บทที่ 1
บทนำ

1.1 หลักการเหตุผลความเป็นมาและความสำคัญของโครงงาน

โรคเบาหวาน เป็นภาวะเรื้อรังที่เกิดจากระดับน้ำตาลในเลือดสูงกว่าปกติ เนื่องจากร่างกายไม่สามารถใช้น้ำตาลได้อย่างมีประสิทธิภาพ ซึ่งเกิดจากความผิดปกติของฮอร์โมนอินซูลิน หรือการตอบสนองต่ออินซูลินที่บกพร่อง โรคนี้สามารถนำไปสู่ภาวะแทรกซ้อนที่รุนแรงหากไม่ได้รับการควบคุมอย่างเหมาะสม เช่น โรคหัวใจ ไตวายเรื้อรัง ตาบอด หรือแผลเรื้อรังที่เท้า ในปัจจุบันหลายประเทศใช้เกณฑ์ระดับน้ำตาลที่มากกว่า 126 มิลลิกรัมต่อเดซิลิตร โดยมีข้อแม้ว่าเป็นค่าของน้ำตาลในน้ำเลือดหลังจากอดอาหารมาอย่างน้อย 8 ชม. แล้ว

การที่ผู้ที่เป็นโรคเบาหวานส่วนใหญ่ไม่สามารถควบคุมระดับน้ำตาลให้อยู่ในเกณฑ์ที่เหมาะสมได้จะส่งผลให้เกิดภาวะแทรกซ้อนต่างๆ ตามมา ปัจจุบันอัตราการเข้าพักรักษาตัวในโรงพยาบาลของผู้ที่เป็นโรคเบาหวานมีแนวโน้มเพิ่มมากขึ้นอย่างต่อเนื่อง หากผู้ที่เป็นโรคเบาหวานไม่ได้รับการดูแลรักษาอย่างถูกต้อง รวมถึงการให้สุขศึกษาและการสนับสนุนการดูแลที่ต่อเนื่อง จะก่อให้เกิดภาวะแทรกซ้อนในหลายระบบของร่างกาย ทั้งภาวะแทรกซ้อนแบบเฉียบพลันและแบบเรื้อรัง นำไปสู่ความพิการและเสียชีวิตก่อนวัยอันควร ส่งผลกระทบต่อการดำรงชีวิต ภาวะเศรษฐกิจ คุณภาพชีวิตของผู้ที่เป็นโรค ครอบครัว รวมทั้งประเทศชาติ

คณะกรรมการผู้เชี่ยวชาญสมาคมเบาหวานแห่งประเทศสหรัฐประชุมกันพิจารณางานวิจัยที่มีทั้งหมดก่อนปีค.ศ.1999 เพื่อหาว่าโรคแทรกซ้อนของโรคเบาหวานจะเริ่มเกิดขึ้นเมื่อระดับน้ำตาลในเลือดสูงขึ้นที่ระดับเท่าใด ผลพบโรคแทรกซ้อนที่เกิดขึ้นเร็วที่สุด คือ จอประสาทตาเสื่อม เกิดขึ้นที่ระดับน้ำตาลในเลือด 126 มิลลิกรัมต่อเดซิลิตร ดังนั้นทุกคนควรจะเริ่มตระหนักว่า ถ้าเรามีระดับน้ำตาลสูงมากกว่า 126 จะมีโอกาสเกิดโรคแทรกซ้อนในระยะยาวของเบาหวาน แม้ว่าที่น้ำตาลระดับนี้จะไม่ทำให้เกิดอาการใดๆ แต่วิธีการจะทำให้ระดับน้ำตาลต่ำลงนั้น ต้องพิจารณากันอีกทีเพราะการควบคุมโรคเบาหวานนั้นมีหลายวิธี อย่างไรก็ตามองค์การอนามัยโลกยังคงใช้เกณฑ์ระดับน้ำตาลที่ 200 มิลลิกรัมต่อเดซิลิตร

การวิเคราะห์ข้อมูลโรคเบาหวาน คือการนำข้อมูลที่ได้ถูกรวบรวมไว้และถูกเผยแพร่ผ่าน เว็บไซต์ Kaggle ซึ่งเป็นเว็บไซต์ที่เก็บรวบรวมชุดข้อมูลต่างๆ เป็นข้อมูลที่เปิดเผยได้ เพื่อให้ผู้ใช้บริการสามารถนำชุดข้อมูลไปศึกษา ซึ่งการนำข้อมูลโรคเบาหวานมาทำการวิเคราะห์นั้น

ต้องการทราบถึงข้อมูลในเชิงลึกของโรคเบาหวานทั้ง เพศ อายุ ประวัติโรคหัวใจ ประวัติการสูบบุหรี่ ค่าดัชนีมวลกาย และการเป็นโรคเบาหวาน ซึ่งการวิเคราะห์ปัจจัยต่างๆ จากข้อมูลที่ได้นั้นถือเป็นเรื่องสำคัญต่อการศึกษาหาผลสรุป แต่ด้วยจำนวนข้อมูลที่มีปริมาณมากนั้นทำให้ยากต่อการนำเสนอให้เข้าใจได้ง่าย ซึ่งถือเป็นข้อจำกัดหากต้องการนำข้อมูลชุดนี้ไปใช้ เพื่อเผยแพร่ให้แก่บุคคลภายนอกหรือกลุ่มผู้ใช้ข้อมูลได้รับรู้

จากปัญหาข้างต้นดังนั้นผู้วิเคราะห์ข้อมูลจึงนำชุดข้อมูลที่ได้มานั้น เข้าสู่กระบวนการจำแนก ข้อมูล จัดการ แก้ไขข้อมูล เพื่อให้ได้ข้อมูลที่เหมาะสมตามความต้องการสำหรับการวิเคราะห์ข้อมูล และวิเคราะห์ข้อมูลด้วย CRISP-DM จากเทคนิคการทำเหมืองข้อมูลแบบ Classification ได้แก่ แผนภูมิต้นไม้ (Decision tree) นาอีฟเบย์ (Naïve Bayes) ป่าสุ่ม (Random Forest) ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM) โครงข่ายประสาทเทียม (Artificial Neural Network: ANN) เป็นโครงสร้างที่ใช้แสดงกฎที่ได้จากเทคนิคการจำแนกประเภทข้อมูล จะช่วยให้เกิดความเข้าใจต่อข้อมูลและสิ่งที่ต้องการทราบอย่างแท้จริง และทางผู้วิเคราะห์ข้อมูลยังได้เผยแพร่ข้อมูลสารสนเทศนี้บน Web browser โดยการใช้ภาษา HTML ในการพัฒนาหน้าเว็บไซต์สำหรับการเปิดเผยข้อมูลการวิเคราะห์ข้อมูลโรคเบาหวาน เพื่อให้ผู้ใช้งานทราบถึงเกี่ยวกับการวิเคราะห์โรคเบาหวาน ซึ่งได้ผลสรุปที่ถูกต้อง สะดวกรวดเร็ว เข้าใจง่าย ทันต่อเวลาที่ต้องการศึกษา

1.2 วัตถุประสงค์

1.2.1 เพื่อศึกษาองค์ประกอบความเสี่ยงของผู้ที่เป็นโรคเบาหวานจากฐานข้อมูลแบบเปิด

1.2.2 เพื่อวิเคราะห์ปัจจัยเสี่ยงหลักที่มีผลต่อการเกิดโรคเบาหวาน เช่น อายุ BMI ความดันโลหิต ประวัติการสูบบุหรี่ และระดับน้ำตาลในเลือด

1.2.3 เพื่อพัฒนาเว็บไซต์สำหรับเผยแพร่องค์ความรู้ที่ได้จากการวิเคราะห์ปัจจัยเสี่ยงหลักที่มีผลต่อการเกิดโรคเบาหวาน

1.3 ประโยชน์ที่จะได้รับ

1.3.1 ได้องค์ประกอบความเสี่ยงของผู้ที่เป็นโรคเบาหวานจากฐานข้อมูลแบบเปิด

1.3.2 ได้ผลการวิเคราะห์ปัจจัยเสี่ยงหลักที่มีผลต่อการเกิดโรคเบาหวาน

1.3.3 ได้เว็บไซต์สำหรับเผยแพร่องค์ความรู้ที่ได้จากการวิเคราะห์ปัจจัยเสี่ยงหลักที่มีผลต่อการเกิดโรคเบาหวาน

1.4 ขอบเขต และเครื่องมือที่ใช้ในการพัฒนาระบบ

1.4.1 ขอบเขตการวิเคราะห์ข้อมูล

1.4.1.1 รวบรวมข้อมูลโรคเบาหวาน เช่น อายุ ความดันโลหิต ประวัติโรคหัวใจ ประวัติการสูบบุหรี่ ค่าดัชนีมวลกาย และการเป็นโรคเบาหวาน เป็นต้น จากฐานข้อมูลแบบเปิด เช่น Kaggle Data.org และศูนย์กลางข้อมูลเปิดภาครัฐ รวมทั้งข้อมูลในระบบอินเตอร์เน็ตที่เชื่อถือได้ เช่น ข้อมูลเว็บไซต์โรงพยาบาลและผู้เชี่ยวชาญเกี่ยวกับโรคเบาหวาน เพื่อนำมาศึกษาองค์ประกอบความเสี่ยงของผู้ที่เป็นโรคเบาหวาน

1.4.1.2 กลั่นกรองข้อมูล (Data Cleaning) ที่รวมรวบได้ทั้งหมด เช่น ข้อมูลที่ผิดรูปแบบ เติมข้อมูลที่ขาดหายไปให้ครบถ้วน ให้ข้อมูลมีความสมบูรณ์ ตัดข้อมูลที่ไม่จำเป็นออก เป็นต้น เพื่อให้ข้อมูลมีความถูกต้องและสามารถนำไปใช้การวิเคราะห์ข้อมูลต่อไปได้

1.4.1.3 ใช้ข้อมูลจากการกลั่นกรองไปวิเคราะห์เพื่อทำนายองค์ประกอบความเสี่ยงหลักของผู้ที่เป็นโรคเบาหวาน โดยใช้เทคนิคการทำเหมืองข้อมูล (Data Mining) ในการจำแนกประเภท (Classification) เช่น ต้นไม้ตัดสินใจ (Decision Tree) ป่าสุ่ม (Random Forest) ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM) เพื่อทำนายว่าแต่ละบุคคลมีองค์ประกอบสำคัญใดบ้างที่มีความเสี่ยงของการเป็นโรคเบาหวาน

1.4.1.4 โครงงานนี้ใช้โปรแกรมด้านการวิเคราะห์ข้อมูล เช่น Rapid Miner เพื่อใช้ในการทำนายองค์ประกอบความเสี่ยงหลักที่มีผลต่อการเป็นโรคเบาหวาน

1.4.2 ขอบเขตผู้ใช้ทั่วไป

1.4.2.1 สามารถดูข้อมูลความรู้และบทความเกี่ยวกับโรคเบาหวานได้

1.4.2.2 สามารถดูแดชบอร์ดข้อมูลสารสนเทศของการโรคเบาหวาน

1.4.2.3 สามารถนำข้อมูลบนแดชบอร์ดไปใช้ในการสรุปผลข้อมูลเบื้องต้นได้

1.4.2.4 สามารถดาวน์โหลดชุดข้อมูลที่กลั่นกรองแล้วไปใช้ต่อได้

1.4.3 ขอบเขตผู้ดูแลระบบ

1.4.3.1 สามารถเข้าสู่ระบบได้ โดยใช้ชื่อผู้ใช้ และรหัสผ่าน

1.4.3.2 สามารถแก้ไขข้อมูลส่วนตัวของผู้ดูแลเว็บไซต์ได้

1.4.3.3 สามารถเพิ่ม ลบ แก้ไขข้อมูลความรู้ทั่วไปเกี่ยวกับโรคเบาหวานทั้งหมดบนเว็บไซต์ได้ เช่น บทความวิชาการที่เกี่ยวกับโรคเบาหวาน ข้อมูลองค์ประกอบความเสี่ยงหลักที่มีต่อโรคเบาหวานที่ได้จากผลการวิเคราะห์ข้อมูล แบบสอบถามเพื่อทำนายความเสี่ยงที่อาจก่อให้เกิดโรคเบาหวานในระดับสูง ปานกลาง ต่ำ เป็นต้น รวมทั้งเผยแพร่เทคนิคการวิเคราะห์ข้อมูลที่มีความแม่นยำในระดับดี

1.4.3.4 นำเทคนิคที่ได้จากผลการทดสอบประสิทธิภาพการทำนายที่ดีที่สุดมาประยุกต์ใช้บนเว็บไซต์เพื่อทำนายความเสี่ยงการเป็นโรคเบาหวานของผู้ที่เข้ามาใช้แบบสอบถาม

1.4.3.5 อัปโหลดไฟล์ข้อมูลที่กลั่นกรองแล้ว หรือบันทึกข้อมูลที่กลั่นกรองแล้วทีละรายการ เพื่อนำมาใช้ในการทำนายความเสี่ยงการเป็นโรคเบาหวาน

1.4.4 เครื่องมือที่ใช้ในการพัฒนาโปรแกรม

1.4.4.1 Hardware

1) โน๊ตบุ๊ค Huawei CPU Intel(R) core i3 1.80 GHz Ram 4 GB Harddisk 500 GB

1.4.4.2 Software

1) โปรแกรม Power BI เพื่อใช้ในการแสดงข้อมูลในรูปแบบของภาพ

2) โปรแกรม Adobe XD ใช้สำหรับการออกแบบหน้าเว็บไซต์

3) โปรแกรม Rapid Miner ใช้ในการสร้างโมเดล Decision tree

4) โปรแกรม Apache เพื่อจำลองเซิร์ฟเวอร์

5) โปรแกรม Visual Studio Code ใช้ในการสร้างเว็บไซต์

6) โปรแกรม Tableau Public เพื่อใช้ในการแสดงข้อมูลในรูปแบบของภาพ

7) Microsoft Office Word 2019 ใช้ในการทำเอกสาร

8) Microsoft Office Excel 2019 ใช้ในการจัดการชุดข้อมูล

9) ระบบปฏิบัติการ Window 10 ใช้ในการเปิดซอฟต์แวร์ต่างๆ

10. สถานที่ใช้ในการดำเนินการศึกษาและรวบรวมข้อมูล

- มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา ตั้งที่อยู่ที่ 128 ถนนห้วยแก้ว ตำบลช้างเผือก อำเภอเมืองเชียงใหม่ จังหวัดเชียงใหม่ 50300
- แหล่งรวบรวมข้อมูล เว็บไซต์ Kaggle หรือ

https://kaggle.com/datasets/priyamchoksi/100000-diabetes-clinical-dataset

11. ระยะเวลาในการดำเนินการ

แผนการดำเนินการ	2568											
	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.	มิ.ย.	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.
1. ศึกษาและกำหนดความต้องการ	◄——►											
2. วิเคราะห์ออกแบบระบบและสร้างฐานข้อมูล		◄——►										
3. เขียนและทดสอบโปรแกรม		◄———	———	———►								
4. ติดตั้ง ทดสอบ และปรับปรุงระบบ					◄———	———►						
5. ตรวจสอบระบบโดยรวม							◄——►					
6. ประเมินการใช้งานระบบ								◄——►				
7. จัดทำคู่มือการใช้งาน	◄———	———	———	———	———	———	———	———	———	———►		
8. จัดทำเอกสารประกอบโครงงาน	◄———	———	———	———	———	———	———	———	———	———►		

12. นิยามศัพท์เฉพาะ

โรคเบาหวาน หมายถึง ภาวะความผิดปกติเรื้อรังที่ร่างกายไม่สามารถนำน้ำตาลกลูโคสจากอาหารไปใช้งานได้อย่างมีประสิทธิภาพ เนื่องจากการขาดฮอร์โมนอินซูลิน หรือร่างกายไม่ตอบสนองต่อฮอร์โมนอินซูลินอย่างเหมาะสม ส่งผลให้ระดับน้ำตาลในเลือดสูงกว่าปกติ และหากระดับน้ำตาลในเลือดสูงเกินขีดความสามารถของไตในการดูดกลับ น้ำตาลส่วนเกินจะถูกขับออกมาทางปัสสาวะ

ภาวะน้ำตาลในเลือดสูง (Impaired Glucose Tolerance , IGT) หมายถึง ระดับน้ำตาลในเลือดหลังดื่มน้ำตาล 75 กรัม สูงกว่าปกติ แต่ยังไม่ถึงระดับเบาหวาน

ภาวะน้ำตาลในเลือดขณะอดอาหารสูงผิดปกติ (Impaired fasting glucose , IFG) หมายถึง ระดับน้ำตาลในเลือดหลังอดอาหาร (Fasting Plasma Glucose – FPG) สูงกว่าปกติ แต่ยังไม่ถึงระดับเบาหวาน

ดัชนีมวลกาย (Body Mass Index: BMI) หมายถึง ค่าดัชนีที่ใช้ชี้วัดความสมดุลของน้ำหนักตัว (กิโลกรัม) และส่วนสูง (เซนติเมตร) ซึ่งสามารถระบุได้ว่า ตอนนี้รูปร่างของคนคนนั้นอยู่ในระดับใด ตั้งแต่อ้วนมากไปจนถึงผอมเกินไป

การเจาะน้ำตาลหลังอดอาหาร (Fasting Blood Sugar: FBS) หมายถึง วิธีการตรวจที่ต้องงดอาหารก่อนการเจาะเลือดอย่างน้อย 8 ชั่วโมง ผลการตรวจคือค่าความเข้มข้นน้ำตาลในเลือด ณ เวลาที่เจาะเลือดภายใต้ระยะเวลาการอดอาหารตามที่กำหนด เป็นวิธีมาตรฐานที่สามารถเปรียบเทียบค่าระหว่างบุคคลได้ว่าร่างกายของใครควบคุมระดับน้ำตาลได้ดีกว่ากัน

น้ำตาลสะสม (Glycated hemoglobin: HbA1c) หมายถึง การตรวจวัดระดับโปรตีนฮีโมโกลบินในเซลล์เม็ดเลือดแดงที่ถูกจับเกาะด้วยน้ำตาลกลูโคส ค่าที่ได้จะสะท้อนถึงระดับความเข้มข้นเฉลี่ยของน้ำตาลในเลือดตลอดช่วงเวลา 2–3 เดือนที่ผ่านมา ซึ่งชนิดและปริมาณของอาหารที่รับประทานในช่วงนั้นจะมีอิทธิพลต่อค่านี้มาก การตรวจด้วยวิธีนี้ใช้ในการวินิจฉัยโรคเบาหวานและติดตามการควบคุมโรคเบาหวานเพื่อประเมินความเสี่ยงของการเกิดโรคแทรกซ้อน การตรวจวิธีนี้ไม่จำเป็นต้องอดอาหารเหมือนวิธีแรก