

บทที่ 4

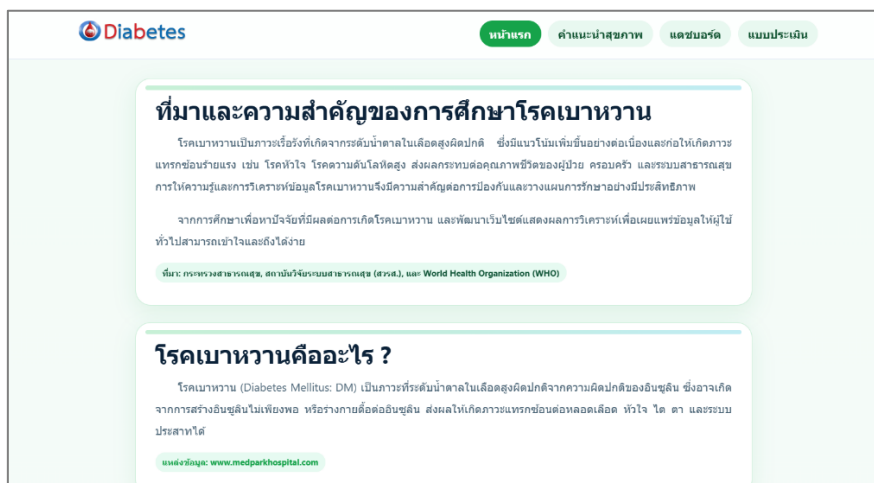
ผลการดำเนินงาน

จากการวิเคราะห์ข้อมูลเพื่อทำนายลักษณะของผู้ป่วยโรคเบาหวานจากชุดข้อมูลทุติยภูมิแบบเปิด โดยข้อมูลที่นำมาวิเคราะห์จำนวน 9 แอตทริบิวต์ ได้แก่ เพศ (Gender) อายุ (Age) ความดันโลหิตสูง (Hypertension) ประวัติโรคหัวใจ (Heart_disease) ประวัติการสูบบุหรี่ (smoking_history) ค่าดัชนีมวลกาย (BMI) ระดับน้ำตาลสะสมเฉลี่ยในเลือด (hbA1c_level) ระดับน้ำตาลในเลือด (blood_glucose_level) และการเป็นโรคเบาหวาน (diabetes) เพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์ โดยมีวัตถุประสงค์วิเคราะห์ปัจจัยเสี่ยงหลักที่มีผลต่อการเกิดโรคเบาหวาน สำหรับเผยแพร่ข้อมูลความเป็นไปได้บนเว็บไซต์ เพื่อให้ได้เว็บไซต์ที่รวมแหล่งความรู้ทั่วไปเกี่ยวกับข้อมูล และแสดงกระบวนการวิเคราะห์ข้อมูลที่สามารถนำไปใช้งานได้ อย่างมีประสิทธิภาพและประสิทธิผล โดยผู้จัดทำโครงการได้ทำการวิเคราะห์ข้อมูล และนำเทคโนโลยีเข้าช่วยในการเสนอข้อมูล ทำให้มีความสะดวกรวดเร็ว และสามารถใช้งานได้จริงจนสำเร็จลุล่วงตามเป้าหมาย

4.1 ผลการดำเนินงาน

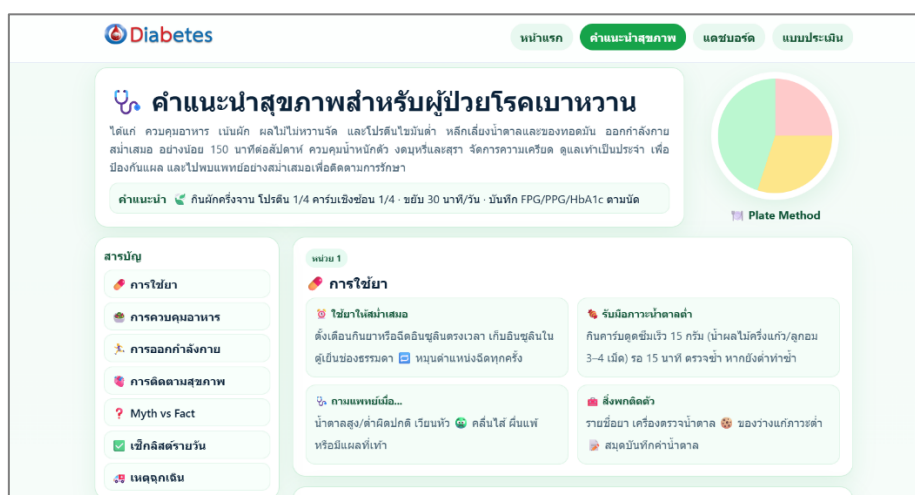
ผู้วิเคราะห์ข้อมูลได้ทำการจัดการกับข้อมูลวิเคราะห์ลักษณะของผู้ป่วยโรคเบาหวานจากชุดข้อมูลทุติยภูมิแบบเปิด ด้วยขั้นตอนกระบวนการวิเคราะห์ข้อมูล CRISP-DM อย่างละเอียดเพื่อให้ข้อมูลมีความสมบูรณ์ และมีความถูกต้องมากที่สุด จากนั้นผู้วิเคราะห์ข้อมูลได้ทำการวิเคราะห์ข้อมูลให้อยู่ในรูปแบบของ Data Mining โดยการจำแนกข้อมูลออกเป็นประเภทต่างๆ ด้วย เทคนิคการทำเหมืองข้อมูลแบบการจำแนกข้อมูล (Classification) ในรูปแบบของเทคนิคต้นไม้ตัดสินใจ เป็นโครงสร้างที่ใช้แสดงกฎที่ได้จากเทคนิคการจำแนกประเภทข้อมูล จะช่วยให้เกิดความเข้าใจต่อข้อมูล โดยมีวัตถุประสงค์เพื่อวิเคราะห์ข้อมูลเพื่อทำนายลักษณะของผู้ป่วยโรคเบาหวานจากชุดข้อมูลทุติยภูมิแบบเปิด สำหรับเผยแพร่ข้อมูลความเป็นไปได้บนเว็บไซต์ผู้วิเคราะห์ข้อมูลจัดทำการนำเสนอข้อมูลสารสนเทศแบบ Visualization ได้เผยแพร่ข้อมูลสารสนเทศนี้บน Web Browser ให้กับผู้ใช้งานทั่วไป ซึ่งทำให้ผู้ใช้งานทั่วไปเข้าใจได้อย่างสะดวกรวดเร็วยิ่งขึ้น รวมถึงมีความถูกต้อง ลดความซ้ำซ้อนจากข้อมูลที่มีจำนวนมากมหาศาล

เพิ่มประสิทธิภาพให้กับการศึกษาค้นคว้าและทันต่อเวลา ผู้วิเคราะห์ข้อมูลจึงได้ทำการวิเคราะห์ข้อมูล และจัดทำเว็บไซต์ตามวัตถุประสงค์ที่ได้กำหนดไว้และมีผลการดำเนินงานดังนี้



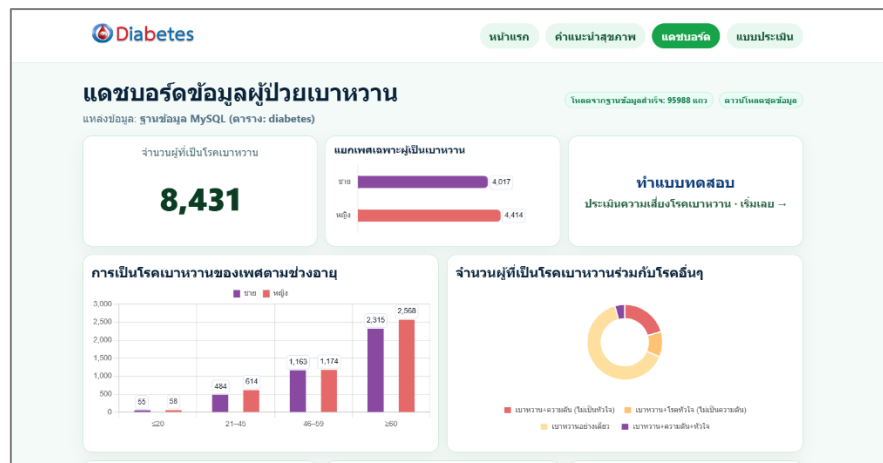
ภาพที่ 4.1 หน้าแรกของเว็บไซต์ และหน้าแสดงข้อมูลในด้านความรู้โรคเบาหวาน

จากภาพที่ 4.1 แสดงข้อมูลความรู้ที่เกี่ยวกับโรคเบาหวาน โดยนำเสนอข้อมูลต่างๆ เช่น ความหมายของโรคเบาหวาน อาการและระดับน้ำตาลในเลือดที่มีผลต่อสุขภาพ รวมถึงการแบ่งประเภทของโรคเบาหวานออกเป็น 4 ประเภท พร้อมคำแนะนำในการป้องกันและรักษาโรค



ภาพที่ 4.2 หน้าคำแนะนำสุขภาพสำหรับผู้ป่วยโรคเบาหวาน

จากภาพที่ 4.2 ภาพแสดงหน้าเว็บแนะนำสุขภาพสำหรับผู้ป่วยโรคเบาหวาน โดยมีคำแนะนำเรื่องการควบคุมอาหาร การใช้ยา และการออกกำลังกายอย่างเหมาะสม พร้อมตัวอย่างสัดส่วนอาหารแบบ Plate Method เพื่อช่วยให้ผู้ป่วยควบคุมระดับน้ำตาลในเลือดได้อย่างมีประสิทธิภาพ



ภาพที่ 4.3 หน้าแสดง Dashboard ข้อมูลผู้ป่วยเบาหวาน

จากภาพที่ 4.3 แสดงภาพแดชบอร์ดข้อมูลผู้ป่วยโรคเบาหวาน พร้อมการเปรียบเทียบของแต่ละแอตทริบิวต์บนกราฟ

The screenshot shows the login page for the Diabetes system. At the top, there's a navigation bar with links: หน้าแรก, คำแนะนำสุขภาพ, **แดชบอร์ด**, แบบประเมิน, and **สำหรับผู้ดูแล**. The main title is "เข้าสู่ระบบผู้จัดการข้อมูล" (System Manager Login).

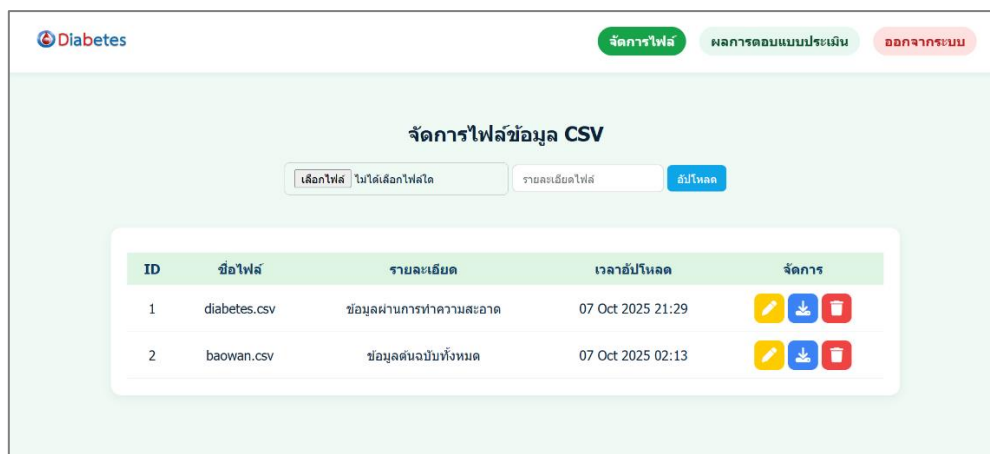
The login form includes:

- ชื่อผู้ใช้ (Username) input field
- รหัสผ่าน (Password) input field
- เข้าสู่ระบบ (Login) button

Below the button, it says "เข้าสู่ระบบเพื่อจัดการข้อมูล Diabetes Info Center".

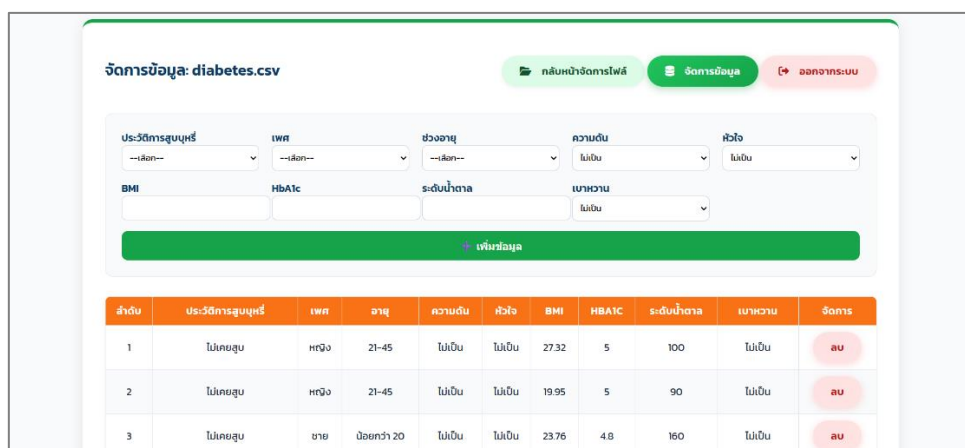
ภาพที่ 4.4 หน้าแสดงการเข้าสู่ระบบ

จากภาพที่ 4.4 แสดงหน้าการเข้าสู่ระบบเพื่อไปจัดการกับข้อมูลภายในเว็บไซต์ โดยจำกัดการเข้าถึงข้อมูลเฉพาะแอดมินเท่านั้น



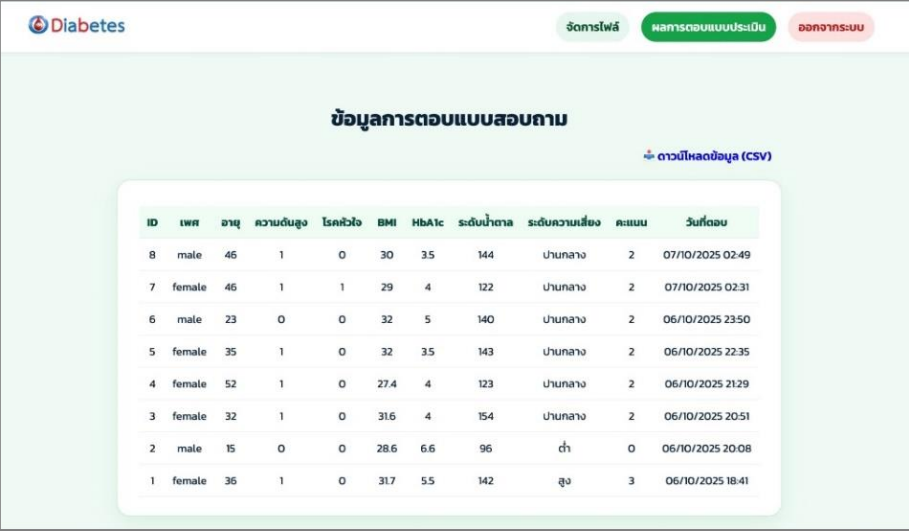
ภาพที่ 4.5 หน้าแสดงการจัดการไฟล์ข้อมูล csv สำหรับแอดมิน

จากภาพที่ 4.5 แสดงหน้าจัดการไฟล์สำหรับข้อมูล csv สามารถทำการแก้ไข ดาวน์โหลด และลบข้อมูลได้



ภาพที่ 4.6 หน้าแสดงการเพิ่มข้อมูล

จากภาพที่ 4.6 แสดงหน้าการเพิ่มข้อมูล สามารถเพิ่มข้อมูลในแถวได้



ID	เพศ	อายุ	ความสูง	น้ำหนัก	BMI	HbA1c	ระดับน้ำตาล	ระดับความเครียด	คะแนน	วันที่ตอบ
8	male	46	1	0	30	3.5	144	ปานกลาง	2	07/10/2025 02:49
7	female	46	1	1	29	4	122	ปานกลาง	2	07/10/2025 02:31
6	male	23	0	0	32	5	140	ปานกลาง	2	06/10/2025 23:50
5	female	35	1	0	32	3.5	143	ปานกลาง	2	06/10/2025 22:35
4	female	52	1	0	27.4	4	123	ปานกลาง	2	06/10/2025 21:29
3	female	32	1	0	31.6	4	154	ปานกลาง	2	06/10/2025 20:51
2	male	15	0	0	28.6	6.6	96	ต่ำ	0	06/10/2025 20:08
1	female	36	1	0	31.7	5.5	142	สูง	3	06/10/2025 18:41

ภาพที่ 4.7 หน้าแสดงข้อมูลการตอบแบบสอบถาม

จากภาพที่ 4.7 แสดงผลการตอบแบบประเมินของผู้ป่วยโรคเบาหวาน พร้อมสามารถดาวน์โหลดข้อมูลตารางออกมาเป็นไฟล์ csv ได้

4.2 การอภิปรายผล

จากวัตถุประสงค์เพื่อวิเคราะห์ข้อมูลเพื่อทำนายลักษณะของผู้ป่วยโรคเบาหวานจากชุดข้อมูลทุติยภูมิแบบเปิด สำหรับเผยแพร่ข้อมูลความเป็นไปได้บนเว็บไซต์ ผู้วิเคราะห์ได้ศึกษาปัญหา และเก็บรวบรวมข้อมูล ซึ่งผู้วิเคราะห์ข้อมูลได้ทำการวิเคราะห์ในรูปแบบของ Data Mining โดยใช้กระบวนการวิเคราะห์ข้อมูลด้วย Cross Industry Standard Process for Data Mining หรือ CRISP-DM เพื่อใช้ในการวิเคราะห์ตามขอบเขตของโครงการดังนี้

4.2.1 อภิปรายผลการดำเนินโครงการ

4.2.1.1 อภิปรายผลจากการวิเคราะห์ข้อมูล

Diabetes

หน้าแรก คำแนะนำสุขภาพ แลชบอร์ด **แบบประเมิน**

แบบประเมินความเสี่ยงโรคเบาหวาน

เพศ
☐ หญิง ☐ ชาย

อายุ (ปี)
 เช่น 35 ปี

ความดันโลหิตสูง
☐ มี ☐ ไม่มี

โรคหัวใจ
☐ มี ☐ ไม่มี

BMI (kg/m²)
 เช่น 24.7 kg/m²

ระดับน้ำตาลในเลือดสะสม HbA1c (% เอเลียม)
 เช่น 6.2 %

ภาพที่ 4.8 หน้าแสดงการทำแบบประเมินความเสี่ยงโรคเบาหวาน

ผลการประเมินความเสี่ยงโรคเบาหวาน

คะแนนของคุณคือ **2.0** คะแนน

ระดับความเสี่ยง: **ปานกลาง**

[พิมพ์/บันทึก PDF](#) [ปิดหน้าต่าง](#)

ภาพที่ 4.9 ตัวอย่างผลแบบประเมินสุขภาพจิต

จากภาพที่ 4.8 และ 4.9 หน้าแสดงการทำแบบทดสอบ ผู้ที่เข้ามาใช้งานเว็บไซต์สามารถทำการแบบประเมินความเสี่ยงโรคเบาหวานได้ โดยมีการเก็บข้อมูล เพศ อายุ ความดันโลหิต รวมไปถึงข้อคำถามที่ได้ทำการตอบแบบทดสอบเอาไว้ และจะมีเกณฑ์แบ่งระดับความเสี่ยงในการเกิดโรคเบาหวานออกเป็น 4 ระดับ

(1) ความเสี่ยงสูงมาก (คะแนน ≥ 3.0): มีความเสี่ยงในระดับอันตราย ผู้ใช้ควรดำเนินการตรวจสอบสุขภาพโดยด่วน เช่น ตรวจเลือดซ้ำ หรือเข้าพบแพทย์ทันที

(2) ความเสี่ยงสูง (คะแนน 2.5 – 2.9): จำเป็นต้องพิจารณาเข้ารับการตรวจยืนยันอย่างละเอียด หรือหากผลตรวจยังปกติ ก็ต้องปรับเปลี่ยนพฤติกรรมเพื่อป้องกันโรค

(3) ความเสี่ยงปานกลาง (คะแนน 2.0 – 2.4): ควรอยู่ในกลุ่มที่ต้องเฝ้าระวัง รับทราบถึงปัจจัยเสี่ยงที่ตนเองมี และควรปรึกษาผู้เชี่ยวชาญเพื่อขอคำแนะนำเพิ่มเติม

(4) ความเสี่ยงต่ำ (คะแนน < 2.0): ไม่เข้าข่ายเบาหวาน ณ ขณะนี้ และควรดูแลสุขภาพโดยรวมต่อไป

1) กระบวนการศึกษาทำความเข้าใจธุรกิจ (Business Understanding) ผลการศึกษาพบว่าปัญหาของข้อมูลคือ ข้อมูลที่เกี่ยวข้องกับปัญหาโรคเบาหวานมีจำนวนมหาศาล โดยมีจำนวนข้อมูลทั้งหมด 100,000 รายการ ทำให้ไม่สามารถทำความเข้าใจกับข้อมูลที่มีความซับซ้อนได้อย่างรวดเร็ว ส่งผลให้การนำเสนอและเผยแพร่ข้อมูลดังกล่าวต่อ

บุคคลภายนอกหรือกลุ่มผู้ใช้ข้อมูลยังไม่สามารถทำได้อย่างมีประสิทธิภาพมากนัก ทั้งนี้ การวิเคราะห์แนวโน้มและการสื่อสารข้อมูลที่ชัดเจนมีความสำคัญอย่างยิ่งในการสร้างความตระหนักรู้เกี่ยวกับปัจจัยเสี่ยงของการเกิดโรคเบาหวาน

2) การทำความเข้าใจข้อมูล (Data Understanding) ผู้วิเคราะห์ข้อมูลทำการรวบรวมข้อมูลเพื่อตรวจสอบรายละเอียด ปริมาณ และความน่าเชื่อถือของข้อมูลปัญหาโรคเบาหวานที่ได้จากเว็บไซต์ [Kaggle – 100000 Diabetes Clinical Dataset](https://www.kaggle.com/datasets/stone-island/diabetes-clinical-dataset) ซึ่งเป็นเว็บไซต์ที่เก็บรวบรวมชุดข้อมูลต่างๆ เป็นแหล่งรวม Datasets หรือชุดข้อมูลสำหรับฝึกสอน Machine Learning ที่ใหญ่ที่สุดในโลกแห่งหนึ่ง มีข้อมูลทุกประเภท มีทั้ง Datasets ในหมวดหมู่ Finance, Business, Physics, Biology, Sports, News ซึ่งเป็นข้อมูลที่เปิดเผยได้ เพื่อให้ผู้ใช้บริการสามารถนำชุดข้อมูลไปศึกษาหรือวิเคราะห์ให้เกิดประโยชน์ต่อไป

3) การเตรียมข้อมูล (Data Preparation) ผู้วิเคราะห์ข้อมูลได้ทำการคัดเลือกข้อมูล และทำการ Data cleaning ข้อมูลเพื่อให้ได้ข้อมูลที่สมบูรณ์

4) การสร้างโมเดล (Modeling) ผู้วิเคราะห์ข้อมูลได้วิเคราะห์ข้อมูลด้วยการทำเหมืองข้อมูลแบบ Data Classification เลือกใช้เทคนิคการสร้างแบบจำลองพยากรณ์ด้วยเทคนิคต้นไม้ตัดสินใจ ซึ่งเป็นแบบจำลองที่ใช้สมการทางคณิตศาสตร์ และสถิติในการคำนวณ โดยใช้โปรแกรมที่ใช้ทำเหมืองข้อมูล ด้วยชุดข้อมูลที่คัดเลือก ประกอบด้วย 9 แอตทริบิวต์ คือ เพศ (Gender) อายุ (Age) ความดันโลหิตสูง (Hypertension) ประวัติโรคหัวใจ (Heart_disease) ประวัติการสูบบุหรี่ (Smoking_history) ค่าดัชนีมวลกาย (BMI) ระดับน้ำตาลสะสมเฉลี่ยในเลือด (hbA1c_level) ระดับน้ำตาลในเลือด (blood_glucose_level) และ การเป็นโรคเบาหวาน (diabetes)

5) การวัดประสิทธิภาพของโมเดล (Evaluation) การทดสอบประสิทธิภาพของแบบจำลองได้ดำเนินการวัดค่าความแม่นยำ (Accuracy) ค่าเรียกกลับ

(Precision) และค่าความถูกต้อง (Recall) เพื่อสร้างแบบจำลองการพยากรณ์จากการเรียนรู้และทดสอบแบบจำลองการพยากรณ์ โดยค่าความถูกต้อง (Accuracy) เท่ากับ 90.69% และได้ถูกจำนวน 10 กฎดังนี้

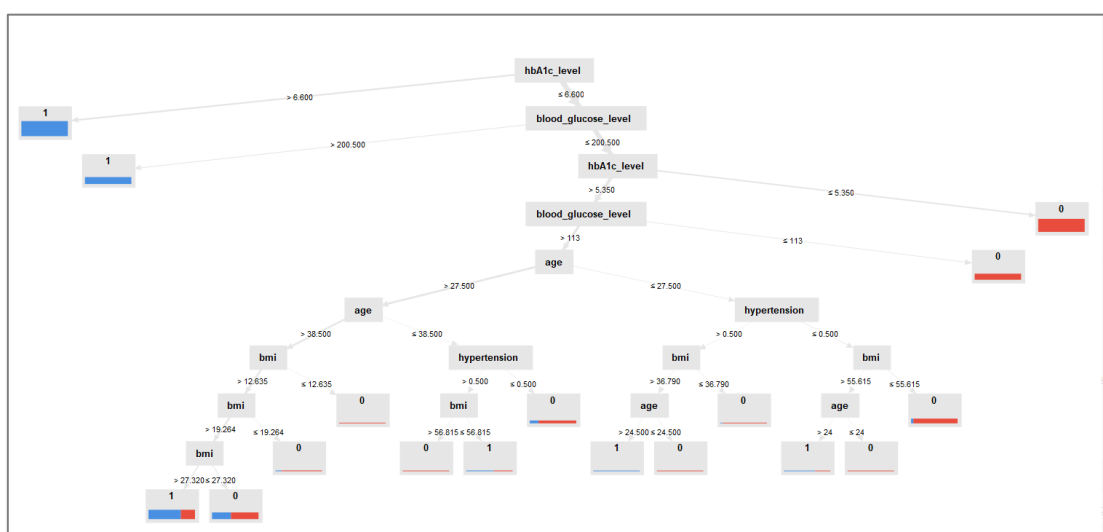
```

Tree

hbA1c_level > 6.600: 1 {1=28385, 0=0}
hbA1c_level ≤ 6.600
|   blood_glucose_level > 200.500: 1 {1=12296, 0=0}
|   blood_glucose_level ≤ 200.500
|   |   hbA1c_level > 5.350
|   |   |   blood_glucose_level > 113
|   |   |   |   age > 27.500
|   |   |   |   |   age > 38.500
|   |   |   |   |   |   bmi > 12.635
|   |   |   |   |   |   |   bmi > 19.264
|   |   |   |   |   |   |   |   bmi > 27.320: 1 {1=11889, 0=5192}
|   |   |   |   |   |   |   |   bmi ≤ 27.320: 0 {1=4826, 0=7032}
|   |   |   |   |   |   |   |   bmi ≤ 19.264: 0 {1=34, 0=234}
|   |   |   |   |   |   |   |   bmi ≤ 12.635: 0 {1=0, 0=2}
|   |   |   |   |   |   |   |   age ≤ 38.500
|   |   |   |   |   |   |   |   |   hypertension > 0.500
|   |   |   |   |   |   |   |   |   |   bmi > 56.815: 0 {1=0, 0=2}
|   |   |   |   |   |   |   |   |   |   bmi ≤ 56.815: 1 {1=117, 0=81}
|   |   |   |   |   |   |   |   |   |   hypertension ≤ 0.500: 0 {1=812, 0=3496}
|   |   |   |   |   |   |   |   |   |   age ≤ 27.500
|   |   |   |   |   |   |   |   |   |   |   hypertension > 0.500
|   |   |   |   |   |   |   |   |   |   |   |   bmi > 36.790
|   |   |   |   |   |   |   |   |   |   |   |   |   age > 24.500: 1 {1=7, 0=0}
|   |   |   |   |   |   |   |   |   |   |   |   |   age ≤ 24.500: 0 {1=0, 0=3}
|   |   |   |   |   |   |   |   |   |   |   |   |   bmi ≤ 36.790: 0 {1=1, 0=18}
|   |   |   |   |   |   |   |   |   |   |   |   |   hypertension ≤ 0.500
|   |   |   |   |   |   |   |   |   |   |   |   |   |   bmi > 55.615
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   age > 24: 1 {1=2, 0=1}
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   age ≤ 24: 0 {1=0, 0=4}
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   bmi ≤ 55.615: 0 {1=393, 0=7648}
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   blood_glucose_level ≤ 113: 0 {1=0, 0=10658}
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   hbA1c_level ≤ 5.350: 0 {1=0, 0=24391}

```

ภาพที่ 4.10 แสดงคำบรรยายลักษณะงาน Decision Tree ในโปรแกรม Rapid Miner



ภาพที่ 4.11 แสดงโมเดล Graph Decision Tree ในโปรแกรม Rapid Miner Studio

จากภาพที่ 4.1 จากผลลัพธ์ต้นไม้ตัดสินใจที่แสดงในภาพสามารถอธิบายได้ว่า ผลลัพธ์ที่ได้จากแต่ละปลายกิ่ง (leaf node) ของโมเดลจะถูกแบ่งออกเป็นสองกลุ่มหลักคือ เลข 0 และ เลข 1 โดย เลข 0 หมายถึง “ไม่เป็นเบาหวาน” และ เลข 1 หมายถึง “เป็นเบาหวาน” ซึ่งโมเดลได้เรียนรู้จากข้อมูลทางคลินิก เช่น ระดับน้ำตาลสะสมในเลือด (HbA1c), ระดับน้ำตาลในเลือดขณะสุ่มตรวจ (Blood Glucose), ค่าดัชนีมวลกาย (BMI), อายุ และภาวะความดันโลหิตสูง (Hypertension) เพื่อตัดสินใจผลลัพธ์นี้ แต่ละโหนดในต้นไม้จะแบ่งข้อมูลตามเงื่อนไขของตัวแปรเหล่านี้อย่างเป็นลำดับขั้น เช่น หากผู้ป่วยมีค่า HbA1c หรือ Blood Glucose สูง โมเดลจะมีแนวโน้มทำนายเป็น เลข 1 (เป็นเบาหวาน) ขณะที่หากค่าตัวแปรอยู่ในระดับต่ำและไม่มีปัจจัยเสี่ยงรวม เช่น BMI ต่ำหรือไม่มีความดันสูง ผลลัพธ์จะอยู่ในกลุ่ม เลข 0 (ไม่เป็นเบาหวาน) การกระจายของสีในกล่องแต่ละใบยังสะท้อนความน่าจะเป็นของการจำแนก เช่น สีแดงเด่นแสดงว่ากลุ่มนั้นส่วนใหญ่เป็นเบาหวาน (1) ส่วนสีน้ำเงินเด่นแสดงว่ากลุ่มนั้นส่วนใหญ่ไม่เป็นเบาหวาน (0) ทำให้เห็นได้ชัดเจนว่าตัวเลข 0 และ 1 คือผลลัพธ์สุดท้ายที่โมเดลใช้ในการจำแนกผู้ป่วยตามความเสี่ยงของโรคเบาหวาน

- กฎข้อที่ 1: IF hbA1c_level > 6.600 ผลลัพธ์ = 1 หมายความว่า หากค่า hbA1c_level สูงกว่า 6.600 จะได้ผลลัพธ์เป็น 1 ถือว่าเข้าสู่ภาวะเบาหวานแน่นอน
- กฎข้อที่ 2: IF hbA1c_level ≤ 6.600 AND blood_glucose_level > 200.500 ผลลัพธ์ = 1 หมายความว่า ถ้าค่า hbA1c_level ไม่เกิน 6.600 แต่กลูโคสในเลือดสูงมาก (>200.500) จะได้ผลลัพธ์เป็น 1 ซึ่งแสดงถึงระดับน้ำตาลในเลือดที่สูงมาก
- กฎข้อที่ 3: IF hbA1c_level ≤ 6.600 AND blood_glucose_level ≤ 200.500 AND hba1c_level > 5.350 AND blood_glucose_level > 113 AND age > 38.5 AND bmi > 27.320 ผลลัพธ์ = 1 หมายความว่า กลุ่มอายุ 38 ปีขึ้นไปที่มีน้ำหนักเกิน (BMI สูงมาก) และมีค่าน้ำตาลในเลือดมากกว่า 113 mg/dL มีแนวโน้มเป็นเบาหวานสูง
- กฎข้อที่ 4: IF hbA1c_level ≤ 6.600 AND blood_glucose_level ≤ 200.500 AND hba1c_level > 5.350 AND blood_glucose_level > 113 AND age > 38.5 AND 19.264 < bmi ≤ 27.320 ผลลัพธ์ = 0 หมายความว่า อายุเกิน 38 ปีแต่น้ำหนักแค่เกินเล็กน้อย-ปานกลาง ส่วนใหญ่ยังไม่เป็นเบาหวาน
- กฎข้อที่ 5: IF hbA1c_level ≤ 6.600 AND blood_glucose_level ≤ 200.500 AND hba1c_level > 5.350 AND blood_glucose_level > 113 AND age > 38.5 AND

$bmi \leq 12.635$ ผลลัพธ์ = 0 หมายความว่า ผู้ที่มีค่า BMI ต่ำมาก (ผอมมาก) แม้อายุมากแต่น้ำตาลในเลือดไม่สูงจัด มักไม่เป็นเบาหวาน

- กฎข้อที่ 6: IF $hbA1c_level \leq 6.600$ AND $blood_glucose_level \leq 200.500$ AND $hba1c_level > 5.350$ AND $blood_glucose_level > 113$ AND $27.5 < age \leq 38.5$ AND $hypertension > 0.5$ ผลลัพธ์ = 1 หมายความว่า กลุ่มวัยกลาง (~30-38 ปี) ที่มีความดันโลหิตสูงร่วมด้วย มักเข้าสู่ภาวะเบาหวาน

- กฎข้อที่ 7: IF $hbA1c_level \leq 6.600$ AND $blood_glucose_level \leq 200.500$ AND $hba1c_level > 5.350$ AND $blood_glucose_level > 113$ AND $27.5 < age \leq 38.5$ AND $hypertension \leq 0.5$ ผลลัพธ์ = 0 หมายความว่า อายุไม่เกิน 38 ปี และไม่มี ความดันโลหิตสูง มีแนวโน้มว่ายังไม่เป็นเบาหวาน

- กฎข้อที่ 8: IF $hbA1c_level \leq 6.600$ AND $blood_glucose_level \leq 200.500$ AND $hba1c_level > 5.350$ AND $blood_glucose_level > 113$ AND $age \leq 27.5$ AND $hypertension > 0.5$ AND $bmi > 36.790$ ผลลัพธ์ = 1 หมายความว่า กลุ่มอายุน้อยกว่า 27 ปีที่มีทั้งความดันโลหิตสูงและน้ำหนักมาก (อ้วนจัด) มีความเสี่ยงสูงเป็นเบาหวาน

- กฎข้อที่ 9: IF $hbA1c_level \leq 6.600$ AND $blood_glucose_level \leq 200.500$ AND $hba1c_level > 5.350$ AND $blood_glucose_level > 113$ AND $age \leq 27.5$ AND $hypertension \leq 0.5$ ผลลัพธ์ = 0 หมายความว่า ผู้วัยหนุ่มสาว ไม่มี ความดันสูง และค่าน้ำตาลเพียงสูงเล็กน้อย ยังไม่เป็นเบาหวาน

- กฎข้อที่ 10: IF $hbA1c_level \leq 5.350$ ผลลัพธ์ = 0 หมายความว่า ค่าฮีโมโกลบิน (HbA1c) ต่ำกว่า 5.35 จัดว่าอยู่ในเกณฑ์ปกติ ไม่เป็นเบาหวาน

จากภาพที่ 4.1 การสร้างโมเดลต้นไม้ตัดสินใจ ได้ใช้การจำแนกประเภทของข้อมูล (Classification) เป็นหนึ่งในเทคนิคที่สำคัญของเหมืองข้อมูล (Data Mining) เพื่อให้ได้ผลการทำนายที่เหมาะสม และผลการคำนวณค่าความถูกต้อง (accuracy) ค่าเรียกกลับ (recall) และค่าความแม่นยำ (precision) จากการทดสอบข้อมูล แสดงไว้ในตารางที่ 4.1

ตารางที่ 4.1 ผลการวัดประสิทธิภาพของโมเดล Classification ของโมเดล Decision Tree

Accuracy = 90.69%

Model	True 0	True 1	Precision	Recall
pred. 0	26,116	0	100.00%	90.69%
pred. 1	2,680	0	0.00%	0.00%

PerformanceVector				
PerformanceVector:				
accuracy: 89.16%				
ConfusionMatrix:				
True:	0	1		
0:	25675	0		
1:	3121	0		

ภาพที่ 4.12 แสดงคำบรรยายลักษณะงาน Naïve Bayes

ตารางที่ 4.2 ผลการวัดประสิทธิภาพของโมเดล Classification ของโมเดล Naïve Bayes

Accuracy = 89.16%

Model	True 0	True 1	Precision	Recall
pred. 0	25,675	0	100.00%	89.16%
pred. 1	3,121	0	0.00%	0.00%

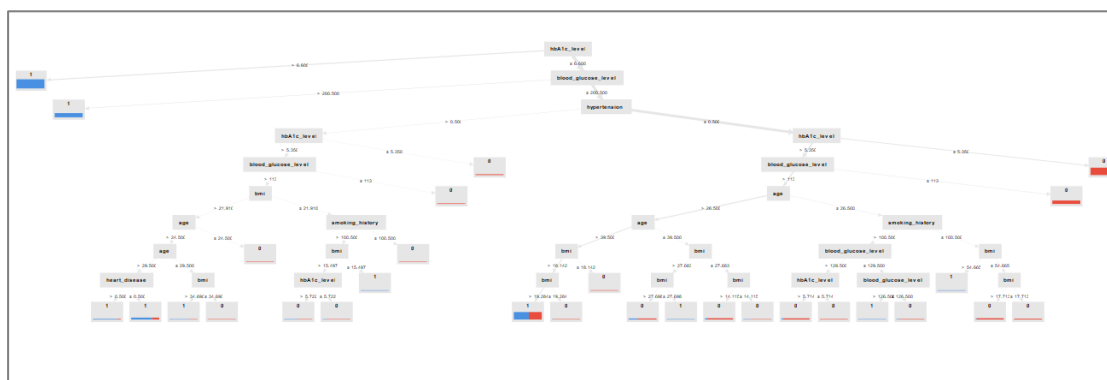
Tree

```

hbA1c_level > 6.600: 1 {1=28402, 0=0}
hbA1c_level ≤ 6.600
| blood_glucose_level > 200.500: 1 {1=12393, 0=0}
| blood_glucose_level ≤ 200.500
| | hypertension > 0.500
| | | hbA1c_level > 5.350
| | | | blood_glucose_level > 113
| | | | | bmi > 21.910
| | | | | | age > 24.500
| | | | | | | age > 29.500
| | | | | | | | heart_disease > 0.500: 1 {1=737, 0=128}
| | | | | | | | heart_disease ≤ 0.500: 1 {1=3635, 0=1239}
| | | | | | | | age ≤ 29.500
| | | | | | | | | bmi > 34.690: 1 {1=7, 0=3}
| | | | | | | | | | bmi ≤ 34.690: 0 {1=0, 0=10}
| | | | | | | | | | age ≤ 24.500: 0 {1=0, 0=10}
| | | | | | | | | | | bmi ≤ 21.910
| | | | | | | | | | | smoking_history > 100.500
| | | | | | | | | | | | bmi > 15.497
| | | | | | | | | | | | | hbA1c_level > 5.722: 0 {1=44, 0=47}
| | | | | | | | | | | | | hbA1c_level ≤ 5.722: 0 {1=3, 0=17}
| | | | | | | | | | | | | | bmi ≤ 15.497: 1 {1=4, 0=0}
| | | | | | | | | | | | | | smoking_history ≤ 100.500: 0 {1=0, 0=5}
| | | | | | | | | | | | | | | blood_glucose_level ≤ 113: 0 {1=0, 0=667}
| | | | | | | | | | | | | | | hbA1c_level ≤ 5.350: 0 {1=0, 0=1353}
| | | | | | | | | | | | | | | hypertension ≤ 0.500
| | | | | | | | | | | | | | | | hbA1c_level > 5.350
| | | | | | | | | | | | | | | | | blood_glucose_level > 113
| | | | | | | | | | | | | | | | | | age > 26.500
| | | | | | | | | | | | | | | | | | | age > 39.500
| | | | | | | | | | | | | | | | | | | | bmi > 18.142

```

ภาพที่ 4.13 แสดงคำบรรยายลักษณะงาน Random Forest



ภาพที่ 4.14 แสดงโมเดล Graph Random Forest

ตารางที่ 4.3 ผลการวัดประสิทธิภาพของโมเดล Classification ของโมเดล Random Forest

Accuracy = 86.76%

Model	True 0	True 1	Precision	Recall
pred. 0	24,982	0	100.00%	86.76%
pred. 1	3,814	0	0.00%	0.00%

```

PerformanceVector

PerformanceVector:
accuracy: 77.96%
ConfusionMatrix:
True:  1      0
1:    40958   0
0:    11577   0

```

ภาพที่ 4.15 แสดงค่าที่ใช้ในการประเมินผลลัพธ์ของโมเดล Support Vector Machine

ตารางที่ 4.4 ผลการวัดประสิทธิภาพของโมเดล Classification ของโมเดล SVM

Accuracy = 77.96%

Model	True 0	True 1	Precision	Recall
pred. 0	40,958	0	0.00%	0.00%
pred. 1	11,577	0	77.96%	100.00%

```

PerformanceVector

PerformanceVector:
accuracy: 89.97%
ConfusionMatrix:
True:  0      1
0:    25908   0
1:    2888    0

```

ภาพที่ 4.16 แสดงค่าที่ใช้ในการประเมินผลลัพธ์ของโมเดล Artificial Neural Network

ตารางที่ 4.5 ผลการวัดประสิทธิภาพของโมเดล Classification ของโมเดล ANN

Accuracy = 89.97%

Model	True 0	True 1	Precision	Recall
pred. 0	25,908	0	100.00%	89.97%
pred. 1	2,888	0	0.00%	0.00%

สรุปผลการประเมินประสิทธิภาพของทั้ง 5 โมเดล:

1. Decision Tree:
 - Accuracy: 90.69%
 - Precision: 100.00% (True 0), 0.00% (True 1)
 - Recall: 90.69% (True 0), 0.00% (True 1)
2. Naïve Bayes:
 - Accuracy: 89.16%
 - Precision: 100.00% (True 0), 0.00% (True 1)
 - Recall: 89.16% (True 0), 0.00% (True 1)
3. Random Forest:
 - Accuracy: 86.76%
 - Precision: 100.00% (True 0), 0.00% (True 1)
 - Recall: 86.76% (True 0), 0.00% (True 1)
4. Support Vector Machine (SVM):
 - Accuracy: 77.96%
 - Precision: 0.00% (True 0), 77.96% (True 1)
 - Recall: 100.00% (True 0), 87.5% (True 1)
5. Artificial Neural Network (ANN):
 - Accuracy: 89.97%
 - Precision: 100.00% (True 0), 0.00% (True 1)
 - Recall: 89.97% (True 0), 0.00% (True 1)

ผลลัพธ์ดังกล่าวแสดงถึงประสิทธิภาพของแต่ละโมเดล โดยมีค่า Precision และ Recall สำหรับทั้งสองคลาส (True 0 และ True 1) ซึ่ง True 0 หมายถึง "ไม่เป็นเบาหวาน" และ True 1 หมายถึง "เป็นเบาหวาน"

6) การนำโมเดลไปใช้งานจริง (Deployment) นำผลลัพธ์ที่ได้จากการวิเคราะห์ข้อมูล ได้ทำการนำเสนอข้อมูลแบบ Visualization ด้วยการแสดงแผนภาพสรุปข้อมูล และเผยแพร่ ข้อมูลสารสนเทศนี้บนเว็บเบราว์เซอร์โดยการใช้ภาษา HTML, CSS และ PHP ร่วมกับการ นำเสนอข้อมูลแบบ Visualization ด้วยการแสดงผลข้อมูลในรูปแบบของภาพโดยใช้โปรแกรม Power BI ซึ่งทางผู้วิเคราะห์ข้อมูลได้ยกตัวอย่างการจัดทำเป็นรูปแบบของรายงาน (Report) หรือแผนภาพ (Dashboard) ในการพัฒนาหน้าเว็บไซต์สำหรับการเปิดเผย และเผยแพร่ข้อมูลการวิเคราะห์ของชุดข้อมูลโรคเบาหวาน เพื่อให้ผู้ใช้งานทราบถึงเกี่ยวกับการวิเคราะห์โรคเบาหวานที่ได้ผลสรุปที่ถูกต้อง สะดวกรวดเร็วและเข้าใจง่ายต่อเวลาที่ต้องการศึกษาเพิ่มเติม แล้วจึงทำการวิเคราะห์ข้อมูลและออกแบบเว็บไซต์ โดยเว็บไซต์มี 7 ส่วนประกอบต่างๆ ดังนี้ 1) ส่วนของหน้าแรกที่รวบรวมเมนูต่างๆไว้ 2) ส่วนของที่มาของการวิเคราะห์ข้อมูล 3) ส่วนของการแสดงผล Dashboard ผ่าน Power Bi 4) ส่วนของการวิเคราะห์ข้อมูลด้วย Decision Tree 5) ส่วนที่แสดงความรู้ในด้านความรู้โรคเบาหวาน 6) ส่วนของการทดสอบแบบประเมินสุขภาพจิต 7) ส่วนของผู้จัดทำ

4.2.1.2 ขอบเขตผู้ใช้งานทั่วไปบนเว็บเบราว์เซอร์

- 1) สามารถดูข้อมูลความรู้และบทความเกี่ยวกับโรคเบาหวาน
- 2) สามารถดูแดชบอร์ดข้อมูลสารสนเทศของการโรคเบาหวาน
- 3) สามารถนำข้อมูลบนแดชบอร์ดไปใช้ในการสรุปผลข้อมูลเบื้องต้นได้
- 4) สามารถดาวน์โหลดชุดข้อมูลที่ถ่วงน้ำหนักแล้วไปใช้ต่อได้

4.3 บทสรุป

จากผลการดำเนินงานโครงการทั้งหมด ผู้วิเคราะห์ข้อมูลได้นำผลการวิเคราะห์เชิงสถิติ และการแสดงข้อมูลในรูปแบบ Visualization ด้วยโปรแกรม Power BI มาจัดทำเป็นระบบสารสนเทศ เพื่อเผยแพร่ให้บุคคลภายนอกหรือกลุ่มผู้ใช้ข้อมูลสามารถเข้าถึงและทำความเข้าใจลักษณะของผู้ป่วยโรคเบาหวานจากชุดข้อมูลทุกมิติแบบเปิดได้อย่างชัดเจน โดยมีการจำแนกผลการทำนายตาม ปัจจัยเสี่ยงที่เกี่ยวข้อง ได้แก่ เพศ อายุ ภาวะความดันโลหิตสูง ประวัติโรคหัวใจ และประวัติการสูบบุหรี่ พร้อมจัดเรียงลำดับจากค่าความเสี่ยงสูงสุดไปยังต่ำสุด

นอกจากนี้ ผลงานดังกล่าวยังทำหน้าที่เป็นแหล่งองค์ความรู้ให้แก่บุคคลทั่วไปและกลุ่มผู้ใช้ข้อมูล เพื่อสร้างความเข้าใจเกี่ยวกับ ปัจจัยเสี่ยงและปัจจัยป้องกัน การเกิดโรคเบาหวาน อีกทั้งยังอธิบาย วิธีการและขั้นตอนในการสร้างโมเดล Decision Tree สำหรับการทำนายลักษณะผู้ป่วยโรคเบาหวานอย่างเป็นระบบ ซึ่งสามารถใช้เป็นแนวทางให้แก่ผู้สนใจศึกษาและค้นคว้าเพิ่มเติม ตลอดจนใช้เป็นต้นแบบในการพัฒนาโปรแกรมประมวลผลที่ให้ผลลัพธ์สอดคล้องกับโมเดลที่สร้างขึ้น เพื่อเป็นเครื่องมือช่วยแนะนำและสนับสนุนการวางแผนจัดการความเสี่ยงของผู้ป่วยเบาหวานได้อย่างมีประสิทธิภาพ และสามารถนำข้อมูลไปใช้ประโยชน์ในการดำเนินชีวิตประจำวันได้อย่างเหมาะสม