

Automated Essay Grading: Daniel Lee & Jun Seob Shim

Description:

This project focuses on the intriguing domain of automated essay grading. The initial approach began with simple machine learning models such as Linear Regression, SVM, and Random Forest. As the project matured, a shift was made to leverage deep learning models, including LSTM, BERT, and Word2Vec, aiming to enhance the accuracy and efficiency of grading.

Repository and Code Structure

Folders:

- DeepML: Contains the AESDeepML.ipynb, which is the notebook that includes implementation and documentation for deep learning models like LSTM, BERT, and Word2Vec.
- Dataset: Folder containing the dataset used for both simple and deep machine learning models.

Files:

- AESSimpleML.ipynb: This notebook includes the implementation and documentation for traditional machine learning models such as Linear Regression, SVM, and Random Forest.
- all_models_results.pdf: A comprehensive report containing the results of running both the simple and deep learning models.

Execution:

To run the code:

Import all files & folders into either google drive for colab or your local machine for jupyter. Open notebooks AESDeepML.ipynb and AESSimpleML.ipynb, then run each cell chronologically.

Results and Observations

The summarized results from running the deep ML models are available in all_models_results.pdf. The report comprises:

- Plots and tables presenting the results, specifically Kappa Score and runtime.
- Detailed information for each epoch for each set, allowing the user to follow along the training process.