

Sommart Aungkaseraneeekul 2012: Automated Thai-Language Essay Scoring. Master of Science (Computer Science), Major Field: Computer Science, Department of Computer Science. Thesis Advisor: Associate Professor Chuleerat Jaruskulchai, D.Sc. 116 pages.

Machine learning research has been applied in the daily life including automatic essay scoring. The concept of automatic essay grading is comparable to the documents clustering or classification according to a set of essays. However, the preparation of the training data needs to include every category of the answer sets which is impractical. The training data usually includes the appropriate essays and it might be available only for some types of data. Additionally, the number of clusters may not be known in advance.

In this research, EM and Cobweb algorithms are used to discover automatically the number of clusters for each question. The similarity score of each of clusters was calculated using cosine function between each representative of clusters and its one solution. We used query expansion and LSA techniques in our work; query expansion technique was deployed in order to solve the problem of synonyms and LSA technique was used to find a latent semantic meaning and reduced data dimension. The results from tests were compared with the learning algorithm such as K-NN, ANN and Bayesian.

The effectiveness of the algorithms is assessed with seven questions of object-oriented programming test with total of fifty-five students' answers and applied only a single solution for assigning score. The experimental results showed that the total average accuracy rate between unsupervised algorithm and supervise algorithm are very close.

---

Student's signature

---

Thesis Advisor's signature