T5 -clickstream data contains a significant amount of information that accurately describes students’ learning processes, which makes it difficult to construct using hand-crafted features . to analyze student behavior effectively, we attempted to capture critical knowledge from the field of natural language processing (NLP) to the field of student performance prediction in Massive Open Online Courses (MOOCs), owing to how closely they resemble each other . our results showed the potential of NLP techniques to improve dropout prediction in MOOCs by extracting informative representations from clickstream data . .

Bart–Click-Based Representation Learning Framework of Student Navigational Behavior in MOOCs. Predictive learning outcomes’ models for online students can provide useful information to instructors to estimate students’ final performance in the early stages of a course. Existing research models that analysed student data havefocused on handcrafted features, but these models have limitations in exploring new behavioral patterns that indicate student performance and how they can be used in online courses. Our results showed the potential of NLP techniques to improve dropout prediction inMOOCs by extracting informative representations from clickstream data, allowing a deeper understanding of student behavior, and facilitating early intervention strategies. The combined approach of skip-gram embeddings with Principal Component Analysis (PCA) achieved the highest accuracy, particularly on the Xutangx dataset, with an accuracy of approximately 72.70% and an F1-score of approximately 81.03%. Furthermore, when applied to the KDDCUP dataset, this methodology exhibited even higher performance, with a accuracy of 80.91% and a F1 score of 87.42%. .

Gemini Existing student performance prediction models in MOOCs rely on handcrafted features, limiting their ability to capture complex behavioral patterns. This paper presents a novel framework utilizing self-supervised learning to automatically generate useful data representations from student clickstream data. The framework comprises two stages: self-supervised clickstream pre-training to model click generation and adjustment of latent representations for supervised learning. Experimental results on real-world datasets demonstrate the effectiveness of the proposed approach, particularly the combination of skip-gram embeddings and Principal Component Analysis (PCA). This framework highlights the potential of natural language processing (NLP) techniques in improving dropout prediction and deepening understanding of student behavior in MOOCs.

Gpt-4

The research paper titled "Click-Based Representation Learning Framework of Student Navigational Behavior in MOOCs" explores the potential of predictive models to forecast student performance in online courses early in their academic journey. Traditional models have relied on manually crafted features to analyze student data, which often fail to uncover new behavioral patterns critical for performance prediction in Massive Open Online Courses (MOOCs). This study introduces a novel framework that leverages clickstream data—detailed logs of student interactions—to enhance prediction outcomes through a self-supervised learning approach.

The framework draws inspiration from natural language processing (NLP) techniques, recognizing the similarities between language and student behavior in digital environments. The proposed method involves a two-step process: initially, a self-supervised pre-training setup models learner click generation, capturing latent behavioral patterns. Subsequently, these representations are refined and applied to a supervised learning task aimed at predicting student outcomes.

The effectiveness of this approach was validated through extensive experiments on two real-world datasets. The integration of skip-gram embeddings with Principal Component Analysis (PCA) yielded the most accurate results. Specifically, on the Xutangx dataset, the model achieved an accuracy of approximately 72.70% and an F1-score of about 81.03%. The methodology demonstrated even superior performance on the KDDCUP dataset, achieving an accuracy of 80.91% and an F1-score of 87.42%.

These results underscore the potential of applying NLP techniques to MOOCs, enabling more nuanced insights into student navigational behaviors. By effectively capturing and analyzing clickstream data, the framework facilitates a deeper understanding of learning processes, thereby aiding in the development of early intervention strategies to reduce dropout rates and enhance learning efficiency. This research marks a significant step forward in the predictive analysis of student performance, offering a robust tool for educators to tailor educational support and improve student engagement in online learning environments.

Gpt-3

The study focuses on developing a novel framework to enhance predictive learning outcomes for online students in Massive Open Online Courses (MOOCs) by leveraging clickstream data. Traditional models analyzing student data have limitations in capturing new behavioral patterns that indicate student performance. To address this, the researchers propose a click-based representation learning framework that integrates natural language processing (NLP) techniques with self-supervised learning. The framework involves a self-supervised clickstream pre-training setup followed by adjusting latent representations for downstream supervised learning tasks. Experimental results on real-world datasets demonstrate the effectiveness of the approach, with the combination of skip-gram embeddings and Principal Component Analysis (PCA) achieving high accuracy rates. Specifically, on the Xutangx dataset, the methodology achieved an accuracy of approximately 72.70% and an F1-score of approximately 81.03%, while on the KDDCUP dataset, it reached an accuracy of 80.91% and an F1-score of 87.42%. These results highlight the potential of NLP techniques in improving dropout prediction in MOOCs by extracting informative representations from clickstream data, enabling a deeper understanding of student behavior and facilitating early intervention strategies. Overall, the framework offers a promising approach to enhance learning efficiency and predict student performance in online courses.