หัวข้อวิจัย : การพัฒนาประสิทธิภาพการจำแนกเว็บเพจโดยใช้เทคนิคการถ่วงน้ำหนักของคำร่วมกับการเรียนรู้ของเครื่อง
ผู้วิจัย : ดร.นิเวศ จิระวิชิตชัย
หน่วยงาน : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีปทุม วิทยาเขตชลบุรี
ปีที่พิมพ์ : พ.ศ. 2556

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาวิธีการคำนวณค่าน้ำหนัก สำหรับสะท้อนอำนาจจำแนกของคำที่สกัดได้ โดยนำเสนอวิธีการคำนวณค่าน้ำหนักชื่อAmbiguity Weighting (AW) เพื่อพัฒนาประสิทธิภาพในการจำแนกเว็บเพจโดยใช้เทคนิคการถ่วงน้ำหนักของคำร่วมกับการเรียนรู้ของเครื่อง ผลจากการทดลองเมื่อวัดประสิทธิภาพด้วยค่าเฉลี่ย F-Measure เมื่อทำการลดคุณลักษณะด้วยวิธีค่าสารสนเทศ (Information Gain) และประมวลผลด้วยเครื่องจักรเรียนรู้ สามารถสรุปได้ว่าเมื่อเรียนรู้ด้วยอัลกอริทึม Support Vector Machine หรือ Naïve-Bayes หรือ Decision Tree การคำนวณค่าน้ำหนักเพื่อสร้างดัชนีด้วยวิธี AW ให้ประสิทธิภาพจำแนกโดยเฉลี่ยออกมาดีที่สุด คือ 99% หากพิจารณาพารามิเตอร์ที่ลดคุณลักษณะลงและ ส่งผลให้ประสิทธิภาพด้านการจำแนกเว็บเพจดีที่สุดด้วยวิธีให้ค่าน้ำหนัก AW พบว่า ที่จำนวน 500 คุณลักษณะ ให้ประสิทธิภาพการจำแนกดีสุดที่ระดับ 99% ผลจากการทดลองสรุปได้ว่าวิธีการคำนวณค่าน้ำหนักให้กับดัชนี AW ที่พัฒนาขึ้นให้ประสิทธิภาพในจำแนกเว็บเพจสูงกว่าวิธีการคำนวณน้ำหนักแบบอื่น ๆ ทุกประเภทและทุกอัลกอริทึมที่นำมาเรียนรู้อย่างชัดเจน

คำสำคัญ: จำแนกเว็บเพจ, การถ่วงน้ำหนักของคำ, การเรียนรู้ของเครื่อง

Research Title : Developing an Effective Webpage Classification Using Term Weighting Technique and Machine Learning Algorithms
Name of Researcher : Dr.Nivet Chirawichitchai
Name of Institution : Graduate School, Sripatum University Chonburi Campus
Year of Publication : 2013

ABSTRACT

The objective of this research is to develop a method for term weighting calculation that reflects the actual discrimination of isolated words. The method is called Ambiguity Weighting (AW), this will improve efficiency of webpage classification framework. The experimental results showed that reducing the feature by Information Gain method and process Ambiguity Weighting (AW) with support vector machines, Naïve-Bayes or Decision Tree will yielded a very high classification with F-Measure equaling 99%. The experiments concluded that the AW weighting method clearly improved the efficiency of webpage classification framework over other weighting methods and algorithms.

Keywords**:** webpage classification, weighting technique, machine learning