Review of: "Enhancing Student Writing Skills: Leveraging Transfer Learning and Fine-tuned Language Models for Automated Essay Structure Recognition"

Fan Zhang[1]

1 University of Delaware

Introduction:

The introduction effectively provides a background and rationale for the study by emphasizing the importance of writing skills and the limitations of traditional writing instruction. But it would be helpful to clearly state the research objective or hypothesis of the study.

Related Works:

The section offers a comprehensive overview of previous research on AWE systems. But it would be beneficial to synthesize and address the limitations or research gaps.

Methodology:

The paper provides a detailed explanation of the dataset used for training and testing the models, along with the steps taken for data cleaning and preprocessing. However, it would be useful to provide more information about the specific features in the PERSUADE corpus and how they relate to the proposed discourse classification task.

Two Models Used - PyTorch BigBird and TensorFlow Longformer:

The choice of BigBird and Longformer models is well-explained, considering their performance and computational efficiency. But it would be helpful to mention the additional information of BigBird and Longformer models to provide readers with a better understanding.

Hyperparameters:

The paper presents the hyperparameters used for model training, such as the number of epochs, batch size, and learning rate scheduler. However, it would be valuable to discuss the reasoning behind these choices and any experimentation or tuning conducted to determine their effectiveness.

Model Evaluation - F1-score:

The use of F1 score as an evaluation metric is appropriate for assessing model performance. But it would be beneficial to provide a concise explanation of precision, recall, and how the F1 score combines these measures to offer a comprehensive evaluation value.

Results:

The results section presents the F1 scores, highlighting the superior performance of the Longformer model. But it would be valuable to provide the specific strengths and weaknesses of each model in capturing different discourse elements or any patterns observed in the classification results.

Discussion:

The discussion section indicates the potential contributions of the proposed approach. However, it would be helpful to further discuss the implications of the findings and their significance in enhancing student writing skills, as well as the study's limitations.

Overall Structure and Clarity:

The paper follows a clear and logical structure. But some lengthy paragraphs can be divided into smaller ones for readability.