

การเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลปัจจัยสนับสนุนการเรียนรู้ ของผู้เรียนระดับอุดมศึกษาในสภาพแวดล้อมการเรียนรู้เสมือนจริง

อุบลรัตน์ ศิริสุขโกคา^{1*} และ จริญญา แสนราช²

บทคัดย่อ

บทความนี้นำเสนอผลการประเมินและเปรียบเทียบประสิทธิภาพการจำแนกปัจจัยสนับสนุนการเรียนรู้ของผู้เรียนระดับอุดมศึกษาเมื่อเรียนรู้ในสภาพแวดล้อมการเรียนรู้เสมือนจริง 3 เทคนิค คือ เทคนิค Decision Tree เทคนิค k-NN และเทคนิค Naive Bayes โดยการนำข้อมูลเกี่ยวกับการลงทะเบียนเรียนและผลการเรียนของผู้เรียนที่เรียนรู้ด้วยการโต้ตอบกับสภาพแวดล้อมการเรียนรู้เสมือนจริง จำนวน 8 แอททริบิวต์ มีข้อมูลจำนวน 32,950 ชุด ซึ่งผ่านการตรวจสอบและปรับให้ข้อมูลมีความสมบูรณ์ โดยแบ่งข้อมูลออกเป็น 2 ส่วน ด้วยวิธี Split Test ประกอบด้วย ข้อมูลส่วนที่หนึ่ง จำนวน 26,360 ชุด คิดเป็น 80% ของข้อมูลทั้งหมด เป็นข้อมูลที่ใช้ในการสร้างโมเดลและข้อมูลส่วนที่สอง จำนวน 6,590 ชุด คิดเป็น 20% ของข้อมูลทั้งหมด เป็นข้อมูลที่ใช้ในการทดสอบความถูกต้องและแม่นยำของโมเดลที่สร้าง ผลการวิจัย พบว่าเทคนิคที่ใช้ในการจำแนกปัจจัยสนับสนุนการเรียนรู้ของผู้เรียนระดับอุดมศึกษาในสภาพแวดล้อมการเรียนรู้เสมือนจริงที่มีประสิทธิภาพสูงสุด คือ k-NN (เมื่อกำหนดให้ $k = 3$) โดยมีค่าความแม่นยำ 97.87% ค่าความระลึกลับ 97.90% ค่าความแม่นยำ 97.93% และค่าถ่วงดุล 0.352

คำสำคัญ: การจำแนกข้อมูล, การแบ่งข้อมูลเพื่อทดสอบโมเดล, ต้นไม้ตัดสินใจ, k-NN, เนอิวเบย์

¹ ผู้ช่วยศาสตราจารย์ สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม

² ผู้ช่วยศาสตราจารย์ ภาควิชาคอมพิวเตอร์ศึกษา คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

* ผู้นิพนธ์ประสานงาน โทร. +669 4469 1651 อีเมล: ubonrat76@gmail.com



A Comparison of the Efficiency of Data Classification in Learning Factors of Tertiary Level Students in a Virtual Learning Environment

Ubonrat Sirisukpoca^{1*} and Charan Sanrach²

Abstract

This research presents the result of evaluation and comparison of the efficiency of data classification in learning support factors of tertiary level students in a virtual learning environment. Three primary techniques were compared in this research, including Decision Tree technique, k-NN technique, and Naive Bayes technique. There were 8 attributes with 32,950 datasets of the data which were included student registration and the learning outcome of the students who interactively learned in a virtual learning environment. The data was divided into 2 parts by Split Test method. The first part which was comprised of 26,360 datasets (approximately 80 percent of the total data) was used to generate a model while the second part which was included 6,590 datasets (approximately 20 percent of the total data) was utilized to examine the precision and accuracy of the model. The results revealed that the most efficient technique, used to classify the learning factors of the tertiary students in a virtual learning environment, was k-NN technique (k = 3) with 97.87 percent of precision, 97.90 percent of recall, 97.93 percent of accuracy, and 0.352 F-measure.

Keywords: Classification, Split Test, Decision Tree, k-NN, naive bayes

¹ Assistance Professor, Department of Computer Science Faculty of Science and Technology, Nakhon Pathom Rajabhat University

² Assistance Professor, Department of Computer Education, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok

* Corresponding Author Tel. +669 4469 1651 email: ubonrat76@gmail.com