Study, Attend and Predict:
Academic Performance Prediction using Transformers

Nicolas Araque, Mario Catapano, Massimo Di Gennaro

NEURAL INFORMATION PROCESSING SYSTEMS

LXAI

UNIVERSIDAD METROPOLITANA

Introduction and Motivation

- College students may have a set of possible courses to enroll, and lack of tutoring could result in failing grades.

- Predicting academic performance and dropout students enables institutions to help students make better decisions.

- The transformer architecture has made it possible to train models on sequence data and obtaining state of the art results.

- The objective of this research is to create a model that uses the transformer architecture for the academic performance prediction task to obtain state of the art results.

Dataset

- We collected a dataset containing 822 different students over a 7 year period from 2011 to 2017 from the System Engineering major of the Universidad Metropolitana.

- This dataset represents the academic performance history, containing the courses and grades obtained by these students in each term.

- The train/test split was carried out with 739 students for the train set and 83 students for the test set. With no overlapping of students from the differents sets
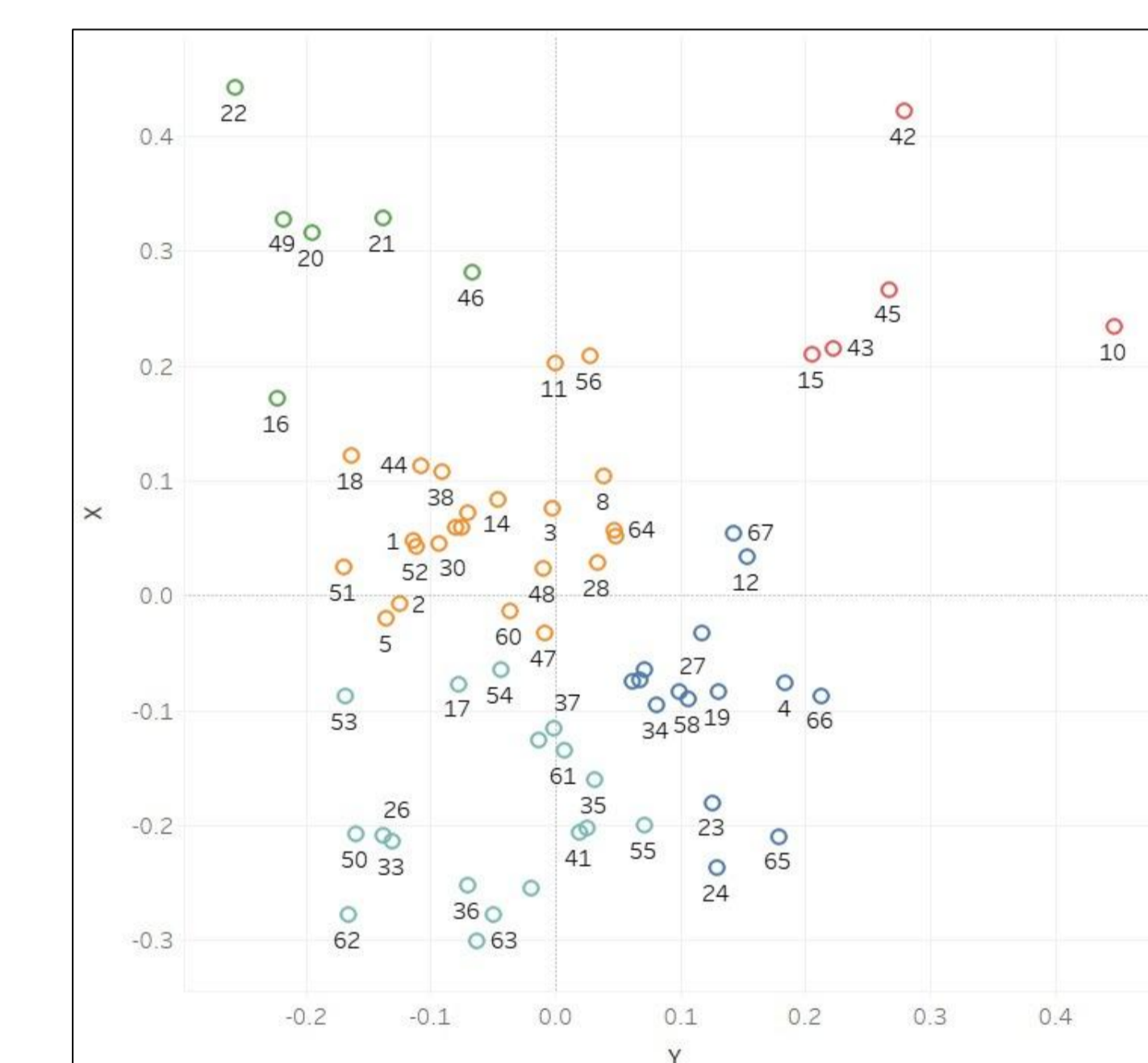
Model

- We use the encoder structure from the Transformers architecture, but discarding the positional encoding.

- We replace the embedding of words with a random embedding of courses and trained them.

Related Work

- M. Sweeney, H. Rangwala, J. Lester, and A. Johri, "Next-Term Student Performance Prediction: A Recommender Systems Approach," 2016. [Online]. Available: https://arxiv.org/pdf/1604.01840.pdf.

- N. Araque, G. Rojas, and M. Vitali, "UniNet: Next Term Course Recommendation using Deep Learning," arXiv.org, 2020. https://arxiv.org/abs/2009.09326).

- C. Wong, "Sequence based course recommender for personalized curriculum planning," inArtifcial Intelligence in Education, vol. 10948 of Lecture Notes in Computer Science, pp. 531–534,Springer International Publishing, 2018

Results

- Accuracy of 73.6 % and an F1 score of 81.26 % for the validation set.

- Confusion Matrix of the results:

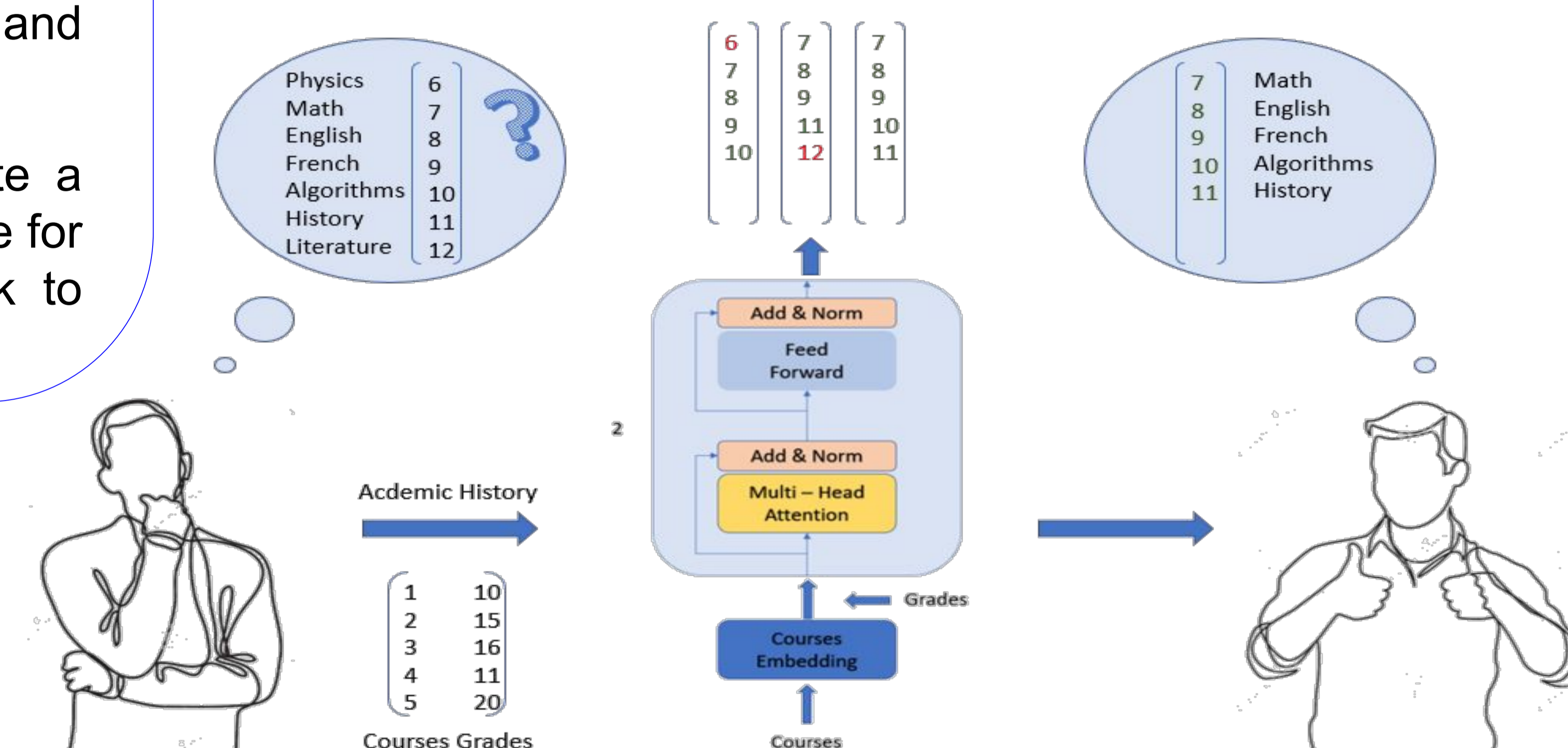|	Passed Grade	Failed Grade
Passed Grade Predicted	1359	423
Failed Grade Predicted	204	390

- 2D PCA of trained course embedding:

Analysis and Future Work

- Tthe Transformer architecture can be used to predict academic performance on a university dataset.

- Future works needs to be done in order to compare with a benchmark, to state the improvement with already existing methods.

- .We plan to incorporate positional encoding from the Transformers architecture to experiment on better accuracy of the model

- We are working in a pre trained course embedding to better represent the course input

-