

ชื่อโครงการ	การวิเคราะห์ข้อมูลเพื่อทำนายลักษณะของผู้ป่วยโรคเบาหวานจากชุดข้อมูลทุดิยภูมิแบบเปิด	
โดย	นางสาวกัญญา ดินนะศรี	รหัส 65541207005-5
	นางสาวสิริกร แสงเมาะ	รหัส 65541207163-2
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร.รัชฎาพร ปุกแก้ว	
หลักสูตร	ระบบสารสนเทศทางธุรกิจ	
	สาขาบริหารธุรกิจ คณะบริหารธุรกิจและศิลปศาสตร์	
ปีการศึกษา	2568	

บทคัดย่อ

โครงการวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์และทำนายลักษณะของผู้ป่วยโรคเบาหวานจากชุดข้อมูลทุดิยภูมิแบบเปิด โดยนำชุดข้อมูลจากเว็บไซต์ Kaggle ซึ่งประกอบด้วยข้อมูลผู้ป่วยจำนวน 100,000 รายการ มาดำเนินการเตรียมข้อมูล แก้ไขค่าที่ขาดหาย และคัดเลือกตัวแปรสำคัญที่เกี่ยวข้องกับปัจจัยเสี่ยง เช่น เพศ ความดันโลหิตสูง และโรคหัวใจ จากนั้นใช้กระบวนการวิเคราะห์ข้อมูลตามขั้นตอนของ CRISP-DM และนำเทคนิคการทำเหมืองข้อมูลแบบ Classification ในรูปแบบของต้นไม้ตัดสินใจ (Decision tree) มาสร้างแบบจำลองเพื่อใช้ในการทำนาย กระบวนการวิจัยเริ่มจากการจัดเตรียมและทำความสะอาดข้อมูล ซึ่งข้อมูลต้นฉบับมีความไม่สมดุลระหว่างกลุ่มผู้ป่วยและผู้ไม่ป่วย จึงนำเทคนิค SMOTE (Synthetic Minority Over-sampling Technique) มาใช้เพื่อปรับสมดุลข้อมูลก่อนการสร้างแบบจำลอง ผลการวิเคราะห์พบว่าโมเดลให้ความแม่นยำสูงสุดมีความถูกต้องเฉลี่ย 90.69% โดยมีตัวแปร hbA1c level และ blood glucose level เป็นปัจจัยสำคัญที่สุดในการจำแนกผลลัพธ์ ส่วนตัวแปร age, bmi และ hypertension มีอิทธิพลในลำดับรองลงมา ผลการศึกษาชี้ให้เห็นว่าเทคนิคการทำเหมืองข้อมูลรวมกับการปรับสมดุลข้อมูลสามารถช่วยเพิ่มประสิทธิภาพในการจำแนกผู้ป่วยโรคเบาหวานได้อย่างมีนัยสำคัญ และผลลัพธ์ที่ได้สามารถนำไปประยุกต์ใช้ในการสร้างระบบสนับสนุนการตัดสินใจทางด้านสาธารณสุข เพื่อช่วยวางแผนการป้องกันและติดตามความเสี่ยงของผู้ป่วยโรคเบาหวานได้อย่างมีประสิทธิภาพ