

ปิยธิดา อินทร์รักษ์ : การประยุกต์ใช้การวิเคราะห์ความหมายแฝงกับการจำแนกประเภทอารมณ์ในข้อความภาษาไทย. (APPLYING LATENT SEMANTIC ANALYSIS TO CLASSIFICATION OF EMOTIONS IN THAI TEXT) อ.ที่ปรึกษาวิทยานิพนธ์หลัก : ผศ. ดร.สุกรี สินธุภิญโญ, 74 หน้า.

ในยุคที่การติดต่อสื่อสารข้อมูลผ่านเครือข่ายอินเทอร์เน็ตเติบโตขึ้นอย่างต่อเนื่อง ข้อมูลประเภทตัวอักษรก็ถูกผลิตขึ้นมาเป็นจำนวนมากเช่นกัน ข้อมูลเหล่านี้สามารถถูกถ่ายทอดออกมาและจำแนกหมวดหมู่ของตัวอักษรได้ การจำแนกด้านอารมณ์ก็เป็นอีกหัวข้อที่น่าสนใจในปัจจุบัน แต่การจำแนกด้านอารมณ์จากตัวอักษรภาษาไทยนั้นยังไม่มีประสิทธิภาพที่ดีพอ หัวข้อวิจัยนี้ได้แบ่งการจำแนกประเภทอารมณ์จากข้อความสั้นภาษาไทยออกมาเป็น ๖ อารมณ์สากลพื้นฐาน ได้แก่ โกรธ ชะแหยง กลัว มีความสุข เศร้า และประหลาดใจ ซึ่งอ้างอิงจากข้อมูลการวิจัยในการวิจัยนี้ได้เปรียบเทียบผลลัพธ์ของ ๒ ตัวแบบที่สร้างมาจากประโยครูปแบบต่างๆ และประยุกต์ใช้กับ ๓ ระเบียบวิธีได้แก่นาอีฟเบย์ส (Naïve Bayes), เครื่องจักรเวกเตอร์สนับสนุน (Support Vector Machine, SVM) และต้นไม้ตัดสินใจ (Decision Tree) โดยตัวแบบที่หนึ่งใช้การจำแนกโดยการวิเคราะห์ความหมายแฝงของคำเดี่ยว ส่วนตัวแบบที่สองใช้การประยุกต์การวิเคราะห์ความหมายแฝงของคำคู่ที่มักปรากฏคู่กันร่วมกับระนาบความหมายของคำเดี่ยว ผลการเปรียบเทียบผลลัพธ์แสดงให้เห็นว่า ตัวแบบที่สองให้ความถูกต้องได้สูงกว่าตัวแบบที่หนึ่ง อ้างอิงจากระเบียบวิธีการจำแนกของนาอีฟเบย์สที่ให้ผลสูงกว่าระเบียบวิธีการอื่น

5171427521 : MAJOR COMPUTER SCIENCE

KEYWORDS: EMOTIONS IN TEXT / LATENT SEMANTIC ANALYSIS / AFFECTIVE COMPUTING

PIYATIDA INRAK : APPLYING LATENT SEMANTIC ANALYSIS TO CLASSIFICATION OF EMOTIONS IN THAI TEXT. THESIS ADVISOR : ASST. PROF. SUKREE SINTHUPINYO, Ph.D., 74 pp.

With a rapid growth of the internet communication, many types of text are produced. They can convey the meanings that can contribute to text categorization. Moreover, emotion classification becomes more interesting, but emotion classification in Thai text is still not able to be correctly classified. Thus, this paper proposes a novel approach that takes advantage of bi-words occurrence to classify emotion hidden in a short sentence. In this paper, we classify Thai text into six basic universal emotions including anger, disgust, fear, happiness, sadness, and surprise based on Latent Semantic Analysis (LSA) approach. We compared the results between two models which construct features from the sentences and applied both models to three classification methods, i.e. Naïve Bayes, SVM, and Decision Tree. The first feature model uses only single word occurrence in the classification. The second model uses single word combined with bi-words occurrence in the classification. The results show that the second model yielded higher accuracy than the first model based on the Naïve Bayes classification method.