### เทคโนโลยีเกี่ยวข้อง

#### การทำเหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูล หรืออาจเรียกว่า การค้นหาความรู้ในฐานข้อมูล (Knowledge Discovery in Database: KDD) กระบวนการที่กระทำกับข้อมูลจำนวนมาก เพื่อค้นหารูปแบบ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น โดยอาศัยหลักการทางสถิติ การรู้จำ การเรียนรู้ของเครื่องจักร และหลักคณิตศาสตร์ ซึ่งความรู้ที่ได้จากการทำเหมืองข้อมูลนั้นมีหลากหลายรูปแบบ ได้แก่

* **กฎความสัมพันธ์ (Association Rule)**

แสดงความสัมพันธ์ของเหตุการณ์หรือวัตถุ ที่เกิดขึ้นพร้อมกัน ตัวอย่างของการประยุกต์ใช้กฎเชื่อมโยง เช่น การวิเคราะห์ข้อมูลการขายสินค้า โดยเก็บข้อมูลจากระบบ ณ จุดขาย (POS) หรือร้านค้าออนไลน์ แล้วพิจารณาสินค้าที่ผู้ซื้อมักจะซื้อพร้อมกัน เช่น ถ้าพบว่าคนที่ซื้อเทปวิดีโอมักจะซื้อเทปกาวด้วย ร้านค้าก็อาจจะจัดร้านให้สินค้าสองอย่างอยู่ใกล้กัน เพื่อเพิ่มยอดขาย หรืออาจจะพบว่าหลังจากคนซื้อหนังสือ ก แล้ว มักจะซื้อหนังสือ ข ด้วย ก็สามารถนำความรู้นี้ไปแนะนำผู้ที่กำลังจะซื้อหนังสือ ก ได้

* **การจำแนกประเภทข้อมูล (Data Classification)**

หากฎเพื่อระบุประเภทของวัตถุจากคุณสมบัติของวัตถุ เช่น หาความสัมพันธ์ระหว่างผลการตรวจร่างกายต่าง ๆ กับการเกิดโรค โดยใช้ข้อมูลผู้ป่วยและการวินิจฉัยของแพทย์ที่เก็บไว้ เพื่อนำมาช่วยวินิจฉัยโรคของผู้ป่วย หรือการวิจัยทางการแพทย์ ในทางธุรกิจจะใช้เพื่อดูคุณสมบัติของผู้ที่จะก่อหนี้ดีหรือหนี้เสีย เพื่อประกอบการพิจารณาการอนุมัติเงินกู้

* **การแบ่งกลุ่มข้อมูล (Data Clustering)**

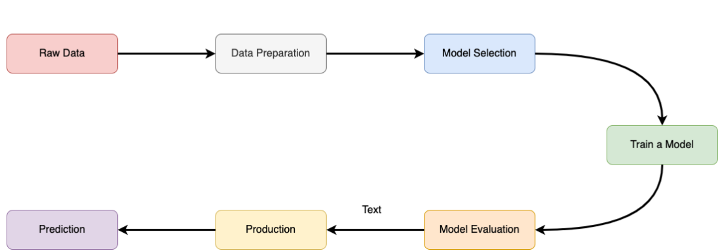
แบ่งข้อมูลที่มีลักษณะคล้ายกันออกเป็นกลุ่ม แบ่งกลุ่มผู้ป่วยที่เป็นโรคเดียวกันตามลักษณะอาการ เพื่อนำไปใช้ประโยชน์ในการวิเคราะห์หาสาเหตุของโรค โดยพิจารณาจากผู้ป่วยที่มีอาการคล้ายคลึงกัน

* **การสร้างมโนภาพ (Visualization)**

สร้างภาพคอมพิวเตอร์กราฟิกที่สามารถนำเสนอข้อมูลมากมายอย่างครบถ้วนแทนการใช้ข้อความนำเสนอข้อมูลที่มากมาย เราอาจพบข้อมูลที่ซ้อนเร้นเมื่อดูข้อมูลชุดนั้นด้วยจินตทัศน์ [14]

ทางผู้จัดทำมีการนำการทำเหมืองข้อมูล (Data Mining) มาประยุกต์ใช้ โดยจะมุ่งเน้นไปที่การทำเหมืองข้อมูลในด้านการจำแนกประเภทข้อมูล (Data Classification) เพื่อดำเนินการจำแนกข้อมูลที่เป็นข้อมูลที่สำคัญหรือข้อมูลส่วนบุคคล

#### การเรียนรู้ของเครื่อง (Machine Learning)



**รูปที่ 2.1** กระบวนการของการเรียนรู้ของเครื่อง

การเรียนรู้ของเครื่อง (Machine Learning) คือ ระบบที่สามารถเรียนรู้ได้จากชุดตัวอย่างข้อมูลด้วยตนเองโดยปราศจากการป้อนคำสั่งของผู้เขียนโปรแกรม ซึ่งระบบนี้ประกอบด้วยข้อมูลและเครื่องมือทางสถิติเพื่อทำนายผลลัพธ์ออกมา เพื่อนำไปใช้ต่อในทางธุรกิจหรือเป้าหมาย [6] โดยการทำการเรียนรู้ของเครื่อง (Machine Learning) จะเริ่มจากการวิเคราะห์รูปแบบของชุดตัวอย่างข้อมูล และทำการเรียนรู้และจดจำสิ่งต่าง ๆ ในชุดตัวอย่างข้อมูลนั้น ยิ่งมีข้อมูลมากเท่าไหร่ ตัวระบบก็จะสามารถเรียนรู้ได้ฉลาดขึ้น โดยเครื่องจักรจะเรียนรู้ผ่านการค้นพบรูปแบบหรือแบบแผนซ้ำ ๆ ส่งผลให้การทำนาย การพยากรณ์ มีความแม่นยำมากขึ้น [7]

เนื่องจากโครงงานนี้เป็นการปกป้องข้อมูลที่ระบุตัวบุคคล ซึ่งต้องมีการฝึกฝนแบบจำลองให้สามารถตรวจจับรูปแบบของข้อมูลส่วนบุคคลได้ จึงต้องมีการนำการเรียนรู้ของเครื่อง (Machine Learning) เข้ามาประยุกต์ใช้ในโครงงานนี้

#### การรู้จำเสียงพูด (Speech Recognition)

#### การประมวลผลภาษาธรรมชาติ (Natural Language Processing: NLP)

#### การระบุคำที่เป็นนิพจน์ระบุนาม (Named-Entity Recognition: NER)

#### spaCy