

# บทที่ 1

## บทนำ

โครงการเรื่อง การวิเคราะห์ข้อมูลโรคหัวใจ สำหรับเผยแพร่ข้อมูลผ่านเว็บไซต์ ในบทนี้ จะเป็นการกล่าวถึงความสำคัญ และที่มาของปัญหา วัตถุประสงค์ ประโยชน์ที่จะได้รับ ขอบเขตงานของคณะผู้จัดทำและผู้ใช้งานทั่วไปบนเว็บไซต์ เครื่องมือที่ใช้ในการพัฒนาเว็บไซต์ สถานที่ใช้ในการดำเนินการศึกษาและรวบรวมข้อมูล และระยะเวลาในการดำเนินการ รวมถึง นิยามศัพท์เฉพาะของโครงการ ดังนี้

### 1.1 ความสำคัญและที่มาของปัญหา

ปัจจุบันโรคหัวใจ นับเป็นโรคที่พบบ่อยในลำดับต้นๆ และเป็นสาเหตุการเสียชีวิตอันดับแรกของประเทศทั่วโลกในปี ค.ศ.1990 มีการทำนายว่าแนวโน้มจะเพิ่มเป็น 25 ล้านคนในปี ค.ศ. 2020 จากข้อมูลของกระทรวงสาธารณสุขในช่วงปี พ.ศ. 2555-2558 อัตราผู้ป่วยโรคหัวใจในประเทศไทยมีแนวโน้มเพิ่มสูงขึ้นอย่างต่อเนื่อง โดยในปี พ.ศ. 2557 มีผู้เสียชีวิตจากโรคหัวใจและหลอดเลือดจำนวน 58,681 คน หรือโดยเฉลี่ยถึงชั่วโมงละ 7 คน และยังมีแนวโน้มเพิ่มมากขึ้นเรื่อย ๆ หากไม่รีบรวมการเสียชีวิตจากอุบัติเหตุ อัตราการเสียชีวิตด้วยโรคหัวใจนับเป็นสาเหตุการเสียชีวิตเป็นอันดับสอง รองจากการเสียชีวิตจากโรคมะเร็ง และการป่วยเป็นโรคหัวใจ ส่งผลให้มีการเปลี่ยนแปลงแบบแผนการดำเนินชีวิต เนื่องจากเป็นโรคที่สัมพันธ์กับสาเหตุของความพิการและ การไร้ความสามารถ โดยเฉพาะความสมบูรณ์ในการทำงาน ดังนั้น แม้ว่าผู้ป่วยโรคหัวใจในปัจจุบันจะมีอัตราการรอดชีวิตเพิ่มสูงขึ้นจากความก้าวหน้าด้านเทคโนโลยีทางการแพทย์ แต่การป่วยด้วยโรคหัวใจก็ทำให้ผู้ป่วยไม่สามารถกลับไปอยู่ในสภาวะปกติได้ ผู้ป่วยจึงมีความทุกข์ทรมานทางร่างกาย จิตใจ สังคม ที่ต้องมารับการรักษาพยาบาลต่อเนื่องเป็นระยะเวลานาน ซึ่งสาเหตุส่วนใหญ่ร้อยละ 80-90 เกิดจากการรับประทานผักผลไม้ไม่เพียงพอ การออกกำลังกายไม่เพียงพอ ภาวะไขมันผิดปกติ โรคความดันโลหิตสูงโรคเบาหวาน การดื่มเครื่องดื่มแอลกอฮอล์มาก การสูบบุหรี่และความเครียด รวมถึงพันธุกรรม เพศและวัย ตลอดจนจนประวัติครอบครัว และปัจจุบันเทคโนโลยีดิจิทัลยังมีบทบาทในการดำรงชีวิตประจำวันของคนเราเป็นอย่างมาก หน่วยงานต่าง ๆ ไม่ว่าจะเป็นภาคธุรกิจองค์กรต่าง ๆ รวมถึงหน่วยงานภาครัฐเองต่างก็มีการนำเทคโนโลยีต่าง ๆ เข้ามาช่วยเพื่อ

ปรับปรุงประสิทธิภาพการดำเนินงาน และยกระดับการให้บริการไปสู่ประชาชนให้สามารถเข้าถึงบริการต่าง ๆ ได้สะดวก และรวดเร็วยิ่งขึ้น จึงทำให้เกิดข้อมูลสารสนเทศขึ้นอย่างมากมายมหาศาล โดยเฉพาะอย่างยิ่งข้อมูลของภาครัฐที่จัดเก็บในรูปแบบของฐานข้อมูลนับเป็นเวลาหลายสิบปี และกำลังเพิ่มขึ้นอีกในปัจจุบันและอนาคต ข้อมูลเหล่านี้สามารถนำมาใช้ประโยชน์ได้ทั้งสำหรับภาครัฐ ภาคเอกชน และประชาชนซึ่งทำให้เกิดเป็นแนวคิด วิธีการรูปแบบทางธุรกิจรวมถึงวิถีชีวิตใหม่ ๆ การแบ่งปันข้อมูลเพื่อนำมาใช้ประโยชน์

การวิเคราะห์ข้อมูลโรคหัวใจ คือการนำข้อมูลที่ได้ถูกรวบรวมไว้และถูกเผยแพร่ผ่านเว็บไซต์ Kaggle ซึ่งเป็นเว็บไซต์ที่เก็บรวบรวมชุด ข้อมูลต่าง ๆ ซึ่งเป็นข้อมูลที่เปิดเผยได้ เพื่อให้ผู้ใช้บริการสามารถนำชุดข้อมูลไปศึกษา ซึ่งการนำข้อมูลโรคหัวใจ มาทำการวิเคราะห์นั้น เนื่องจากต้องการทราบถึงข้อมูลในเชิงลึกของโรคหัวใจทั้ง เพศ อากาศเจ็บหน้าอก ความดันโลหิต ไขมันในเลือด น้ำตาลในเลือด ผลคลื่นไฟฟ้าอัตราการเต้นของหัวใจสูงสุด ความชันคลื่นหัวใจ รอยโรคในหลอดเลือดหัวใจ การรักษาคอนไชน์ ซึ่งการวิเคราะห์ปัจจัยต่าง ๆ จากข้อมูลที่ได้มานั้นถือเป็นเรื่องสำคัญต่อการศึกษาหาผลสรุป แต่ด้วยจำนวนข้อมูลที่มีปริมาณ มากนั้นทำให้ยากต่อการนำเสนอให้เข้าใจได้ง่าย ซึ่งถือเป็นข้อจำกัดหากต้องการนำข้อมูลชุดนี้ไปใช้ เพื่อเผยแพร่ให้แก่บุคคลภายนอกหรือกลุ่มผู้ใช้ข้อมูลได้รับรู้

จากปัญหาข้างต้นดังนั้นผู้วิเคราะห์ข้อมูลจึงได้นำชุดข้อมูลที่ได้มานั้นเข้าสู่กระบวนการจำแนก ข้อมูล จัดการ แก่ไขข้อมูล เพื่อให้ได้ข้อมูลที่เหมาะสมตามความต้องการสำหรับการวิเคราะห์ข้อมูล และวิเคราะห์ข้อมูลด้วย CRISP-DM จากเทคนิคการทำเหมืองข้อมูลแบบ Classification ในรูปแบบของแผนภูมิต้นไม้ หรือที่เรียกว่า Decision tree เป็นโครงสร้างที่ใช้แสดงกฎที่ได้จากเทคนิคการจำแนกประเภทข้อมูล จะช่วยให้เกิดความเข้าใจต่อข้อมูลและสิ่งที่ต้องการทราบอย่างแท้จริง และทางผู้วิเคราะห์ข้อมูลยังได้ทำการนำเสนอข้อมูลแบบ Visualization ด้วยการแสดงผลข้อมูลในรูปแบบของภาพ ด้วยโปรแกรม Tableau Public และเผยแพร่ข้อมูลสารสนเทศนี้บน Web browser โดยการใช้ภาษา HTML ในการพัฒนาหน้าเว็บไซต์ สำหรับการ เปิดเผยข้อมูลการวิเคราะห์ข้อมูลโรคหัวใจ เพื่อให้ผู้ใช้งานทราบถึงเกี่ยวกับการวิเคราะห์โรคหัวใจ ได้ผลสรุปที่ ถูกต้อง สะดวก รวดเร็ว เข้าใจง่าย ทันต่อเวลาที่ต้องการศึกษา

## 1.2 วัตถุประสงค์

- 1.2.1 เพื่อทำการวิเคราะห์ข้อมูลโรคหัวใจ
- 1.2.2 เพื่อเผยแพร่ผลการวิเคราะห์ข้อมูลโรคหัวใจ บนเว็บไซต์

## 1.3 ประโยชน์ที่จะได้รับ

- 1.3.1 ได้ผลสรุปของการวิเคราะห์ข้อมูลโรคหัวใจ
- 1.3.2 ได้เผยแพร่ผลการวิเคราะห์ข้อมูลโรคหัวใจ บนเว็บไซต์

## 1.4 ขอบเขต

### 1.4.1 ขอบเขตผู้วิเคราะห์ข้อมูล

1.4.1.1 ผู้วิเคราะห์ข้อมูลได้รวบรวมข้อมูลโรคหัวใจที่ได้จากเว็บไซต์ Kaggle เพื่อนำมาวิเคราะห์ข้อมูลที่ได้นำมาวิเคราะห์ ได้แก่ อาการเจ็บหน้าอก ความดันโลหิต ไขมันในเลือด น้ำตาลในเลือด อัตราการเต้นของหัวใจสูงสุด ความชันคลื่นหัวใจ รอยโรคในหลอดเลือดหัวใจ และการรักษาคนไข้

1.4.1.2 ผู้วิเคราะห์ข้อมูลวิเคราะห์ข้อมูลด้วยเทคนิคทางตาต้า ไม่นิ่ง เพื่อใช้ทำนายแนวโน้มการเกิดขึ้นของปัจจัยที่ก่อให้เกิดโรคหัวใจ และใช้เทคนิคการจำแนกประเภทข้อมูล ด้วยการสร้างโมเดล Decision Tree โดยการใช้โมเดล Decision Trees หากมีการตัดสินใจ

1.4.1.3 ผู้วิเคราะห์ข้อมูลได้ทำการประมวลผลโดยใช้โปรแกรม Rapid Miner

1.4.1.4 ผู้วิเคราะห์ข้อมูลได้นำเสนอข้อมูลแบบ Visualization ด้วยการแสดงผลข้อมูลในรูปแบบของภาพ ด้วยโปรแกรม Tableau Public

1.4.1.5 ผู้วิเคราะห์ข้อมูลได้ทำการนำเสนอเผยแพร่ข้อมูลสารสนเทศนี้บน Web browser ซึ่งเว็บไซต์ผู้วิเคราะห์สามารถอัปเดตข้อมูลใหม่เข้าไปได้และสามารถ ลบ แก้ไข ตาวนั้โหลดข้อมูลออกมาได้ โดยการใช้ภาษา HTML และชุดคำสั่ง CSS3

### 1.4.2 ขอบเขตผู้ใช้งานทั่วไปบนเว็บเบราว์เซอร์

1.4.2.1 สามารถดูข้อมูลเกี่ยวกับโรคหัวใจ ที่เผยแพร่ผ่านทางเว็บไซต์

1.4.2.2 สามารถดูข้อมูลการวิเคราะห์ข้อมูลโรคหัวใจ ในลักษณะรูปแบบของแผนภูมิชนิดต่าง ๆ จาก โปรแกรม Tableau Public ได้

1.4.2.3 สามารถกรองข้อมูลการวิเคราะห์ข้อมูลโรคหัวใจ ผ่านระบบของ Tableau Public บน Web browser ได้

1.4.2.4 สามารถทำแบบสอบถามประเมินความเสี่ยงของโรคหัวใจได้

1.4.2.5 สามารถดาวน์โหลดข้อมูลที่เกี่ยวข้องกับข้อมูลโรคหัวใจไปศึกษาต่อได้

## 1.5 เครื่องมือที่ใช้ในการพัฒนาโปรแกรม

### 1.5.1 Hardware

#### 1.5.1.1 เครื่องคอมพิวเตอร์โน้ตบุ๊ก

- หน่วยประมวลผลกลาง Intel(R) Core™i5-8300H @ 2.30GHz
- หน่วยประมวลผลกราฟิก NVIDIA® GeForce® GTX 1050TI 6 GB
- หน่วยความจำหลัก RAM 16 GB
- หน่วยความจำ HardDisk 1 TB

#### 1.5.1.2 เครื่องคอมพิวเตอร์โน้ตบุ๊ก

- หน่วยประมวลผลกลาง Intel(R) Core™i5-8250U @ 1.60GHz
- หน่วยประมวลผลกราฟิก NVIDIA® GeForce® GTX 1050TI 6 GB
- หน่วยความจำหลัก RAM 16 GB
- หน่วยความจำ HardDisk 1 TB

### 1.5.2 Software

1.5.2.1 โปรแกรม Tableau Public ใช้ในการแสดงผลข้อมูลในรูปแบบของภาพ

1.5.2.2 ชุดคำสั่งภาษา HTML4, HTML5 (Hyper Text Markup Language)

1.5.2.3 ชุดคำสั่ง CSS3 (Cascading Style Sheets)

1.5.2.4 โปรแกรม Adobe XD ใช้สำหรับการออกแบบหน้าเว็บไซต์

1.5.2.5 โปรแกรม Dreamweaver CS6 ใช้สำหรับในการทำเว็บไซต์

1.5.2.6 โปรแกรม Rapid Miner ใช้ในการสร้างโมเดล Decision tree

1.5.2.7 ระบบปฏิบัติการ : Window 10 ใช้ในการเปิดซอฟต์แวร์ต่าง ๆ

1.5.2.8 Microsoft office Word 2016 ใช้ในการทำเอกสารต่าง ๆ

1.5.2.9 Microsoft office Excel 2016 ใช้ในการ Clean Data กับชุดข้อมูล

## 1.6 สถานที่ใช้ในการดำเนินการศึกษาและรวบรวมข้อมูล

### 1.6.1 สถานที่ใช้ในการดำเนินการศึกษา

มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนาภาคพายัพเชียงใหม่ 128 ถนนห้วยแก้ว ตำบลช้างเผือก อำเภอเมือง จังหวัดเชียงใหม่ 50300

### 1.6.2 แหล่งรวบรวมข้อมูล

เว็บไซต์ Kaggle หรือ <https://www.kaggle.com/datasets>

## 1.7 ระยะเวลาในการดำเนินการ

ตารางที่ 1.1 ระยะเวลาในการดำเนินการ

แผนการดำเนินการ	2563				
	พ.ค.	มิ.ย.	ก.ค.	ส.ค.	ก.ย.
1. ศึกษาและกำหนดความต้องการ	←→				
2. ตรวจสอบความถูกต้องข้อมูล		←→			
3. วิเคราะห์ข้อมูลจำแนกตามกลุ่ม		←→			
4. ออกแบบระบบ			←→		
5. เขียนและทดสอบระบบ			←→	←→	
6. ตรวจสอบระบบโดยรวม				←→	
7. ประเมินการใช้งานระบบ				←→	←→
8. จัดทำเอกสารประกอบ โครงการ	←→	←→	←→	←→	←→

## 1.8 นิยามศัพท์เฉพาะ

1.8.1 โรคหัวใจ (Heart Disease) หมายถึง โรคต่าง ๆ ที่ส่งผลกระทบต่อการทำงานของหัวใจโดยโรคหัวใจสามารถแบ่งย่อยได้เป็นหลายกลุ่มโรค เช่น โรคหลอดเลือดหัวใจ โรคกล้ามเนื้อหัวใจ โรคหัวใจเต้นผิดจังหวะ โรคลิ้นหัวใจ โรคหัวใจพิการแต่กำเนิด และโรคติดเชื้อบริเวณหัวใจ เป็นต้น

1.8.2 การวิเคราะห์ข้อมูล หมายถึง เป็นการจัดระเบียบแยกแยะส่วนต่าง ๆ ของหลักฐานหรือข้อมูลที่ได้ออกเป็นหมวดหมู่ เพื่อหาคำตอบตามความมุ่งหมาย และตามสมมติฐานที่ได้กำหนดไว้การวิเคราะห์ข้อมูลนี้เป็นขั้นตอนการทำงานที่ต่อเนื่องมาจากการวัดการนับ และจัดเรียงลำดับข้อมูลส่วนใหญ่ เป็นเรื่องเกี่ยวกับการนำเอาวิธีการทางสถิติมาวิเคราะห์หาค่าตัวแปรหรือหาลักษณะของตัวแปร

1.8.3 เว็บไซต์ (Website) หมายถึง หน้าเว็บเพจที่จัดทำขึ้น เพื่อนำเสนอข้อมูลต่าง ๆ ผ่านทางคอมพิวเตอร์และอินเทอร์เน็ต โดยจะมีหน้าเว็บเพจหลาย ๆ หน้า que เชื่อมโยงเข้ากับไฮเปอร์ลิงค์ เพื่อให้สามารถเปิดไปยังหน้าเพจต่าง ๆ ได้อย่างง่ายดายและถูกจัดเก็บไว้ใน www.(เวิลด์ไวด์เว็บ) โดยเว็บไซต์ส่วนใหญ่นั้นก็มีทั้งเว็บไซต์ที่เปิดให้เข้าชมได้ฟรี และเว็บไซต์ที่ต้องสมัครสมาชิกและเสียค่าบริการ จึงจะเข้าใช้งานเว็บได้ ซึ่งข้อมูลในเว็บก็จะมีหลากหลาย

รูปแบบ ขึ้นอยู่กับความต้องการนำเสนอของเจ้าของเว็บไซต์ การเรียกดูเว็บไซต์จะเรียกดูผ่านทางซอฟต์แวร์ ในลักษณะของเบราว์เซอร์

1.8.4 อาการเจ็บหน้าอก (Chest pain) หมายถึง อาการเจ็บปวดที่เกิดขึ้นบริเวณหน้าอก ซึ่งเกิดขึ้นตั้งแต่บริเวณไหล่ลงมาถึงช่วงล่างของซี่โครง สาเหตุของอาการเจ็บหน้าอกมีหลายอย่าง อาจเป็นผลมาจากการป่วยด้วยโรคที่ร้ายแรง เช่น โรคหัวใจ เป็นต้น

1.8.5 ความดันโลหิต (Hypertension) หมายถึง แรงดันของกระแสเลือดที่กระทบต่อผนังหลอดเลือดแดง ซึ่งเกิดจากการสูบฉีดของหัวใจ (คล้ายแรงลมที่ดันผนังของยางรถเมื่อเราสูบลมเข้า) สามารถวัดได้โดยการใช้เครื่องวัดความดัน (Sphygmomanometer)

1.8.6 ไขมันในเลือด (Dyslipidemia I) หมายถึง ภาวะที่มีระดับคอเลสเตอรอลรวม (Total Cholesterol, Chol) หรือมีระดับแอลดีแอลคอเลสเตอรอลสูง (Low Density Lipoprotein Cholesterol, LDL-C) และ / หรือมีระดับเฮชดีแอลคอเลสเตอรอลต่ำ (High Density Lipoprotein, HDL-C) และ / หรือมีระดับไตรกลีเซอไรด์สูง (Triglyceride, TG) ซึ่งเป็นหนึ่งในปัจจัยเสี่ยงของการเกิดโรคหัวใจ และหลอดเลือด รวมไปถึงพฤติกรรมการใช้ชีวิตประจำวัน และปัจจัยเสี่ยงที่สัมพันธ์กับการเกิดโรคหัวใจ และหลอดเลือด

1.8.7 น้ำตาลในเลือด (Hyperglycemia) หมายถึง ความเข้มข้นของน้ำตาลในเลือด หรือ ระดับกลูโคสในเลือด ที่มีอยู่ในเลือดของมนุษย์หรือสัตว์ โดยปกติในสัตว์เลี้ยงลูกด้วยนม ร่างกายจะควบคุมระดับกลูโคสในเลือดให้อยู่ที่ 3.6 - 5.8 mM (mmol/L, เช่น millimoles/liter) หรือ 64.8 - 104.4 mg/dL โดยธรรมชาติแล้วร่างกายมนุษย์จะควบคุมระดับน้ำตาลในเลือดอย่างเข้มงวด ซึ่งเป็นส่วนหนึ่งของการรักษาสมดุลของร่างกาย (homeostasis)

1.8.8 อัตราการเต้นของหัวใจสูงสุด (Maximum Heart Rate) หมายถึง การทำงานของหัวใจในการสูบฉีดเลือด เพื่อนำออกซิเจนและสารอาหารไปหล่อเลี้ยงทั่วร่างกาย กระบวนการนี้ หรือระบบไหลเวียนเลือดจะเริ่มทำงานในทุก ๆ ครั้งที่หัวใจมีการบีบตัว

1.8.9 คลื่นไฟฟ้าหัวใจ (Electrocardiogram) หมายถึง การทดสอบสัญญาณไฟฟ้าของหัวใจในแต่ละจังหวะการเต้นของหัวใจ ซึ่งจะถูกลบออกและส่งผ่านไปทั่วทั้งหัวใจ ส่งผลให้กล้ามเนื้อหัวใจมีการบีบตัวอย่างสมบูรณ์ในการส่งเลือดไปเลี้ยงส่วนต่าง ๆ ของร่างกาย

1.8.10 การฉีดสีดูเส้นเลือด (Coronary angiogram) หมายถึง การสอดสายตรวจขนาดเล็ก ผ่านเข้าทางหลอดเลือดแดงที่ข้อมือ หรือขาหนีบ จนกระทั่งปลายสายไปถึงหลอดเลือดหัวใจ แล้วฉีดสารทึบรังสีเพื่อดูช่องทางเดินของหลอดเลือดเลี้ยงกล้ามเนื้อหัวใจ พร้อมกับใช้เอกซเรย์บันทึกภาพของหลอดเลือดหัวใจแต่ละเส้นไว้

## 1.9 บทสรุป

จากบทนำที่ได้กล่าวมาในข้างต้นทั้งหมดนั้น ผู้วิเคราะห์ข้อมูลได้สังเกตเห็นความสำคัญของการนำเทคโนโลยีมาประยุกต์ใช้ในการวิเคราะห์ข้อมูลโรคหัวใจ ด้วยกระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM จากเทคนิคการทำเหมืองข้อมูลแบบ Classification ในรูปแบบของแผนภูมิต้นไม้ หรือที่เรียกว่า Decision tree และเผยแพร่ข้อมูลสารสนเทศบน Web browser โดยใช้ภาษา HTML ในการพัฒนาหน้าเว็บไซต์สำหรับการเปิดเผยข้อมูลจากเว็บไซต์ Kaggle เพื่อให้ผู้ใช้งานทราบถึงเกี่ยวกับการวิเคราะห์ข้อมูลโรคหัวใจ ได้ผลสรุปที่ถูกต้อง สะดวก รวดเร็ว เข้าใจง่าย ทันต่อเวลาที่ต้องการศึกษา