

นฤป รัชงาม 2550: ระบบวิเคราะห์คำถาม และตอบปัญหาการใช้งานอินเทอร์เน็ต ด้วยเทคนิคการทำเหมืองข้อมูล ปริญญาวิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์) สาขาวิทยาการคอมพิวเตอร์ ภาควิชาวิทยาการคอมพิวเตอร์ ประธานกรรมการที่ปรึกษา: รองศาสตราจารย์อนงค์นาฏ ศรีวิหค, Ph.D. 63 หน้า

งานวิจัยนี้นำเสนอระบบการวิเคราะห์คำถามอัตโนมัติ เพื่อสร้างระบบที่สามารถทำนายสาเหตุของปัญหาการใช้งานอินเทอร์เน็ตของลูกค้า และนำเสนอแนวทางแก้ไข โดยประยุกต์ใช้เทคนิคการสกัดคำถาม (Data Extraction) ใช้ร่วมกับเทคนิคทางการทำเหมืองข้อมูล (Data mining) เทคนิคการสกัดคำถามนั้นใช้สำหรับสกัดข้อมูลที่สัมพันธ์กันจากคำถามที่เกี่ยวข้องกับการใช้งานอินเทอร์เน็ตของลูกค้า เพื่อใช้เป็นข้อมูลนำเข้าในขั้นตอนการทำเหมืองข้อมูลต่อไป ขั้นตอนของการทำเหมืองข้อมูลของงานวิจัยนี้แบ่งเป็น 2 กรณี คือ (1) กรณีที่สามารถเชื่อมต่ออินเทอร์เน็ตได้ (online) จะใช้การจัดกลุ่มข้อมูลแบบ 2 ขั้นตอน (SOM และ K-Means อัลกอริทึม) ในการจัดกลุ่มของลูกค้าและใช้ OLAP Cube ในการวิเคราะห์เพื่อหาความน่าจะเป็นของสาเหตุที่ทำให้เกิดปัญหา (2) กรณีที่ไม่สามารถเชื่อมต่ออินเทอร์เน็ตได้ (offline) จะใช้โปรแกรมตัวแทน (Software Agent) ในการตรวจสอบปัญหาและใช้ต้นไม้ตัดสินใจ (Decision Tree) เพื่อหาคำตอบของปัญหาที่เกิดขึ้นในงานวิจัยนี้ใช้ค่า Precision ในการวัดความแม่นยำ โดยผลที่ได้จากการใช้ OLAP Cube จะให้ค่าความแม่นยำร้อยละ 70.46 ส่วนการใช้ Decision tree ได้ให้ค่าความแม่นยำร้อยละ 87.56 ผลลัพธ์ที่ได้จากการใช้ระบบนี้นอกจากจะช่วยลดปริมาณการโทรสอบถามมายังพนักงาน Call center แล้ว ยังเอื้อประโยชน์แก่ลูกค้าทั่วไปที่ไม่มีความรู้ในเรื่องอินเทอร์เน็ตให้สามารถตรวจสอบแก้ไขปัญหาด้วยตนเองได้

Narroup Ruk-ngam 2007: Question Analysis and Answering System for the Internet Usage by Using Data Mining Techniques. Master of Science (Computer Science), Major Field: Computer Science, Department of Computer Science. Thesis Advisor: Associate Professor Anongnart Srivihok, Ph.D. 63 pages.

This study proposes a system for an automatic question analysis and answering which applies information extraction techniques for extracting the relevant information from questions of the Internet connections. The system uses data extractions and data mining techniques. User profiles and their questions on the Internet usage are used as input for the system. The system processes are divided in two categories due to the type of internet connection cases. First, online case: two steps clustering by SOM and K-Means algorithms are used to segment customer data. Then outputs from clustering and question extraction are used for OLAP to analyze for cause of the Internet problems. Second, offline case: software agent is applied; this case a decision tree is built which each node contains cause of problems. The system will give the answers for solving the Internet connections in both online and offline cases. The results from this study reveal that it is possible to develop an automatic question answering system. The proposed system had been tested with sample questions from the Internet service provider: True corporation Call Center. Precision of the system is about 70.46% for online and 87.56% for offline which is good. This study offers useful information regarding the areas of data mining for Call Center or Web Based Support Systems.