ที่กล่าวมา	94	2.90
รวม	3,241	100.00

จากตารางที่ 4.6 พบว่า ผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจส่วนใหญ่มีอาชีพรับจ้าง/ผู้ใช้แรงงาน/พนักงานคมนาคม/เกษตรกร มีจำนวน 1,459 คน คิดเป็นร้อยละ 45.02 รองลงมาคือ ไม่มีอาชีพ จำนวน 947 คน คิดเป็นร้อยละ 29.22 ส่วนผู้ที่มีอาชีพเป็นพนักงานออฟฟิศ/พนักงานโรงงาน มีจำนวน 280 คน คิดเป็นร้อยละ 8.64 ซึ่งมากกว่าผู้ที่มีอาชีพเป็นพ่อค้า/แม่ค้า ที่มีจำนวน 253 คน คิดเป็นร้อยละ 7.81 เพียง 27 คน ผู้ที่เป็นนักเรียนมีจำนวนใกล้เคียงกับผู้ที่มีอาชีพอื่น ๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

นอกเหนือจากที่กล่าวมา ซึ่งมีจำนวน 144 คน และ 94 คน คิดเป็นร้อยละ 4.44 และ 2.90 ตามลำดับ ส่วนผู้ที่เป็นข้าราชการ/รัฐวิสาหกิจ มีจำนวนน้อยที่สุดเพียง 64 คน คิดเป็นร้อยละ 1.97

ตารางที่ 4.7 จำนวนและร้อยละของรายได้ต่อเดือนผู้ติดยาเสพติดแบบระบบสมัครใจ

รายได้ต่อเดือน	จำนวน	ร้อยละ
น้อยกว่า หรือเท่ากับ 5,000 บาท	256	7.90
5,001 ถึง 10,000 บาท	777	23.97
10,001 ถึง 15,000 บาท	872	26.91
มากกว่า หรือเท่ากับ 15,001 บาท	458	14.13
ไม่มีรายได้	878	27.09
รวม	3,241	100.00

จากตารางที่ 4.7 พบว่า ผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจส่วนใหญ่คือ ผู้ที่ไม่มีรายได้ ผู้ที่มีรายได้อยู่ในช่วง 10,001 ถึง 15,000 บาท และผู้ที่มีรายได้ในช่วง 5,001 ถึง 10,000 บาท มีจำนวน 878 คน 872 คน และ 777 คน คิดเป็นร้อยละ 27.09 26.91 และ 23.97 ตามลำดับ ส่วนผู้ที่มีรายได้มากกว่า หรือเท่ากับ 15,001 บาท มีจำนวน 458 คน คิดเป็นร้อยละ 14.13 ส่วนผู้ที่มีรายได้น้อยกว่า หรือเท่ากับ 5,000 บาท มีจำนวนน้อยที่สุดเพียง 256 คน คิดเป็นร้อยละ 7.90

ตารางที่ 4.8 จำนวนและร้อยละของการอาศัยอยู่กับใคร (ในช่วง 30 วันที่ผ่านมา)
ผู้ติดยาเสพติดแบบระบบสมัครใจ

อาศัยอยู่กับใคร (ในช่วง 30 วันที่ผ่านมา)	จำนวน	ร้อยละ
อาศัยอยู่คนเดียว	234	7.22
อยู่กับพ่อ แม่ และคนอื่น ๆ	220	6.79
อยู่กับพ่อหรือแม่ และคนอื่น ๆ	282	8.70
อยู่กับพ่อและแม่	791	24.41
อยู่กับพ่อหรือแม่	472	14.56
อยู่กับเพื่อน	94	2.90
อยู่กับญาติ (ที่ไม่ใช่พ่อแม่)	412	12.71
อยู่กับคู่สมรสและคนอื่น ๆ	232	7.16
อยู่กับคู่สมรสหรือบุตร	504	15.55
รวม	3,241	100.00

จากตารางที่ 4.8 พบว่า ผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจส่วนใหญ่อาศัยอยู่กับพ่อและแม่ มีจำนวน 791 คน คิดเป็นร้อยละ 24.41 รองลงมาคือ ผู้ที่อาศัยอยู่กับคู่สมรสหรือบุตร มีจำนวน 504 คน คิดเป็นร้อยละ 15.55 ซึ่งมีจำนวนมากกว่าผู้ที่อาศัยอยู่กับพ่อหรือแม่ที่มีจำนวน 472 คน คิดเป็นร้อยละ 14.56 เพียง 32 คน ผู้ที่อาศัยอยู่กับญาติ (ที่ไม่ใช่พ่อแม่) มีจำนวน 412 คน คิดเป็นร้อยละ 12.71 ส่วนผู้ที่อาศัยอยู่กับพ่อหรือแม่ และคนอื่น ๆ มีจำนวนใกล้เคียงกับผู้ที่อาศัยอยู่กับคู่สมรสหรือบุตร

ไม่ว่าการณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อาศัยอยู่คนเดียว ผู้ที่อาศัยอยู่กับคู่สมรสและคนอื่น ๆ และผู้ที่อาศัยอยู่กับพ่อ แม่ และคนอื่น ๆ ซึ่งมีจำนวน 282 คน 234 คน 232 คน และ 220 คน คิดเป็นร้อยละ 8.70 7.22 7.16 และ 6.79 ตามลำดับ ส่วนผู้ที่อาศัยอยู่กับเพื่อน มีจำนวนน้อยที่สุดเพียง 94 คน คิดเป็นร้อยละ 2.90

ตารางที่ 4.9 จำนวนและร้อยละของความสัมพันธ์ในครอบครัวผู้ติดยาเสพติดแบบระบบสมัครใจ

ความสัมพันธ์ในครอบครัว	จำนวน	ร้อยละ
หย่าร้าง/แยกกันอยู่	888	27.40
พ่อและแม่เสียชีวิต	179	5.52
พ่อหรือแม่เสียชีวิต	517	15.95
อยู่ร่วมกันอย่างไม่ราบรื่น	41	1.27
อยู่ร่วมกันอย่างราบรื่น	1,616	49.86
รวม	3,241	100.00

จากตารางที่ 4.9 พบว่า ประมาณครึ่งหนึ่งของผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจ มีความสัมพันธ์ในครอบครัวแบบอยู่ร่วมกันอย่างราบรื่น มีจำนวน 1,616 คน คิดเป็นร้อยละ 49.86 รองลงมาคือ มีความสัมพันธ์ในครอบครัวแบบหย่าร้าง/แยกกันอยู่ มีจำนวน 888 คน คิดเป็นร้อยละ 27.40 ส่วนผู้ที่มีความสัมพันธ์ในครอบครัวแบบพ่อหรือแม่เสียชีวิต มีจำนวน 517 คน คิดเป็นร้อยละ 15.95 ผู้ที่มีความสัมพันธ์ในครอบครัวแบบพ่อและแม่เสียชีวิต มีจำนวน 179 คน คิดเป็นร้อยละ 5.52 และผู้ที่มีความสัมพันธ์ในครอบครัวแบบอยู่ร่วมกันอย่างไม่ราบรื่น มีจำนวนน้อยที่สุดเพียง 41 คน คิดเป็นร้อยละ 1.27

ตารางที่ 4.10 จำนวนและร้อยละของอายุที่ใช้ยาเสพติดครั้งแรกผู้ติดยาเสพติดแบบระบบสมัครใจ

อายุที่ใช้ยาเสพติดครั้งแรก	จำนวน	ร้อยละ
10 - 19 ปี	2,206	68.07
20 - 29 ปี	847	26.13
30 - 39 ปี	145	4.47
40 - 49 ปี	33	1.02
ตั้งแต่ 50 ปีขึ้นไป	10	0.31
รวม	3,241	100.00

จากตารางที่ 4.10 พบว่า ผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจส่วนใหญ่ จำนวน 2,206 คน คิดเป็นร้อยละ 68.07 อยู่ในช่วงอายุ 10 - 19 ปี รองลงมาคืออยู่ในช่วงอายุ 20 - 29 ปี มีจำนวน 847 คน คิดเป็นร้อยละ 26.13 ส่วนผู้ที่อยู่ในช่วงอายุ 30 - 39 ปี มีจำนวน 145 คน คิดเป็นร้อยละ 4.47 และผู้ที่อยู่ในช่วงอายุ 40 - 49 ปี มีจำนวน 33 คน คิดเป็นร้อยละ 1.02 มีจำนวนน้อยใกล้เคียงกับผู้ที่อยู่ในช่วงอายุตั้งแต่ 50 ปีขึ้นไป ที่มีจำนวน 10 คน คิดเป็นร้อยละ 0.31 ซึ่งมีจำนวนน้อยที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.11 จำนวนและร้อยละของยาเสพติดที่ใช้ครั้งแรกผู้ติดยาเสพติดแบบระบบสมัครใจ

ยาเสพติดที่ใช้ครั้งแรก	จำนวน	ร้อยละ
ยาบ้า หรือ แอมเฟตามีน	995	30.70
กัญชา	973	30.02
เฮโรอีน	615	18.98
ยาไอซ์	284	8.76
ยาเค	153	4.72
กระท่อม	74	2.28
ยาอื่น ๆ นอกเหนือจากที่กล่าวมา	147	4.54
รวม	3,241	100.00

จากตารางที่ 4.11 พบว่า ผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจส่วนใหญ่ใช้ยาบ้า หรือ แอมเฟตามีน และ กัญชา เป็นยาเสพติดที่ใช้ครั้งแรก มีจำนวน 995 คน และ 973 คน คิดเป็นร้อยละ 30.70 และ 30.02 ตามลำดับ รองลงมาคือใช้เฮโรอีนเป็นยาเสพติดที่ใช้ครั้งแรก มีจำนวน 615 คน คิดเป็นร้อยละ 18.98 ส่วนผู้ที่ใช้ยาไอซ์เป็นยาเสพติดที่ใช้ครั้งแรก มีจำนวน 284 คน คิดเป็นร้อยละ 8.76 ผู้ที่ใช้ยาเคเป็นยาเสพติดที่ใช้ครั้งแรก มีจำนวน 153 คน คิดเป็นร้อยละ 4.72 ซึ่งมากกว่าผู้ที่ใช้อื่น ๆ นอกเหนือจากที่กล่าวมาซึ่งมีจำนวน 147 คน คิดเป็นร้อยละ 4.54 เพียง 6 คน และผู้ที่ใช้กระท่อมเป็นยาเสพติดที่ใช้ครั้งแรก มีจำนวนน้อยที่สุดเพียง 74 คน คิดเป็นร้อยละ 2.28

ตารางที่ 4.12 จำนวนและร้อยละของสาเหตุการใช้ยาเสพติดครั้งแรกผู้ติดยาเสพติดแบบระบบสมัครใจ

สาเหตุการใช้ยาเสพติดครั้งแรก	จำนวน	ร้อยละ
ทำให้หายใจ	121	3.73
เพื่อนชักชวน	1,094	33.76
เพื่อความสนุก	145	4.47
ช่วยในการทำงาน	44	1.36
ไม่สบายใจ	73	2.25
อยากลอง	1,764	54.43
รวม	3,241	100.00

จากตารางที่ 4.12 พบว่า ประมาณครึ่งหนึ่งของผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจ มีสาเหตุในการใช้ยาเสพติดครั้งแรกคืออยากลอง มีจำนวน 1,764 คน คิดเป็นร้อยละ 54.43 รองลงมาคือ เพื่อนชักชวน มีจำนวน 1,094 คน คิดเป็นร้อยละ 33.76 ส่วนผู้ที่ใช้ยาเสพติดครั้งแรกเพื่อความสนุก มีจำนวน 145 คน คิดเป็นร้อยละ 4.47 ใกล้เคียงกับผู้ที่ใช้ยาเสพติดครั้งแรกเพื่อให้หายใจมีจำนวน 121 คน คิดเป็นร้อยละ 3.73 ส่วนผู้ที่ใช้ยาเสพติดครั้งแรกเพราะความไม่สบายใจ มีจำนวน 73 คน คิดเป็นร้อยละ 2.25 และผู้ที่ใช้ยาเสพติดครั้งแรกเพื่อช่วยในการทำงานมีจำนวนน้อยที่สุดเพียง 44 คน คิดเป็นร้อยละ 1.36

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.13 จำนวนและร้อยละของยาเสพติดที่ใช้ในปัจจุบันผู้ติดยาเสพติดแบบระบบสมัครใจ

ยาเสพติดที่ใช้ในปัจจุบัน	จำนวน	ร้อยละ
ยาบ้า หรือ แอมเฟตามีน	634	19.56
กัญชา	214	6.60
เฮโรอีน	1,690	52.14
ยาไอซ์	464	14.32
ยาเค	94	2.90
กระท่อม	87	2.68
ยาอื่น ๆ นอกเหนือจากที่กล่าวมา	58	1.80
รวม	3,241	100.00

จากตารางที่ 4.13 พบว่า ประมาณครึ่งหนึ่งของผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจใช้เฮโรอีนเป็นยาเสพติดที่ใช้ในปัจจุบัน มีจำนวน 1,690 คน คิดเป็นร้อยละ 52.14 รองลงมาคือผู้ใช้ยาบ้า หรือ แอมเฟตามีนเป็นยาเสพติดที่ใช้ในปัจจุบัน จำนวน 634 คน คิดเป็นร้อยละ 19.56 ซึ่งมีจำนวนใกล้เคียงกับผู้ใช้ยาไอซ์ เป็นยาเสพติดที่ใช้ในปัจจุบัน มีจำนวน 464 คน คิดเป็นร้อยละ 14.32 ส่วนผู้ใช้กัญชา มีจำนวน 214 คน คิดเป็นร้อยละ 6.60 ผู้ที่ใช้ยาเคมีจำนวน 94 คน คิดเป็นร้อยละ 2.90 ซึ่งมากกว่าผู้ใช้กระท่อมที่มีจำนวน 87 คน คิดเป็นร้อยละ 2.68 เพียง 7 คน และผู้ใช้ยาอื่น ๆ นอกเหนือจากที่กล่าวมา มีจำนวนน้อยที่สุดเพียง 58 คน คิดเป็นร้อยละ 1.80

ตารางที่ 4.14 จำนวนและร้อยละของระยะเวลาที่ใช้ยาเสพติดผู้ติดยาเสพติดแบบระบบสมัครใจ

ระยะเวลาที่ใช้ยาเสพติด	จำนวน	ร้อยละ
น้อยกว่า หรือเท่ากับ 1 ปี	984	30.36
มากกว่า 1 ปี ถึง 5 ปี	1,519	46.87
มากกว่า 5 ปี ถึง 10 ปี	474	14.63
มากกว่า 10 ปี ถึง 15 ปี	91	2.80
มากกว่า 15 ปี	173	5.34
รวม	3,241	100.00

จากตารางที่ 4.14 พบว่า ประมาณครึ่งหนึ่งของผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจหรือมีจำนวน 1,519 คน คิดเป็นร้อยละ 46.87 มีระยะเวลาการใช้ยาเสพติดในช่วงมากกว่า 1 ปี ถึง 5 ปี รองลงมาคือผู้ที่มีระยะเวลาการใช้ยาเสพติดในช่วงน้อยกว่า หรือเท่ากับ 1 ปี มีจำนวน 984 คน คิดเป็นร้อยละ 30.36 ส่วนผู้ที่มีระยะเวลาการใช้ยาเสพติดในช่วงมากกว่า 5 ปี ถึง 10 ปี มีจำนวน 474 คน คิดเป็นร้อยละ 14.63 มีจำนวนมากกว่าผู้ที่มีระยะเวลาการใช้ยาเสพติดมากกว่า 15 ปี จำนวน 173 คน คิดเป็นร้อยละ 5.34 ส่วนผู้ที่มีระยะเวลาการใช้ยาเสพติดมากกว่า 10 ปี ถึง 15 ปี มีจำนวนน้อยที่สุดเพียง 91 คน คิดเป็นร้อยละ 2.80

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.15 จำนวนและร้อยละของเคยรับการบำบัดรักษาหรือไม่ผู้ติดยาเสพติดแบบระบบสมัครใจ

เคยรับการบำบัดรักษาหรือไม่	จำนวน	ร้อยละ
เคย	1,367	42.18
ไม่เคย	1,874	57.82
รวม	3,241	100.00

จากตารางที่ 4.15 พบว่า ผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจที่ไม่เคยเข้ารับการบำบัดรักษามาก่อน มีจำนวน 1,874 คน คิดเป็นร้อยละ 57.82 และผู้ที่เคยเข้ารับการบำบัดรักษามาก่อน มีจำนวน 1,367 คน คิดเป็นร้อยละ 42.18

ตารางที่ 4.16 จำนวนและร้อยละของสาเหตุที่เข้ารับการบำบัดรักษาผู้ติดยาเสพติดแบบระบบสมัครใจ

สาเหตุที่เข้ารับการบำบัดรักษา	จำนวน	ร้อยละ
เงื่อนไขของสถานประกอบการ	86	2.65
ไม่มีเงน	23	0.71
ครอบครัวบังคับหรือขอร้อง	34	1.05
มีปัญหาด้านสุขภาพจิต	276	8.52
มีปัญหาด้านสุขภาพกาย	65	2.01
เงื่อนไขของสถานศึกษา	20	0.61
อยากรเลิก	2,737	84.45
รวม	3,241	100.00

จากตารางที่ 4.16 พบว่า ผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจส่วนใหญ่มีสาเหตุที่เข้ารับการบำบัดรักษาคืออยากรเลิก มีจำนวนมากถึง 2,737 คน คิดเป็นร้อยละ 84.45 รองลงมาคือ ผู้ที่มีปัญหาด้านสุขภาพจิต จำนวน 276 คน คิดเป็นร้อยละ 8.52 ส่วนผู้ที่เข้ารับการบำบัดรักษาด้วยสาเหตุเงื่อนไขของสถานประกอบการ มีจำนวน 86 คน คิดเป็นร้อยละ 2.65 ซึ่งใกล้เคียงกับผู้ที่มีปัญหาด้านสุขภาพกาย มีจำนวน 65 คน คิดเป็นร้อยละ 2.01 ส่วนผู้ที่ครอบครัวบังคับหรือขอร้อง มีจำนวน 34 คน คิดเป็นร้อยละ 1.05 มีจำนวนมากกว่าผู้ที่ไม่มีเงนที่มีจำนวน 23 คน คิดเป็นร้อยละ 0.71 เพียง 11 คน และผู้ที่เข้ารับการบำบัดรักษาด้วยสาเหตุเงื่อนไขของสถานศึกษา มีจำนวนน้อยที่สุดเพียง 20 คน คิดเป็นร้อยละ 0.61

4.1.2 วิเคราะห์ผลโดยใช้สถิติเชิงอนุมาน (Inferential Statistics)

4.1.2.1 ผลการตรวจสอบความสัมพันธ์ระหว่างตัวแปรอิสระทุกตัวกับตัวแปรตาม

ผู้วิจัยใช้ตารางสองทางในการแสดงจำนวน ร้อยละ และค่าสถิติต่าง ๆ ของตัวแปรแต่ละตัว ได้แก่ ค่าไคสแควร์, ค่า p-value และสหสัมพันธ์คราเมอร์ (Cramer's V : V) เพื่อวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม ดังตารางที่ 4.17 – 4.31

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.17 ค่าสถิติต่าง ๆ ของเพศผู้ติดยาเสพติดแบบระบบสมัครใจ

เพศ	บำบัดรักษา ไม่ครบกำหนด		บำบัดรักษา ครบกำหนด		รวม		χ^2	p- value	V
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ			
ชาย	1,792	89.24	1,088	88.24	2,880	88.86	0.78	0.38	0.02
หญิง	216	10.76	145	11.76	361	11.14			
รวม	2,008	100.00	1,233	100.00	3,241	100.00			

จากตารางที่ 4.17 พบว่า ผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจเป็นเพศชาย จำนวน 2,880 คน คิดเป็นร้อยละ 88.86 และเพศหญิง จำนวน 361 คน คิดเป็นร้อยละ 11.14 แบ่งเป็นบำบัดไม่ครบกำหนดเป็นเพศชาย จำนวน 1,792 คน คิดเป็นร้อยละ 89.24 เพศหญิง จำนวน 216 คน คิดเป็นร้อยละ 10.76 และบำบัดครบกำหนดเป็นเพศชาย จำนวน 1,088 คน คิดเป็นร้อยละ 88.24 และเพศหญิง จำนวน 145 คน คิดเป็นร้อยละ 11.76

ตารางที่ 4.18 ค่าสถิติต่าง ๆ ของอายุผู้ติดยาเสพติดแบบระบบสมัครใจ

อายุ	บำบัดรักษา ไม่ครบกำหนด		บำบัดรักษา ครบกำหนด		รวม		χ^2	p- value	V
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ			
10 - 19 ปี	255	12.70	117	9.49	372	11.48	77.55	0.00	0.15
20 - 29 ปี	1,154	57.47	564	45.74	1,718	53.01			
30 - 39 ปี	351	17.48	333	27.01	684	21.10			
40 - 49 ปี	192	9.56	182	14.76	374	11.54			
ตั้งแต่ 50 ปี ขึ้นไป	56	2.79	37	3.00	93	2.87			
รวม	2,008	100.00	1,233	100.00	3,241	100.00			

จากตารางที่ 4.18 พบว่า ประมาณครึ่งหนึ่งของผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจ หรือมีจำนวน 1,718 คน คิดเป็นร้อยละ 53.01 แบ่งเป็น บำบัดรักษาไม่ครบกำหนด จำนวน 1,154 คน คิดเป็นร้อยละ 57.47 อยู่ในช่วงอายุ 20 – 29 ปี รองลงมาคืออยู่ในช่วงอายุ 30 – 39 ปี มีจำนวน 684 คน คิดเป็นร้อยละ 21.10 ส่วนผู้ที่อยู่ในช่วงอายุ 40 – 49 ปี มีจำนวน 374 คน คิดเป็นร้อยละ 11.54 มีจำนวนมากกว่าผู้ที่อยู่ในช่วงอายุ 10 – 19 ปี ที่มีจำนวน 372 คน คิดเป็นร้อยละ 11.48 เพียง 2 คน และผู้ที่อยู่ในช่วงอายุตั้งแต่ 50 ปีขึ้นไป มีจำนวนน้อยที่สุดเพียง 93 คน คิดเป็นร้อยละ 2.87

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.19 ค่าสถิติต่าง ๆ ของสถานภาพสมรสผู้ติดยาเสพติดแบบระบบสมัครใจ

สถานภาพสมรส	บำบัดรักษาไม่ครบกำหนด		บำบัดรักษาครบกำหนด		รวม		χ^2	p-value	V
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ			
หย่า/หม้าย/แยกกันอยู่	189	9.41	106	8.60	295	9.11	2.27	0.32	0.03
สมรส	677	33.72	393	31.87	1,070	33.01			
โสด	1,142	56.87	734	59.53	1,876	57.88			
รวม	2,008	100.00	1,233	100.00	3,241	100.00			

จากตารางที่ 4.19 พบว่าผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจส่วนใหญ่มีสถานะโสดซึ่งมีจำนวนทั้งหมด 1,876 คน คิดเป็นร้อยละ 57.88 รองลงมาคือมีสถานะสมรสจำนวน 1,070 คน คิดเป็นร้อยละ 33.01 และสถานะหย่า/หม้าย/แยกกันอยู่มีจำนวนน้อยที่สุดคือ 295 คน คิดเป็นร้อยละ 9.11 แบ่งเป็นบำบัดไม่ครบกำหนด สถานะโสด จำนวน 1,142 คน คิดเป็นร้อยละ 56.87 สถานะสมรส จำนวน 677 คน คิดเป็นร้อยละ 33.72 สถานะหย่า/หม้าย/แยกกันอยู่ จำนวน 189 คน คิดเป็นร้อยละ 9.41 และแบ่งเป็นบำบัดครบกำหนด สถานะโสด จำนวน 734 คน คิดเป็นร้อยละ 59.53 สถานะสมรส จำนวน 393 คน คิดเป็นร้อยละ 31.87 สถานะหย่า/หม้าย/แยกกันอยู่ จำนวน 106 คน คิดเป็นร้อยละ 8.60

ตารางที่ 4.20 ค่าสถิติต่าง ๆ ของการศึกษาของผู้ติดยาเสพติดแบบระบบสมัครใจ

การศึกษา	บำบัดรักษาไม่ครบกำหนด		บำบัดรักษาครบกำหนด		รวม		χ^2	p-value	V
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ			
ประถมศึกษา	574	28.59	287	23.28	861	26.57	14.52	0.01	0.07
มัธยมศึกษาตอนต้น	710	35.36	442	35.85	1,152	35.54			
มัธยมศึกษาตอนปลาย	344	17.13	248	20.11	592	18.27			
อนุปริญญา/ปวช.,ปวส.	229	11.40	150	12.17	379	11.69			
ตั้งแต่ปริญญาตรีขึ้นไป	122	6.08	91	7.38	213	6.57			
ไม่มีการศึกษา	29	1.44	15	1.22	44	1.36			
รวม	2,008	100.00	1,233	100.00	3,241	100.00			

จากตารางที่ 4.20 พบว่า ผู้ติดยาเสพติดแบบระบบสมัครใจส่วนใหญ่ศึกษาอยู่ในระดับมัธยมศึกษาตอนต้น จำนวน 1,152 คน คิดเป็นร้อยละ 35.54 แบ่งเป็นบำบัดรักษาไม่ครบกำหนดเอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อการใช้งานเท่านั้น เมื่อผู้ใช้งานเห็นเป็นประโยชน์ในการศึกษาไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จำนวน 710 คน คิดเป็นร้อยละ 35.36 และบำบัดรักษาครบกำหนด จำนวน 442 คน คิดเป็นร้อยละ 35.85 รองลงมาคือ ระดับประถมศึกษาและมัธยมศึกษาตอนปลาย จำนวน 861 คน และ 592 คน คิดเป็นร้อยละ 26.57 และ 18.27 ตามลำดับ

ตารางที่ 4.21 ค่าสถิติต่าง ๆ ของอาชีพผู้ติดยาเสพติดแบบระบบสมัครใจ

อาชีพ	บำบัดรักษา ไม่ครบกำหนด		บำบัดรักษา ครบกำหนด		รวม		χ^2	p- value	V
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ			
ข้าราชการ/ รัฐวิสาหกิจ	25	1.25	39	3.16	64	1.97	52.53	1.46 $\times 10^{-9}$	0.13
พ่อค้า/แม่ค้า	171	8.52	82	6.65	253	7.81			
พนักงาน ออฟฟิศ/ พนักงาน โรงงาน	133	6.62	147	11.92	280	8.64			
นักเรียน	88	4.38	56	4.54	144	4.44			
รับจ้าง/ผู้ใช้ แรงงาน/ พนักงาน คมนาคม/ เกษตรกร	924	46.02	535	43.39	1,459	45.02			
ไม่มีอาชีพ	595	29.63	352	28.55	947	29.22			
อาชีพอื่น ๆ นอกเหนือ จากที่กล่าวมา	72	3.59	22	1.78	94	2.90			
รวม	2,008	100.00	1,233	100.00	3,241	100.00			

จากตารางที่ 4.21 พบว่า ผู้ติดยาเสพติดแบบระบบสมัครใจส่วนใหญ่มีอาชีพรับจ้าง/ผู้ใช้แรงงาน/พนักงานคมนาคม/เกษตรกร จำนวน 1,459 คน คิดเป็นร้อยละ 45.02 แบ่งเป็นบำบัดรักษาไม่ครบกำหนด จำนวน 924 คน คิดเป็นร้อยละ 46.02 และบำบัดรักษาครบกำหนด จำนวน 535 คน คิดเป็นร้อยละ 43.39 รองลงมาคือ ไม่มีอาชีพ จำนวน 947 คน คิดเป็นร้อยละ 29.22

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.22 ค่าสถิติต่าง ๆ ของรายได้ต่อเดือนของผู้ติดยาเสพติดแบบระบบสมัครใจ

รายได้ ต่อเดือน	บำบัดรักษา ไม่ครบกำหนด		บำบัดรักษา ครบกำหนด		รวม		χ^2	p- value	V
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ			
น้อยกว่า หรือเท่ากับ 5,000 บาท	142	7.07	114	9.25	256	7.90	10.02	0.04	0.06
5,001 ถึง 10,000 บาท	492	24.50	285	23.11	777	23.97			
10,001 ถึง 15,000 บาท	520	25.90	352	28.55	872	26.91			
มากกว่า หรือเท่ากับ	286	14.24	172	13.95	458	14.13			
ไม่มีรายได้	568	28.29	310	25.14	878	27.09			
รวม	2,008	100.00	1,233	100.00	3,241	100.00			

จากตารางที่ 4.22 พบว่า ผู้ติดยาเสพติดแบบระบบสมัครใจส่วนใหญ่ไม่มีรายได้ต่อเดือน มีรายได้อยู่ที่ 10,001 ถึง 15,000 บาท และ 5,001 ถึง 10,000 บาท จำนวน 878 คน 872 คน และ 256 คน คิดเป็นร้อยละ 27.09 26.91 และ 23.97 ตามลำดับ แบ่งเป็นบำบัดรักษาไม่ครบกำหนด จำนวน 568 คน 520 คน และ 492 คน คิดเป็นร้อยละ 28.29 25.90 และ 24.50 ตามลำดับ และบำบัดรักษาครบกำหนด จำนวน 310 คน 352 คน และ 285 คน คิดเป็นร้อยละ 25.14 28.55 และ 23.11 ตามลำดับ

ตารางที่ 4.23 ค่าสถิติต่าง ๆ ของการอาศัยอยู่กับใคร (ในช่วง 30 วันที่ผ่านมา)

ผู้ติดยาเสพติดแบบระบบสมัครใจ

อาศัยอยู่ กับใคร (ในช่วง 30 วันที่ผ่าน มา)	บำบัดรักษา ไม่ครบกำหนด		บำบัดรักษา ครบกำหนด		รวม		χ^2	p- value	V
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ			
อาศัยอยู่ คนเดียว	142	7.07	92	7.46	234	7.22	99.11	6.49 $\times 10^{-18}$	0.17
อยู่กับพ่อ แม่ และ คนอื่น ๆ	141	7.02	79	6.41	220	6.79			
อยู่กับพ่อ หรือแม่ และคนอื่น ๆ	198	9.86	84	6.81	282	8.70			

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.23 (ต่อ) ค่าสถิติต่าง ๆ ของการอาศัยอยู่กับใคร (ในช่วง 30 วันที่ผ่านมา)
ผู้ติดยาเสพติดแบบระบบสมัครใจ

อาศัยอยู่กับใคร (ในช่วง 30 วันที่ผ่านมา)	บำบัดรักษา ไม่ครบกำหนด		บำบัดรักษา ครบกำหนด		รวม		χ^2	p-value	V
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ			
อยู่กับพ่อและแม่	385	19.17	406	32.93	791	24.41			
อยู่กับพ่อหรือแม่	335	16.68	137	11.11	472	14.56			
อยู่กับเพื่อน	66	3.30	28	2.27	94	2.90			
อยู่กับญาติ (ที่ไม่ใช่พ่อแม่)	289	14.39	123	9.98	412	12.71			
อยู่กับคู่สมรสและคนอื่น ๆ	146	7.27	86	6.97	232	7.16			
อยู่กับคู่สมรสหรือบุตร	306	15.24	198	16.06	504	15.55			
รวม	2,008	100.00	1,233	100.00	3,241	100.00			

จากตาราง 4.23 พบว่าผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจส่วนใหญ่อาศัยอยู่กับพ่อและแม่ จำนวน 791 คน คิดเป็นร้อยละ 24.41 แบ่งเป็นบำบัดรักษาไม่ครบกำหนดจำนวน 385 คน คิดเป็นร้อยละ 19.17 และบำบัดรักษาครบกำหนด จำนวน 406 คน คิดเป็นร้อยละ 32.93

ตารางที่ 4.24 ค่าสถิติต่าง ๆ ของความสัมพันธ์ในครอบครัวของผู้ติดยาเสพติดแบบระบบสมัครใจ

ความสัมพันธ์ ในครอบครัว	บำบัดรักษา ไม่ครบกำหนด		บำบัดรักษา ครบกำหนด		รวม		χ^2	p-value	V
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ			
หย่าร้าง/ แยกกันอยู่	666	33.17	222	18.00	888	27.40	153.95	2.90 $\times 10^{-32}$	0.21
พ่อและแม่ เสียชีวิต	110	5.48	69	5.60	179	5.52			
พ่อหรือแม่ เสียชีวิต	370	18.43	147	11.92	517	15.95			

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับกรใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.24 (ต่อ) ค่าสถิติต่าง ๆ ของความสัมพันธ์ในครอบครัวของผู้ติดยาเสพติด
แบบระบบสมัครใจ

ความสัมพันธ์ ในครอบครัว	บำบัดรักษา ไม่ครบกำหนด		บำบัดรักษา ครบกำหนด		รวม		χ^2	p- value	V
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ			
อยู่ร่วมกัน อย่างไม่ ราบรื่น	20	1.00	21	1.70	41	1.27			
อยู่ร่วมกัน อย่างราบรื่น	842	41.93	774	62.77	1,616	49.86			
รวม	2,008	100.00	1,233	100.00	3,241	100.00			

จากตารางที่ 4.24 พบว่าความสัมพันธ์ในครอบครัวของผู้ติดยาเสพติดแบบระบบสมัครใจ ส่วนใหญ่อยู่ร่วมกันอย่างราบรื่น จำนวน 1,616 คน คิดเป็นร้อยละ 49.86 แบ่งเป็นบำบัดรักษาไม่ครบกำหนด จำนวน 842 คน คิดเป็นร้อยละ 41.93 และบำบัดรักษาครบกำหนด จำนวน 774 คน คิดเป็นร้อยละ 62.77 รองลงมาคือหย่าร้าง/แยกกันอยู่ จำนวน 888 คน คิดเป็นร้อยละ 27.40 และความสัมพันธ์ในครอบครัวแบบอยู่ร่วมกันอย่างไม่ราบรื่น มีจำนวนน้อยที่สุดคือ 41 คน คิดเป็นร้อยละ 1.27

ตารางที่ 4.25 ค่าสถิติต่าง ๆ ของอายุที่ใช้ยาเสพติดครั้งแรกของผู้ติดยาเสพติดแบบระบบสมัครใจ

อายุที่ใช้ ยาเสพติด ครั้งแรก	บำบัดรักษา ไม่ครบกำหนด		บำบัดรักษา ครบกำหนด		รวม		χ^2	p- value	V
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ			
10 - 19 ปี	1,538	76.59	668	54.18	2,206	68.07	190.47	4.21	0.24
20 - 29 ปี	407	20.27	440	35.69	847	26.13		$\times 10^{-40}$	
30 - 39 ปี	49	2.44	96	7.79	145	4.47			
40 - 49 ปี	11	0.55	22	1.78	33	1.02			
ตั้งแต่ 50 ปี ขึ้นไป	3	0.15	7	0.57	10	0.31			
รวม	2,008	100.00	1,233	100.00	3,241	100.00			

จากตารางที่ 4.25 พบว่า อายุที่ใช้ยาเสพติดครั้งแรกของผู้ติดยาเสพติดที่เข้ารับ การบำบัดรักษาระบบสมัครใจส่วนใหญ่ คืออายุ 10 - 19 ปี มีจำนวน 2,206 คน คิดเป็นร้อยละ 68.07 แบ่งเป็น บำบัดรักษาไม่ครบกำหนด จำนวน 1,538 คน คิดเป็นร้อยละ 76.59 รองลงมาคืออยู่ในช่วง อายุ 20 - 29 ปี มีจำนวน 847 คน คิดเป็นร้อยละ 26.13 ส่วนผู้ที่อยู่ในช่วงอายุ 30 - 39 ปี, 40 - 49 ปี และผู้ที่อยู่ในช่วงอายุตั้งแต่ 50 ปีขึ้นไป มีจำนวนน้อยมาก ซึ่งมีจำนวน 145 คน 33 คน และ 10 คน คิดเป็นร้อยละ 4.47, 1.02 และ 0.31 ตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีกรนำไปใช้

ตารางที่ 4.26 ค่าสถิติต่าง ๆ ของยาเสพติดที่ใช้ครั้งแรกของผู้ติดยาเสพติดแบบระบบสมัครใจ

ยาเสพติด ที่ใช้ครั้งแรก	บำบัดรักษา ไม่ครบกำหนด		บำบัดรักษา ครบกำหนด		รวม		χ^2	p- value	V
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ			
ยาบ้า หรือ แอมเฟตามีน	407	20.27	588	47.69	995	30.70	362.75	2.82 $\times 10^{-75}$	0.32
กัญชา	676	33.67	297	24.09	973	30.02			
เฮโรอีน	520	25.90	95	7.70	615	18.98			
ยาไอซ์	157	7.82	127	10.30	284	8.76			
ยาเค	113	5.63	40	3.24	153	4.72			
กระท่อม	38	1.89	36	2.92	74	2.28			
ยาอื่น ๆ นอกเหนือ จากที่กล่าวมา	97	4.83	50	4.06	147	4.54			
รวม	2,008	100.00	1,233	100.00	3,241	100.00			

จากตาราง 4.26 พบว่า ยาเสพติดที่ใช้ครั้งแรกของผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาแบบสมัครใจส่วนใหญ่ คือ ยาบ้าหรือแอมเฟตามีน และกัญชา มีจำนวน 995 คน และ 973 คน คิดเป็นร้อยละ 30.70 และ 30.02 ตามลำดับ แต่มีผลการบำบัดรักษาต่างกันคือ ยาบ้ามีผู้บำบัดรักษาไม่ครบกำหนด จำนวน 407 คน คิดเป็นร้อยละ 20.27 และบำบัดรักษาครบกำหนด จำนวน 588 คน คิดเป็นร้อยละ 47.69 กัญชามีผู้บำบัดรักษาไม่ครบกำหนด จำนวน 676 คน คิดเป็นร้อยละ 33.67 และบำบัดรักษาครบกำหนด จำนวน 297 คน คิดเป็นร้อยละ 24.09

ตารางที่ 4.27 ค่าสถิติต่าง ๆ ของสาเหตุการใช้ยาเสพติดครั้งแรกของผู้ติดยาเสพติดแบบระบบสมัครใจ

สาเหตุ การใช้ ยาเสพติด ครั้งแรก	บำบัดรักษา ไม่ครบกำหนด		บำบัดรักษา ครบกำหนด		รวม		χ^2	p- value	V
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ			
ทำให้หาย ป่วย	84	4.18	37	3.00	121	3.73	41.47	7.55 $\times 10^{-8}$	0.11
เพื่อน ชักชวน	601	29.93	493	39.98	1,094	33.76			
เพื่อ ความสนุก	98	4.88	47	3.81	145	4.47			
ช่วยใน การทำงาน	22	1.10	22	1.78	44	1.36			
ไม่สบายใจ	43	2.14	30	2.43	73	2.25			

เอกสารนี้เป็นทรัพย์สินของสำนักงานตำรวจแห่งชาติเพื่อการศึกษาเท่านั้น ไม่สามารถนำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.27 (ต่อ) ค่าสถิติต่าง ๆ ของสาเหตุการใช้ยาเสพติดครั้งแรกของผู้ติดยาเสพติดแบบระบบสมัครใจ

สาเหตุการใช้ยาเสพติดครั้งแรก	บำบัดรักษาไม่ครบกำหนด		บำบัดรักษาครบกำหนด		รวม		χ^2	p-value	V
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ			
อยากลอง	1,160	57.77	604	48.99	1,764	54.43			
รวม	2,008	100.00	1,233	100.00	3,241	100.00			

จากตารางที่ 4.27 พบว่าประมาณครึ่งหนึ่งของสาเหตุการใช้ยาเสพติดครั้งแรกของผู้ติดยาเสพติดแบบระบบสมัครใจ คือ อยากลอง จำนวน 1,764 คน คิดเป็นร้อยละ 54.43 แบ่งเป็นบำบัดรักษาไม่ครบกำหนด จำนวน 1,160 คน คิดเป็นร้อยละ 57.77 และบำบัดรักษาครบกำหนดจำนวน 604 คน คิดเป็นร้อยละ 48.99 รองลงมาคือ เพื่อนชักชวน จำนวน 1,094 คน คิดเป็นร้อยละ 33.76 แบ่งเป็นบำบัดรักษาไม่ครบกำหนด จำนวน 601 คน คิดเป็นร้อยละ 29.93 และบำบัดรักษาครบกำหนดจำนวน 493 คน คิดเป็นร้อยละ 39.98 ส่วนสาเหตุเพื่อความสนุก ทำให้หายป่วย ไม่สบายใจ และช่วยในการทำงาน มีจำนวนค่อนข้างน้อย มีจำนวนเพียง 145 คน 121 คน 73 คน และ 44 คน คิดเป็นร้อยละ 4.47 3.73 2.25 และ 1.36 ตามลำดับ

ตารางที่ 4.28 ค่าสถิติต่าง ๆ ของยาเสพติดที่ใช้ในปัจจุบันของผู้ติดยาเสพติดแบบระบบสมัครใจ

ยาเสพติดที่ใช้ในปัจจุบัน	บำบัดรักษาไม่ครบกำหนด		บำบัดรักษาครบกำหนด		รวม		χ^2	p-value	V
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ			
ยาบ้า หรือแอมเฟตามีน	167	8.32	467	37.88	634	19.56	825.61	0.00	0.45
กัญชา	58	2.89	156	12.65	214	6.60			
เฮโรอีน	1,410	70.22	280	22.71	1,690	52.14			
ยาไอซ์	224	11.16	240	19.46	464	14.32			
ยาเค	68	3.39	26	2.11	94	2.90			
กระท่อม	49	2.44	38	3.08	87	2.68			
ยาอื่น ๆ นอกเหนือจากที่กล่าวมา	32	1.59	26	2.11	58	1.80			
รวม	2,008	100.00	1,233	100.00	3,241	100.00			

จากตารางที่ 4.28 พบว่า ครึ่งหนึ่งของยาเสพติดที่ใช้ในปัจจุบันของผู้ติดยาเสพติดแบบระบบสมัครใจ คือ เฮโรอีน จำนวน 1,690 คน คิดเป็นร้อยละ 52.14 แบ่งเป็นบำบัดรักษาไม่ครบกำหนดจำนวน 1,410 คน คิดเป็นร้อยละ 70.22 และบำบัดรักษาครบกำหนด จำนวน 280 คน คิดเป็นร้อยละ 22.71

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รองลงมาคือยาบ้าหรือแอมเฟตามีน จำนวน 634 คน คิดเป็นร้อยละ 19.56 และยาไอซ์ จำนวน 464 คน คิดเป็นร้อยละ 14.32

ตารางที่ 4.29 ค่าสถิติต่าง ๆ ของระยะเวลาที่ใช้ยาเสพติดของผู้ติดยาเสพติดแบบระบบสมัครใจ

ระยะเวลาที่ใช้ยาเสพติด	บำบัดรักษาไม่ครบกำหนด		บำบัดรักษาครบกำหนด		รวม		χ^2	p-value	V
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ			
น้อยกว่าหรือเท่ากับ 1 ปี	575	28.64	409	33.17	984	30.36	30.11	4.65 $\times 10^{-6}$	0.10
มากกว่า 1 ปี ถึง 5 ปี	1,015	50.55	504	40.88	1,519	46.87			
มากกว่า 5 ปี ถึง 10 ปี	265	13.20	209	16.95	474	14.63			
มากกว่า 10 ปี ถึง 15 ปี	50	2.49	41	3.33	91	2.80			
มากกว่า 15 ปี	103	5.13	70	5.68	173	5.34			
รวม	2,008	100.00	1,233	100.00	3,241	100.00			

จากตารางที่ 4.29 พบว่าผู้ติดยาเสพติดแบบระบบสมัครใจ ส่วนใหญ่มีระยะเวลาในการใช้ยาเสพติดอยู่ที่ มากกว่า 1 ปี ถึง 5 ปี มีจำนวน 1,519 คน คิดเป็นร้อยละ 46.87 แบ่งเป็นบำบัดรักษาไม่ครบกำหนด จำนวน 1,015 คน คิดเป็นร้อยละ 50.55 และบำบัดรักษาครบกำหนด จำนวน 504 คน คิดเป็นร้อยละ 40.88 รองลงมาคือ ใช้เวลาน้อยกว่าหรือเท่ากับ 1 ปี จำนวน 984 คน คิดเป็นร้อยละ 30.36 และระยะเวลามากกว่า 10 ปี ถึง 15 ปี เป็นระยะเวลาที่มีจำนวนน้อยที่สุด คือ 91 คน คิดเป็นร้อยละ 2.80

ตารางที่ 4.30 ค่าสถิติต่าง ๆ ของการเคยรับการบำบัดรักษาของผู้ติดยาเสพติดแบบระบบสมัครใจ

เคยรับการบำบัดรักษาหรือไม่	บำบัดรักษาไม่ครบกำหนด		บำบัดรักษาครบกำหนด		รวม		χ^2	p-value	V
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ			
เคย	1,066	53.09	301	24.41	1,367	42.18	257.57	0.00	0.27
ไม่เคย	942	46.91	932	75.59	1,874	57.82			
รวม	2,008	100.00	1,233	100.00	3,241	100.00			

จากตารางที่ 4.30 พบว่า ผู้ติดยาเสพติดที่เข้ารับการบำบัดรักษาระบบสมัครใจไม่เคยรับการบำบัดรักษามาก่อน จำนวน 1,874 คน คิดเป็นร้อยละ 57.82 เคยรับการบำบัดรักษามาก่อน จำนวน 1,367 คน คิดเป็นร้อยละ 42.18 แบ่งเป็นบำบัดไม่ครบกำหนดเป็นผู้ที่ไม่เคยรับการบำบัดรักษามาก่อน จำนวน 942 คน คิดเป็นร้อยละ 46.91 ผู้ที่เคยรับการบำบัดรักษามาก่อน

จำนวน 1,066 คน คิดเป็นร้อยละ 53.09 และบำบัดครบกำหนดเป็นผู้ที่ไม่เคยรับการบำบัดรักษามาก่อน จำนวน 301 คน คิดเป็นร้อยละ 24.41 และบำบัดครบกำหนดเป็นผู้ที่เคยรับการบำบัดรักษามาก่อน จำนวน 634 คน คิดเป็นร้อยละ 52.19

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาค้นคว้าเท่านั้น มิใช่ผู้ใดเห็นไปใช้ประโยชน์ในการทำ
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จำนวน 932 คน คิดเป็นร้อยละ 75.59 ผู้ที่เคยรับการบำบัดรักษามาก่อน จำนวน 301 คน คิดเป็นร้อยละ 24.41

ตารางที่ 4.31 ค่าสถิติต่าง ๆ ของสาเหตุที่เข้ารับการบำบัดรักษาของผู้ติดยาเสพติด แบบระบบสมัครใจ

สาเหตุ ที่เข้ารับ การบำบัด รักษา	บำบัดรักษา ไม่ครบกำหนด		บำบัดรักษา ครบกำหนด		รวม		χ^2	p- value	V
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ			
เงินไม่ ของสถาน ประกอบการ	21	1.05	65	5.27	86	2.65	516.23	0.00	0.37
ไม่มีเงิน	19	0.95	4	0.32	23	0.71			
ครอบครัว บังคับหรือ ขอร้อง	25	1.25	9	0.73	34	1.05			
มีปัญหาด้าน สุขภาพจิต	11	0.55	265	21.49	276	8.52			
มีปัญหาด้าน สุขภาพกาย	50	2.49	15	1.22	65	2.01			
เงินไม่ของ สถานศึกษา	6	0.30	14	1.14	20	0.61			
อยากเลิก	1,876	93.43	861	69.83	2,737	84.45			
รวม	2,008	100.00	1,233	100.00	3,241	100.00			

จากตารางที่ 4.31 พบว่า ส่วนใหญ่สาเหตุที่เข้ารับการบำบัดรักษาของผู้ติดยาเสพติด แบบระบบสมัครใจ คือ อยากเลิก จำนวน 2,737 คน คิดเป็นร้อยละ 84.45 แบ่งเป็นบำบัดรักษา ไม่ครบกำหนด จำนวน 1,876 คน คิดเป็นร้อยละ 93.43 และบำบัดรักษาครบกำหนด จำนวน 861 คน คิดเป็นร้อยละ 69.83

จากตารางที่ 4.17 ถึง ตารางที่ 4.31 ทำการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระและ ตัวแปรตามโดยใช้ไคสแควร์ พบว่า จากตัวแปรอิสระทั้งหมด 15 ตัวแปร มี 13 ตัวแปร ได้แก่ อายุ, การศึกษา, อาชีพ, รายได้ต่อเดือน, อาศัยอยู่กับใครในช่วง 30 วันที่ผ่านมา, ความสัมพันธ์ในครอบครัว, อายุที่ใช้ยาเสพติดครั้งแรก, ยาเสพติดที่ใช้ครั้งแรก, สาเหตุการเข้ายาเสพติดครั้งแรก, ยาเสพติดที่ใช้ในปัจจุบัน, ระยะเวลาที่ใช้ยาเสพติด, เคยรับการบำบัดรักษาหรือไม่, สาเหตุที่เข้ารับการบำบัดรักษา มีค่า p-value น้อยกว่า 0.05 ซึ่งหมายความว่ามีความสัมพันธ์อย่างมีนัยสำคัญทางสถิติกับตัวแปรตาม นั้นคือผลการบำบัดรักษาผู้ติดยาเสพติดแบบระบบสมัครใจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มีความสัมพันธ์ระหว่างกันสูงจนทำให้เกิดปัญหาตัวแปรอิสระสัมพันธ์กันเอง (Multicollinearity) อ้างอิงจาก ที่กล่าวว่า ค่าสหสัมพันธ์ที่มีค่ามากกว่า 0.80 แสดงว่า ตัวแปรอิสระคู่หนึ่งมีความสัมพันธ์ระหว่างกันสูง (อัครพงษ์ อ้นทอง ,2550)

4.1.3 วิเคราะห์โดยใช้การเรียนรู้ของเครื่อง (Machine Learning)

4.1.3.1 ผลการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระที่ผ่านการคัดเลือกกับตัวแปรตามโดยใช้การถดถอยโลจิสติกอย่างง่าย (Univariate Logistic Regression)

ตารางที่ 4.33 ผลการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระที่ผ่านการคัดเลือกกับตัวแปรตามโดยใช้การถดถอยโลจิสติกอย่างง่าย (Univariate Logistic Regression)

ตัวแปร	Coefficients	S.E.	p-value
เพศ			
เพศชาย	0.1108	0.1345	0.4103
อายุ			
อายุ 10 ถึง 19	-0.0108	0.1458	0.9410
อายุ 30 ถึง 39	-0.6878	0.1101	4.11×10^{-10}
อายุ 40 ถึง 49	-0.6246	0.1410	9.40×10^{-6}
อายุตั้งแต่ 50 ปีขึ้นไป	-0.3850	0.2688	0.1520
สถานภาพสมรส			
โสด	-0.0780	0.1541	0.6127
สมรส	0.0927	0.1626	0.5689
การศึกษา			
อนุปริญญา หรือ ปวช.,ปวส	0.0470	0.2054	0.8191
มัธยมศึกษาตอนต้น	0.1439	0.1810	0.4267
ประถมศึกษา	0.3927	0.1867	0.0354
มัธยมศึกษาตอนปลาย	0.0008	0.1925	0.9968
ไม่มีการศึกษา	0.5935	0.4499	0.1870
อาชีพ			
พ่อค้า/แม่ค้า	1.2638	0.3307	0.0001
พนักงานออฟฟิศ/ พนักงานโรงงาน	0.5059	0.3188	0.1125
อาชีพอื่น ๆ นอกเหนือจากที่กล่าวมา	1.8051	0.4228	1.96×10^{-5}

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.33 (ต่อ) ผลการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระที่ผ่านการคัดเลือกกับตัวแปรตาม โดยใช้การถดถอยโลจิสติกอย่างง่าย (Univariate Logistic Regression)

ตัวแปร	Coefficients	S.E.	p-value
นักเรียน	0.7714	0.3495	0.0273
ไม่มีอาชีพ	0.9003	0.2979	0.0025
รับจ้าง/ผู้ใช้แรงงาน/ พนักงานคมนาคม/ เกษตรกร	0.9929	0.2941	0.0007
รายได้ต่อเดือน			
น้อยกว่า หรือเท่ากับ 5,000 บาท	-0.2481	0.1714	0.1480
5,001 ถึง 10,000 บาท	0.1431	0.1214	0.2380
มากกว่า หรือเท่ากับ 15,001 บาท	0.1538	0.1425	0.2800
ไม่มีรายได้	0.0925	0.1186	0.4360
อาศัยอยู่กับใครในช่วง 30 วัน ที่ผ่านมา			
อยู่กับพ่อ แม่ และ คนอื่น ๆ	0.1399	0.2346	0.5511
อยู่กับพ่อหรือแม่ และ คนอื่น ๆ	0.4791	0.2287	0.0362
อยู่กับพ่อและแม่	-0.5928	0.1861	0.0014
อยู่กับพ่อหรือแม่	0.2179	0.2011	0.2786
อยู่กับเพื่อน	0.3572	0.3139	0.2551
อยู่กับญาติ (ที่ไม่ใช่พ่อ แม่)	0.3881	0.2095	0.0640
อยู่กับคู่สมรสและ คนอื่น ๆ	-0.0571	0.2375	0.8099
อยู่กับคู่สมรสหรือบุตร	-0.0190	0.1977	0.9235
ความสัมพันธ์ ในครอบครัว			
พ่อและแม่เสียชีวิต	-0.3908	0.2172	0.0720
พ่อหรือแม่เสียชีวิต	-0.1711	0.1448	0.2373
อยู่ร่วมกันอย่างไม่ ราบรื่น	-0.8082	0.3975	0.0421
อยู่ร่วมกันอย่างราบรื่น	-0.9434	0.1078	$< 2 \times 10^{-16}$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.33 (ต่อ) ผลการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระที่ผ่านการคัดเลือกกับตัวแปรตาม โดยใช้การถดถอยโลจิสติกอย่างง่าย (Univariate Logistic Regression)

ตัวแปร	Coefficients	S.E.	p-value
อายุที่ใช้ยาเสพติดครั้งแรก			
อายุ 10 ถึง 19 ปี	0.9315	0.0986	$< 2 \times 10^{-16}$
อายุ 30 ถึง 39 ปี	-0.4254	0.2204	0.0536
อายุ 40 ถึง 49 ปี	-0.1917	0.4484	0.6689
อายุตั้งแต่ 50 ปีขึ้นไป	-13.4701	218.5807	0.9509
ยาเสพติดที่ใช้ครั้งแรก			
กัญชา	1.1500	0.1133	$< 2 \times 10^{-16}$
เฮโรอีน	2.0605	0.1549	$< 2 \times 10^{-16}$
ยาไอซ์	0.4890	0.1598	0.0022
ยาเค	1.5923	0.2440	6.7900×10^{-11}
กระท่อม	0.3138	0.2852	0.2712
ยาอื่น ๆ นอกเหนือจาก ที่กล่าวมา	1.2811	0.2288	2.1400×10^{-8}
สาเหตุการใช้ยาเสพติดครั้งแรก			
เพื่อนชักชวน	-0.6257	0.2493	0.0121
เพื่อความสนุก	-0.1655	0.3200	0.6050
ช่วยในการทำงาน	-1.0528	0.4327	0.0150
ไม่สบายใจ	-0.3944	0.3619	0.2759
อยากลอง	-0.2239	0.2461	0.3629
ยาเสพติดที่ใช้ในปัจจุบัน			
กัญชา	-0.0245	0.2107	0.9070
เฮโรอีน	2.6375	0.1322	$< 2 \times 10^{-16}$
ยาไอซ์	0.8204	0.1548	1.17×10^{-7}
ยาเค	2.2413	0.3116	6.38×10^{-13}
กระท่อม	1.3509	0.2770	1.07×10^{-6}
ยาอื่น ๆ นอกเหนือจาก ที่กล่าวมา	1.5779	0.3502	6.62×10^{-6}
ระยะเวลาที่ใช้ยาเสพติด			
น้อยกว่าหรือเท่ากับ 1 ปี	-0.4282	0.1014	2.41×10^{-5}
มากกว่า 5 ปี ถึง 10 ปี	-0.5698	0.1268	7.03×10^{-6}
มากกว่า 10 ปี ถึง 15 ปี	-0.7747	0.2496	0.0019

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.33 (ต่อ) ผลการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระที่ผ่านการคัดเลือกกับตัวแปรตาม โดยใช้การถดถอยโลจิสติกอย่างง่าย (Univariate Logistic Regression)

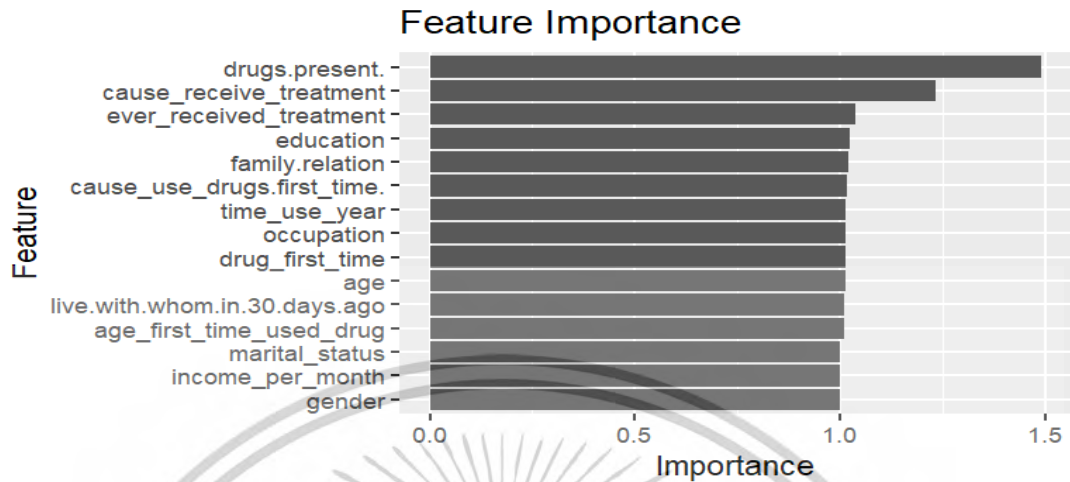
ตัวแปร	Coefficients	S.E.	p-value
มากกว่า 15 ปี	-0.2602	0.2114	0.2184
เคยรับกาบำบัดรักษาหรือไม่			
เคยรับการบำบัดรักษา	1.2106	0.0952	$< 2 \times 10^{-16}$
สาเหตุที่เข้ารับกาบำบัดรักษา			
ไม่มีเงิน	2.9429	0.8055	0.0003
ครอบครัวบังคับหรือขอร้อง	2.1518	0.5839	0.0002
มีปัญหาด้านสุขภาพจิต	-1.9734	0.4331	5.20×10^{-6}
มีปัญหาด้านสุขภาพกาย	2.1677	0.4464	1.20×10^{-6}
เงื่อนไขขอสถานศึกษา	0.0525	0.6045	0.9308
อยากเลิก	1.6902	0.2907	6.07×10^{-9}

จากตารางที่ 4.33 การวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม โดยใช้สถิติการถดถอยโลจิสติกอย่างง่าย พบว่า จากตัวแปรอิสระทั้งหมด 15 ตัวแปร มี 12 ตัวแปร ได้แก่ อายุ, การศึกษา, อาชีพ, อาศัยอยู่กับใครในช่วง 30 วันที่ผ่านมา, ความสัมพันธ์ในครอบครัว, อายุที่ใช้ยาเสพติดครั้งแรก, ยาเสพติดที่ใช้ครั้งแรก, สาเหตุการใช้ยาเสพติดครั้งแรก, ยาเสพติดที่ใช้ในปัจจุบัน, ระยะเวลาที่ใช้ยาเสพติด, เคยรับการบำบัดรักษาหรือไม่, สาเหตุที่เข้ารับกาบำบัดรักษา มีค่า p-value น้อยกว่า 0.05 ซึ่งหมายความว่ามีความสัมพันธ์อย่างมีนัยสำคัญทางสถิติกับตัวแปรตาม นั่นคือ ผลการบำบัดรักษาผู้ติดยาเสพติดแบบระบบสมัครใจ และ 3 ตัวแปร ได้แก่ เพศ, สถานภาพสมรส และรายได้ต่อเดือน ไม่มีความสัมพันธ์อย่างมีนัยสำคัญทางสถิติกับตัวแปรตาม

4.1.3.2 การหาค่า Feature Importance

ผู้วิจัยทำการหาค่าคะแนนความสำคัญของแต่ละตัวแปรอิสระ เพื่อดูว่าสำหรับแต่ละตัวแบบตัวแปรอิสระใดมีความสำคัญและส่งผลต่อผลการบำบัดรักษา ได้ผลลัพธ์ดังนี้

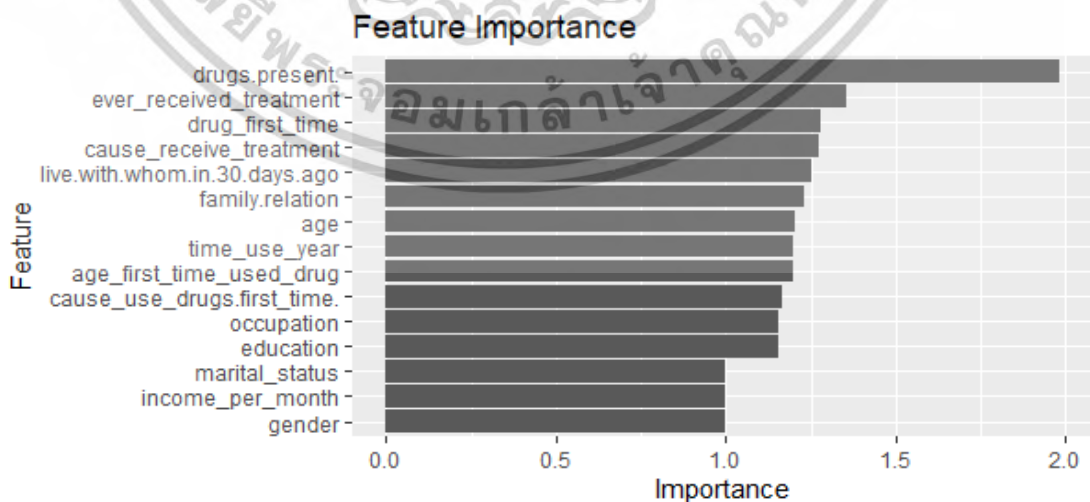
4.1.3.2.1 ตัวแบบการถดถอยโลจิสติก



รูปที่ 4.1 คะแนนความสำคัญของตัวแปรของตัวแบบการถดถอยโลจิสติก

จากรูปที่ 4.1 พบว่า สำหรับตัวแบบการถดถอยโลจิสติก ตัวแปรยาเสพติดที่ใช้ในปัจจุบันเป็นตัวแปรที่ส่งผลกับผลการบำบัดรักษามากที่สุดเนื่องจากมีคะแนนความสำคัญมากที่สุด รองลงมาคือ ตัวแปรสาเหตุที่เข้ารับการบำบัดรักษา เนื่องจากมีคะแนนความสำคัญรองลงมา ตัวแปรเพศ, ตัวแปรสถานภาพสมรส และตัวแปรรายได้ต่อเดือนมีคะแนนความสำคัญน้อยที่สุดเท่ากันทั้ง 3 ตัวแปร ทำให้ทั้ง 3 ตัวแปรนี้ส่งผลกับผลการบำบัดรักษาน้อยที่สุด ส่วนตัวแปรอื่น ๆ นอกเหนือจากตัวแปรเหล่านี้ เช่น ตัวแปรความสัมพันธ์ในครอบครัว, ตัวแปรอาชีพ มีคะแนนความสำคัญใกล้เคียงกัน หมายความว่า ส่งผลกับผลการบำบัดรักษาใกล้เคียงกัน โดยที่ตัวแปรที่มีค่าความสำคัญมากกว่าจะส่งผลกับผลการบำบัดรักษา

4.1.3.2.2 ตัวแบบป่าสุ่ม

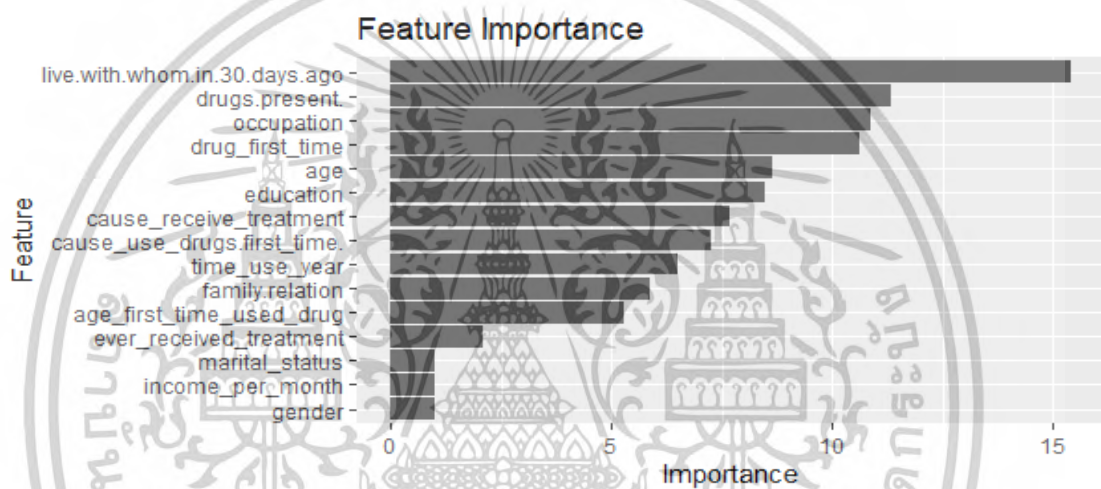


รูปที่ 4.2 คะแนนความสำคัญของตัวแปรของตัวแบบป่าสุ่ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.2 พบว่า สำหรับตัวแบบป่าสุ่ม ตัวแปรยาเสพติดที่ใช้ในปัจจุบันเป็นตัวแปรที่ส่งผลกับผลการบำบัดรักษามากที่สุดเนื่องจากมีคะแนนความสำคัญมากที่สุด รองลงมาคือ ตัวแปรเคยรับการบำบัดรักษาหรือไม่ เนื่องจากมีคะแนนความสำคัญรองลงมา ตัวแปรเพศ, ตัวแปรสถานภาพสมรส และตัวแปรรายได้ต่อเดือนมีคะแนนความสำคัญน้อยที่สุดเท่ากันทั้ง 3 ตัวแปร ทำให้ทั้ง 3 ตัวแปรนี้ส่งผลกับผลการบำบัดรักษาน้อยที่สุด ส่วนตัวแปรอื่น ๆ นอกเหนือจากตัวแปรเหล่านี้ เช่น ตัวแปรความสัมพันธ์ในครอบครัว, ตัวแปรอาชีพ มีคะแนนความสำคัญใกล้เคียงกัน หมายความว่า ส่งผลกับผลการบำบัดรักษาใกล้เคียงกัน โดยที่ตัวแปรที่มีค่าความสำคัญมากกว่าจะส่งผลกับผลการบำบัดรักษา

4.1.3.2.3 ตัวแบบซัพพอร์ตเวกเตอร์แมชชีน



รูปที่ 4.3 คะแนนความสำคัญของตัวแปรของตัวแบบซัพพอร์ตเวกเตอร์แมชชีน

จากรูปที่ 4.3 พบว่า สำหรับตัวแบบซัพพอร์ตเวกเตอร์แมชชีน ตัวแปรอาศัยอยู่กับใครในช่วง 30 วันที่ผ่านมาเป็นตัวแปรที่ส่งผลกับผลการบำบัดรักษามากที่สุดเนื่องจากมีคะแนนความสำคัญมากที่สุด รองลงมาคือ ตัวแปรยาเสพติดที่ใช้ในปัจจุบัน เนื่องจากมีคะแนนความสำคัญรองลงมา ตัวแปรเพศ, ตัวแปรสถานภาพสมรส และตัวแปรรายได้ต่อเดือนมีคะแนนความสำคัญน้อยที่สุดเท่ากันทั้ง 3 ตัวแปร ทำให้ทั้ง 3 ตัวแปรนี้ส่งผลกับผลการบำบัดรักษาน้อยที่สุด ส่วนตัวแปรอื่น ๆ นอกเหนือจากตัวแปรเหล่านี้ เช่น ตัวแปรการศึกษา, ตัวแปรอาชีพ มีคะแนนความสำคัญใกล้เคียงกัน หมายความว่า ส่งผลกับผลการบำบัดรักษาใกล้เคียงกัน โดยที่ตัวแปรที่มีค่าความสำคัญมากกว่าจะส่งผลกับผลการบำบัดรักษา

4.2 ผลการสร้างตัวแบบ

ผู้วิจัยได้ดำเนินการสร้างตัวแบบการจำแนกจากทั้ง 3 เทคนิค ในรูปแบบที่แตกต่างกัน ได้ผลลัพธ์ ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.1 ผลการสร้างตัวแบบโดยใช้ตัวแปรอิสระครบทุกตัว (Full Model)

ผู้วิจัยนำตัวแปรอิสระทั้ง 12 ตัวแปรที่มีความสัมพันธ์อย่างมีนัยสำคัญทางสถิติกับตัวแปรตามมาทำการสร้างตัวแบบการถดถอยโลจิสติก ตัวแบบป่าสุ่ม และตัวแบบซัพพอร์ตเวกเตอร์แมชชีน โดยถือว่าทั้ง 12 ตัวแปรนี้ เป็นตัวแปรเริ่มต้นในการสร้างตัวแบบ จึงใช้ชื่อว่าตัวแบบที่ใช้ตัวแปรอิสระครบทุกตัว

4.2.1.1 ผลการสร้างตัวแบบการถดถอยโลจิสติก

ตัวแบบการถดถอยโลจิสติกนั้นเป็นตัวแบบที่สามารถเขียนสมการการถดถอยได้ จึงแสดงสมการตัวอย่าง ดังนี้ ส่วนสมการเต็มสามารถดูได้ในภาคผนวก

$$\begin{aligned} \log(\text{odds}) = & -0.831 + 0.088(\text{age01}) - 0.01(\text{age03}) - 0.21(\text{age04}) \\ & + 0.377(\text{age05}) - 0.197(\text{edu01}) - 0.119(\text{edu02}) - 0.119(\text{edu03}) \\ & - 0.326(\text{edu04}) + 0.377(\text{edu06}) + \dots + 0.489(\text{evertreatment01}) \\ & + 1.1289(\text{causetreatment01}) + 1.645(\text{causetreatment02}) \\ & - 2.081(\text{causetreatment03}) + 0.903(\text{causetreatment04}) \\ & - 1.266(\text{causetreatment05}) + 0.443(\text{causetreatment07}) \end{aligned} \quad (4.1)$$

โดย $\beta_0, \beta_1, \beta_2, \dots, \beta_{11}, \beta_{12}$ คือ สัมประสิทธิ์การถดถอย

และยังสามารถแสดงค่า Odds Ratio p-value และค่าช่วงความเชื่อมั่นที่ 95% ได้ดังตารางที่ 4.35 พร้อมทั้งสร้างตาราง confusion matrix เพื่อดูผลการทำนาย ดังตารางที่ 4.36 ดังนี้

ตารางที่ 4.34 ค่า Odds Ratio ตัวแปรอิสระต่อการไม่ครบกำหนดการรักษา

ตัวแปร	OR	p-value	95%CI	
อายุ				
อายุ 10 ถึง 19	1.0916	0.6676	-0.3078	0.4930
อายุ 30 ถึง 39	0.9900	0.9498	-0.3205	0.3030
อายุ 40 ถึง 49	0.8106	0.3318	-0.6331	0.2157
อายุตั้งแต่ 50 ปีขึ้นไป	1.4572	0.3693	-0.4330	1.2163
การศึกษา				
อนุปริญญา หรือปวช.,ปวส	0.7219	0.2143	-0.8428	0.1869
มัธยมศึกษาตอนต้น	0.8877	0.6184	-0.5912	0.3472
ประถมศึกษา	0.8210	0.4286	-0.6890	0.2887
มัธยมศึกษาตอนปลาย	0.8876	0.6283	-0.6042	0.3615
ไม่มีการศึกษา	1.4586	0.5136	-0.7154	1.5623
อาชีพ				
พ่อค้า/แม่ค้า	1.6281	0.2168	-0.2783	1.2726
พนักงานออฟฟิศ/ พนักงานโรงงาน	1.0623	0.8703	-0.6579	0.7974
อาชีพอื่น ๆ	2.9610	0.0298	0.1276	2.0916
นอกเหนือจากที่กล่าวมา				
นักเรียน	1.3782	0.4726	-0.5456	1.2087

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้เฉพาะในเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่โดยไม่ได้รับอนุญาตจากทางผู้จัดทำ

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.34 (ต่อ) ค่า Odds Ratio ตัวแปรอิสระต่อการไม่ครบกำหนดการรักษา

ตัวแปร	OR	p-value	95%CI	
ไม่มีอาชีพ	1.1950	0.6168	-0.5124	0.8881
รับจ้าง/ผู้ใช้แรงงาน/ พนักงานคมนาคม/ เกษตรกร	1.168853	0.6525	-0.5156	0.8481
อาศัยอยู่กับใครในช่วง 30 วัน ที่ผ่านมา				
อยู่กับพ่อ แม่ และคนอื่น ๆ	0.8248	0.5317	-0.7955	0.4126
อยู่กับพ่อหรือแม่ และ คนอื่น ๆ	0.7640	0.3341	-0.8165	0.2770
อยู่กับพ่อและแม่	0.8322	0.4695	-0.6842	0.3124964
อยู่กับพ่อหรือแม่	0.7703	0.2955	-0.7533	0.2255
อยู่กับเพื่อน	1.4798	0.3011	-0.3432	1.1457
อยู่กับญาติ (ที่ไม่ใช่พ่อแม่)	1.0144	0.9556	-0.4902	0.5153
อยู่กับคู่สมรสและคนอื่น ๆ	0.7248	0.2626	-0.8855	0.2416
อยู่กับคู่สมรสหรือบุตร	0.8235	0.4180	-0.6672	0.2738
ความสัมพันธ์ในครอบครัว				
พ่อแม่เสียชีวิต	0.6871	0.1917	-0.9342	0.1942
พ่อหรือแม่เสียชีวิต	0.8719	0.4240	-0.4722	0.2004
อยู่ร่วมกันอย่างไม่ราบรื่น	1.1802	0.7369	-0.8014	1.1460
อยู่ร่วมกันอย่างราบรื่น	0.6606	0.0100	-0.7307	-0.0992
อายุที่ใช้ยาเสพติดครั้งแรก				
อายุ 10 ถึง 19 ปี	1.2420	0.1203	-0.0575	0.4895
อายุ 30 ถึง 39 ปี	0.8831	0.6593	-0.6821	0.4243
อายุ 40 ถึง 49 ปี	1.8925	0.2403	-0.4521	1.6955
อายุตั้งแต่ 50 ปีขึ้นไป	1.3707×10^{-6}	0.9700	NA	20.7064
ยาเสพติดที่ใช้ครั้งแรก				
กัญชา	1.3414	0.1039	-0.0622	0.6465
เฮโรอีน	1.7050	0.0148	0.1050	0.9638
ยาไอซ์	1.1796	0.4463	-0.2586	0.5922
ยาเค	1.1044	0.7629	-0.5334	0.7611
กระท่อม	0.4804	0.0941	-1.5992	0.1258
ยาอื่น ๆ นอกเหนือ จากที่กล่าวมา	1.6885	0.0805	-0.0550	1.1227
สาเหตุการใช้ยาเสพติดครั้งแรก				
เพื่อนชักชวน	0.8422	0.5765	-0.7885	0.4201
เพื่อความสนุก	1.1166	0.7785	-0.6633	0.8773

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.34 (ต่อ) ค่า Odds Ratio ตัวแปรอิสระต่อการไม่ครบกำหนดการรักษา

ตัวแปร	OR	p-value	95%CI	
ช่วยในการทำงาน	0.6722	0.4414	-1.4198	0.6095
ไม่สบายใจ	1.0426	0.9252	-0.8295	0.9144
อยากลอง	1.0452	0.8839	-0.5638	0.6258
ยาเสพติดที่ใช้ในปัจจุบัน				
กัญชา	0.7433	0.2876	-0.8485	0.2466
เฮโรอีน	5.8971	$< 2 \times 10^{-16}$	1.3703	2.1861
ยาไอซ์	1.6223	0.0152	0.0936	0.8757
ยาเค	7.7776	7.25×10^{-7}	1.2582	2.8870
กระท่อม	4.7013	0.0004	0.7082	2.4352
ยาอื่น ๆ นอกเหนือจากที่กล่าวมา	2.5098	0.0339	0.0850	1.7941
ระยะเวลาที่ใช้ยาเสพติด				
น้อยกว่า หรือเท่ากับ 1 ปี	0.6419	0.0005	-0.6914	-0.1953
มากกว่า 5 ปี ถึง 10 ปี	1.1098	0.5372	-0.2250	0.4371
มากกว่า 10 ปี ถึง 15 ปี	1.3396	0.3929	-0.3758	0.9689
มากกว่า 15 ปี	1.2010	0.5498	-0.4131	0.7892
เคยรับการบำบัดรักษาหรือไม่				
เคยรับการบำบัดรักษา	1.6313	6.39×10^{-5}	0.2497	0.7298
สาเหตุที่เข้ารับการบำบัดรักษา				
ไม่มีเงิน	3.6278	0.1242	-0.1819	3.2500
ครอบครัวบังคับหรือขอร้อง	5.1799	0.0087	0.4690	2.9573
มีปัญหาด้านสุขภาพจิต	0.1248	1.13×10^{-5}	-3.0374	-1.1661
มีปัญหาด้านสุขภาพกาย	2.4676	0.0694	-0.0453	1.9124
เงื่อนไขของสถานศึกษา	0.2820	0.0865	-2.7715	0.1472
อยากเลิก	1.5570	0.1730	-0.1782	1.1018

จากตารางที่ 4.34 พบว่า ตัวแปรยาเสพติดที่ใช้ในปัจจุบันมีค่า p-value น้อยที่สุดในประเภทเฮโรอีน มีค่า p-value $< 2 \times 10^{-16}$ รองลงมา คือ ยาเค มีค่า p-value เท่ากับ 7.25×10^{-7} จึงถือว่าตัวแปรยาเสพติดที่ใช้ในปัจจุบันมีความสัมพันธ์ต่อการผลการบำบัดรักษาไม่ครบกำหนดมากที่สุด และในตัวแปรสาเหตุการใช้ยาเสพติดครั้งแรกมีค่า p-value มากจึงถือว่าความสัมพันธ์ต่อการผลการบำบัดรักษาไม่ครบกำหนดน้อยที่สุด และสำหรับค่า Odds Ratio ตัวแปรยาเสพติดที่ใช้ในปัจจุบันประเภทยาเค มีค่า Odds Ratio มากที่สุดคือ 7.7776 หมายความว่า ผู้เข้ารับการบำบัดรักษาที่ใช้ยาเสพติดที่ใช้ในปัจจุบันประเภทยาเคมีความเสี่ยงที่จะบำบัดรักษาไม่ครบกำหนด 7.7776 เท่าของผู้เข้ารับการบำบัดรักษาที่ใช้ยาเสพติดที่ใช้ในปัจจุบันประเภทยาบ้าหรือแอมเฟตามีน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้นำไปเผยแพร่โดยไม่หวังกำไร

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รองลงมาคือ ตัวแปรยาเสพติดที่ใช้ในปัจจุบันประเภทเฮโรอีน มีค่า Odds Ratio เท่ากับ 5.8971 หมายความว่า ผู้เข้ารับการรักษาที่ใช้อยาเสพติดที่ใช้ในปัจจุบันประเภทยาเฮโรอีนมีความเสี่ยงที่จะบำบัดรักษาไม่ครบกำหนด 5.8971 เท่า ของผู้เข้ารับการรักษาที่ใช้อยาเสพติดที่ใช้ในปัจจุบันประเภทยาบ้าหรือแอมเฟตามีน จากการที่ตัวแปรยาเสพติดที่ใช้ในปัจจุบันมีค่า Odds Ratio ส่วนใหญ่มีค่า Odds Ratio มากกว่า 1 แสดงให้เห็นว่ายาเสพติดที่ใช้ในปัจจุบันเป็นปัจจัยเสี่ยงที่ทำให้เกิดผลการบำบัดรักษาไม่ครบกำหนด

ตารางที่ 4.35 ผลการทำนายด้วยตัวแบบการถดถอยโลจิสติก

ค่าทำนาย	ค่าจริง	
	ไม่ครบกำหนด	ครบกำหนด
ไม่ครบกำหนด	518	141
ครบกำหนด	85	229

จากตารางที่ 4.35 พบว่า ตัวแบบการถดถอยโลจิสติกสามารถทำนายผลการบำบัดรักษาได้ถูกต้องตรงกับผลการรักษาจริงอยู่ที่ 747 คน โดยแบ่งเป็นบำบัดรักษาไม่ครบกำหนดจำนวน 518 คน และบำบัดรักษาครบกำหนดจำนวน 229 คน จากข้อมูลชุดทดสอบจำนวน 973 คน แสดงให้เห็นว่าตัวแบบมีค่าความถูกต้องในการทำนายอยู่ที่ 76.77% และมีค่าคะแนน F1 เท่ากับ 82.09%

4.2.1.2 ผลการสร้างตัวแบบป่าสุ่ม

ผู้วิจัยทำการสร้างตัวแบบป่าสุ่มและนำตัวแบบมาทำนายข้อมูลชุดทดสอบและสร้างตาราง confusion matrix เพื่อดูผลการทำนาย ได้ผลลัพธ์ดังนี้

ตารางที่ 4.36 ผลการทำนายด้วยตัวแบบป่าสุ่ม

ค่าทำนาย	ค่าจริง	
	ไม่ครบกำหนด	ครบกำหนด
ไม่ครบกำหนด	585	96
ครบกำหนด	18	274

จากตารางที่ 4.36 พบว่า ตัวแบบป่าสุ่มสามารถทำนายผลการบำบัดรักษาได้ถูกต้องตรงกับผลการรักษาจริงอยู่ที่ 859 คน โดยแบ่งเป็นบำบัดรักษาไม่ครบกำหนดจำนวน 585 คน และบำบัดรักษาครบกำหนดจำนวน 274 คน จากข้อมูลชุดทดสอบจำนวน 973 คน แสดงให้เห็นว่าตัวแบบมีค่าความถูกต้องในการทำนายอยู่ที่ 88.28% และมีค่าคะแนน F1 เท่ากับ 91.12%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.1.3 ผลการสร้างตัวแบบซัพพอร์ตเวกเตอร์แมชชีน

ผู้วิจัยทำการสร้างตัวแบบซัพพอร์ตเวกเตอร์แมชชีนและนำตัวแบบมาทำนายข้อมูลชุดทดสอบและสร้างตาราง confusion matrix เพื่อดูผลการทำนาย ได้ผลลัพธ์ดังนี้

ตารางที่ 4.37 ผลการทำนายด้วยตัวแบบซัพพอร์ตเวกเตอร์แมชชีน

ค่าทำนาย	ค่าจริง	
	ไม่ครบกำหนด	ครบกำหนด
ไม่ครบกำหนด	603	14
ครบกำหนด	0	356

จากตารางที่ 4.37 พบว่า ตัวแบบซัพพอร์ตเวกเตอร์แมชชีนสามารถทำนายผลการบำบัดรักษาได้ถูกต้องตรงกับผลการรักษาจริงอยู่ที่ 959 คน โดยแบ่งเป็นบำบัดรักษาไม่ครบกำหนดจำนวน 603 คน และบำบัดรักษาครบกำหนดจำนวน 356 คน จากข้อมูลชุดทดสอบจำนวน 973 คน แสดงให้เห็นว่า ตัวแบบมีค่าความถูกต้องในการทำนายอยู่ที่ 98.56% และมีค่าคะแนน F1 เท่ากับ 98.85%

4.2.2 ผลการสร้างตัวแบบโดยใช้ตัวแปรอิสระไม่ครบทุกตัวจากการคัดเลือกตัวแปร

4.2.2.1 ตัวแบบการถดถอยโลจิสติก

เริ่มจากผลการคัดเลือกตัวแปรในตัวแบบการถดถอยโลจิสติกที่ทำการคัดเลือกตัวแปรด้วยวิธี stepwise selection เพื่อทราบถึงจำนวนตัวแปรอิสระและตัวแปรอิสระที่ใช้ในตัวแบบ และเมื่อทราบจำนวนตัวแปรเรียบร้อยแล้วจึงทำการคัดเลือกตัวแปรด้วยค่า Feature Importance ได้ผลการคัดเลือกตัวแปรดังตารางที่ 4.38 จากนั้นสร้างตาราง confusion matrix เพื่อดูผลการทำนาย ได้ผลดังตารางที่ 4.39 และ 4.40

ตารางที่ 4.38 ผลการคัดเลือกตัวแปรของตัวแบบการถดถอยโลจิสติก

ตัวแบบ	ตัวแปร	VIF
ตัวแบบการถดถอยโลจิสติกที่ทำการคัดเลือกตัวแปรด้วย stepwise selection (Logistic Regression with Feature Selection by stepwise selection)	1. ความสัมพันธ์ในครอบครัว	1.10
	2. อายุที่ใช้ยาเสพติดครั้งแรก	1.19
	3. ยาเสพติดที่ใช้ครั้งแรก	1.26
	4. ระยะเวลาที่ใช้ยาเสพติด	1.13
	5. เคยรับการบำบัดรักษาหรือไม่	1.07
	6. สาเหตุที่เข้ารับการบำบัดรักษา	1.16
ตัวแบบการถดถอยโลจิสติกที่ทำการคัดเลือกตัวแปรด้วยค่า Feature Importance (Logistic Regression with Feature Selection by Feature Importance)	1. ยาเสพติดที่ใช้ในปัจจุบัน	1.48
	2. สาเหตุที่เข้ารับการบำบัดรักษา	1.28
	3. เคยรับการบำบัดรักษาหรือไม่	1.15
	4. การศึกษา	1.24
	5. ความสัมพันธ์ในครอบครัว	1.15
	6. สาเหตุการใช้ยาเสพติดครั้งแรก	1.18

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อผู้จัดทำเห็นไปใช้ประโยชน์ด้านการศึกษาไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตารางที่ 4.38 สำหรับการสร้างตัวแบบการถดถอยโลจิสติกทำการคัดเลือกตัวแปร โดยวิธี Stepwise Selection ด้วยคำสั่ง stats::step พบว่า ได้ตัวแปรอิสระที่จะนำไปสร้างตัวแบบ ทั้งหมด 6 ตัวแปร ได้แก่ ความสัมพันธ์ในครอบครัว, อายุที่ใช้ยาเสพติดครั้งแรก, ยาเสพติดที่ใช้ครั้งแรก, ระยะเวลาที่ใช้ยาเสพติด, เคยรับการบำบัดรักษาหรือไม่ และสาเหตุที่เข้ารับการบำบัดรักษา ส่วนการคัดเลือกตัวแปรด้วยค่า Feature Importance ได้ตัวแปรอิสระดังนี้ ยาเสพติดที่ใช้ในปัจจุบัน, สาเหตุที่เข้ารับการบำบัดรักษา, เคยรับการบำบัดรักษาหรือไม่, การศึกษา, ความสัมพันธ์ในครอบครัว และสาเหตุการใช้ยาเสพติดครั้งแรก

ตารางที่ 4.39 ผลการจำแนกด้วยตัวแบบการถดถอยโลจิสติกที่ทำการคัดเลือกตัวแปร ด้วยวิธี stepwise selection

ค่าทำนาย	ค่าจริง	
	ไม่ครบกำหนด	ครบกำหนด
ไม่ครบกำหนด	534	157
ครบกำหนด	69	213

จากตารางที่ 4.39 พบว่า ตัวแบบการถดถอยโลจิสติกที่ทำการคัดเลือกตัวแปรด้วยวิธี stepwise selection สามารถทำนายผลการบำบัดรักษาได้ถูกต้องตรงกับผลการรักษาจริงอยู่ที่ 747 คน โดยแบ่งเป็นบำบัดรักษาไม่ครบกำหนดจำนวน 534 คน และบำบัดรักษาครบกำหนดจำนวน 213 คน จากข้อมูลชุดทดสอบจำนวน 973 คน แสดงให้เห็นว่า ตัวแบบมีค่าความถูกต้องในการทำนาย อยู่ที่ 76.77% และมีค่าคะแนน F1 เท่ากับ 82.53%

ตารางที่ 4.40 ผลการจำแนกด้วยตัวแบบการถดถอยโลจิสติกที่ทำการคัดเลือกตัวแปรด้วย ค่า Feature Importance

ค่าทำนาย	ค่าจริง	
	ไม่ครบกำหนด	ครบกำหนด
ไม่ครบกำหนด	510	136
ครบกำหนด	93	234

จากตารางที่ 4.40 พบว่า ตัวแบบการถดถอยโลจิสติกที่ทำการคัดเลือกตัวแปร ด้วยค่า Feature Importance สามารถทำนายผลการบำบัดรักษาได้ถูกต้องตรงกับผลการรักษาจริง อยู่ที่ 744 คน โดยแบ่งเป็นบำบัดรักษาไม่ครบกำหนดจำนวน 510 คน และบำบัดรักษาครบกำหนด จำนวน 234 คน จากข้อมูลชุดทดสอบจำนวน 973 คน แสดงให้เห็นว่า ตัวแบบมีค่าความถูกต้อง ในการทำนายอยู่ที่ 76.46% และมีค่าคะแนน F1 เท่ากับ 81.67%

จากตาราง 4.39 และ ตาราง 4.40 พบว่าตัวแบบการถดถอยโลจิสติกที่ทำการคัดเลือก ตัวแปรด้วยวิธี stepwise selection มีประสิทธิภาพมากกว่าตัวแบบการถดถอยโลจิสติกที่ทำการคัดเลือกตัวแปรด้วยค่า Feature Importance เนื่องจากมีค่าคะแนน F1 มากกว่า และมีสมการ

การถดถอยโลจิสติก คือ
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\begin{aligned} \log(\text{odds}) = & 0.377 + 0.595(\text{fam04}) + 1.533(\text{agefirst01}) + 2.369(\text{drugfirst01}) + 5.532(\text{drugfirst02}) \\ & + 1.715(\text{drugfirst03}) + 4.324(\text{drugfirst04}) + 2.780(\text{drugfirst06}) + 0.637(\text{timeuse01}) \\ & + 2.196(\text{eventreatment01}) + 8.124(\text{causetreatment01}) + 6.274(\text{causetreatment02}) \\ & + 0.131(\text{causetreatment03}) + 3.829(\text{causetreatment04}) + 2.394(\text{causetreatment06}) \end{aligned} \quad (4.2)$$

4.2.2.2 ตัวแบบป่าสุ่ม

สำหรับตัวแบบป่าสุ่ม ทำการคัดเลือกตัวแปร 2 แบบ แบบที่ 1 คัดเลือกตัวแปรโดยใช้ตัวแปรเหมือนกับตัวแบบการถดถอยโลจิสติก และแบบที่ 2 คัดเลือกตัวแปรด้วยค่า Feature Importance ได้ผลดังตารางที่ 4.41 จากนั้นสร้างตัวแบบและสร้างตาราง confusion matrix เพื่อดูผลการทำนาย ได้ผลดังตารางที่ 4.42 และตารางที่ 4.43 ตามลำดับ

ตารางที่ 4.41 ผลการคัดเลือกตัวแปรของตัวแบบป่าสุ่ม

ตัวแบบ	ตัวแปร
ตัวแบบป่าสุ่มที่ใช้ตัวแปรอิสระเหมือนกับตัวแบบการถดถอยโลจิสติก (Random Forest with Feature Selection same as Logistic Regression)	<ol style="list-style-type: none"> 1. ความสัมพันธ์ในครอบครัว 2. อายุที่ใช้ยาเสพติดครั้งแรก 3. ยาเสพติดที่ใช้ครั้งแรก 4. ระยะเวลาที่ใช้ยาเสพติด 5. เคยรับการบำบัดรักษาหรือไม่ 6. สาเหตุที่เข้ารับการบำบัดรักษา
ตัวแบบป่าสุ่มที่ทำกรคัดเลือกตัวแปรด้วยค่า Feature Importance (Random Forest with Feature Selection by Feature Importance)	<ol style="list-style-type: none"> 1. ยาเสพติดที่ใช้ในปัจจุบัน 2. เคยรับการบำบัดรักษาหรือไม่ 3. ยาเสพติดที่ใช้ครั้งแรก 4. อาศัยอยู่กับใครในช่วง 30 วันที่ผ่านมา 5. ความสัมพันธ์ในครอบครัว 6. สาเหตุที่เข้ารับการบำบัดรักษา

จากตารางที่ 4.41 พบว่า ตัวแบบป่าสุ่มที่ใช้ตัวแปรอิสระเหมือนกับตัวแบบการถดถอยโลจิสติก มีตัวแปรอิสระที่นำมาใช้ในการสร้างตัวแบบ ดังนี้ ความสัมพันธ์ในครอบครัว, อายุที่ใช้ยาเสพติดครั้งแรก, ยาเสพติดที่ใช้ครั้งแรก, ระยะเวลาที่ใช้ยาเสพติด, เคยรับการบำบัดรักษาหรือไม่ และ สาเหตุที่เข้ารับการบำบัดรักษา ส่วนตัวแบบป่าสุ่มทำการคัดเลือกตัวแปรด้วยค่า Feature Importance มีตัวแปรอิสระที่นำมาใช้ในการสร้างตัวแบบ ดังนี้ ยาเสพติดที่ใช้ในปัจจุบัน, เคยรับการบำบัดรักษาหรือไม่, ยาเสพติดที่ใช้ครั้งแรก, อาศัยอยู่กับใครในช่วง 30 วันที่ผ่านมา, ความสัมพันธ์ในครอบครัว และสาเหตุที่เข้ารับการบำบัดรักษา

ตารางที่ 4.42 ผลการทำนายด้วยตัวแบบป่าสุ่มที่ทำการคัดเลือกตัวแปรโดยใช้ตัวแปรเหมือนกับ
ตัวแบบการถดถอยโลจิสติก

ค่าทำนาย	ค่าจริง	
	ไม่ครบกำหนด	ครบกำหนด
ไม่ครบกำหนด	561	140
ครบกำหนด	42	230

จากตารางที่ 4.42 พบว่า ตัวแบบป่าสุ่มที่ทำการคัดเลือกตัวแปรโดยใช้ตัวแปรเหมือนกับตัวแบบการถดถอยโลจิสติกที่ทำการคัดเลือกตัวแปรสามารถทำนายผลการบำบัดรักษาได้ถูกต้องตรงกับผลการรักษาจริงอยู่ที่ 791 คน โดยแบ่งเป็นบำบัดรักษาไม่ครบกำหนดจำนวน 561 คน และบำบัดรักษาครบกำหนดจำนวน 230 คน จากข้อมูลชุดทดสอบจำนวน 973 คน แสดงให้เห็นว่า ตัวแบบมีค่าความถูกต้องในการทำนายอยู่ที่ 81.29% และมีค่าคะแนน F1 เท่ากับ 86.04%

ตารางที่ 4.43 ผลการทำนายด้วยตัวแบบป่าสุ่มที่ทำการคัดเลือกตัวแปรด้วย ค่า Feature Importance

ค่าทำนาย	ค่าจริง	
	ไม่ครบกำหนด	ครบกำหนด
ไม่ครบกำหนด	556	114
ครบกำหนด	47	256

จากตารางที่ 4.43 พบว่า ตัวแบบป่าสุ่มที่ทำการคัดเลือกตัวแปรโดยใช้จำนวนตัวแปรเท่ากับตัวแบบการถดถอยโลจิสติก สามารถทำนายผลการบำบัดรักษาได้ถูกต้องตรงกับผลการรักษาจริงอยู่ที่ 812 คน โดยแบ่งเป็นบำบัดรักษาไม่ครบกำหนดจำนวน 556 คน และบำบัดรักษาครบกำหนดจำนวน 256 คน จากข้อมูลชุดทดสอบจำนวน 973 คน แสดงให้เห็นว่า ตัวแบบมีค่าความถูกต้องในการทำนายอยู่ที่ 83.45% และมีค่าคะแนน F1 เท่ากับ 87.35%

4.2.2.3 ตัวแบบซัพพอร์ตเวกเตอร์

สำหรับตัวแบบซัพพอร์ตเวกเตอร์แมชชีน ทำการคัดเลือกตัวแปร 2 แบบ แบบที่ 1 คัดเลือกตัวแปรโดยใช้ตัวแปรเหมือนกับตัวแบบการถดถอยโลจิสติก และแบบที่ 2 คัดเลือกตัวแปรด้วยค่า Feature Importance ได้ผลดังตารางที่ 4.44 จากนั้นสร้างตัวแบบและสร้างตาราง confusion matrix เพื่อดูผลการทำนาย ได้ผลดังตารางที่ 4.45 และตารางที่ 4.46 ตามลำดับ

ตารางที่ 4.44 ผลการคัดเลือกตัวแปรของตัวแบบซัพพอร์ตเวกเตอร์แมชชีน

ตัวแบบ	ตัวแปร
ตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ใช้ตัวแปรอิสระเหมือนกับตัวแบบการถดถอยโลจิสติก (Support Vector Machine with Feature Selection same as Logistic Regression)	<ol style="list-style-type: none"> 1. ความสัมพันธ์ในครอบครัว 2. อายุที่ใช้ยาเสพติดครั้งแรก 3. ยาเสพติดที่ใช้ครั้งแรก 4. ระยะเวลาที่ใช้ยาเสพติด 5. เคยรับการบำบัดรักษาหรือไม่ 6. สาเหตุที่เข้ารับการบำบัดรักษา
ตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ทำการคัดเลือกตัวแปรด้วยค่า Feature Importance (Support Vector Machine with Feature Selection by Feature Importance)	<ol style="list-style-type: none"> 1. อาศัยอยู่กับใครในช่วง 30 วันที่ผ่านมา 2. ยาเสพติดที่ใช้ในปัจจุบัน 3. อาชีพ 4. ยาเสพติดที่ใช้ครั้งแรก 5. อายุ 6. การศึกษา

จากตารางที่ 4.44 พบว่า ตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ใช้ตัวแปรอิสระเหมือนกับตัวแบบการถดถอยโลจิสติกมีตัวแปรอิสระที่นำมาใช้ในการสร้างตัวแบบ ดังนี้ ความสัมพันธ์ในครอบครัว, อายุที่ใช้ยาเสพติดครั้งแรก, ยาเสพติดที่ใช้ครั้งแรก, ระยะเวลาที่ใช้ยาเสพติด, เคยรับการบำบัดรักษาหรือไม่ และสาเหตุที่เข้ารับการบำบัดรักษา ส่วนตัวแบบซัพพอร์ตเวกเตอร์แมชชีนทำการคัดเลือกตัวแปรด้วยค่า Feature Importance มีตัวแปรอิสระที่นำมาใช้ในการสร้างตัวแบบ ดังนี้ อาศัยอยู่กับใครในช่วง 30 วันที่ผ่านมา, ยาเสพติดที่ใช้ในปัจจุบัน, อาชีพ, ยาเสพติดที่ใช้ครั้งแรก, อายุ และการศึกษา

ตารางที่ 4.45 ผลการทำนายด้วยตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ทำการคัดเลือกตัวแปรโดยใช้ตัวแปรเหมือนกับตัวแบบการถดถอยโลจิสติก

ค่าทำนาย	ค่าจริง	
	ไม่ครบกำหนด	ครบกำหนด
ไม่ครบกำหนด	576	125
ครบกำหนด	27	245

จากตารางที่ 4.45 พบว่า ตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ทำการคัดเลือกตัวแปรโดยใช้ตัวแปรเหมือนกับตัวแบบการถดถอยโลจิสติกที่ทำการคัดเลือกตัวแปรด้วยค่าสิ่ง stats::step สามารถทำนายผลการบำบัดรักษาได้ถูกต้องตรงกับผลการรักษาจริงอยู่ที่ 821 คน โดยแบ่งเป็นบำบัดรักษาไม่ครบกำหนดจำนวน 576 คน และบำบัดรักษาครบกำหนดจำนวน 245 คน จากข้อมูลชุดทดสอบจำนวน 973 คน แสดงให้เห็นว่า ตัวแบบมีค่าความถูกต้องในการทำนายอยู่ที่ 84.38% และมีค่าคะแนน F1 เท่ากับ 88.34%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.46 ผลการทำนายด้วยตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ทำการคัดเลือกตัวแปร ด้วยค่า Feature Importance

ค่าทำนาย	ค่าจริง	
	ไม่ครบกำหนด	ครบกำหนด
ไม่ครบกำหนด	594	59
ครบกำหนด	9	311

จากตารางที่ 4.46 พบว่า ตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ทำการคัดเลือกตัวแปร โดยใช้จำนวนตัวแปรเท่ากับตัวแบบการถดถอยโลจิสติก สามารถทำนายผลการบำบัดรักษาได้ถูกต้องตรงกับผลการรักษาจริงอยู่ที่ 905 คน โดยแบ่งเป็นบำบัดรักษาไม่ครบกำหนดจำนวน 594 คน และบำบัดรักษาครบกำหนดจำนวน 311 คน จากข้อมูลชุดทดสอบจำนวน 973 คน แสดงให้เห็นว่าตัวแบบมีค่าความถูกต้องในการทำนายอยู่ที่ 93.01% และมีค่าคะแนน F1 เท่ากับ 94.59%

4.2.3 ผลการสร้างตัวแบบโดยใช้ตัวแปรอิสระไม่ครบทุกตัวจากการคัดเลือกตัวแปร ร่วมกับการปรับค่าไฮเปอร์พารามิเตอร์

ผู้วิจัยได้ทำการปรับค่าไฮเปอร์พารามิเตอร์ในตัวแบบที่ทำการคัดเลือกตัวแปรเพื่อเป็นการเพิ่มประสิทธิภาพให้กับตัวแบบ ได้ผลการปรับค่าไฮเปอร์พารามิเตอร์ ดังตารางที่ 4.47 และสร้างตาราง confusion matrix เพื่อดูผลการทำนายได้ผลดังตารางที่ 4.48 - 4.52

ตารางที่ 4.47 ค่าพารามิเตอร์หลังจากมีการปรับค่าไฮเปอร์พารามิเตอร์กับตัวแบบ

ตัวแบบที่ใช้ตัวแปรเหมือนกับตัวแบบการถดถอยโลจิสติก	
ตัวแบบ	พารามิเตอร์
ป่าสุ่ม	mtry: 3 trees: 1250
ซัพพอร์ตเวกเตอร์แมชชีน	C: 1 Rbf_sigma: 0.01
ตัวแบบที่ทำการคัดเลือกตัวแปรด้วยค่า Feature Importance	
ตัวแบบ	พารามิเตอร์
ป่าสุ่ม	mtry: 4 trees: 1500
ซัพพอร์ตเวกเตอร์แมชชีน	C: 1 Rbf_sigma: 0.01

จากตารางที่ 4.47 ทำการปรับค่าไฮเปอร์พารามิเตอร์กับตัวแบบ 2 แบบ ตัวแบบที่ทำการปรับค่าไฮเปอร์พารามิเตอร์กับตัวแบบป่าสุ่มและตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ใช้ตัวแปรอิสระเหมือนกับตัวแบบการถดถอยโลจิสติก พบว่า ตัวแบบป่าสุ่ม ได้ค่าพารามิเตอร์ mtry เป็น 3 trees เป็น 1250 และตัวแบบซัพพอร์ตเวกเตอร์แมชชีนได้ค่าพารามิเตอร์ C เป็น 1 Rbf_sigma เป็น 0.01 แบบที่สอง คือ ตัวแบบที่ทำการปรับค่าไฮเปอร์พารามิเตอร์กับตัวแบบป่าสุ่มและตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่คัดเลือกตัวแปรด้วยค่า Feature Importance พบว่า ตัวแบบป่าสุ่มได้ค่าพารามิเตอร์ mtry เป็น 4 trees เป็น 1500 และตัวแบบซัพพอร์ตเวกเตอร์แมชชีนได้ค่าพารามิเตอร์ C เป็น 1 Rbf_sigma เป็น 0.01

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.48 ผลการทำนายด้วยตัวแบบป่าสุ่มที่ทำการคัดเลือกตัวแปรโดยใช้ตัวแปรเหมือนกับตัวแบบการถดถอยโลจิสติกพร้อมกับการปรับค่าไฮเปอร์พารามิเตอร์

ค่าทำนาย	ค่าจริง	
	ไม่ครบกำหนด	ครบกำหนด
ไม่ครบกำหนด	559	44
ครบกำหนด	148	222

จากตารางที่ 4.48 พบว่า ตัวแบบป่าสุ่มที่ทำการคัดเลือกตัวแปรโดยใช้ตัวแปรเหมือนกับตัวแบบการถดถอยโลจิสติกพร้อมกับการปรับค่าไฮเปอร์พารามิเตอร์ สามารถทำนายผลการบำบัดรักษาได้ถูกต้องตรงกับผลการรักษาจริงอยู่ที่ 781 คน โดยแบ่งเป็นบำบัดรักษาไม่ครบกำหนดจำนวน 559 คน และบำบัดรักษาครบกำหนดจำนวน 222 คน จากข้อมูลชุดทดสอบจำนวน 973 คน แสดงให้เห็นว่า ตัวแบบมีค่าความถูกต้องในการทำนายอยู่ที่ 80.58% และมีค่าคะแนน F1 เท่ากับ 85.52%

ตารางที่ 4.49 ผลการทำนายด้วยตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ทำการคัดเลือกตัวแปรโดยตัวแปรใช้เหมือนกับตัวแบบการถดถอยโลจิสติกพร้อมกับการปรับค่าไฮเปอร์พารามิเตอร์

ค่าทำนาย	ค่าจริง	
	ไม่ครบกำหนด	ครบกำหนด
ไม่ครบกำหนด	531	169
ครบกำหนด	44	201

จากตารางที่ 4.49 พบว่า ตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ทำการคัดเลือกตัวแปรโดยใช้ตัวแปรเหมือนกับตัวแบบการถดถอยโลจิสติกพร้อมกับการปรับค่าไฮเปอร์พารามิเตอร์ สามารถทำนายผลการบำบัดรักษาได้ถูกต้องตรงกับผลการรักษาจริงอยู่ที่ 760 คน โดยแบ่งเป็นบำบัดรักษาไม่ครบกำหนดจำนวน 531 คน และบำบัดรักษาครบกำหนดจำนวน 201 คน จากข้อมูลชุดทดสอบจำนวน 973 คน แสดงให้เห็นว่า ตัวแบบมีค่าความถูกต้องในการทำนายอยู่ที่ 78.11% และมีค่าคะแนน F1 เท่ากับ 84.00%

ตารางที่ 4.50 ผลการทำนายด้วยตัวแบบป่าสุ่มที่ทำการคัดเลือกตัวแปรด้วยค่า Feature Importance พร้อมกับการปรับค่าไฮเปอร์พารามิเตอร์

ค่าทำนาย	ค่าจริง	
	ไม่ครบกำหนด	ครบกำหนด
ไม่ครบกำหนด	559	44
ครบกำหนด	119	251

จากตารางที่ 4.50 พบว่า ตัวแบบป่าสุ่มที่ทำการคัดเลือกตัวแปรด้วยค่า Feature Importance พร้อมกับการปรับค่าไฮเปอร์พารามิเตอร์ สามารถทำนายผลการบำบัดรักษา
เอกสารนี้เป็นเอกสารลิขสิทธิ์ไว้สำหรับการแข่งขันเพื่อการศึกษาเท่านั้น เมื่อผู้เห็นใบสมัครแข่งขันใด ๆ
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ได้ถูกต้องตรงกับผลการรักษาจริงอยู่ที่ 810 คน โดยแบ่งเป็นบำบัดรักษาไม่ครบกำหนดจำนวน 559 คน และบำบัดรักษาครบกำหนดจำนวน 251 คน จากข้อมูลชุดทดสอบจำนวน 973 คน แสดงให้เห็นว่า ตัวแบบมีค่าความถูกต้องในการทำนายอยู่ที่ 83.14% และมีค่าคะแนน F1 เท่ากับ 87.21%

ตารางที่ 4.51 ผลการทำนายด้วยตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ทำการคัดเลือกตัวแปร ด้วยค่า Feature Importance พร้อมกับทำการปรับค่าไฮเปอร์พารามิเตอร์

ค่าทำนาย	ค่าจริง	
	ไม่ครบกำหนด	ครบกำหนด
ไม่ครบกำหนด	531	156
ครบกำหนด	72	214

จากตารางที่ 4.51 พบว่า ตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ทำการคัดเลือกตัวแปร ด้วยค่า Feature Importance พร้อมกับทำการปรับค่าไฮเปอร์พารามิเตอร์ สามารถทำนายผลการบำบัดรักษาได้ถูกต้องตรงกับผลการรักษาจริงอยู่ที่ 745 คน โดยแบ่งเป็นบำบัดรักษาไม่ครบกำหนดจำนวน 531 คน และบำบัดรักษาครบกำหนดจำนวน 214 คน จากข้อมูลชุดทดสอบจำนวน 973 คน แสดงให้เห็นว่า ตัวแบบมีค่าความถูกต้องในการทำนายอยู่ที่ 76.57% และมีค่าคะแนน F1 เท่ากับ 82.33%

4.3 ผลการประเมินประสิทธิภาพตัวแบบ

หลังจากสร้างตัวแบบ ผู้วิจัยนำทุกตัวแบบมาทำการประเมินประสิทธิภาพ โดยการทำนายข้อมูลชุดทดสอบ สร้างตาราง confusion matrix และทำการประเมินประสิทธิภาพ โดยให้ความสนใจกับคะแนน F1 และค่าความถูกต้อง (Accuracy) ดังตารางที่ 4.53 - 4.57 และเปรียบเทียบตัวแบบที่มีประสิทธิภาพดีที่สุดในแต่ละวิธีการจำแนก ดังตารางที่ 4.58

ตารางที่ 4.52 ประสิทธิภาพของตัวแบบที่ใช้ตัวแปรอิสระครบทุกตัว (Full Model)

ตัวแบบที่	ตัวแบบ	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
1	ตัวแบบการถดถอยโลจิสติก	76.77	78.60	85.90	82.09
2	ตัวแบบป่าสุ่ม	88.28	97.01	85.90	91.12
3	ตัวแบบซัพพอร์ตเวกเตอร์แมชชีน	98.56	100.00	97.73	98.85

จากตารางที่ 4.52 การประเมินประสิทธิภาพตัวแบบแต่ละตัวแบบโดยใช้ตัวแปรอิสระครบทุกตัว พบว่า ตัวแบบการถดถอยโลจิสติกมีคะแนน F1 82.09% ตัวแบบป่าสุ่มมีคะแนน F1 91.12% และตัวแบบซัพพอร์ตเวกเตอร์แมชชีนมีคะแนน F1 98.85% ซึ่งมียุทธศาสตร์ที่ดีที่สุด ใน 3 ตัวแบบ

ตารางที่ 4.53 ประสิทธิภาพของตัวแบบที่ใช้ตัวแปรอิสระเหมือนกัน

ตัวแบบ ที่	ตัวแบบ	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
4	ตัวแบบการถดถอยโลจิสติก	76.77	77.28	88.56	82.53
5	ตัวแบบป่าสุ่ม	81.29	93.03	80.03	86.04
6	ตัวแบบซัพพอร์ตเวกเตอร์แมชชีน	84.38	95.52	82.17	88.34

จากตารางที่ 4.53 การประเมินประสิทธิภาพตัวแบบแต่ละตัวแบบโดยใช้ตัวแปรอิสระเหมือนกัน พบว่า ตัวแบบการถดถอยโลจิสติกมีคะแนน F1 82.53% ตัวแบบป่าสุ่มมีคะแนน F1 86.04% และตัวแบบซัพพอร์ตเวกเตอร์แมชชีนมีคะแนน F1 88.34% ซึ่งมีค่ามากที่สุด ใน 3 ตัวแบบ และยังพบว่า การคัดเลือกตัวแปรทำให้คะแนน F1 เพิ่มขึ้นจากตัวแบบที่ใช้ตัวแปรอิสระครบทุกตัว ในตัวแบบการถดถอยโลจิสติกซึ่งหมายความว่าตัวแบบมีประสิทธิภาพมากขึ้น ส่วนตัวแบบป่าสุ่มและตัวแบบซัพพอร์ตเวกเตอร์แมชชีนมีคะแนน F1 น้อยลงเมื่อเทียบกับตัวแบบที่ใช้ตัวแปรอิสระครบทุกตัว ซึ่งหมายความว่าตัวแบบมีประสิทธิภาพลดลง

ตารางที่ 4.54 ประสิทธิภาพของตัวแบบที่ใช้ตัวแปรอิสระตามค่า Feature Importance

ตัวแบบ ที่	ตัวแบบ	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
7	ตัวแบบการถดถอยโลจิสติก	76.46	78.95	84.58	81.67
8	ตัวแบบป่าสุ่ม	83.45	92.20	82.99	87.35
9	ตัวแบบซัพพอร์ตเวกเตอร์แมชชีน	93.01	98.51	90.96	94.59

จากตารางที่ 4.54 การประเมินประสิทธิภาพตัวแบบแต่ละตัวแบบโดยใช้จำนวนตัวแปรอิสระตามค่า Feature Importance พบว่า ตัวแบบการถดถอยโลจิสติกมีคะแนน F1 81.67% ตัวแบบป่าสุ่มมีคะแนน F1 87.35% และตัวแบบซัพพอร์ตเวกเตอร์แมชชีนมีคะแนน F1 94.59% ซึ่งมีค่ามากที่สุด ใน 3 ตัวแบบ และยังพบว่า การคัดเลือกตัวแปรด้วยการใช้ค่า Feature Importance ทำให้คะแนน F1 ลดน้อยลงเมื่อเทียบกับตัวแบบที่ใช้ตัวแปรอิสระครบทุกตัวในทั้ง 3 ตัวแบบ

ตารางที่ 4.55 ประสิทธิภาพของตัวแบบโดยใช้ตัวแปรอิสระเหมือนกันพร้อมกับมีการปรับค่าไฮเปอร์พารามิเตอร์

ตัวแบบ ที่	ตัวแบบ	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
10	ตัวแบบการถดถอยโลจิสติก	76.88	77.16	89.05	82.68
11	ตัวแบบป่าสุ่ม	80.58	92.54	79.49	85.52
12	ตัวแบบซัพพอร์ตเวกเตอร์แมชชีน	78.11	92.73	76.79	84.00

จากตารางที่ 4.55 การประเมินประสิทธิภาพตัวแบบแต่ละตัวแบบโดยใช้ตัวแปรอิสระเหมือนกันพร้อมกับทำการปรับค่าไฮเปอร์พารามิเตอร์ พบว่า ตัวแบบการถดถอยโลจิสติกมีคะแนน F1 82.68% เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้ในการศึกษาเท่านั้น เมื่ออนุญาตให้เผยแพร่ไปยังบุคคลอื่นโดยไม่ได้รับอนุญาต ไม่ว่าจะโดยวิธีใดก็ตาม อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวแบบป่าสุ่มมีคะแนน F1 85.52% และตัวแบบซัพพอร์ตเวกเตอร์แมชชีนมีคะแนน F1 84.00% และยิ่งพบว่า การปรับค่าไฮเปอร์พารามิเตอร์ทำให้คะแนน F1 เพิ่มขึ้นจากตัวแบบที่ใช้ตัวแปรอิสระเหมือนกันโดยที่ยังไม่ทำการปรับค่าไฮเปอร์พารามิเตอร์ในตัวแบบการถดถอยโลจิสติก ส่วนตัวแบบป่าสุ่มและตัวแบบซัพพอร์ตเวกเตอร์แมชชีนมีคะแนน F1 น้อยลงเมื่อเทียบกับตัวแบบที่ใช้ตัวแปรอิสระเหมือนกันโดยที่ยังไม่ทำการปรับค่าไฮเปอร์พารามิเตอร์

ตารางที่ 4.56 ประสิทธิภาพของตัวแบบที่ใช้ตัวแปรอิสระตามค่า Feature Importance พร้อมกับมีการปรับค่าไฮเปอร์พารามิเตอร์

ตัวแบบที่	ตัวแบบ	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
13	ตัวแบบการถดถอยโลจิสติก	76.46	78.95	84.58	81.67
14	ตัวแบบป่าสุ่ม	83.14	92.70	82.33	87.21
15	ตัวแบบซัพพอร์ตเวกเตอร์แมชชีน	76.57	88.06	77.29	82.33

จากตารางที่ 4.56 การประเมินประสิทธิภาพตัวแบบแต่ละตัวแบบที่ใช้ตัวแปรอิสระตามค่า Feature Importance พร้อมกับทำการปรับค่าไฮเปอร์พารามิเตอร์ พบว่า ตัวแบบการถดถอยโลจิสติกมีคะแนน F1 81.67% ตัวแบบป่าสุ่มมีคะแนน F1 87.21% และตัวแบบซัพพอร์ตเวกเตอร์แมชชีนมีคะแนน F1 82.33% ซึ่งตัวแบบที่มีคะแนน F1 มากที่สุดคือตัวแบบป่าสุ่ม และยิ่งพบว่า การปรับค่าไฮเปอร์พารามิเตอร์ทำให้คะแนน F1 น้อยลงในทั้ง 3 ตัวแบบ เมื่อเทียบกับตัวแบบที่ใช้ตัวแปรอิสระตามค่า Feature Importance โดยที่ยังไม่ทำการปรับค่าไฮเปอร์พารามิเตอร์

ตารางที่ 4.57 เปรียบเทียบประสิทธิภาพที่ดีที่สุดของแต่ละตัวแบบ

ตัวแบบ	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
ตัวแบบที่ 10 : ตัวแบบการถดถอยโลจิสติกที่ใช้ตัวแปรเหมือนกันพร้อมกับการปรับค่าไฮเปอร์พารามิเตอร์	76.88	77.16	89.05	82.68
ตัวแบบที่ 2 : ตัวแบบป่าสุ่มที่ใช้ตัวแปรอิสระครบทุกตัว (Full model)	88.28	97.01	85.90	91.12
ตัวแบบที่ 3 : ตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ใช้อิสระครบทุกตัว (Full Model)	98.56	100.00	97.73	98.85

จากตารางที่ 4.57 พบว่า ตัวแบบการถดถอยโลจิสติกที่มีประสิทธิภาพที่ดีที่สุด คือ ตัวแบบที่ 10 : ตัวแบบการถดถอยโลจิสติกที่ใช้ตัวแปรเหมือนกันพร้อมกับการปรับค่าไฮเปอร์พารามิเตอร์โดยมีคะแนน F1 82.68% ตัวแบบป่าสุ่มที่มีประสิทธิภาพที่ดีที่สุด คือ ตัวแบบที่ 2 : ตัวแบบป่าสุ่มที่ใช้ตัวแปร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้นำไปเผยแพร่โดยไม่ผ่านการอนุมัติจากเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อิสระทุกตัว มีคะแนน F1 91.12% ตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่มีประสิทธิภาพดีที่สุด คือ ตัวแบบที่ 3 :ตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ใช้ตัวแปรอิสระทุกตัว มีคะแนน F1 98.85% และจากทั้ง 3 ตัวแบบนี้ ตัวแบบที่มีประสิทธิภาพดีที่สุด คือ ตัวแบบที่ 3 : ตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ใช้ตัวแปรอิสระทุกตัว มีคะแนน F1 98.85%

นอกจากนี้ตัวแบบการถดถอยโลจิสติกที่มีประสิทธิภาพดีที่สุด คือ ตัวแบบการถดถอยโลจิสติกที่ใช้ตัวแปรเหมือนกัน พร้อมกับมีการปรับค่าไฮเปอร์พารามิเตอร์ มีสมการดังนี้

$$\begin{aligned} \log(\text{odds}) = & 0.377 + 0.595(\text{fam04}) + 1.533(\text{agefirst01}) + 2.369(\text{drugfirst01}) + 5.532(\text{drugfirst02}) \\ & + 1.715(\text{drugfirst03}) + 4.324(\text{drugfirst04}) + 2.780(\text{drugfirst06}) + 0.637(\text{timeuse01}) \\ & + 2.196(\text{eventreatment01}) + 8.124(\text{causetreatment01}) + 6.274(\text{causetreatment02}) \\ & + 0.131(\text{causetreatment03}) + 3.829(\text{causetreatment04}) + 2.394(\text{causetreatment06}) \end{aligned} \quad (4.2)$$



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สรุปผลการวิจัยและข้อเสนอแนะ

ในบทนี้จะกล่าวถึงผลสรุปและข้อเสนอแนะที่ได้จากการวิเคราะห์ข้อมูลการเปรียบเทียบตัวแบบการจำแนกและปัจจัยที่ส่งผลต่อกลุ่มการยุติการบำบัดรักษาผู้ติดยาเสพติดแบบระบบสมัครใจในพื้นที่กรุงเทพมหานคร จำนวนทั้งหมด 3,241 ข้อมูล ด้วยการสร้างตัวแบบการจำแนกทั้งหมด 3 ตัวแบบ ได้แก่ ตัวแบบการถดถอยโลจิสติก ตัวแบบป่าสุ่ม และตัวแบบซัพพอร์ตเวกเตอร์แมชชีนพร้อมทั้งทำการคัดเลือกตัวแปรร่วมกับการปรับค่าไฮเปอร์พารามิเตอร์ และทำการประเมิน ประสิทธิภาพของตัวแบบด้วยคะแนน F1 ซึ่งสามารถสรุปผลการวิจัยและข้อเสนอแนะ ได้ดังนี้

5.1 สรุปผลการวิจัย

1. จากการศึกษาวิธีการเปรียบเทียบและสร้างตัวแบบการจำแนกกลุ่มการยุติการบำบัดรักษาผู้ติดยาเสพติด โดยใช้การเรียนรู้ของเครื่อง 3 วิธี คือ วิธีการถดถอยโลจิสติก (Logistic Regression) วิธีป่าสุ่ม (Random Forest) และวิธีซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) สรุปได้ว่าตัวแบบที่มีประสิทธิภาพมากที่สุด คือ ตัวแบบซัพพอร์ตเวกเตอร์แมชชีนแบบที่ใช้ตัวแปรอิสระครบทุกตัว

2. จากการศึกษาและจำแนกข้อมูลผู้ติดยาเสพติดแบบระบบสมัครใจในพื้นที่กรุงเทพมหานคร ด้วยตัวแบบการจำแนกที่ทำการคัดเลือกตัวแปรทั้ง 2 วิธีในแต่ละตัวแบบ ได้แก่ การคัดเลือกตัวแปรโดยใช้ตัวแปรอิสระเหมือนกันซึ่งเป็นการใช้ตัวแปรอิสระที่ได้จากการคัดเลือกตัวแปรในตัวแบบการถดถอยโลจิสติกด้วยวิธี stepwise selection และการคัดเลือกตัวแปรโดยใช้ค่า Feature Importance สรุปได้ว่า สำหรับตัวแบบการถดถอยโลจิสติก การคัดเลือกตัวแปรโดยใช้ตัวแปรอิสระเหมือนกันเป็นวิธีการคัดเลือกตัวแปรที่ดีกว่าการคัดเลือกตัวแปรโดยใช้ค่า Feature Importance เนื่องจากมีประสิทธิภาพมากกว่า ส่วนตัวแบบป่าสุ่มและตัวแบบซัพพอร์ตเวกเตอร์แมชชีน การคัดเลือกตัวแปรโดยใช้ค่า Feature Importance เป็นวิธีการคัดเลือกตัวแปรที่ดีกว่าการคัดเลือกตัวแปรโดยใช้ตัวแปรอิสระเหมือนกัน เนื่องจากมีประสิทธิภาพมากกว่า

5.2 ข้อเสนอแนะ

1) การคัดเลือกตัวแปรด้วยค่า Feature Importance และใช้ตัวแปรอิสระเหมือนกับตัวแบบการถดถอยโลจิสติกในตัวแบบป่าสุ่มและตัวแบบซัพพอร์ตเวกเตอร์แมชชีนทำให้ตัวแบบมีประสิทธิภาพน้อยลงเมื่อเทียบกับประสิทธิภาพของตัวแบบป่าสุ่มและตัวแบบซัพพอร์ตเวกเตอร์แมชชีนที่ใช้ตัวแปรอิสระครบทุกตัว ทั้งนี้อาจมีการคัดเลือกตัวแปรแบบอื่นที่เหมาะสมกับตัวแบบทั้งสองเพื่อที่จะเพิ่มประสิทธิภาพในการทำนายได้อย่างถูกต้องแม่นยำมากขึ้น

2) สามารถนำตัวแบบไปใช้ทำนายลักษณะผู้เข้ารับการบำบัดรักษาในระบบข้อมูลการบำบัดรักษาและฟื้นฟูผู้ติดยาเสพติดของประเทศ (บสต.) ในอนาคตได้ หากตัวแบบทำนายว่าผู้เข้ารับการบำบัดมีโอกาสยุติการเข้าร่วมในโปรแกรมบำบัดยาเสพติด สถานบำบัดและหน่วยงานที่เกี่ยวข้องควรให้ความสนใจและเอาใจใส่ผู้เข้ารับการบำบัดรายนั้นมากขึ้น เพื่อเพิ่มโอกาสในการบำบัดครบกำหนดมากยิ่งขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

- กนกลักษณ์ ศิริรุ่งวัฒนากุล, “ประสิทธิผลและปัจจัยที่มีผลต่อการบำบัดรักษาผู้ติดยาและสารเสพติด โดยใช้รูปแบบการบำบัดความคิดและพฤติกรรมแบบประยุกต์ของโรงพยาบาลสวรรค์ประชารักษ์”, *สวรรค์ประชารักษ์เวชสาร*, ปีที่ 16, ฉบับที่ 1, น. 23-33, 2562.
- กมล บุขบา, มทนาลัย สุดสายสาคร, ภรณ์จิรา เตชะพรสิน, และ สรัลพร โกมลทองทิพย์. การเปรียบเทียบ การทำนายการต่ออายุกรรมธรรม์ประกันภัยรถยนต์ประเภท 1 ระหว่างตัวแบบป่าสุ่ม และ ตัวแบบบูสต์ติงปรับได้. *Thai Journal of Science and Technology*, ปีที่ 10, ฉบับที่ 2, น. 124-134, 2564.
- กัลยา วานิชย์บัญชา, การวิเคราะห์สถิติขั้นสูงด้วย SPSS FOR WINDOWS, (พิมพ์ครั้งที่ 14), กรุงเทพฯ : บริษัท สามลดา จำกัด, 2564
- กัลยา วานิชย์บัญชา, สถิติสำหรับงานวิจัย, (พิมพ์ครั้งที่ 15), กรุงเทพฯ : บริษัท สามลดา จำกัด, 2567
- จรรยาตรี โคมพุดชา และชัยภัทร อีร์ชาญไชย, “ปัจจัยที่มีความสัมพันธ์กับการบำบัดครบกำหนดของผู้ใช้ สารเสพติดในระบบสมัครใจ จังหวัดนครราชสีมา ปีงบประมาณ 2555-2556”, *วารสาร วิทยาศาสตร์และเทคโนโลยี (สทวท.)*, ปีที่ 2, ฉบับที่ 1, น. 21-34, 2558.
- ชลอวัฒน์ อินปา และพิศิษฐ์ ศรีประเสริฐ, “ปัจจัยที่ส่งผลต่อการเข้ารับการรักษาครบกำหนด ตามเกณฑ์ของผู้ป่วยยาเสพติดในจังหวัดตาก”, *วารสารสาธารณสุขล้านนา*, ปีที่ 17, ฉบับที่ 2, น. 91-100, 2564.
- ภรณ์ยา ปาลวิสุทธิ, “การเพิ่มประสิทธิภาพเทคนิคต้นไม้ตัดสินใจบนข้อมูลที่ไม่สมดุลโดยวิธีการสุ่มเพิ่ม ตัวอย่างกลุ่มน้อยสำหรับข้อมูลการเป็นโรคติดเชื้อไวรัสเอดส์”, *วารสารเทคโนโลยีสารสนเทศ*, ปีที่ 12, ฉบับที่ 1, น. 54-63, 2559.
- ศรัณย์พร กาญจนสุนทร. 2566. “การจำแนกประเภทความรู้สึกของรีวิวกู้บริการแอปสโอดีฟาย ด้วยวิธีการเรียนรู้ด้วยเครื่อง” วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิทยาการข้อมูล, มหาวิทยาลัยศรีนครินทรวิโรฒ.
- ศราววุฒิ ทองเชื้อ. วิธีการเรียนรู้แบบกลุ่มด้วยเทคนิค Stacking สามารถเพิ่มความแม่นยำของการ ทำนายสัญญาณการซื้อขายสกุลเงินดิจิทัลได้หรือไม่ (Doctoral dissertation, จุฬาลงกรณ์ มหาวิทยาลัย), 2566.
- ศูนย์อำนวยการป้องกันและปราบปรามยาเสพติด กระทรวงสาธารณสุข “แผนยุทธศาสตร์ด้านการ บำบัดรักษาและฟื้นฟู ผู้ป่วยยาเสพติดประจำปีงบประมาณ พ.ศ. 2564 – 2566 และ แผนปฏิบัติการประจำปี 2565”, 2565. แหล่งข้อมูล :https://ncmc.moph.go.th/home/upload/web_download/hzm6y1e59m8scog04o.pdf
- สมพร สุวรรณมาโจ, ดวงใจ ดวงฤทธิ์, ณิชฎณิชา คมกล้า, วราพร พัฒนะโชติ และอุไรวรรณ สถาพร , “ปัจจัยทำนายในการคงอยู่ในระบบระยะเวลา 1 ปี ของผู้บำบัดด้วยเมทาโดนระยะยาว ในคลินิกเมทาโดน ณ สถานบำบัดรักษาฟื้นฟูผู้ติดยาเสพติดแห่งชาติบรมราชชนนี”, *วารสาร กรมการแพทย์*, ปีที่ 44, ฉบับที่ 4, น. 119-124, 2562.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- สุรวุฑ ศรีเปารยะ และสายชล สนิสมบุญทอง, “การเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่มการเป็นโรคไตเรื้อรัง : กรณีศึกษาโรงพยาบาลแห่งหนึ่งในประเทศไทย”, *วารสารวิทยาศาสตร์และเทคโนโลยี*, ปีที่ 25, ฉบับที่ 5, น.839-853, 2560.
- สำนักงานเลขาธิการคณะกรรมการกำกับรักษาและฟื้นฟูผู้ติดยาเสพติด กระทรวงสาธารณสุข “คู่มือแนวทางการดำเนินงานบำบัดรักษาและฟื้นฟูสมรรถภาพผู้ติดยาเสพติดตามประมวลกฎหมายยาเสพติด(ฉบับปรับปรุง)”, 2565. แหล่งข้อมูล : https://ncmc.moph.go.th/home/upload/web_download/wpnrvtv5buow00088.pdf
- เสาวนีย์ พันธุ์พัฒนกุล และณัฐณีย์ มีมนต์, “ปัจจัยที่สัมพันธ์กับการบรรลุผลกระบวนการบำบัดรักษาฟื้นฟูสมรรถภาพและติดตามของผู้รับการรักษาติดยาเสพติด จังหวัดพะเยา”, *วารสารวิชาการป้องกันควบคุมโรค สคร.9 พิษณุโลก*, ปีที่ 3, ฉบับที่ 1, น. 71-83, 2558.
- อังกูร ภัทรการ และนันทา ชัยพิชิตพันธ์, “การเปรียบเทียบผลการบำบัดฟื้นฟูสมรรถภาพแบบเข้มข้นทางสายใหม่ (FAST Model) ระหว่างผู้ป่วยเสติดระบบสมัครใจกับระบบบังคับบำบัด:กรณีศึกษาสถาบันธัญญารักษ์”, *วารสารสมาคมจิตแพทย์แห่งประเทศไทย*, ปีที่ 60, ฉบับที่ 1, น. 71-81, 2558
- อารี สุภาวงศ์, “ประสิทธิผลของการบำบัดรักษาผู้ป่วยยาเสพติดตามรูปแบบจิตสังคมบำบัดแบบประยุกต์ โรงพยาบาลทุ่งสง”, *วารสารพยาบาลสงขลานครินทร์*, ปีที่ 36, ฉบับพิเศษ, น. 160-170, 2559
- Abhishek Jain, “Everything about Random Forest”, 2024, <https://medium.com/@abhishekjainindore24/everything-about-random-forest-90c106d63989>
- An, Qi, Saifur Rahman, Jingwen Zhou, and James Jin Kang. "A comprehensive review on machine learning in healthcare industry: classification, restrictions, opportunities and challenges." *Sensors* 2023, Vol. 23, No. 9, 2023.
- Anubha Singh, “Cross-Validation Techniques”, 2020, <https://medium.com/analytics-vidhya/cross-validation-techniques-a925782517e8>
- Bagnall, Anthony, and Gavin C. Cawley. "On the use of default parameter settings in the empirical evaluation of classification algorithms." *arXiv preprint arXiv:1703.06777*, 2017.
- Jakkula Vikramaditya. Tutorial on support vector machine (svm). School of EECS, Washington State University, Vol. 37, No. 2.5, 2006
- K. M. Rashedul et al. “A Comparative Machine Learning Study to Predict Drug Addiction in Bangladesh.” 2021 15th International Conference on Application of Information and Communication Technologies (AICT). IEEE, 2021
- Osborne, J. W., *Regression & linear modeling: Best practices and modern methods*. Sage Publications, 2016.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก

ตาราง ก1 ประสิทธิภาพชุดสอน (Training Data) และชุดทดสอบ (Testing Data) ของทุกตัวแบบ

ตัวแบบ ที่	ตัวแบบ	Accuracy (%)		Precision (%)		Recall (%)		F1-Score (%)	
		train	test	train	test	train	test	train	test
ตัวแบบที่ใช้ตัวแปรอิสระครบทุกตัว (Full Model)									
1	ตัวแบบการ ถดถอยโลจิสติก	78.44	76.77	80.49	78.60	86.05	85.90	83.18	82.09
2	ตัวแบบป่าสุ่ม	87.90	88.28	95.87	97.01	86.18	85.90	90.77	91.12
3	ตัวแบบซัพพอร์ต เวกเตอร์แมชชีน	98.41	98.56	99.93	100.00	97.57	97.73	98.73	98.85
ตัวแบบที่ใช้ตัวแปรอิสระเหมือนกัน									
4	ตัวแบบการ ถดถอยโลจิสติก	75.26	76.77	75.83	77.28	88.19	88.56	81.54	82.53
5	ตัวแบบป่าสุ่ม	78.79	81.29	93.38	93.03	77.18	80.03	84.51	86.04
6	ตัวแบบซัพพอร์ต เวกเตอร์แมชชีน	81.66	84.38	93.67	95.52	80.10	82.17	86.35	88.34
ตัวแบบโดยใช้จำนวนตัวแปรอิสระตามค่า Feature Importance									
7	ตัวแบบการ ถดถอยโลจิสติก	78.40	76.46	80.97	78.95	85.12	84.58	83.00	81.67
8	ตัวแบบป่าสุ่ม	82.09	83.45	90.18	92.20	82.54	82.99	86.19	87.35
9	ตัวแบบซัพพอร์ต เวกเตอร์แมชชีน	90.61	93.01	96.37	98.51	89.31	90.96	92.71	94.59
ตัวแบบโดยใช้ตัวแปรอิสระเหมือนกัน พร้อมกับการปรับค่าไฮเปอร์พารามิเตอร์									
10	ตัวแบบการ ถดถอยโลจิสติก	75.26	76.88	75.54	77.16	88.83	89.05	81.65	82.68
11	ตัวแบบป่าสุ่ม	78.17	80.58	92.81	92.54	76.80	79.49	84.05	85.52
12	ตัวแบบซัพพอร์ต เวกเตอร์แมชชีน	75.93	78.11	90.46	92.73	75.52	76.79	82.32	84.00
ตัวแบบโดยใช้จำนวนตัวแปรอิสระตามค่า Feature Importance พร้อมกับการปรับค่าไฮเปอร์พารามิเตอร์									
13	ตัวแบบการ ถดถอยโลจิสติก	78.35	76.46	80.96	78.95	85.05	84.58	82.96	81.67
14	ตัวแบบป่าสุ่ม	82.05	83.14	90.04	92.70	82.57	82.33	86.14	87.21
15	ตัวแบบซัพพอร์ต เวกเตอร์แมชชีน	77.51	76.57	80.50	88.06	82.74	77.29	81.60	82.33

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข

คำสั่งโปรแกรมอาร์ทีใช้ในงานวิจัย

```

Logistic
library(dplyr)
library(tidymodels)
library(caret)
library(glmnet)
library(car)
getwd()
setwd("c:/")
df= read.csv("drug_treatment(new)1.csv")
head(df)
str(df)

df = df %>%
  mutate(Evaluation_treatment =
    recode(Evaluation_treatment,
           "uncomplete" = 1,
           "complete" = 0)
  )

#####Split Data#####
set.seed(999)
split <- initial_split(df, prop = 0.7, strata = Evaluation_treatment)
train <- split %>%
  training()
test <- split %>%
  testing()

#Create all model 12 variables#
model <- glm(Evaluation_treatment ~
age+education+occupation+live.with.whom.in.30.days.ago +family.relation
+age_first_time_used_drug+drug_first_time
+cause_use_drugs.first_time.+drugs.present.+time_use_year+ever_received_treatmen
t+
cause_receive_treatment,family = 'binomial',data = train)

```

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อการศึกษาเท่านั้น เมื่อผู้ใดเห็นประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

summary(model)

# Extract the coefficients
coefficients = coef(model)
coefficients

# Calculate the odds ratios
odds_ratios = exp(coefficients)
odds_ratios

#Calculate confidence interval #
confint(model)

# Training Set #
train$Evaluation_treatment <- as.factor(train$Evaluation_treatment)
modelpred <- predict(model, train, type = 'response')
modelpreds <- factor(ifelse(modelpred>0.5, 1, 0))
confusionMatrix(modelpreds,train$Evaluation_treatment,
  dnn = c("Predict",'Actual'),
  positive = '1',
  mode = 'prec_recall')

# Predict and Evaluate all model with test set #
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)
test_modelpred <- predict(model, test, type = 'response')
test_modelpreds <- factor(ifelse(test_modelpred>0.5, 1, 0))
confusionMatrix(test_modelpreds,test$Evaluation_treatment,
  dnn = c("Predict",'Actual'),
  positive = '1',
  mode = 'prec_recall')

#Create model with feature selection#
model2 <- glm(Evaluation_treatment ~ education +family.relation
  +cause_use_drugs.first_time.+drugs.present.+ever_received_treatment+
  cause_receive_treatment, family = 'binomial',data = train)
vif(model2)

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

train$Evaluation_treatment <- as.factor(train$Evaluation_treatment)
modelpred2 <- predict(model2, train, type = 'response')
modelpreds2 <- factor(ifelse(modelpred2>0.5, 1, 0))
confusionMatrix(modelpreds2,train$Evaluation_treatment,
                 dnn = c("Predict",'Actual'),
                 positive = '1',
                 mode = 'prec_recall')

# Predict and Evaluate model with test set #
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)
test_modelpred2 <- predict(model2, test, type = 'response')
test_modelpreds2 <- factor(ifelse(test_modelpred2>0.5, 1, 0))
confusionMatrix(test_modelpreds2,test$Evaluation_treatment,
                 dnn = c("Predict",'Actual'),
                 positive = '1',
                 mode = 'prec_recall')
# Hyperparameter Tuning model with feature selection 6 variables #
hyper_model1 <- train(Evaluation_treatment~education +family.relation
                    +cause_use_drugs.first_time.+drugs.present.+ever_received_treatment+
                    cause_receive_treatment, data = train,
                    method = 'glmnet',
                    metric = 'Accuracy',
                    trControl = trainControl(method = 'cv',
                                             number = 10,
                                             search = 'grid',
                                             verboseIter = T))

train$Evaluation_treatment <- as.factor(train$Evaluation_treatment)
hyper_model1pred <- predict(hyper_model1, train, type = 'prob')
hyper_model1preds <-
factor(colnames(hyper_model1pred)[max.col(hyper_model1pred)])
hyper_model1preds <- as.factor(hyper_model1preds)
confusionMatrix(hyper_model1preds,train$Evaluation_treatment,
                 dnn = c("Predict",'Actual'),
                 positive = '1',
                 mode = 'prec_recall')

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

# Predict and Evaluate model with test set #
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)
test_hyper_model1pred <- predict(hyper_model1, test, type = 'prob')
test_hyper_model1preds <-
factor(colnames(test_hyper_model1pred)[max.col(test_hyper_model1pred)])
test_hyper_model1preds <- as.factor(test_hyper_model1preds)
confusionMatrix(test_hyper_model1preds,test$Evaluation_treatment,
                dnn = c("Predict",'Actual'),
                positive = '1',
                mode = 'prec_recall')

# Stepwise selection #
stats::step(model)
model4 <- glm(formula = Evaluation_treatment ~ family.relation +
age_first_time_used_drug +
                drug_first_time + drugs.present. + time_use_year +
ever_received_treatment +
                cause_receive_treatment, family = "binomial", data = train)
vif(model4)

# Create model with stepwise selection 6 variables #
#remove drugs.present.
model5 <- glm(formula = Evaluation_treatment ~ family.relation +
age_first_time_used_drug +
                drug_first_time+ time_use_year + ever_received_treatment +
                cause_receive_treatment, family = "binomial", data = train)
vif(model5)

train$Evaluation_treatment <- as.factor(train$Evaluation_treatment)
modelpred5 <- predict(model5, train, type = 'response')
modelpreds5 <- factor(ifelse(modelpred5>0.5, 1, 0))
confusionMatrix(modelpreds5,train$Evaluation_treatment,
                dnn = c("Predict",'Actual'),
                positive = '1',
                mode = 'prec_recall')

# Predict and Evaluate model with test set #
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

test_modelpred5 <- predict(model5, test, type = 'response')
test_modelpreds5 <- factor(ifelse(test_modelpred5>0.5, 1, 0))
confusionMatrix(test_modelpreds5,test$Evaluation_treatment,
                 dnn = c("Predict",'Actual'),
                 positive = '1',
                 mode = 'prec_recall')

### Hyperparameter tuning with 10-fold CV###
# Create model with stepwise selection 6 variables #
hyper_model5 <- train(Evaluation_treatment~family.relation +
age_first_time_used_drug +
                    drug_first_time + time_use_year + ever_received_treatment +
                    cause_receive_treatment, data = train,
                    method = 'glmnet',
                    metric = 'Accuracy',
                    trControl = trainControl(method = 'cv',
                                             number = 10,
                                             search = 'grid',
                                             verboseIter = T))

train$Evaluation_treatment <- as.factor(train$Evaluation_treatment)
hyper_model5pred <- predict(hyper_model5, train, type = 'prob')
hyper_model5preds <-
factor(colnames(hyper_model5pred)[max.col(hyper_model5pred)])
hyper_model5preds <- as.factor(hyper_model5preds)
confusionMatrix(hyper_model5preds,train$Evaluation_treatment,
                 dnn = c("Predict",'Actual'),
                 positive = '1',
                 mode = 'prec_recall')

```

```

# Predict and Evaluate model with test set #
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)
test_hyper_model5pred <- predict(hyper_model5, test, type = 'prob')
test_hyper_model5preds <-
factor(colnames(test_hyper_model5pred)[max.col(test_hyper_model5pred)])
test_hyper_model5preds <- as.factor(test_hyper_model5preds)
confusionMatrix(test_hyper_model5preds,test$Evaluation_treatment,
                 dnn = c("Predict",'Actual'),

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

positive = '1',
mode = 'prec_recall')

```

```

RandomForest
library(tidymodels)
library(vip)
library(tidyverse)
library(dplyr)
library(doParallel)
library(caret)
library(iml)

setwd("c:/")
df= read.csv("drug_treatment(new)1.csv")
head(df)
str(df)
df = df %>%
  mutate(Evaluation_treatment =
    recode(Evaluation_treatment,
      "uncomplete" = 1,
      "complete" = 0)
  )

df$Evaluation_treatment = as.factor(df$Evaluation_treatment)

#Split Data
set.seed(999)
split <- initial_split(df, prop = 0.7, strata = Evaluation_treatment)
train <- split %>%
  training()
test <- split %>%
  testing()
train$Evaluation_treatment <- as.factor(train$Evaluation_treatment)

# Create all model 12 variables #
rf_recipe_based = recipe(Evaluation_treatment ~
  age+education+occupation+live.with.whom.in.30.days.ago +family.relation

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

+age_first_time_used_drug+drug_first_time
+cause_use_drugs.first_time.+drugs.present.+time_use_year+ever_received_treatmen
t+
      cause_receive_treatment, data = train)%>%
step_dummy(all_nominal_predictors(), one_hot = TRUE)

rf_model_based <- rand_forest() %>%
  set_args(mtry = 4, trees = 500) %>%
  set_engine("ranger", importance = "impurity") %>%
  set_mode("classification")

workflow_based <- workflow() %>%
  add_recipe(rf_recipe_based) %>%
  add_model(rf_model_based)

# Training set #
# Make predictions #
rf_based_fit <- workflow_based %>%
  fit(data = train)

default_train_prediction <- predict(rf_based_fit,
  new_data = train)

# Create confusion matrix #
all_train_predicted <- default_train_prediction$.pred_class
train_actual <- train$Evaluation_treatment

confusion_matrix <- table(train_actual,all_train_predicted)
confusion_matrix

# Calculate accuracy, precision, recall, f1_score #
accuracy_train = confusionMatrix(confusion_matrix)$overall['Accuracy']
precision_train = posPredValue(confusion_matrix,
  positive = "1")
recall_train = sensitivity(confusion_matrix,
  positive = "1")
f1_score_train = 2*(precision_train*recall_train)/(precision_train + recall_train)

metrics_train = c(accuracy_train, precision_train, recall_train, f1_score_train)

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้拿去ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

metrics_train

# Test set #
# Make predictions
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)
rf_based_fittest <- workflow_based %>%
  fit(data = test)

allt_test_predictions <- predict(rf_based_fittest,
  new_data = test)

# Create confusion matrix
all_test_predicted = allt_test_predictions$.pred_class
test_actual = test$Evaluation_treatment

confusion_matrix_test = table(test_actual, all_test_predicted)
confusion_matrix_test

# Calculate accuracy, precision, recall, and F1 score
accuracy_test = confusionMatrix(confusion_matrix_test)$overall['Accuracy']

precision_test = posPredValue(confusion_matrix_test,
  positive = "1")
recall_test = sensitivity(confusion_matrix_test,
  positive = "1")
f1_score_test = 2 * (precision_test * recall_test) / (precision_test + recall_test)

# Combine metrics into a single vector
metrics_test = c(accuracy_test, precision_test, recall_test, f1_score_test)
metrics_test

rbind(metrics_train, metrics_test) %>%
  `colnames<-`(c("Accuracy", "Precision", "Recall", "F1-score"))

# Extract VIP
trainX = train[,-c(16)]

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

explainer <- Predictor$new(rf_based_fit, data = trainX, y =
train$Evaluation_treatment)

# Compute permutation feature importance
importance <- FeatureImp$new(explainer, loss = "ce") # "ce" stands for classification
error

# Plot the variable importance
plot(importance)

#Print VIP
importance$results
importance_df = as.data.frame(importance$results)

# Create a VIP plot using ggplot2
ggplot(importance_df, aes(x = reorder(feature, importance), y = importance)) +
  geom_bar(stat = "identity") +
  coord_flip() +
  xlab("Feature") +
  ylab("Importance") +
  ggtitle("Feature Importance")

# Create with feature selection 6 variables #
rf_recipe_based1 = recipe(Evaluation_treatment ~
drugs.present+ever_received_treatment+drug_first_time
+live.with.whom.in.30.days.ago+family.relation+cause_receive_treatment, data =
train)%>%
  step_dummy(all_nominal_predictors(), one_hot = TRUE)

rf_model_based1 <- rand_forest() %>%
  set_args(mtry = 4, trees = 500) %>%
  set_engine("ranger", importance = "impurity") %>%
  set_mode("classification")

workflow_based1 <- workflow() %>%
  add_recipe(rf_recipe_based1) %>%
  add_model(rf_model_based1)

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

# Training set #
# Make predictions #
rf_based_fit1 <- workflow_based1 %>%
  fit(data = train)

train_prediction1 <- predict(rf_based_fit1,
                             new_data = train)

# Create confusion matrix #
train_predicted1 <- train_prediction1$.pred_class
train_actual <- train$Evaluation_treatment

confusion_matrix1 <- table(train_actual,train_predicted1)
confusion_matrix1

# Calculate accuracy, precision, recall, f1_score #
accuracy_train = confusionMatrix(confusion_matrix1)$overall['Accuracy']

precision_train = posPredValue(confusion_matrix1,
                                positive = "1")
recall_train = sensitivity(confusion_matrix1,
                           positive = "1")
f1_score_train = 2*(precision_train*recall_train)/(precision_train + recall_train)

metrics_train1 = c(accuracy_train, precision_train, recall_train, f1_score_train)
metrics_train1

# Test set #
# Make predictions
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)
test_rf_based_fit1 <- workflow_based1 %>%
  fit(data = test)

test_predictions1 <- predict(test_rf_based_fit1,
                              new_data = test)

# Create confusion matrix
test_predicted1 = test_predictions1$.pred_class
test_actual = test$Evaluation_treatment

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

confusion_matrix_test1 = table(test_actual, test_predicted1)
confusion_matrix_test1

# Calculate accuracy, precision, recall, and F1 score
accuracy_test = confusionMatrix(confusion_matrix_test1)$overall["Accuracy"]

precision_test = posPredValue(confusion_matrix_test1,
                              positive = "1")
recall_test = sensitivity(confusion_matrix_test1,
                          positive = "1")
f1_score_test = 2 * (precision_test * recall_test) / (precision_test + recall_test)

# Combine metrics into a single vector
metrics_test1 = c(accuracy_test, precision_test, recall_test, f1_score_test)
metrics_test1

rbind(metrics_train1, metrics_test1) %>%
  `colnames`->c("Accuracy", "Precision", "Recall", "F1-score")

#Tuning#
# set independent variables and create dummy for categorical variable
# Defined hyperparameter to be tuned and others
hyper_rf_model_based1 <- rand_forest() %>%
  set_args(mtry = tune(), trees = tune()) %>%
  set_engine("ranger", importance = "impurity") %>%
  set_mode("classification")

# Combined rf_recipe and rf_model
hyper_rf_workflow_based1 <- workflow() %>%
  add_recipe(rf_recipe_based1) %>%
  add_model(hyper_rf_model_based1)

# Set searching grid
rf_grid_based<- expand.grid(
  mtry= c(1:(ncol(train)-1)),
  trees = c(250,500,750,1000,1250,1500))

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

# Set Parallel computation to save time in tuning
all_cores = parallel::detectCores(logical = FALSE)
cl = makePSOCKcluster(all_cores)
registerDoParallel(cl)

# Set sample for each fold
set.seed(999)
train_cv = vfold_cv(train,v = 10,strata = Evaluation_treatment)

# Run hyperparameter tuning
set.seed(999)
hyper_rf_results1 = hyper_rf_workflow_based1 %>%
  tune_grid(resamples = train_cv,
            grid = rf_grid_based,
            metrics = metric_set(accuracy,roc_auc))

# Print best set of hyperparameter based on accuracy
best_params_acc_based1 = hyper_rf_results1 %>%
  select_best(metric = "accuracy")
best_params_acc_based1
# 4 1500 Preprocessor1_Model79

# Create the random forest model from optimal hyperparameter based on accuracy
opt_rf_model_based1 <- rand_forest() %>%
  set_args(mtry = 4, trees = 1500) %>%
  set_engine("ranger", importance = "impurity") %>%
  set_mode("classification")

opt_workflow_based1 <- workflow() %>%
  add_recipe(rf_recipe_based1) %>%
  add_model(opt_rf_model_based1)

# Training set #
# Make predictions #
opt_rf_based_fit1 <- opt_workflow_based1 %>%
  fit(data = train)

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

opt_train_predictions1 = predict(opt_rf_based_fit1,
                                new_data = train)

# Create confusion matrix
opt_train_predicted1 = opt_train_predictions1$.pred_class
train_actual = train$Evaluation_treatment

opt_confusion_matrix1 = table(train_actual, opt_train_predicted1)
opt_confusion_matrix1

# Calculate accuracy, precision, recall, f1_score
accuracy_train = confusionMatrix(opt_confusion_matrix1)$overall['Accuracy']

precision_train = posPredValue(opt_confusion_matrix1,
                               positive = "1")
recall_train = sensitivity(opt_confusion_matrix1,
                           positive = "1")
f1_score_train = 2*(precision_train*recall_train)/(precision_train + recall_train)

opt_metrics_train1 = c(accuracy_train, precision_train, recall_train, f1_score_train)
opt_metrics_train1

# Test set #
# Make predictions
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)
opt_test_rf_based_fit1 <- opt_workflow_based1 %>%
  fit(data = test)

opt_test_predictions1 = predict(opt_test_rf_based_fit1,
                                new_data = test)

# Create confusion matrix
opt_test_predicted1 = opt_test_predictions1$.pred_class
test_actual = test$Evaluation_treatment

opt_confusion_matrix_test1 = table(test_actual, opt_test_predicted1)
opt_confusion_matrix_test1

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

# Calculate accuracy, precision, recall, and F1 score
accuracy_test = confusionMatrix(opt_confusion_matrix_test1)$overall['Accuracy']

precision_test = posPredValue(opt_confusion_matrix_test1,
                              positive = "1")

recall_test = sensitivity(opt_confusion_matrix_test1,
                          positive = "1")

f1_score_test = 2 * (precision_test * recall_test) / (precision_test + recall_test)

# Combine metrics into a single vector
opt_metrics_test1 = c(accuracy_test, precision_test, recall_test, f1_score_test)
opt_metrics_test1

rbind(opt_metrics_train1, opt_metrics_test1) %>%
  `colnames<-`(c("Accuracy", "Precision", "Recall", "F1-score"))

# Create model with feature selection 6 variables same as logistic regression #
rf_recipe_based2 = recipe(Evaluation_treatment ~ family.relation+drug_first_time
                          +drugs.present.+time_use_year+ever_received_treatment+
                          cause_receive_treatment, data = train)%>%
  step_dummy(all_nominal_predictors(), one_hot = TRUE)

rf_model_based2 <- rand_forest() %>%
  set_args(mtry = 4, trees = 500) %>%
  set_engine("ranger", importance = "impurity") %>%
  set_mode("classification")

workflow_based2 <- workflow() %>%
  add_recipe(rf_recipe_based2) %>%
  add_model(rf_model_based2)

# Training set #
# Make predictions #
rf_based_fit2 <- workflow_based2 %>%
  fit(data = train)

choose_train_prediction2 <- predict(rf_based_fit2,
                                   new_data = train)

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

# Create confusion matrix #
choose_train_predicted2 <- choose_train_prediction2$.pred_class
train_actual <- train$Evaluation_treatment

choose_confusion_matrix2 <- table(train_actual,choose_train_predicted2)
choose_confusion_matrix2

# Calculate accuracy, precision, recall, f1_score #
accuracy_train = confusionMatrix(choose_confusion_matrix2)$overall['Accuracy']

precision_train = posPredValue(choose_confusion_matrix2,
                               positive = "1")
recall_train = sensitivity(choose_confusion_matrix2,
                           positive = "1")
f1_score_train = 2*(precision_train*recall_train)/(precision_train + recall_train)
metrics_train2 = c(accuracy_train, precision_train, recall_train, f1_score_train)
metrics_train2

# Test set #
# Make predictions
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)
test_rf_based_fit2 <- workflow_based2 %>%
  fit(data = test)

choose_test_predictions2 <- predict(test_rf_based_fit2,
                                   new_data = test)

# Create confusion matrix
choose_test_predicted2 = choose_test_predictions2$.pred_class
test_actual = test$Evaluation_treatment

choose_confusion_matrix_test2 = table(test_actual, choose_test_predicted2)
choose_confusion_matrix_test2

# Calculate accuracy, precision, recall, and F1 score
accuracy_test = confusionMatrix(choose_confusion_matrix_test2)$overall['Accuracy']

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

precision_test = posPredValue(choose_confusion_matrix_test2,
                             positive = "1")
recall_test = sensitivity(choose_confusion_matrix_test2,
                          positive = "1")
f1_score_test = 2 * (precision_test * recall_test) / (precision_test + recall_test)

# Combine metrics into a single vector
metrics_test2 = c(accuracy_test, precision_test, recall_test, f1_score_test)
metrics_test2

rbind(metrics_train2, metrics_test2) %>%
  `colnames<-`(c("Accuracy", "Precision", "Recall", "F1-score"))

#Tuning#
# set independent variables and create dummy for categorical variable
# Defined hyperparameter to be tuned and others
hyper_rf_model_based2 <- rand_forest() %>%
  set_args(mtry = tune(), trees = tune()) %>%
  set_engine("ranger", importance = "impurity") %>%
  set_mode("classification")

# Combined rf_recipe and rf_model
hyper_rf_workflow_based2 <- workflow() %>%
  add_recipe(rf_recipe_based2) %>%
  add_model(hyper_rf_model_based2)

# Set searching grid
rf_grid_based<- expand_grid(
  mtry= c(1:(ncol(train)-1)),
  trees = c(250,500,750,1000,1250,1500))

# Set Parallel computation to save time in tuning
all_cores = parallel::detectCores(logical = FALSE)
cl = makePSOCKcluster(all_cores)
registerDoParallel(cl)

# Set sample for each fold
set.seed(999)

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

train_cv = vfold_cv(train,v = 10,strata = Evaluation_treatment)

# Run hyperparameter tuning
set.seed(999)
hyper_rf_results2 = hyper_rf_workflow_based2 %>%
  tune_grid(resamples = train_cv,
            grid = rf_grid_based,
            metrics = metric_set(accuracy,roc_auc))

# Print best set of hyperparameter based on accuracy
best_params_acc_based2 = hyper_rf_results2 %>%
  select_best(metric = "accuracy")
best_params_acc_based2
# 2 250 Preprocessor1_Model02

# Create the random forest model from optimal hyperparameter based on accuracy
opt_rf_model_based2 <- rand_forest() %>%
  set_args(mtry = 2, trees = 250) %>%
  set_engine("ranger", importance = "impurity") %>%
  set_mode("classification")

opt_workflow_based2 <- workflow() %>%
  add_recipe(rf_recipe_based2) %>%
  add_model(opt_rf_model_based2)

# Training set #
# Make predictions #
opt_rf_based_fit2 <- opt_workflow_based2 %>%
  fit(data = train)

opt_train_predictions2 = predict(opt_rf_based_fit2,
                                new_data = train)

# Create confusion matrix
opt_train_predicted2 = opt_train_predictions2$.pred_class
train_actual = train$Evaluation_treatment

opt_confusion_matrix2 = table(train_actual, opt_train_predicted2)

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

opt_confusion_matrix2

# Calculate accuracy, precision, recall, f1_score
accuracy_train = confusionMatrix(opt_confusion_matrix2)$overall['Accuracy']

precision_train = posPredValue(opt_confusion_matrix2,
                               positive = "1")
recall_train = sensitivity(opt_confusion_matrix2,
                           positive = "1")
f1_score_train = 2*(precision_train*recall_train)/(precision_train + recall_train)

opt_metrics_train2 = c(accuracy_train, precision_train, recall_train, f1_score_train)
opt_metrics_train2

# Test set #
# Make predictions
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)
opt_test_rf_based_fit2 <- opt_workflow_based2 %>%
  fit(data = test)

opt_test_predictions2 = predict(opt_test_rf_based_fit2,
                                new_data = test)

# Create confusion matrix
opt_test_predicted2 = opt_test_predictions2$.pred_class
test_actual = test$Evaluation_treatment

opt_confusion_matrix_test2 = table(test_actual, opt_test_predicted2)
opt_confusion_matrix_test2

# Calculate accuracy, precision, recall, and F1 score
accuracy_test = confusionMatrix(opt_confusion_matrix_test2)$overall['Accuracy']

precision_test = posPredValue(opt_confusion_matrix_test2,
                              positive = "1")
recall_test = sensitivity(opt_confusion_matrix_test2,
                          positive = "1")

```

```

f1_score_test = 2 * (precision_test * recall_test) / (precision_test + recall_test)

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

# Combine metrics into a single vector
opt_metrics_test2 = c(accuracy_test, precision_test, recall_test, f1_score_test)
opt_metrics_test2

rbind(opt_metrics_train2, opt_metrics_test2) %>%
  `colnames<-`(c("Accuracy", "Precision", "Recall", "F1-score"))

# Create model with feature selection 6 variables same as logistic regression
stat::step#
rf_recipe_based3 = recipe(Evaluation_treatment ~ family.relation +
  age_first_time_used_drug +
    drug_first_time + time_use_year + ever_received_treatment +
    cause_receive_treatment, data = train)%>%
  step_dummy(all_nominal_predictors(), one_hot = TRUE)

rf_model_based3 <- rand_forest() %>%
  set_args(mtry = 4, trees = 500) %>%
  set_engine("ranger", importance = "impurity") %>%
  set_mode("classification")

workflow_based3 <- workflow() %>%
  add_recipe(rf_recipe_based3) %>%
  add_model(rf_model_based3)

# Training set #
# Make predictions #
rf_based_fit3 <- workflow_based3 %>%
  fit(data = train)

choose_train_prediction3 <- predict(rf_based_fit3,
  new_data = train)

# Create confusion matrix #
choose_train_predicted3 <- choose_train_prediction3$.pred_class
train_actual <- train$Evaluation_treatment

choose_confusion_matrix3 <- table(train_actual, choose_train_predicted3)
choose_confusion_matrix3

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

# Calculate accuracy, precision, recall, f1_score #
accuracy_train = confusionMatrix(choose_confusion_matrix3)$overall['Accuracy']

precision_train = posPredValue(choose_confusion_matrix3,
                               positive = "1")
recall_train = sensitivity(choose_confusion_matrix3,
                           positive = "1")
f1_score_train = 2*(precision_train*recall_train)/(precision_train + recall_train)

metrics_train3 = c(accuracy_train, precision_train, recall_train, f1_score_train)
metrics_train3

# Test set #
# Make predictions
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)
test_rf_based_fit3 <- workflow_based3 %>%
  fit(data = test)

choose_test_predictions3 <- predict(test_rf_based_fit3,
                                   new_data = test)

# Create confusion matrix
choose_test_predicted3 = choose_test_predictions3$.pred_class
test_actual = test$Evaluation_treatment

choose_confusion_matrix_test3 = table(test_actual, choose_test_predicted3)
choose_confusion_matrix_test3

# Calculate accuracy, precision, recall, and F1 score
accuracy_test = confusionMatrix(choose_confusion_matrix_test3)$overall['Accuracy']

precision_test = posPredValue(choose_confusion_matrix_test3,
                              positive = "1")
recall_test = sensitivity(choose_confusion_matrix_test3,
                          positive = "1")
f1_score_test = 2 * (precision_test * recall_test) / (precision_test + recall_test)

```

```

# Combine metrics into a single vector

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
metrics_test3 = c(accuracy_test, precision_test, recall_test, f1_score_test)
metrics_test3
```

```
rbind(metrics_train3, metrics_test3) %>%
  `colnames<-`(c("Accuracy", "Precision", "Recall", "F1-score"))
```

```
#Tuning#
```

```
# set independent variables and create dummy for categorical variable
```

```
# Defined hyperparameter to be tuned and others
```

```
hyper_rf_model_based3 <- rand_forest() %>%
  set_args(mtry = tune(), trees = tune()) %>%
  set_engine("ranger", importance = "impurity") %>%
  set_mode("classification")
```

```
# Combined rf_recipe and rf_model
```

```
hyper_rf_workflow_based3 <- workflow() %>%
  add_recipe(rf_recipe_based3) %>%
  add_model(hyper_rf_model_based3)
```

```
# Set searching grid
```

```
rf_grid_based<- expand_grid(
  mtry= c(1:(ncol(train)-1)),
  trees = c(250,500,750,1000,1250,1500))
```

```
# Set Parallel computation to save time in tuning
```

```
all_cores = parallel::detectCores(logical = FALSE)
cl = makePSOCKcluster(all_cores)
registerDoParallel(cl)
```

```
# Set sample for each fold
```

```
set.seed(999)
```

```
train_cv = vfold_cv(train,v = 10,strata = Evaluation_treatment)
```

```
# Run hyperparameter tuning
```

```
set.seed(999)
```

```
hyper_rf_results3 = hyper_rf_workflow_based3 %>%
```

```
  tune_grid(resamples = train_cv,
```

```
    grid = rf_grid_based,
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

metrics = metric_set(accuracy,roc_auc)

# Print best set of hyperparameter based on accuracy
best_params_acc_based3 = hyper_rf_results3 %>%
  select_best(metric = "accuracy")
best_params_acc_based3
# 3 1250 Preprocessor1_Model63

# Create the random forest model from optimal hyperparameter based on accuracy
opt_rf_model_based3 <- rand_forest() %>%
  set_args(mtry = 3, trees = 1250) %>%
  set_engine("ranger", importance = "impurity") %>%
  set_mode("classification")

opt_workflow_based3 <- workflow() %>%
  add_recipe(rf_recipe_based3) %>%
  add_model(opt_rf_model_based3)

# Training set #
# Make predictions #
opt_rf_based_fit3 <- opt_workflow_based3 %>%
  fit(data = train)

opt_train_predictions3 = predict(opt_rf_based_fit3,
  new_data = train)

# Create confusion matrix
opt_train_predicted3 = opt_train_predictions3$pred_class
train_actual = train$Evaluation_treatment

opt_confusion_matrix3 = table(train_actual, opt_train_predicted3)
opt_confusion_matrix3

# Calculate accuracy, precision, recall,f1_score
accuracy_train = confusionMatrix(opt_confusion_matrix3)$overall[1,'Accuracy']

precision_train = posPredValue(opt_confusion_matrix3,
  positive = "1")

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

recall_train = sensitivity(opt_confusion_matrix3,
                           positive = "1")
f1_score_train = 2*(precision_train*recall_train)/(precision_train + recall_train)

opt_metrics_train3 = c(accuracy_train, precision_train, recall_train, f1_score_train)
opt_metrics_train3

# Test set #
# Make predictions
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)
opt_test_rf_based_fit3 <- opt_workflow_based3 %>%
  fit(data = test)

opt_test_predictions3 = predict(opt_test_rf_based_fit3,
                                new_data = test)

# Create confusion matrix
opt_test_predicted3 = opt_test_predictions3$.pred_class
test_actual = test$Evaluation_treatment

opt_confusion_matrix_test3 = table(test_actual, opt_test_predicted3)
opt_confusion_matrix_test3

# Calculate accuracy, precision, recall, and F1 score
accuracy_test = confusionMatrix(opt_confusion_matrix_test3)$overall['Accuracy']

precision_test = posPredValue(opt_confusion_matrix_test3,
                              positive = "1")
recall_test = sensitivity(opt_confusion_matrix_test3,
                          positive = "1")
f1_score_test = 2 * (precision_test * recall_test) / (precision_test + recall_test)

# Combine metrics into a single vector
opt_metrics_test3 = c(accuracy_test, precision_test, recall_test, f1_score_test)
opt_metrics_test3

rbind(opt_metrics_train3, opt_metrics_test3) %>%
  `colnames`<-`c("Accuracy", "Precision", "Recall", "F1-score")`

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

SVM
library(tidymodels)
library(vip)
library(tidyverse)
library(dplyr)
library(doParallel)
library(caret)
library(iml)
getwd()
setwd("c:/")
df= read.csv("drug_treatment(new)1.csv")

head(df)
str(df)

df = df %>%
  mutate(Evaluation_treatment =
    recode(Evaluation_treatment,
           "uncomplete" = 1,
           "complete" = 0)
  )
table(df$Evaluation_treatment)

#####Split Data#####
set.seed(999)
split <- initial_split(df, prop = 0.7, strata = Evaluation_treatment)
train <- split %>%
  training()
test <- split %>%
  testing()
train$Evaluation_treatment <- as.factor(train$Evaluation_treatment)

# Create all model 12 variables #
svm_recipe = recipe(Evaluation_treatment ~
  age+education+occupation+live.with.whom.in.30.days.ago +family.relation
  +age_first_time_used_drug+drug_first_time
  +cause_use_drugs.first_time.+drugs.present.+time_use_year+ever_received_treatmen
  t+

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

      cause_receive_treatment, data = train)%>%
step_dummy(all_nominal_predictors(), one_hot = TRUE)

#Create the svm model
svm_model <- svm_rbf() %>%
  set_args(cost = 1 , rbf_sigma = 0.1) %>%
  set_mode("classification") %>%
  set_engine("kernlab")

# Combined svm_recipe and svm_model
svm_workflow <- workflow() %>%
  add_recipe(svm_recipe) %>%
  add_model(svm_model)

# Training set #
# Make predictions #
svm_based_fit <- svm_workflow %>%
  fit(data = train)

# Extract VIP
trainX = train[,-c(16)]

explainer <- Predictor$new(svm_based_fit, data = trainX, y =
train$Evaluation_treatment)

# Compute permutation feature importance
importance <- FeatureImp$new(explainer, loss = "ce") # "ce" stands for classification
error

# Plot the variable importance
plot(importance)

#Print VIP
importance$results

importance_df = as.data.frame(importance$results)

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
# Create a VIP plot using ggplot2
ggplot(importance_df, aes(x = reorder(feature, importance), y = importance)) +
  geom_bar(stat = "identity") +
  coord_flip() +
  xlab("Feature") +
  ylab("Importance") +
  ggtitle("Feature Importance")
```

```
# Training set #
```

```
# Make predictions #
```

```
train$Evaluation_treatment <- as.factor(train$Evaluation_treatment)
```

```
svm_based_fit <- svm_workflow %>%
```

```
  fit(data = train)
```

```
train_predictions = predict(svm_based_fit,
  new_data = train)
```

```
# Create confusion matrix
```

```
train_predicted = train_predictions$.pred_class
```

```
train_actual = train$Evaluation_treatment
```

```
confusion_matrix = table(train_actual, train_predicted)
```

```
confusion_matrix
```

```
# Calculate accuracy, precision, recall, f1_score
```

```
accuracy_train = confusionMatrix(confusion_matrix)$overall["Accuracy"]
```

```
precision_train = posPredValue(confusion_matrix,
  positive = "1")
```

```
recall_train = sensitivity(confusion_matrix,
  positive = "1")
```

```
f1_score_train = 2*(precision_train*recall_train)/(precision_train + recall_train)
```

```
metrics_train = c(accuracy_train, precision_train, recall_train, f1_score_train)
```

```
metrics_train
```

```
# Test set #
```

```
# Make predictions
```

```
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)
```

```
svm_based_fit_test <- svm_workflow %>%
```

```
  fit(data = test)
```

```
test_predictions = predict(svm_based_fit_test,
  new_data = test)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

# Create confusion matrix
test_predicted = test_predictions$.pred_class
test_actual = test$Evaluation_treatment
confusion_matrix_test = table(test_actual, test_predicted)
confusion_matrix_test

# Calculate accuracy, precision, recall, and F1 score
accuracy_test = confusionMatrix(confusion_matrix_test)$overall['Accuracy']
precision_test = posPredValue(confusion_matrix_test,
                              positive = "1")
recall_test = sensitivity(confusion_matrix_test,
                          positive = "1")
f1_score_test = 2 * (precision_test * recall_test) / (precision_test + recall_test)

# Combine metrics into a single vector
metrics_test = c(accuracy_test, precision_test, recall_test, f1_score_test)
metrics_test
rbind(metrics_train, metrics_test) %>%
  `colnames<-`(c("Accuracy", "Precision", "Recall", "F1-score"))

# Create model with feature selection, 6 variables ##
svm_recipe1 = recipe(Evaluation_treatment ~
  live.with.whom.in.30.days.ago+drugs.present.+occupation+
  drug_first_time+age+education , data = train)%>%
  step_dummy(all_nominal_predictors(), one_hot = TRUE)

#Create the svm model
svm_model1 <- svm_rbf() %>%
  set_args(cost = 1 , rbf_sigma = 0.1) %>%
  set_mode("classification") %>%
  set_engine("kernlab")

# Combined svm_recipe and svm_model
svm_workflow1 <- workflow() %>%
  add_recipe(svm_recipe1) %>%
  add_model(svm_model1)

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

# Training set #
# Make predictions #
train$Evaluation_treatment <- as.factor(train$Evaluation_treatment)

svm_based_fit1 <- svm_workflow1 %>%
  fit(data = train)
train_predictions1 = predict(svm_based_fit1,
                             new_data = train)
# Create confusion matrix
train_predicted1 = train_predictions1$.pred_class
train_actual = train$Evaluation_treatment
confusion_matrix1 = table(train_actual, train_predicted1)
confusion_matrix1

# Calculate accuracy, precision, recall, f1_score
accuracy_train = confusionMatrix(confusion_matrix1)$overall['Accuracy']
precision_train = posPredValue(confusion_matrix1,
                               positive = "1")
recall_train = sensitivity(confusion_matrix1,
                           positive = "1")
f1_score_train = 2*(precision_train*recall_train)/(precision_train + recall_train)
metrics_train1 = c(accuracy_train, precision_train, recall_train, f1_score_train)
metrics_train1

# Test set #
# Make predictions #
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)
test_svm_based_fit1 <- svm_workflow1 %>%
  fit(data = test)
test_predictions1 = predict(test_svm_based_fit1,
                             new_data = test)
# Create confusion matrix
test_predicted1 = test_predictions1$.pred_class
test_actual = test$Evaluation_treatment
confusion_matrix_test1 = table(test_actual, test_predicted1)
confusion_matrix_test1

# Calculate accuracy, precision, recall, and F1 score
accuracy_test = confusionMatrix(confusion_matrix_test1)$overall['Accuracy']

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

precision_test = posPredValue(confusion_matrix_test1,
                              positive = "1")
recall_test = sensitivity(confusion_matrix_test1,
                          positive = "1")
f1_score_test = 2 * (precision_test * recall_test) / (precision_test + recall_test)
# Combine metrics into a single vector
metrics_test1 = c(accuracy_test, precision_test, recall_test, f1_score_test)
metrics_test1
rbind(metrics_train1, metrics_test1) %>%
  `colnames<-`(c("Accuracy", "Precision", "Recall", "F1-score"))

#Tuning#
# set independent variables and create dummy for categorical variable
# Defined hyperparameter to be tuned and others
hyper_svm_model1 <- svm_rbf(cost = tune(), rbf_sigma = tune()) %>%
  set_mode("classification") %>%
  set_engine("kernlab")

# Combined svm_recipe and svm_model
hyper_svm_workflow1 <- workflow() %>%
  add_recipe(svm_recipe1) %>%
  add_model(hyper_svm_model1)

# Set searching grid
svm_grid <- expand_grid(
  cost = c(0.01,0.1,1,10,100,1000),
  rbf_sigma = c(0.01,0.1,1,10,100))

# Set Parallel computation to save time in tuning
all_cores = parallel::detectCores(logical = FALSE)
cl = makePSOCKcluster(all_cores)
registerDoParallel(cl)

# Set sample for each fold
set.seed(999)
train_cv = vfold_cv(train,v = 10,strata = Evaluation_treatment)

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

# classification metric 'accuracy'
set.seed(999)
roc_res <- metric_set(roc_auc)

# Run hyperparameter tuning
set.seed(999)
hyper_svm_results1 = hyper_svm_workflow1 %>%
  tune_grid(resamples = train_cv,
            grid = svm_grid,
            metrics = metric_set(accuracy,roc_auc))

# Print best set of hyperparameter based on accuracy
best_params_acc_based = hyper_svm_results1 %>%
  select_best(metric = "accuracy")
best_params_acc_based
# 1 0.01 Preprocessor1_Model03

# Create the svm model from optimal hyperparameter based on accuracy
opt_svm_model1 <- svm_rbf() %>%
  set_args(cost = 1, rbf_sigma = 0.01 ) %>%
  set_engine("kernlab") %>%
  set_mode("classification")

opt_workflow1 <- workflow() %>%
  add_recipe(svm_recipe1) %>%
  add_model(opt_svm_model1)

# Training set #
# Make predictions #
svm_optimal_fit1 <- opt_workflow1 %>%
  fit(data = train)

opt_train_predictions1 = predict(svm_optimal_fit1,
                                new_data = train)

# Create confusion matrix
opt_train_predicted1 = opt_train_predictions1$.pred_class
train_actual = train$Evaluation_treatment

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

opt_confusion_matrix1 = table(train_actual, opt_train_predicted1)
opt_confusion_matrix1

# Calculate accuracy, precision, recall, f1_score
accuracy_train = confusionMatrix(opt_confusion_matrix1)$overall['Accuracy']

precision_train = posPredValue(opt_confusion_matrix1,
                               positive = "1")
recall_train = sensitivity(opt_confusion_matrix1,
                           positive = "1")
f1_score_train = 2*(precision_train*recall_train)/(precision_train + recall_train)

opt_metrics_train1 = c(accuracy_train, precision_train, recall_train, f1_score_train)
opt_metrics_train1

# Test set #
# Make predictions
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)
test_svm_optimal_fit1 <- opt_workflow1 %>%
  fit(data = test)

opt_test_predictions1 = predict(test_svm_optimal_fit1,
                                new_data = test)

# Create confusion matrix
opt_test_predicted1 = opt_test_predictions1$.pred_class
test_actual = test$Evaluation_treatment

opt_confusion_matrix_test1 = table(test_actual, opt_test_predicted1)
opt_confusion_matrix_test1

# Calculate accuracy, precision, recall, and F1 score
accuracy_test = confusionMatrix(opt_confusion_matrix_test1)$overall['Accuracy']

precision_test = posPredValue(opt_confusion_matrix_test1,
                              positive = "1")
recall_test = sensitivity(opt_confusion_matrix_test1,
                          positive = "1")

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

f1_score_test = 2 * (precision_test * recall_test) / (precision_test + recall_test)

# Combine metrics into a single vector
opt_metrics_test1 = c(accuracy_test, precision_test, recall_test, f1_score_test)
opt_metrics_test1

rbind(opt_metrics_train1, opt_metrics_test1) %>%
  `colnames<-`(c("Accuracy", "Precision", "Recall", "F1-score"))

# Create model with feature selection 6 variables same as logistic regression #
svm_recipe2 = recipe(Evaluation_treatment ~family.relation+drug_first_time
  +drugs.present.+time_use_year+ever_received_treatment+
  cause_receive_treatment, data = train)%>%
  step_dummy(all_nominal_predictors(), one_hot = TRUE)

#Create the svm model
svm_model2 <- svm_rbf() %>%
  set_args(cost = 1 , rbf_sigma = 0.1) %>%
  set_mode("classification") %>%
  set_engine("kernlab")

# Combined svm_recipe and svm_model
svm_workflow2 <- workflow() %>%
  add_recipe(svm_recipe2) %>%
  add_model(svm_model2)

# Training set #
# Make predictions #
train$Evaluation_treatment <- as.factor(train$Evaluation_treatment)
svm_based_fit2 <- svm_workflow2 %>%
  fit(data = train)
train_predictions2 = predict(svm_based_fit2,
  new_data = train)

# Create confusion matrix
train_predicted2 = train_predictions2$.pred_class
train_actual = train$Evaluation_treatment
confusion_matrix2 = table(train_actual, train_predicted2)
confusion_matrix2

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

# Calculate accuracy, precision, recall, f1_score
accuracy_train = confusionMatrix(confusion_matrix2)$overall['Accuracy']
precision_train = posPredValue(confusion_matrix2,
                               positive = "1")
recall_train = sensitivity(confusion_matrix2,
                           positive = "1")
f1_score_train = 2*(precision_train*recall_train)/(precision_train + recall_train)
metrics_train2 = c(accuracy_train, precision_train, recall_train, f1_score_train)
metrics_train2

# Test set #
# Make predictions #
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)
test_svm_based_fit2 <- svm_workflow2 %>%
  fit(data = test)
test_predictions2 = predict(test_svm_based_fit2,
                             new_data = test)
# Create confusion matrix
test_predicted2 = test_predictions2$.pred_class
test_actual = test$Evaluation_treatment
confusion_matrix_test2 = table(test_actual, test_predicted2)
confusion_matrix_test2
# Calculate accuracy, precision, recall, and F1 score
accuracy_test = confusionMatrix(confusion_matrix_test2)$overall['Accuracy']
precision_test = posPredValue(confusion_matrix_test2,
                              positive = "1")
recall_test = sensitivity(confusion_matrix_test2,
                          positive = "1")
f1_score_test = 2 * (precision_test * recall_test) / (precision_test + recall_test)
# Combine metrics into a single vector
metrics_test2 = c(accuracy_test, precision_test, recall_test, f1_score_test)
metrics_test2
rbind(metrics_train2, metrics_test2) %>%
  `colnames<-`(c("Accuracy", "Precision", "Recall", "F1-score"))

```

```
#Tuning#
```

```
# set independent variables and create dummy for categorical variable
```

```
# Defined hyperparameter to be tuned and others
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

hyper_svm_model2 <- svm_rbf(cost = tune(), rbf_sigma = tune()) %>%
  set_mode("classification") %>%
  set_engine("kernlab")

# Combined svm_recipe and svm_model
hyper_svm_workflow2 <- workflow() %>%
  add_recipe(svm_recipe2) %>%
  add_model(hyper_svm_model2)

# Set searching grid
svm_grid <- expand_grid(
  cost = c(0.01,0.1,1,10,100,1000),
  rbf_sigma = c(0.01,0.1,1,10,100))

# Set Parallel computation to save time in tuning
all_cores = parallel::detectCores(logical = FALSE)
cl = makePSOCKcluster(all_cores)
registerDoParallel(cl)

# Set sample for each fold
set.seed(999)
train_cv = vfold_cv(train,v = 10,strata = Evaluation_treatment)

# classification metric 'accuracy'
set.seed(999)
roc_res <- metric_set(roc_auc)

# Run hyperparameter tuning
set.seed(999)
hyper_svm_results2 = hyper_svm_workflow2 %>%
  tune_grid(resamples = train_cv,
            grid = svm_grid,
            metrics = metric_set(accuracy,roc_auc))

# Print best set of hyperparameter based on accuracy
best_params_acc_based2 = hyper_svm_results2 %>%
  select_best(metric = "accuracy")
best_params_acc_based2

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

# 0.1 0.01 Preprocessor1_Model02

# Create the svm model from optimal hyperparameter based on accuracy
opt_svm_model2 <- svm_rbf() %>%
  set_args(cost = 0.1, rbf_sigma = 0.01 ) %>%
  set_engine("kernlab") %>%
  set_mode("classification")

opt_workflow2 <- workflow() %>%
  add_recipe(svm_recipe2) %>%
  add_model(opt_svm_model2)

# Training set #
# Make predictions #
svm_optimal_fit2 <- opt_workflow2 %>%
  fit(data = train)

opt_train_predictions2 = predict(svm_optimal_fit2,
  new_data = train)

# Create confusion matrix
opt_train_predicted2 = opt_train_predictions2$.pred_class
train_actual = train$Evaluation_treatment

opt_confusion_matrix2 = table(train_actual, opt_train_predicted2)
opt_confusion_matrix2

# Calculate accuracy, precision, recall, f1_score
accuracy_train = confusionMatrix(opt_confusion_matrix2)$overall[1, 'Accuracy']

precision_train = posPredValue(opt_confusion_matrix2,
  positive = "1")
recall_train = sensitivity(opt_confusion_matrix2,
  positive = "1")
f1_score_train = 2*(precision_train*recall_train)/(precision_train + recall_train)

opt_metrics_train2 = c(accuracy_train, precision_train, recall_train, f1_score_train)
opt_metrics_train2

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

# Test set #
# Make predictions
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)
test_svm_optimal_fit2 <- opt_workflow2 %>%
  fit(data = test)

opt_test_predictions2 = predict(test_svm_optimal_fit2,
  new_data = test)

# Create confusion matrix
opt_test_predicted2 = opt_test_predictions2$.pred_class
test_actual = test$Evaluation_treatment

opt_confusion_matrix_test2 = table(test_actual, opt_test_predicted2)
opt_confusion_matrix_test2

# Calculate accuracy, precision, recall, and F1 score
accuracy_test = confusionMatrix(opt_confusion_matrix_test2)$overall[["Accuracy"]]

precision_test = posPredValue(opt_confusion_matrix_test2,
  positive = "1")
recall_test = sensitivity(opt_confusion_matrix_test2,
  positive = "1")
f1_score_test = 2 * (precision_test * recall_test) / (precision_test + recall_test)

# Combine metrics into a single vector
opt_metrics_test2 = c(accuracy_test, precision_test, recall_test, f1_score_test)
opt_metrics_test2

rbind(opt_metrics_train2, opt_metrics_test2) %>%
  `colnames<-`(c("Accuracy", "Precision", "Recall", "F1-score"))

# Create model with feature selection 6 variables same as logistic regression
stat::step#
svm_recipe3 = recipe(Evaluation_treatment ~family.relation +
  age_first_time_used_drug +
    drug_first_time + time_use_year + ever_received_treatment +
    cause_receive_treatment, data = train)%>%

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

step_dummy(all_nominal_predictors(), one_hot = TRUE)

#Create the svm model
svm_model3 <- svm_rbf() %>%
  set_args(cost = 1 , rbf_sigma = 0.1) %>%
  set_mode("classification") %>%
  set_engine("kernlab")

# Combined svm_recipe and svm_model
svm_workflow3 <- workflow() %>%
  add_recipe(svm_recipe3) %>%
  add_model(svm_model3)

# Training set #
# Make predictions #
train$Evaluation_treatment <- as.factor(train$Evaluation_treatment)
svm_based_fit3 <- svm_workflow3 %>%
  fit(data = train)
train_predictions3 = predict(svm_based_fit3,
                             new_data = train)
# Create confusion matrix
train_predicted3 = train_predictions3$.pred_class
train_actual = train$Evaluation_treatment
confusion_matrix3 = table(train_actual, train_predicted3)
confusion_matrix3

# Calculate accuracy, precision, recall, f1_score
accuracy_train = confusionMatrix(confusion_matrix3)$overall['Accuracy']
precision_train = posPredValue(confusion_matrix3,
                               positive = "1")
recall_train = sensitivity(confusion_matrix3,
                           positive = "1")
f1_score_train = 2*(precision_train*recall_train)/(precision_train + recall_train)
metrics_train3 = c(accuracy_train, precision_train, recall_train, f1_score_train)
metrics_train3

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

# Test set #
# Make predictions #
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)
test_svm_based_fit3 <- svm_workflow3 %>%
  fit(data = test)
test_predictions3 = predict(test_svm_based_fit3,
                             new_data = test)

# Create confusion matrix
test_predicted3 = test_predictions3$.pred_class
test_actual = test$Evaluation_treatment
confusion_matrix_test3 = table(test_actual, test_predicted3)
confusion_matrix_test3
# Calculate accuracy, precision, recall, and F1 score
accuracy_test = confusionMatrix(confusion_matrix_test3)$overall['Accuracy']
precision_test = posPredValue(confusion_matrix_test3,
                               positive = "1")
recall_test = sensitivity(confusion_matrix_test3,
                           positive = "1")
f1_score_test = 2 * (precision_test * recall_test) / (precision_test + recall_test)
# Combine metrics into a single vector
metrics_test3 = c(accuracy_test, precision_test, recall_test, f1_score_test)
metrics_test3
rbind(metrics_train3, metrics_test3) %>%
  `colnames<-`(c("Accuracy", "Precision", "Recall", "F1-score"))

#Tuning#
# set independent variables and create dummy for categorical variable
# Defined hyperparameter to be tuned and others
hyper_svm_model3 <- svm_rbf(cost = tune(), rbf_sigma = tune()) %>%
  set_mode("classification") %>%
  set_engine("kernlab")

# Combined svm_recipe and svm_model
hyper_svm_workflow3 <- workflow() %>%
  add_recipe(svm_recipe3) %>%
  add_model(hyper_svm_model3)

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

# Set searching grid
svm_grid <- expand.grid(
  cost = c(0.01,0.1,1,10,100,1000),
  rbf_sigma = c(0.01,0.1,1,10,100))

# Set Parallel computation to save time in tuning
all_cores = parallel::detectCores(logical = FALSE)
cl = makePSOCKcluster(all_cores)
registerDoParallel(cl)

# Set sample for each fold
set.seed(999)
train_cv = vfold_cv(train,v = 10,strata = Evaluation_treatment)

# classification metric 'accuracy'
set.seed(999)
roc_res <- metric_set(roc_auc)

# Run hyperparameter tuning
set.seed(999)
hyper_svm_results3 = hyper_svm_workflow3 %>%
  tune_grid(resamples = train_cv,
            grid = svm_grid,
            metrics = metric_set(accuracy,roc_auc))

# Print best set of hyperparameter based on accuracy
best_params_acc_based3 = hyper_svm_results3 %>%
  select_best(metric = "accuracy")
best_params_acc_based3
# 1 0.01 Preprocessor1_Model03

# Create the svm model from optimal hyperparameter based on accuracy
opt_svm_model3 <- svm_rbf() %>%
  set_args(cost = 1, rbf_sigma = 0.01 ) %>%
  set_engine("kernlab") %>%
  set_mode("classification")

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

opt_workflow3 <- workflow() %>%
  add_recipe(svm_recipe3) %>%
  add_model(opt_svm_model3)

# Training set #
# Make predictions #
svm_optimal_fit3 <- opt_workflow3 %>%
  fit(data = train)

opt_train_predictions3 = predict(svm_optimal_fit3,
  new_data = train)

# Create confusion matrix
opt_train_predicted3 = opt_train_predictions3$.pred_class
train_actual = train$Evaluation_treatment

opt_confusion_matrix3 = table(train_actual, opt_train_predicted3)
opt_confusion_matrix3

# Calculate accuracy, precision, recall, f1_score
accuracy_train = confusionMatrix(opt_confusion_matrix3)$overall['Accuracy']

precision_train = posPredValue(opt_confusion_matrix3,
  positive = "1")
recall_train = sensitivity(opt_confusion_matrix3,
  positive = "1")
f1_score_train = 2*(precision_train*recall_train)/(precision_train + recall_train)

opt_metrics_train3 = c(accuracy_train, precision_train, recall_train, f1_score_train)
opt_metrics_train3

# Test set #
# Make predictions
test$Evaluation_treatment <- as.factor(test$Evaluation_treatment)
test_svm_optimal_fit3 <- opt_workflow3 %>%
  fit(data = test)

```

opt_test_predictions3 = predict(test_svm_optimal_fit3,
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

new_data = test)

# Create confusion matrix
opt_test_predicted3 = opt_test_predictions3$.pred_class
test_actual = test$Evaluation_treatment

opt_confusion_matrix_test3 = table(test_actual, opt_test_predicted3)
opt_confusion_matrix_test3

# Calculate accuracy, precision, recall, and F1 score
accuracy_test = confusionMatrix(opt_confusion_matrix_test3)$overall['Accuracy']

precision_test = posPredValue(opt_confusion_matrix_test3,
                              positive = "1")
recall_test = sensitivity(opt_confusion_matrix_test3,
                          positive = "1")
f1_score_test = 2 * (precision_test * recall_test) / (precision_test + recall_test)

# Combine metrics into a single vector
opt_metrics_test3 = c(accuracy_test, precision_test, recall_test, f1_score_test)
opt_metrics_test3

rbind(opt_metrics_train3, opt_metrics_test3) %>%
  `colnames<-`(c("Accuracy", "Precision", "Recall", "F1-score"))

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ค

สมการเต็มของตัวแบบการถดถอยโลจิสติกแบบที่ใช้ตัวแปรอิสระครบทุกตัว คือ

$$\begin{aligned} \log(\text{odds}) = & -0.831 + 0.088(\text{age01}) - 0.01(\text{age03}) - 0.21(\text{age04}) + 0.377(\text{age05}) \\ & - 0.197(\text{edu01}) - 0.119(\text{edu02}) - 0.119(\text{edu03}) - 0.326(\text{edu04}) + 0.377(\text{edu06}) \\ & + 0.487(\text{occupa02}) + 0.060(\text{occupa03}) + 0.321(\text{occupa04}) + 0.156(\text{occupa05}) \\ & + 0.178(\text{occupa06}) + 1.086(\text{occupa07}) - 0.193(\text{live02}) - 0.269(\text{live03}) - 0.184(\text{live04}) \\ & - 0.261(\text{live05}) + 0.392(\text{live06}) + 0.014(\text{live07}) - 0.322(\text{live08}) - 0.194(\text{live09}) \\ & - 0.375(\text{fam02}) - 0.137(\text{fam03}) - 0.415(\text{fam04}) + 0.166(\text{fam05}) + 0.217(\text{agefirst01}) \\ & - 0.124(\text{agefirst03}) + 0.638(\text{agefirst04}) - 13.500(\text{agefirst05}) + 0.294(\text{drugfirst01}) \\ & + 0.534(\text{drugfirst02}) + 0.165(\text{drugfirst03}) + 0.099(\text{drugfirst04}) - 0.733(\text{drugfirst05}) \\ & + 0.524(\text{drugfirst06}) - 0.172(\text{causeuse02}) + 0.110(\text{causeuse03}) - 0.397(\text{causeuse04}) \\ & + 0.042(\text{causeuse05}) + 0.044(\text{causeuse06}) - 0.297(\text{drugpresent01}) + 1.77(\text{drugpresent02}) \\ & + 0.484(\text{drugpresent03}) + 2.051(\text{drugpresent04}) + 0.920(\text{drugpresent06}) \\ & - 0.443(\text{timeuse01}) + 0.104(\text{timeuse03}) + 0.292(\text{timeuse04}) + 0.183(\text{timeuse05}) \\ & + 0.489(\text{evertreatment01}) + 1.1289(\text{causetreatment01}) + 1.645(\text{causetreatment02}) \\ & - 2.081(\text{causetreatment03}) + 0.903(\text{causetreatment04}) - 1.266(\text{causetreatment05}) \\ & + 0.443(\text{causetreatment07}) \end{aligned}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ	นางสาวสุกัลยา เพชรชื่น
วัน เดือน ปีเกิด	16 พฤษภาคม 2541
ที่อยู่ปัจจุบัน	6 ซอยนนทบุรี 17 ต.บางกระสอ อ.เมืองนนทบุรี จ.นนทบุรี 11000
ประวัติการศึกษา	(2563) วิทยาศาสตรบัณฑิต สาขาคณิตศาสตร์ เกรตเฉลี่ย 2.87 (จุฬาลงกรณ์มหาวิทยาลัย) (2568) วิทยาศาสตรมหาบัณฑิต สาขาสถิติประยุกต์ เกรตเฉลี่ย 3.94 (สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง)
ผลงานทางวิชาการ	การศึกษาปัจจัยที่ส่งผลต่อกลุ่มการยุติการบำบัดรักษาผู้ติดยาเสพติด แบบระบบสมัครใจในพื้นที่ กรุงเทพมหานครโดยประยุกต์ใช้การเรียนรู้ ของเครื่อง, สุกัลยา เพชรชื่น ยุวดี กล่อมวิเศษ และอัศวิน วงศ์วิวัฒน์ ในการประชุมวิชาการการวิจัยดำเนินงานแห่งชาติ ประจำปี พ.ศ. 2568 ครั้งที่ 22 (OR-NET 2025) ณ วันที่ 6 เมษายน 2568
สถานที่ทำงาน	สำนักงานสถิติจังหวัดนครปฐม
ประสบการณ์การทำงาน	ธันวาคม 2565 ถึง พฤษภาคม 2568 นักวิชาการสถิติ ปฏิบัติการ สำนักงานป้องกันและบำบัดการติดยาเสพติด สำนักอนามัย มิถุนายน 2568 ถึงปัจจุบัน นักวิชาการสถิติ ปฏิบัติการ กลุ่มวิชาการสถิติและวางแผน สำนักงานสถิติจังหวัดนครปฐม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้