



COMPUTER SCIENCE

# RESEARCH AND DEVELOPMENT

of Sentiment Analysis System from Products Reviews.

LEARN MORE A black circular icon containing a white right-pointing triangle, representing a play button or video control.

# MEET OUR TEAM



นายณัฐวัฒน์ หมายบุญ  
643020045-6



นายอภิสิทธิ์ น้อยหวาน  
643021257-6

# ความเป็นมาและความสำคัญของปัญหา

ปัญหาที่เกี่ยวข้องกับการทราบความคิดเห็นและอารมณ์จากข้อความ (Sentiment Analysis) เป็นเรื่องที่มีความความสำคัญอย่างมากในการวิเคราะห์ข้อมูล และเข้าใจความรู้สึกของผู้ใช้หรือลูกค้าต่อสินค้าหรือบริการต่าง ๆ เนื่องจากปัญหา ความถ่วงของภาษาและภาษาไทยที่มีโครงสร้างที่แตกต่างจากภาษาอังกฤษ

การพัฒนาระบบที่สามารถวิเคราะห์ความคิดเห็นและอารมณ์จากข้อความเพื่อ ใช้ในการปรับปรุงสินค้าหรือบริการของธุรกิจเพื่อเพิ่มประสิทธิภาพในด้านการตลาด และบริการลูกค้าได้อย่างมากขึ้นในอนาคต

# วัตถุประสงค์

1. เพื่อสำรวจความพึงพอใจของบทวิจารณ์สินค้า และนำมาปรับปรุงสินค้าต่อไป
2. เพื่อวิเคราะห์ความคิดเห็นบนร้านค้าออนไลน์ และจำแนกความคิดเห็นเป็นประเภทต่าง ๆ
3. เพื่อศึกษาความแม่นยำของโมเดลในการท่านายความคิดเห็นและอารมณ์จากข้อความ

## เป้าหมายและขอบเขต

1. เนื่องจากความซับซ้อนของภาษาอาจทำให้เกิดข้อผิดพลาดในการวิเคราะห์ความรู้สึกหรือความคิดเห็นจากบทวิจารณ์สินค้าออนไลน์
2. ระบบฐานข้อมูลรองรับการเก็บชุดข้อมูลความรู้สึกหรือความคิดเห็นจากบทวิจารณ์สินค้าออนไลน์เป็นภาษาไทยและภาษาอังกฤษเท่านั้น โดยที่ข้อมูลจะอยู่ในช่วงปี 2020-2023
3. ระบบสามารถจำแนกความรู้สึกหรือความคิดเห็นจากบทวิจารณ์สินค้าจากแพลตฟอร์มร้านค้าออนไลน์ได้อย่างแม่นยำ

# ประโยชน์ที่คาดว่าจะได้รับ

1. เข้าใจความรู้สึกและความคิดเห็นของคนและช่วยให้เข้าใจความต้องการของลูกค้า  
ความคิดเห็นของคนทั่วไป หรือกระแสสังคม
2. นำไปใช้ในการปรับปรุงสินค้าและบริการให้ตรงกับความต้องการของลูกค้ามากขึ้น
3. พัฒนากลยุทธ์การตลาด เพื่อนำไปใช้ในการวางแผนการตลาดและการสื่อสารให้มี  
ประสิทธิภาพมากขึ้น

# ທາງສູງທີ່ເກີຍວ່ອນ

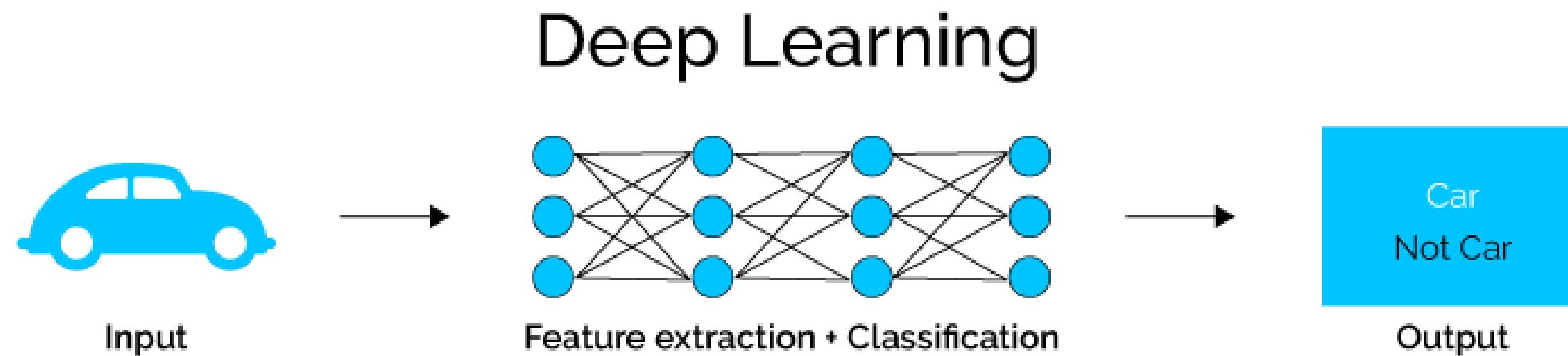
LEARN MORE



# ทฤษฎีที่เกี่ยวข้อง

## DEEP LEARNING

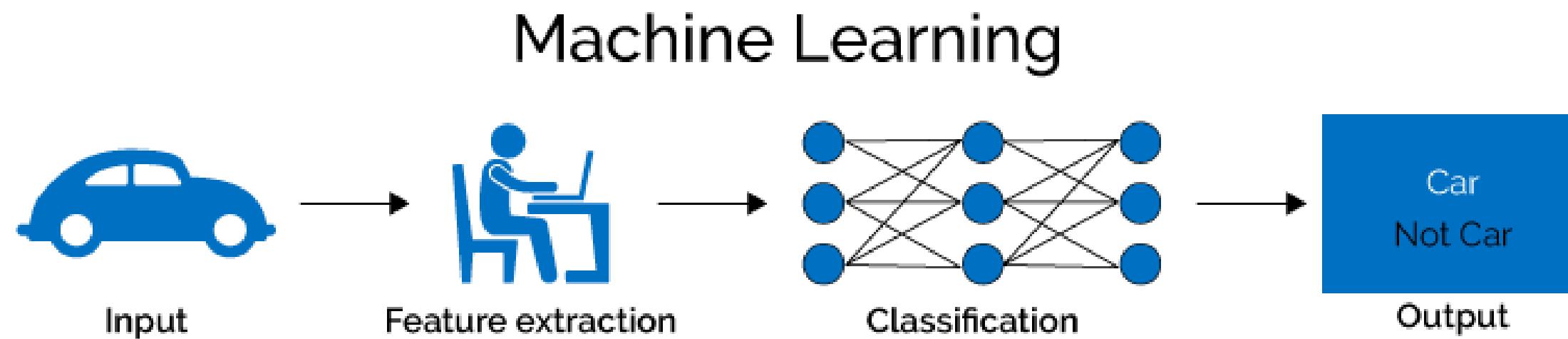
คือวิธีการเรียนรู้แบบอัตโนมัติด้วยการ เลียนแบบการทำงานของโครงข่ายประสาทของมนุษย์ (Neurons) โดยนำระบบโครงข่ายประสาท (Neural Network) มาซ้อนกัน หลายชั้น (Layer) และทำการเรียนรู้ข้อมูลตัวอย่าง



# ทฤษฎีที่เกี่ยวข้อง

## MACHINE LEARNING

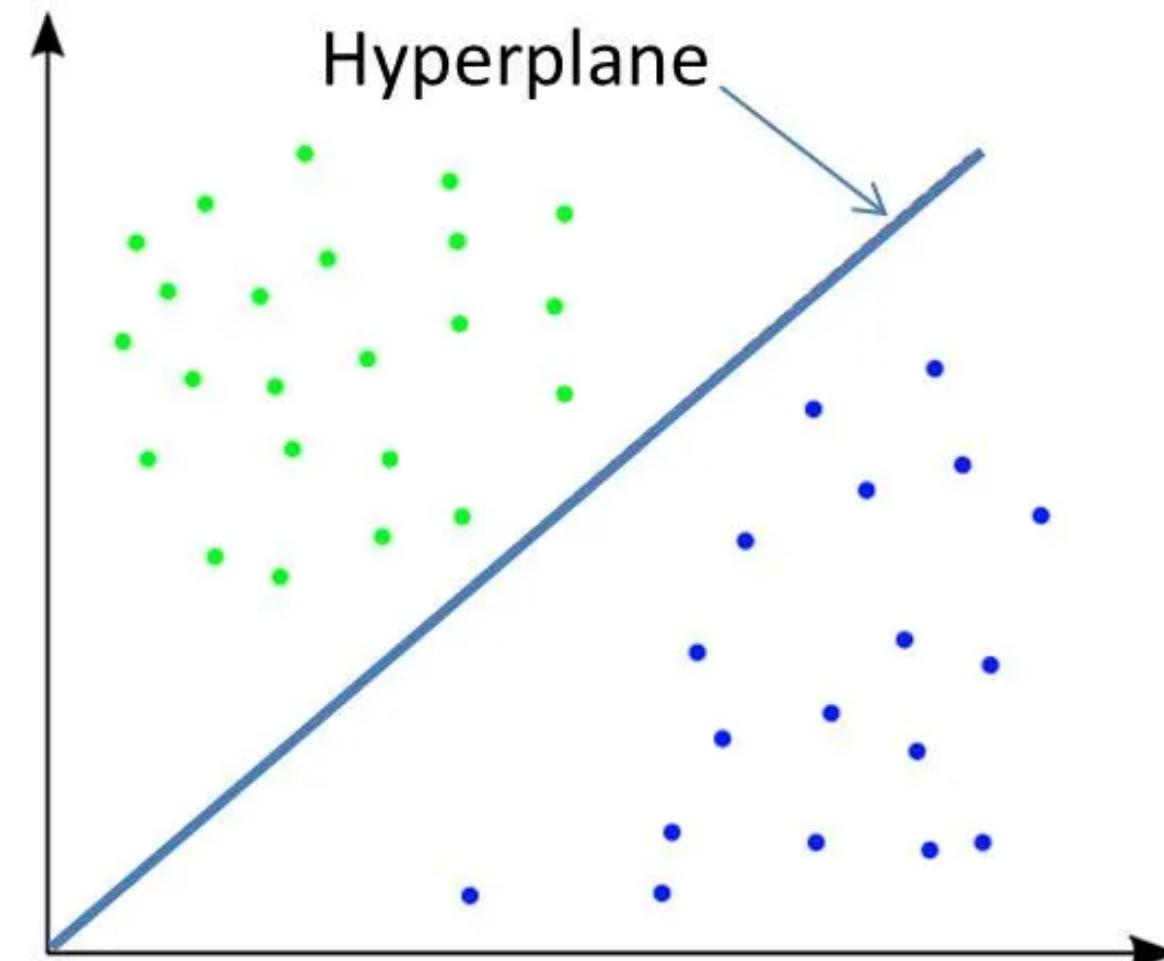
คือ “การทำให้ระบบคอมพิวเตอร์สามารถเรียนรู้ได้ด้วยตนเองโดยใช้ข้อมูล” จุดประสงค์คือ เพื่อใช้ในการทำงานให้มีประสิทธิภาพมากกว่ามนุษย์ในงานบางประเภท โดยการทำให้ ฉลาดขึ้น สามารถพัฒนา และเรียนรู้ได้ด้วยตนเอง



# ทฤษฎีเกี่ยวข้อง

## SUPPORT VECTOR MACHINE

เป็นอัลกอริทึมที่สามารถนำมาช่วยแก้ปัญหาการจำแนกข้อมูล ที่ใช้ในการวิเคราะห์ข้อมูล และจำแนกข้อมูล โดยอาศัยหลักการของการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งแยกกลุ่มข้อมูลที่ถูกป้อนเข้าสู่กระบวนการสอนให้ระบบเรียนรู้ โดยเน้นไปยังเส้นแบ่งแยกกลุ่มข้อมูลให้ได้ดีที่สุด

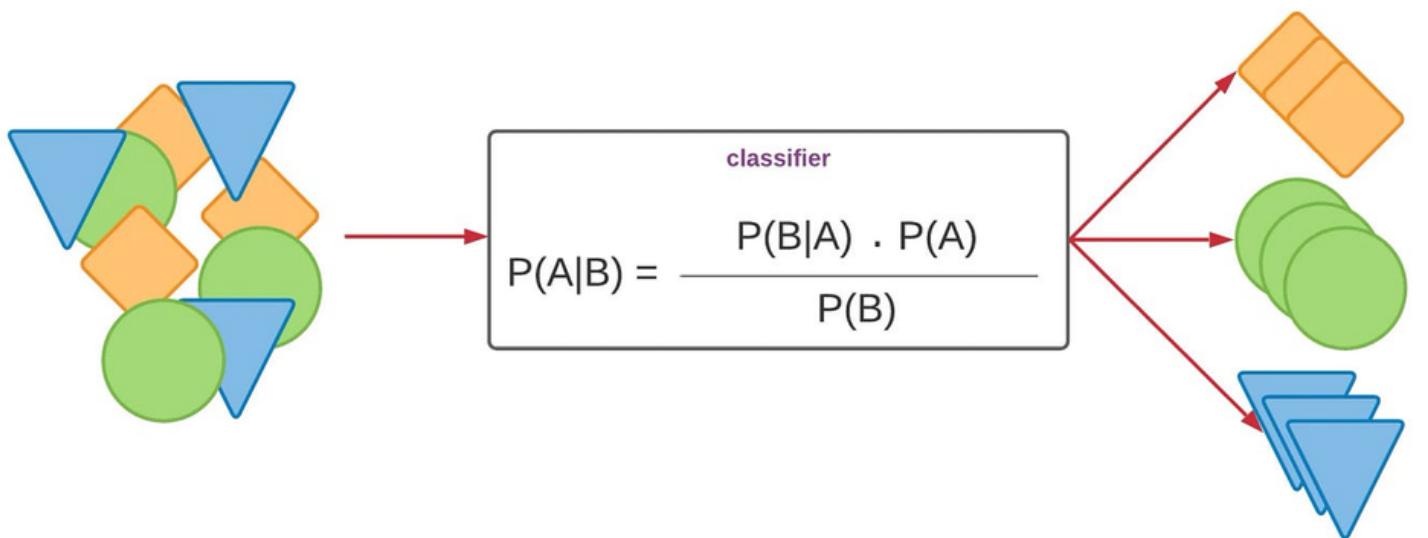


# ทฤษฎีที่เกี่ยวข้อง

## NAÏVE BAYES

ใช้ในการแบ่งกลุ่มหรือหาเหตุการณ์ที่จะเกิดขึ้นโดยการอิงทฤษฎีความน่าจะเป็นของ Bayes ซึ่ง Target ของโมเดลจะมีความคล้ายคลึงกับ Logistic Regression ว่าจะเกิดเหตุการณ์นั้น หรือไม่โดยจะเพิ่มโอกาสในการเกิดเหตุการณ์เข้าไปด้วย โดยมักจะใช้ในการวิเคราะห์ข้อมูลที่มีความต่อเนื่องของเหตุการณ์ (Dependent Event)

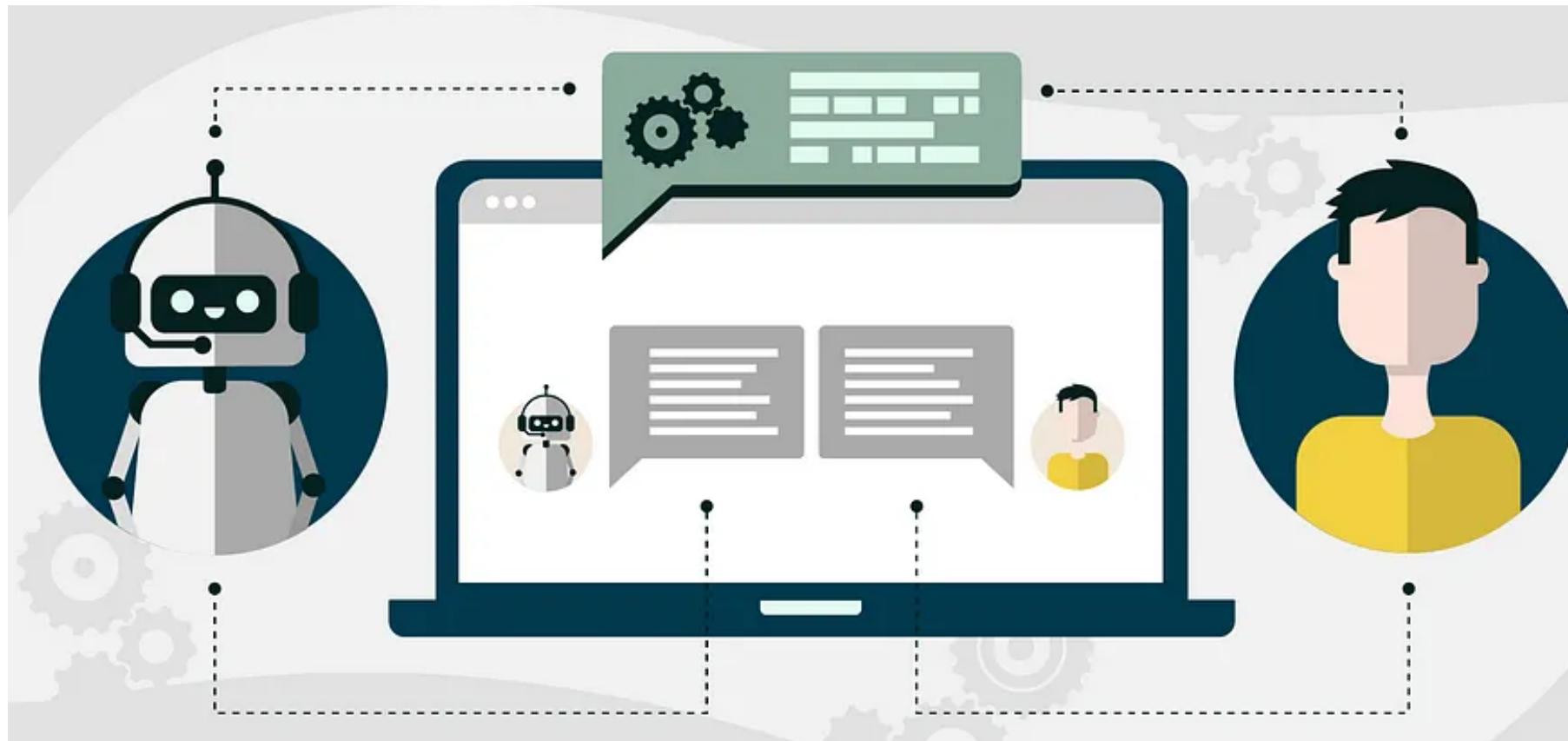
### Naive Bayes Classifier



# ทฤษฎีที่เกี่ยวข้อง

## NATURAL LANGUAGE PROCESSING

สาขางานปัญญาประดิษฐ์ที่มุ่งเน้นการประมวลผลและเข้าใจภาษาธรรมชาติของมนุษย์ โดยใช้คอมพิวเตอร์และเทคโนโลยี การพัฒนา NLP มีวัตถุประสงค์เพื่อทำให้เครื่องจักร สามารถเข้าใจและแปลงข้อมูลภาษามนุษย์



# ทฤษฎีเกี่ยวข้อง

## TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY

เป็นเทคนิคที่ใช้ในการประมวลผลข้อความเพื่อจัดลำดับความสำคัญของคำศัพท์ในเอกสาร โดยคำนวณความถี่ของคำในเอกสาร (TF) และความหนาแน่นของคำในคลังข้อมูล (IDF) และนำมาคูณกัน เพื่อให้ได้ค่า TF-IDF

$$\text{tf}("The", \text{document1}) \rightarrow 2/8 = 0.25$$

$$\text{tf}("sky", \text{document1}) \rightarrow 2/8 = 0.25$$

$$\text{tf}("is", \text{document1}) \rightarrow 2/8 = 0.25$$

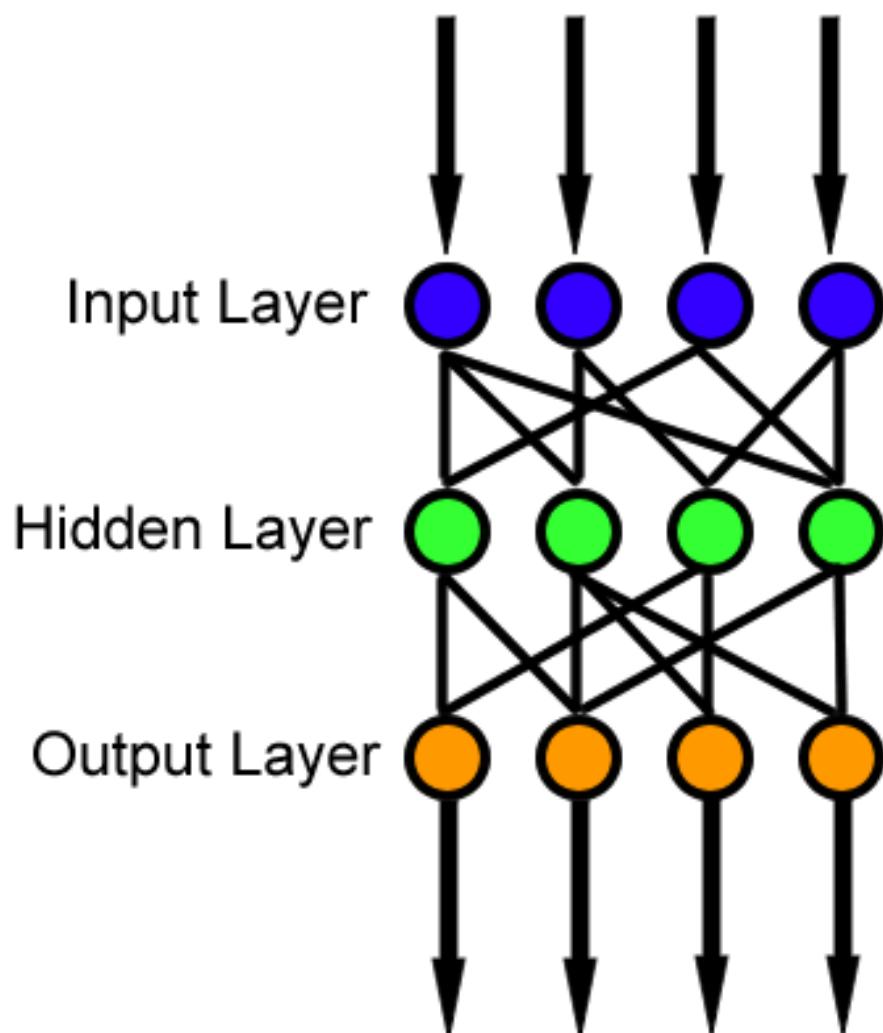
$$\text{tf}("blue", \text{document1}) \rightarrow 1/8 = 0.125$$

$$\text{tf}("beautiful", \text{document1}) \rightarrow 1/8 = 0.125$$

# ทฤษฎีที่เกี่ยวข้อง

## FEED FORWARD NEURAL NETWORK

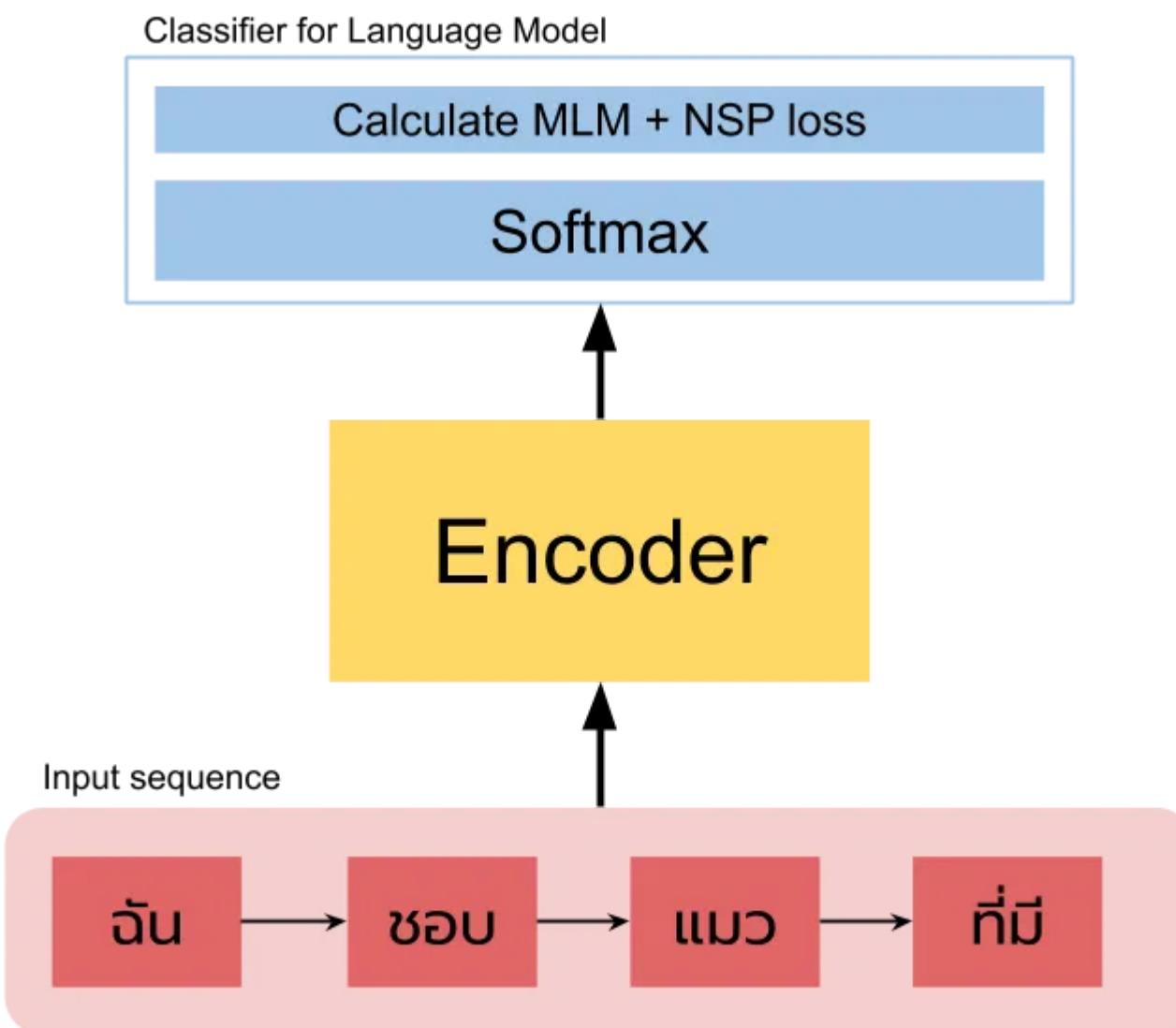
ถือเป็นโมเดลที่มีโครงสร้างที่เรียบง่ายที่สุด เพราะว่า การดำเนินการของข้อมูลจะเป็นไป ในทิศทางเดียว ก็คือ รับข้อมูลจาก input layer และส่งไปต่อไปยัง hidden layer เลื่อยๆ จน กระแทกถึง output layer ก็จะหยุด



# ทฤษฎีที่เกี่ยวข้อง

## BERT

เครื่องมือที่ทำให้สามารถเรียนรู้ภาษาของมนุษย์ Natural Language Processing (NLP) ได้ สามารถทำความเข้าใจภาษามนุษย์ได้ และเรียนรู้แบบ Deep Learning จะช่วยเพิ่มทักษะให้กับระบบ Search Engine ได้ดี



# งานวิจัยที่เกี่ยวข้อง

LEARN MORE



# งานวิจัยที่เกี่ยวข้อง

## RESEARCH ON SENTIMENT ANALYSIS MODEL OF ONLINE SHOPPING

นำเสนอเกี่ยวกับอัลกอริธึมสำหรับการวิเคราะห์ความรู้สึกของข้อความจากประเมินสินค้า โดย Dataset ได้มาจากแพลตฟอร์มช้อปปิ้งออนไลน์ โดยเฉพาะ JD, Sunning และ Taobao หลังจากนั้นก็ทำการทดสอบข้อมูล โดยใช้วิธีจำแนกแบบ Bayesian classification และได้ความแม่นของการวิเคราะห์อารมณ์คิดเป็นร้อยละ 98.003 (Sunning) และ 98.931 (Taobao) ตามลำดับ

# งานวิจัยที่เกี่ยวข้อง

## OVERALL AND FEATURE LEVEL SENTIMENT ANALYSIS OF AMAZON PRODUCT

นำเสนอการวิเคราะห์อารมณ์หรือการขุดเจาะความคิดเห็น (Sentiment Analysis หรือ Opinion mining) โดยรวบรวมข้อมูลจาก Amazon ทั้งหมด 142.8 ล้านข้อมูลหลังจากนั้นทำการทำความสะอาดข้อมูลเพื่อนำไปใช้ฝึกโมเดล Random Forest, Multionmial Naïve Bayes, Complement Naïve Bayes, Bernoulli Naïve Bayes จึงได้ความแม่นยำของการวิเคราะห์อารมณ์คิดเป็นร้อยละ 83.28 82.69 82.28 75.75 ตามลำดับ

# งานวิจัยที่เกี่ยวข้อง

## ENTIMENT ANALYSIS OF MOBILE PHONE PRODUCTS REVIEWS

วิเคราะห์ความรู้สึกและความคิดเห็นจากข้อมูล Kaggle จำนวน 82815 ข้อมูล และแบ่งเป็น Positive, Neutral, และ Negative ซึ่งใช้ในการฝึกกับโมเดล Naïve Bayes, SVM, Maximum Entropy, Decision Tree, และ KNN เพื่อคาดการณ์แนวโน้มและความประทับใจของข้อมูล โดยจะแยกเป็นเชิงบวกหรือเชิงลบ ให้เห็นถึงความรู้สึกของผู้ใช้ต่อเนื้อหาต่าง ๆ ในข้อมูลที่วิเคราะห์ได้

# งานวิจัยที่เกี่ยวข้อง

## **TWITTER-BASED SENTIMENT ANALYSIS FOR INDONESIAN DRUG PRODUCTS**

ข้อมูลจาก Twitter เกี่ยวกับผลิตภัณฑ์ยาของอินโดนีเซียถูกจำแนกเป็น Positive หรือ Negative โดยอัตโนมัติโดยใช้ Key-Word ในการค้นหาคำ เมื่อทำการฝึกกับโมเดล Naïve Bayes, Support Vector Machine, Logistic Regression, Decision Tree, และ Random Forest ผลลัพธ์ในการจำแนกมีค่าเป็นร้อยละ 0.87, 0.89, 0.85, 0.92, และ 0.90 ตามลำดับ

# งานวิจัยที่เกี่ยวข้อง

## **SENTIMENT ANALYSIS ALGORITHM BASED ON BERT AND CNN**

นำเสนอวิธีการวิเคราะห์ความรู้สึกที่รวมกันระหว่าง BERT และ Convolutional Neural Network (CNN) เพื่อการดึงและจำแนกคุณลักษณะของข้อมูล การใช้โมเดลนี้ช่วยให้ได้ผลลัพธ์ที่ดีที่สุดในการแยกแยะคุณลักษณะ โดยมีค่า F1 ที่ 0.903 ซึ่งสูงกว่าโมเดลอื่นๆ เช่น word2vec-svm (0.812), word2vec-cnn (0.850), และ word2vec-Att-cnn (0.872)

# งานวิจัยที่เกี่ยวข้อง

## SENTIMENT ANALYSIS USING SEMEVAL AND AMAZON DATASETS

นำเสนอวิธีการปรับปรุงประสิทธิภาพการวิเคราะห์ความรู้สึกตามแง่มุม Aspect-based sentiment analysis บนชุดข้อมูลขนาดเล็กโดยการเปรียบเทียบ เครื่อข่ายประสาทที่เกิดซ้ำ (Recurrent Neural Network) และพื้ดไปข้างหน้า (Feed Forward Neural Network) การทดลองแสดงให้เห็นว่าการใช้เวกเตอร์ร่วมกับข้อความที่มีป้ายกำกับจะช่วยเพิ่มความแม่นยำในการทำนายความรู้สึกตามแง่มุมสำหรับทวีจารณ์

# วิธีการดำเนินการวิจัย

LEARN MORE



# วิธีการดำเนินการวิจัย

## 1.1 เก็บรวบรวมข้อมูลเนื้อหาจากแหล่งข้อมูลต่าง ๆ และทำ SYNTHESIS MATRIC

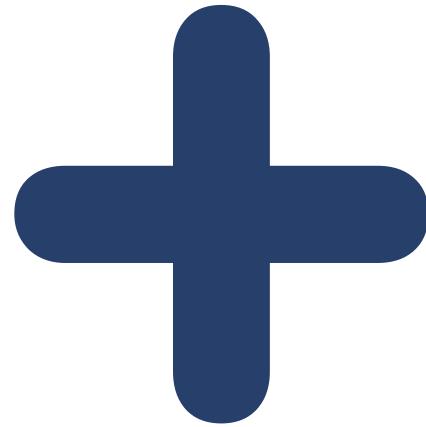
Author & Date	Propose	Method	Sample	Result/Finding	Weakness	Strong
P. Yuvraj, S. K. Dilip, (2023)	ประเมินความรู้สึกจากบทวิจารณ์	โดยใช้โมเดล PAM, AbSA, RNN และ CNN ในการจัดประเภท	Dataset จากบทวิจารณ์ของ Amazon มากกว่า 14,000 ความคิดเห็น	วิจารณ์ คิดเป็นร้อยละ 90.6 91.7 92.32 93.52 94.91 ตามลำดับ	จำกัดในการปรับเข้ากับสถานการณ์และความแม่นยำทางภาษา	เข้าใจความรู้สึกของผู้บริโภคได้ตรงประการ
X. Yifan, R. Yong, (2023)	การวิเคราะห์ความรู้สึกของข้อความจากประมีนสินค้าไทยกันที่	โดยใช้อัลกอริทึมที่จะแก้ไขปัญหาความคลาสสิกกันระหว่างการประเมินและการให้คะแนน	Dataset จากร้านค้าออนไลน์ JD, Sunning และ Taobao	คิดเป็นร้อยละ 98.003 (Sunning) และ 98.931 (Taobao) ตามลำดับ	แหล่งข้อมูลที่ให้มาไม่ระบุโดยชัดเจนถืออ่อนช่องในความคลาสสิกของความรู้สึกที่ถูกนำเสนอในงานวิจัยนี้	ไม่เคลื่อนไหวที่ความรู้สึกที่นำเสนอ ด้วยเหตุผลนี้ จึงเป็นระบบที่ง่ายและมีประสิทธิภาพในเคสที่สามารถทำนายความรู้สึกของจริง ผลิตภัณฑ์ที่ขายออนไลน์ได้ดี มีความแม่นยำสูงถึง 99.17% ขึ้นไป.
S. N. Shailendra, S. Twinkle, (2020)	สำหรับการวิเคราะห์ความรู้ของวิจารณ์ผลิตภัณฑ์	โดยใช้อัลกอริทึม Random Forest การจำแนก	Dataset จากเว็บไซต์ e-commerce	ผลิตภัณฑ์ที่นักวิจัยได้ทดสอบคิดเห็นที่มีความแม่นยำ	การให้คุณภาพความคลาสสิกไม่ทุกครั้งสอดคล้องกับรูปที่เขียนโดยอุตสาหะ ทำให้ยากต่อการวัดความรู้สึกที่แม่นยำตามคะแนนเท่านั้น	ระบบให้ผลลัพธ์ในรูปแบบบัญชีที่บอกผลลัพธ์ที่ต้องการในแต่ละด้าน ความแม่นยำในการอ่านรีวิวทั้งหมดเพื่อวิเคราะห์ผลิตภัณฑ์
V. R. Welgamage, et al., (2020)	การวิเคราะห์อารมณ์หรือการอุตสาหะความคิดเห็น	โดยใช้โมเดลฝึกหัด Random Forest, Multinomial Naive Bayes, Complement Naive Bayes, Bernoulli Naive Bayes	Dataset จากเว็บไซต์ Amazon 142.8 ล้านความคิดเห็น	คิดเป็นร้อยละ 83.28 82.69 82.28 75.75 ตามลำดับ	อาจใช้เวลาในการฝึกคุณพิเศษในการให้ความเข้าใจ	การวิเคราะห์ความรู้สึกสามารถจัดประเภทไว้ก่อนแล้ว แต่ต้องมีความแม่นยำในการจัดระเบียบข้อมูลบริษัทมาก
M. A. M. Salem, A. Y. A. Maghari, (2022)	การวิเคราะห์ความรู้สึกหรือความคิดเห็นและการอุตสาหะความบูรณาการที่แสดงในข้อมูลที่วิเคราะห์และจัดประเภทเป็นเชิงบทหรือเชิงลบ	โดยทำการแบ่ง Positive และ Negative ของ Dataset เพื่อนำไปฝึกหัดโมเดล Naive Bayes, SVM, Maximum Entropy, Decision Tree, KNN Naive Bayes, SVM, Maximum Entropy, Decision Tree, KNN	Dataset จาก kaggle จำนวน 82815 ข้อมูล	คิดเป็นร้อยละ 82.1 80.7 82.7 73.6 76.8 ตามลำดับ	ความพยายามในการหาเว็บไซต์ที่เกี่ยวข้องและสรุปข้อมูล	สามารถจัดการความรู้สึกที่ถูกแสดงในสื่อสังคมออนไลน์และสาขาอื่น ๆ ได้อย่างแม่นยำ
K. Purwandari, et al., (2022)	การจำแนกประเภทความรู้สึกของผู้คนที่พูดถึงในโซเชียลมีเดีย Twitter โดยอัตโนมัติเพื่อระบุความรู้สึกหรือความคิดเห็นเชิงบทหรือเชิงลบ	ทำการจำแนกและอัดข้อมูลและทำการแยก Positive และ Negative เพื่อนำไปฝึกหัดโมเดล Naive Bayes, Support Vector Machine, Logistic Regression, Decision Tree, Random Forest	Dataset จาก Twitter	จำแนกเป็นร้อยละ 0.87 0.89 0.85 0.92 0.90 ตามลำดับ	ไม่ครอบคลุมในตัวเลือกใน (Decision Tree) และ Random Forest ที่ใช้สำหรับการจำแนกความรู้สึกแสดงประสมต้องการให้ข้อมูลที่มีค่าเทียบกับความคิดเห็นและความเชื่อมของผู้บริโภคต่อผลิตภัณฑ์ที่พูดถึง	การวิเคราะห์ความรู้สึกห้องนอนด้วยเครื่องมือที่แสดงความรู้สึกแรงในตัวพิเศษ ซึ่งสามารถให้ข้อมูลที่มีค่าเทียบกับความคิดเห็นและความเชื่อมของผู้บริโภคต่อผลิตภัณฑ์ที่พูดถึง

# วิธีการดำเนินการวิจัย

## 1.1 เก็บรวบรวมข้อมูลเนื้อหาจากแหล่งข้อมูลต่าง ๆ และทำ SYNTHESIS MATRIC

K. Purwandari, et al., (2022)	การจำแนกประบاهความรู้สึกของผู้ดูถูกที่เขียนข้ออันในเนื้อหาเพื่อจัดประบاهชั้นอุปถัมภ์ Twitter โดยอัตโนมัติ เพื่อระบุความรู้สึกหรือความคิดเห็นเชิงบวกหรือเชิงลบ	ทำการจำแนกความสะอาดข้อมูล และทำการแยก Positive และ Negative เพื่อนำไปฝึก กับโมเดล Naive Bayes, Support Vector Machine, Logistic Regression, Decision Tree, Random Forest	Dataset จาก Twitter	จำแนกเป็นร้อยละ 0.87 0.89 0.85 0.92 0.90 ตามลำดับ	โมเดลต้นไม้ตัดสินใจ (Decision Tree) และ Random Forest ที่สำหรับ การจำแนกความรู้สึกและประศิทธิภาพหากำกว่าในการทำนายความรู้สึกของ ความเชื่อของผู้ใช้โดย ต่อผลก็ต้องพิจารณา	การวิเคราะห์ความรู้สึก ที่งบนดังนักเน้นที่วิธีที่แสดง ความรู้สึกแรงในทิศทางเดียว , ซึ่งสามารถให้ข้อมูลที่มีค่า เกี่ยวกับความคิดเห็นและ ความเชื่อของผู้ใช้โดย ต่อผลก็ต้องพิจารณา
S. A. Rahin, T. Hasib, M. Hassan(2022)	วิศวกรรมคือของการวิจัยเพื่อ ให้ปรับปรุงประสิทธิภาพ ของการวิเคราะห์ความรู้สึก ความเชื่อใน (ABSA) บันทึก ข้อมูลขนาดเล็กโดยการ ประยุกต์ Recurrent network และ Feed forward neural	เปลี่ยนเพิ่บไกรช่วย ประสิทธิภาพ Recurrent และ แบบ Feed forward สำหรับ การวิเคราะห์ความรู้สึกความ เชื่อใน (ABSA) และพบว่า โครงสร้างประสิทธิภาพเพิ่มแบบ ที่ดีไปข้างหน้าทำางานได้ดี กว่า ทว่าเชื่อมข้อมูลอินพุต ที่มุ่งเน้นที่สร้างขึ้นโดยใช้ เครื่องมือ NLP ต่างๆ ที่ พร้อมใช้งาน เช่น แท็กตัว ของคำทุกๆ (POS) และการ พิมพ์พิมพ์	Dataset Amazon Review Data และ Laptop train Data	recurrent network สำหรับ Precision Aspect count 79.7%, Aspects 54.32%, Polarities 59.8% และ Feed-forward network สำหรับ Precision Aspect count 94.84%, Aspects 69.61 %, Polarities 64.17%	จำนวนบทวิจารณ์ที่สำหรับ ตัวแยกประบاهการฝึกมี จำนวนน้อยมาก และมีค่า overfitting เยอะเกินไป	การวิเคราะห์ความรู้สึกความ เชื่อใน (ABSA) ผ่านวิธีที่จะ ทำความเข้าใจความคิดเห็น ของผู้คนในค้านได้ด้านหนึ่ง โดยอุปหะ ให้ไปใช้ข้อมูลเชิง ลึกที่มีคุณค่าเกี่ยวกับบท วิจารณ์ของลูกค้าและค่าติดตาม โครงสร้างประสิทธิภาพเพิ่ม รวมถึงโครงสร้างประสิทธิภาพเพิ่ม แบบ feed forward และ recurrent ถูก บันทึกใช้เพื่อจัดการกับงาน ABSA เพื่อปรับปรุงประสิทธิภาพ ของการช่วยประสิทธิภาพ ของโครงสร้างประสิทธิภาพ ใน ABSA
R. Man, K. Lin (2021)	การวิเคราะห์ความรู้สึกเพื่อ ทดสอบความคิดเห็น สาธารณะ	โดยใช้แบบจำลอง BERT สำหรับการถึงข้อมูล คุณลักษณะ และ CNN สำหรับการจำแนกประบاه	Dataset จาก Weibo	ติดเป็นร้อยละ 0.903	ประสิทธิภาพของอัลกอริทึม อาจแตกต่างกันไปขึ้นอยู่กับ ข้อมูลและบริบทเฉพาะ เนื่องจากผลการทดลองขึ้นอยู่กับชนิดข้อมูลสาธารณะ	เข้าใจถูกமนจะอุทิราของ ข้อมูลได้ดีเมื่อขึ้น อัลกอริทึมได้รับความแม่นยำ ถูก ค่า recall และค่า F1 โดย ค่า F1 สูงถึง 0.903

# วิธีการดำเนินการวิจัย



1.3 ทำการกำหนด Label ของ  
คำจากข้อมูลที่คัดแยกไว้เพื่อ<sup>+</sup>  
นำเข้าโมเดลสำหรับทดสอบ

1.5 รวบรวมผลการทดสอบ  
โมเดล เพื่อใช้ในการพัฒนา

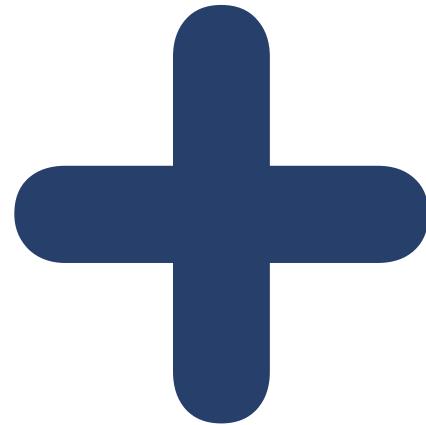


1.2 จัดเตรียมข้อมูลจากบทวิจารณ์สินค้าจาก  
แพลตฟอร์มร้านค้าออนไลน์ Lazada และ  
Amazon เพื่อไปคัดกรองว่าเป็น Positive,  
Neutrally หรือ Negative โดยที่ Positive  
จะเก็บที่ 4 – 5 ดาว ส่วน Neutrally จะเก็บที่  
3 ดาว และ Negative จะเก็บที่ 1 – 2 ดาว  
ก่อนนำไปทดสอบโมเดล

1.4 สรุปผลลัพธ์ที่ได้หลัง  
จากการทดสอบโมเดล

1.6 พัฒนาโมเดลในการวิเคราะห์  
ความรู้สึกและความคิดเห็น

# วิธีการดำเนินการวิจัย



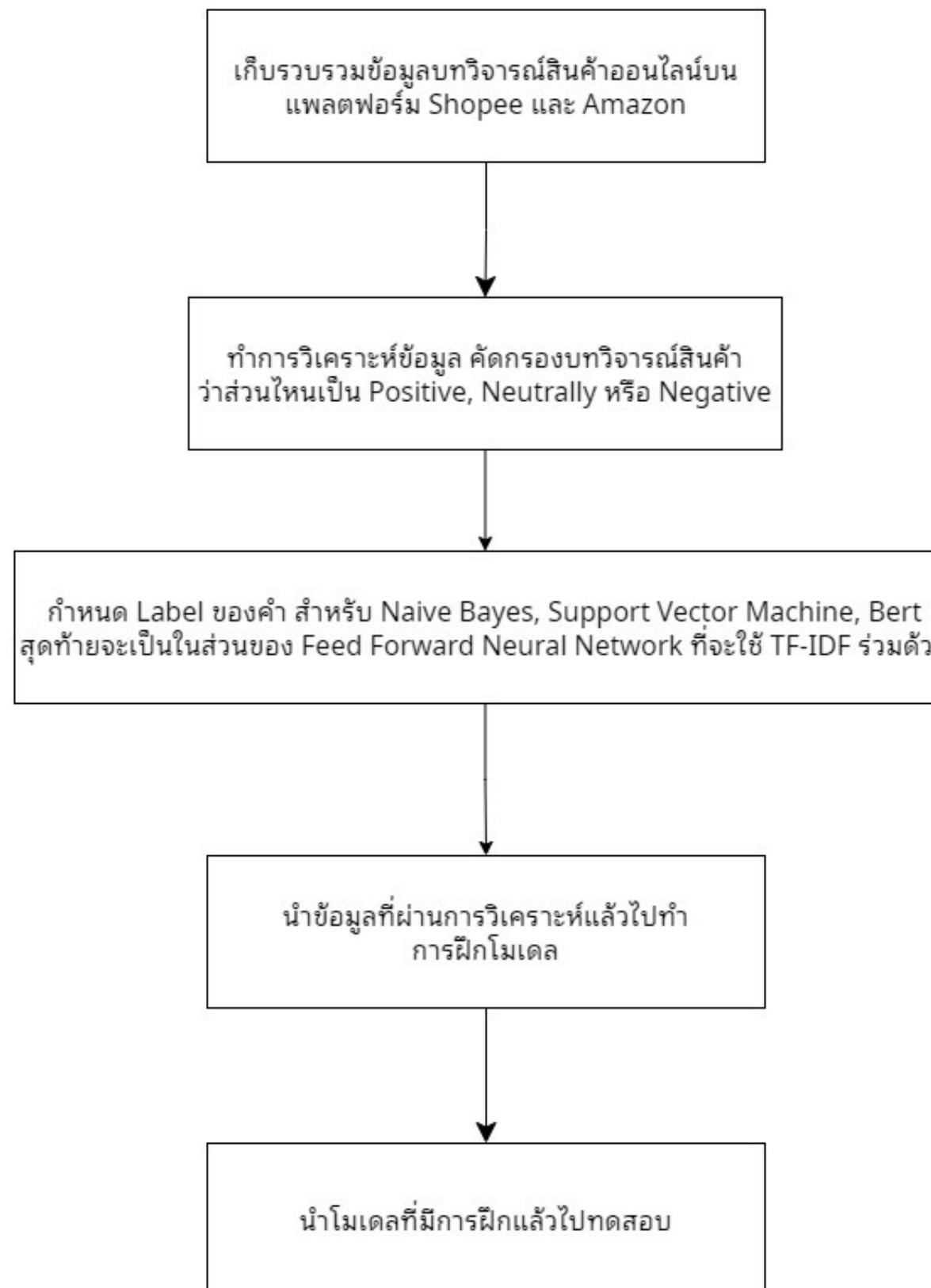
1.7 ทดสอบโมเดลและประเมินผลลัพธ์ที่ได้

1.9 นำโมเดลไปใช้งาน



1.8 วิเคราะห์โมเดลและปรับปรุงโมเดลให้มีประสิทธิภาพ

# วิธีการดำเนินการวิจัย



งานวิจัยนี้แบ่งขั้นตอนในการออกแบบโมเดลโดยการใช้

- **FEED FORWARD NEURAL NETWORK**
- **NAIVE BAYES**
- **SUPPORT VECTOR MACHINE**
- **BERT**

# วิธีการดำเนินการวิจัย

## ข้อมูลบทวิจารณ์สินค้าที่นำเข้า

ข้อมูลที่นำเข้าเป็นภาษาไทยรวมทั้งหมด 4,500 ความคิดเห็นและการแยกเป็น 3 ส่วนคือ 1,500 ความคิดเห็นและจะต้องสามารถทำการแยกความคิดเห็นนั้นได้ว่าเป็น POSITIVE, NEUTRALLY หรือ NEGATIVE โดยที่ POSITIVE จะเก็บที่ 4 – 5 ดาว ส่วน NEUTRALLY จะเก็บที่ 3 ดาว และ NEGATIVE จะเก็บที่ 1 – 2 ดาว

## การคัดกรองบทวิจารณ์สินค้า

ทำการแยก POSITIVE, NEUTRALLY หรือ NEGATIVE ของบทวิจารณ์สินค้า หลังจากนั้นทำการลดข้อมูลด้วยการลบอิโมจิและคำชำต่างๆ

## กำหนดค่าตัวเลขให้กับคำ

ในการทำกำหนดค่าตัวเลขให้กับคำจะใช้ LABEL ในส่วนของ NAÏVE BAYES, SUPPORT VECTOR MACHINE, BERT, สุดท้ายจะเป็นในส่วนของ FEED FORWARD NEURAL NETWORK ที่จะใช้ TF-IDF ร่วมด้วย

## การฝึกโมเดล

ฝึกโมเดลด้วย NAÏVE BAYES, SUPPORT VECTOR MACHINE, BERT บุলะ FEED FORWARD NEURAL NETWORK โดยจะนำข้อมูลที่ทำการคัดกรองแล้วไปใส่ในโมเดล

# ข้อจำกัดของระบบปัญหา อุปสรรค และแนวทางการแก้ไข

LEARN MORE 

# ข้อจำกัดของระบบปัญหาอุปสรรค และแนวทางการแก้ไข

## ข้อจำกัดของระบบ

1. จำแนกความคิดเห็นของ Positive, Neutrally หรือ Negative ได้ทั้งหมด 3 ประเภท
2. สามารถจำแนกความคิดเห็นภาษาไทยเท่านั้น

## ปัญหาอุปสรรค

1. ข้อมูลมีไม่มากพอสำหรับการฝึกโมเดลเพื่อให้ได้ประสิทธิภาพเพิ่มขึ้น
2. การตัดคำที่ยังไม่สมบูรณ์ทำให้เกิดข้อผิดพลาดในการฝึกโมเดลได้
3. การคัดกรองคำส่วนของ Positive, Neutrally หรือ Negative ให้มีความละเอียด



# ข้อจำกัดของระบบปัญหา อุปสรรค และแนวทางการแก้ไข

## แนวทางการแก้ไข

1. ทำการเก็บรวมรวบข้อมูลจากแหล่งต่างๆ ที่มีคนสร้างเอาไว้และนำมา  
รวมกับที่ตัวเองเก็บ
2. เปลี่ยนวิธีการตัดคำใหม่ เพื่อให้ได้ความหมายที่ชัดเจน
3. ทำการจำแนกประเภท Positive, Neutrally หรือ Negative เป็นส่วน  
ย่อยๆ เพื่อให้มีความละเอียดในส่วนของการจำแนกประเภท

## ข้อเสนอแนะ

1. เพื่อความแม่นยำของโมเดล
2. การทำความสะอาดข้อมูลให้ดียิ่งขึ้น



# สรุปผลการวิจัย

LEARN MORE



# สรุปผลการวิจัย

ข้อมูลที่ใช้ในการฝึกเป็นภาษาไทยรวมทั้งหมด 4,500 ความคิดเห็นและทำการแยกเป็น 3 ส่วนคือ Positive, Negative, Neutrally ที่มีส่วนละ 1,500 ความคิดเห็น

พบว่า โมเดลของ Feed Forward Neural Network ได้ผลลัพธ์การทำนายความคิดเห็นภาษาไทยอยู่ที่ร้อยละ 0.695 และเป็นวิธีที่ดีที่สุดสำหรับการวิเคราะห์ความรู้สึกและความคิดเห็น ทั้งนี้เมื่อพิจารณาผลลัพธ์การทำนายความคิดเห็นภาษาไทยแต่ละ โมเดลจากมากไปน้อยพบว่า Feed Forward Neural Network ได้ร้อยละเท่ากับ 0.695, Support Vector Machine ได้ร้อยละเท่ากับ 0.691, Navie Bayes ได้ร้อยละเท่ากับ 0.64, Bert ได้ร้อยละเท่ากับ 0.60 ตามลำดับ



# THANK YOU



COMPUTER SCIENCE