

## บทคัดย่อ

ชื่อโครงการ	กระบวนการสำหรับการจำแนกความรู้สึกที่มีข้อมูลไม่สมดุล
ผู้จัดทำ	61011212107 นายพีระวัฒน์ บุญบ้านจิว
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร.จันทิมา พลพินิจ
หลักสูตร	วิทยาศาสตร์บัณฑิต (สาขาวิชาวิทยาการคอมพิวเตอร์)
คณะ	วิทยาการสารสนเทศ
มหาวิทยาลัย	มหาวิทยาลัยมหาสารคาม
ปีที่พิมพ์	2563

การจำแนกความรู้สึก (Sentiment Classification) คือการจำแนกเอกสารตามชั้นความรู้สึกซึ่งโดยทั่วไปอาจจำแนกเป็นความรู้สึกที่เป็นบวก (Positive) ความรู้สึกที่เป็นลบ (Negative) และความรู้สึกที่เป็นกลาง (Neutral) โดยการจำแนกความรู้สึกนั้น ได้รับการศึกษามาอย่างต่อเนื่อง เพราะการประยุกต์ใช้ในหลายลักษณะ แต่โดยทั่วไปมักจะนิยมใช้ในการจำแนกความรู้สึกที่มีการแสดงไว้ในรูปแบบข้อความ (Text) เช่น ประยุกต์ใช้ในการจัดอันดับความรู้สึกจากข้อความแสดงความคิดเห็นของผู้คนที่ติดต่อสินค้าและบริการ การประยุกต์ใช้เพื่อวิเคราะห์ความรู้สึกของผู้เรียน การประยุกต์ใช้เพื่อวิเคราะห์ความรู้สึกของผู้คนในเรื่องการเมือง เป็นต้น ซึ่งปัญหาความไม่สมดุลของข้อมูลในคลาสนั้น เกิดจากกลุ่มตัวอย่างที่ใช้ในการเรียนรู้มีข้อมูลไม่สมดุลกัน โดยกลุ่มที่มีข้อมูลมากกว่าจะเรียกว่า “ข้อมูลกลุ่มหลัก (Majority Class)” ขณะที่กลุ่มตัวอย่างที่มีข้อมูลจำนวนน้อยกว่าจะเรียกว่า “ข้อมูลกลุ่มรอง (Minority Class)” เมื่อนำเอาชุดข้อมูลในลักษณะนี้ไปเรียนรู้เพื่อสร้างตัวจำแนกความรู้สึก (Sentiment Classifier) ข้อมูลใหม่ๆ ที่อ่านเข้ามาเพื่อวิเคราะห์เพื่อจำแนกกลุ่มด้วยตัวจำแนกความรู้สึกดังกล่าว ก็มีแนวโน้มที่จะทำนายกลุ่มของข้อมูลนั้นไปยังทิศทางของข้อมูลกลุ่มหลักที่ใช้ในการเรียนรู้ตัวจำแนกความรู้สึก ดังนั้น ในโครงการปริญญาโทฉบับนี้ จึงได้นำเสนอการศึกษาการแก้ปัญหาความไม่สมดุลของข้อมูลในการจำแนกความรู้สึกด้วยเทคนิคการให้น้ำหนักค่า 5 เทคนิค คือ TF-IDF, Delta TF-IDF, TF-IDF-ICF, TF-RF และ TF-IGM ร่วมกับแมชชีนเลิร์นนิง 3 ตัว คือ Naïve Bayes, K-Nearest Neighbor และสุดท้าย Convolution Neural Network

**คำสำคัญ:** การจำแนกเอกสาร, การให้น้ำหนักค่า, ข้อมูลไม่สมดุล, ซัพพอร์ตเวกเตอร์แมชชีน