

หัวข้อการค้นคว้าอิสระ	การประยุกต์ใช้โมเดลเบิร์ตเพื่อจำแนกประเภทอาการแพ้จากเนื้อหารีวิวของผู้บริโภคผลิตภัณฑ์ไฮโปอัลเลอร์เจนิก
ชื่อผู้เขียน	อังศุมาลินร์ เสือวิจิตร
ชื่อปริญญา	วิทยาศาสตรมหาบัณฑิต (ระบบสารสนเทศเพื่อการจัดการ)
สาขาวิชา	ระบบสารสนเทศเพื่อการจัดการ
คณะ	พาณิชยศาสตร์และการบัญชี
มหาวิทยาลัย	มหาวิทยาลัยธรรมศาสตร์
อาจารย์ที่ปรึกษาการค้นคว้าอิสระ	ผู้ช่วยศาสตราจารย์ ดร.วินัย นาคี
ปีการศึกษา	2567

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อประยุกต์ใช้โมเดลภาษาเบิร์ต BERT (Bidirectional Encoder Representations from Transformers) ในการจำแนกอาการแพ้จากข้อความรีวิวของผู้บริโภคที่ใช้ผลิตภัณฑ์สกินแคร์ประเภทไฮโปอัลเลอร์เจนิก (Hypoallergenic) บนแพลตฟอร์มเซโฟร่า (Sephora) ระหว่างปี พ.ศ. 2561–2568 โดยมีเป้าหมายในการพัฒนาระบบที่สามารถวิเคราะห์และระบุอาการแพ้ได้อย่างแม่นยำจากรีวิวที่ผู้บริโภคเขียน เพื่อส่งเสริมการตัดสินใจเลือกใช้ผลิตภัณฑ์ที่ปลอดภัยและเหมาะสม

ในการวิจัย ผู้วิจัยได้รวบรวมข้อมูลจากชุด Sephora Products and Skincare Reviews Dataset บนเว็บไซต์ Kaggle และนำมาผ่านกระบวนการกรองและติดป้ายกำกับอาการแพ้ในรูปแบบการจำแนกหลายป้าย (Multi-label Classification) ขั้นตอนการวิจัยประกอบด้วย 5 ขั้นตอนหลัก ได้แก่ (1) การรวบรวมข้อมูลรีวิว (2) การทำความสะอาดและเตรียมข้อมูล (3) การกำหนดป้ายกำกับอาการแพ้ เช่น รอยแดง ผื่น คัน สิว เป็นต้น (4) การฝึกโมเดลเบิร์ต (BERT) (5) การประเมินผลด้วยค่าประสิทธิภาพ

ผลการศึกษาพบว่า โมเดลเบิร์ต (BERT) ที่ได้รับการปรับแต่งโมเดล (Fine-tuning) สามารถจำแนกอาการแพ้จากรีวิวได้อย่างมีประสิทธิภาพ สามารถระบุอาการที่เกี่ยวข้องได้อย่างถูกต้องและครอบคลุม อีกทั้งยังสามารถนำไปใช้ตรวจสอบแนวโน้มและการคาดการณ์อาการแพ้จากข้อมูลรีวิวได้ ซึ่งสะท้อนถึงศักยภาพในการนำไปประยุกต์ใช้เชิงพาณิชย์

คำสำคัญ: โมเดลเบิร์ต, การจำแนกอาการแพ้, การเรียนรู้เชิงลึก, ผลิตภัณฑ์ไฮโปแอลเลอร์เจนิก, การประมวลผลภาษาธรรมชาติ, รีวิฟู้บริโกล, การวิเคราะห์ข้อความ, การจำแนกหลายป้ายกำกับ, อาการแพ้ผิวหนัง, เครื่องสำอางสกินแคร์



Independent Study Title	APPLICATION OF THE BERT MODEL FOR CLASSIFYING TYPES OF ALLERGIC REACTIONS IN CONSUMER REVIEWS OF HYPOALLERGENIC PRODUCTS
Author	Ungsumalin Sueawijit
Degree	Master of Science (Management Information Systems)
Major Field	Management Information Systems
Faculty	Commerce and Accountancy
University	Thammasat University
Independent Study Advisor	Assistant Professor Winai Nadee, Ph.D
Academic Year	2024

ABSTRACT

This research primarily aims to apply the Bidirectional Encoder Representations from Transformers (BERT) model for classifying allergic reactions from consumer reviews of Hypoallergenic skincare products on the Sephora platform from 2018 to 2025. The main goal of this research is to develop a system capable of accurately identifying allergic reactions described in textual reviews, enabling consumers to make safer and more informed decisions when selecting skincare products.

The research methodology began with collecting consumer review data from Sephora via the publicly available Sephora Products and Skincare Reviews Dataset on Kaggle. The dataset underwent a comprehensive data preprocessing stage to clean and standardize the textual data, followed by data labeling where allergic reactions were categorized using multi-label classification into various reaction groups such as redness, rash, acne, burning sensations, and other related symptoms. The next critical phase involved fine-tuning the BERT model, a powerful transformer-based language model renowned for its effectiveness in textual analysis. Finally, the model

was rigorously tested and evaluated using key performance metrics including Accuracy, Precision, Recall, and F1-Score.

The findings revealed that the fine-tuned BERT model effectively classified allergic reactions from consumer reviews with high accuracy, successfully identifying and categorizing consumer-reported symptoms associated with skincare product usage. Additionally, the model demonstrated the ability to analyze and predict trends from newer review datasets, highlighting its utility in commercial and public health applications. The outcomes of this research emphasize the importance of transparent and accurate product labeling, particularly regarding hypoallergenic claims. This awareness can help skincare manufacturers enhance product safety and quality standards, ultimately benefiting consumers and the industry alike.

Keywords: BERT model, allergy classification, deep learning, hypoallergenic products, natural language processing (NLP), consumer reviews, text analysis, multi-label classification, skin allergic reactions, skincare cosmetics