

ABSTRACT

Assessing students' essay answers is a crucial process but requires significant time and effort from teachers. This study aims to develop an automated essay scoring system for Social Studies using a deep learning approach. Two models are implemented in the system: a Long Short-Term Memory (LSTM) classifier for categorizing student answers as correct or incorrect, and a Manhattan LSTM (MaLSTM) model for measuring semantic similarity between student answers and teacher-provided key answers. The dataset consists of questions and labeled student answers provided by educators. The data underwent preprocessing steps including text cleaning, tokenization, FastText embedding, and was split into training and testing sets. Each model was trained and evaluated separately before being integrated into a web-based application using the Flask framework. The evaluation results show that the LSTM classifier achieved higher accuracy compared to MaLSTM for binary classification. The system allows automatic and real-time essay evaluation, providing classification outcomes alongside similarity scores to assist teachers in understanding the rationale behind each prediction.

Keywords: Automated Scoring, Essay Question, LSTM, MaLSTM, Semantic Similarity