公立千歳科学技術大学紀要第 4 巻第 1 号 (2023)
大学院理工学研究科理工学専攻修士論文英文要旨

Analysis of tacit knowledge-based evaluation for learner competencies using machine learning modeling

Allen MOMIJI（紅葉 亜練）

In traditional educational institutions, human resource development has been performed with an awareness of cognitive abilities, concentrating on dealing with knowledge such as fundamental academic skills and specialized knowledge for learning academic areas. In recent years, however, it has become more critical to be aware of noncognitive abilities that cannot be evaluated by tests, such as learners' capacity for problem-solving, continuous learning, independence, and cooperation. The development of these abilities requires a series of interactions in which learners reflect on their learning and teachers assess the content and provide feedback Numerical evaluation and visualization are challenging since evaluation is reliant on tacit information that depends on teachers' instructional expertise. While there are studies on competency requirements, there are few studies on teachers' tacit knowledge as a basis for judging learners' competencies. In contrast, the goal of this work is to determine which aspects of learning outcomes have an impact on the evaluation using machine learning models, presuming that tacit knowledge is relevant to the overall judgment of learning outcomes with reflection. This project worked with the mathematics faculty at various colleges to perform a questionnaire survey and labeling to get data for machine learning. Labels are referred to as scores and are rated on a scale of 1, 2, or 3. The surveys were given at the beginning and end of the spring and fall semesters, with a "pretest" at the beginning of the semester and a "post-test" at the end of the semester. There were two response formats. The two response formats were rubric and open-ended. All data were graded, with score 1 being 1605, score 2 being 2840, and score 3 being 1744, for a total of 6189 data points obtained. In this study, criteria for teachers to assess learners' abilities were taken into consideration, including rubrics, word counts, the frequency of keywords, and contextual information. Two types of keywords were defined from the learners' reflective text: science and math words and ability words. Mathematics-related terms include science and math terms as well as ability-related terms. Using these candidate characteristics, the XGBoost model was developed. The accuracy of this model was about 60%. To verify the extent to which the candidate features used affect the inference accuracy, we performed an analysis using Random Forest and SHAP. The results demonstrated that the rubric and BERT had a particularly strong influence on the inference results, while the other features had little influence. The accuracy increased to roughly 65% when the model was built with only the rubric and contextual information as features, indicating that it is possible to statistically capture teachers' tacit knowledge. Since the number of occurrences of keywords had little effect on inference accuracy, we investigated the possibility that keywords were not picked up from the words in the sentence, so we calculated word similarity using Word2Vec and increased the number of keywords. Before the keyword expansion, there were 388 words linked to science and mathematics

and 437 words related to ability. After the keyword increase, there were 2,683 words related to science and mathematics and 1,685 words related to ability. The influence of the number of keywords on inferential accuracy was not altered after the increase in keywords when feature analysis was carried out again. Next, based on the idea that sentiment in a sentence can also be a new feature, sentiment analysis was performed. For the sentiment analysis, we employed an existing trained model. Sentiment in sentences was introduced as a new candidate feature, and when we reran the analysis, we discovered that sentiment in sentences also significantly impacted the results. However, the accuracy did not change at all; in some cases, it went up and in other cases, it went down. Based on these findings, we separated the data by institution and examined how each factor affected the data. The results demonstrated that tacit knowledge varies from faculty member to faculty member, as each institution has various strong and weak characteristics. The mathematics faculty at each institution was informed of the findings and given a survey to complete. Four of the seven responses showed that the analysis results were valid and helpful. One professor remarked that the findings were in line with the regulation in effect at the time of the evaluation.

In this study, we verified that tacit knowledge varies from institution to institution and assessed whether the findings of the analysis were consistent with the faculty member's tacit knowledge. As a result, we think we were able to uncover some of the faculty members’ unique tacit knowledge.