Early Forecasting of Student Performance in STEM Using Large Language Models

Helen Martinez, Dr. Mohammad R. Hasan & Ahatsham Hayat
Department of Electrical & Computer Engineering, University of Nebraska–Lincoln

Introduction

Retention Rates
- More than ½ of STEM students leave their fields before graduation (Leary, et. al)
- Graduation rate is 20% below their counterparts in non-STEM majors (Leary, et. al)

Advising Centers
- Provide support after issues have been identified
- Traditional advising centers need a more tailored approach
- Need to address the individual needs of a student

Importance of STEM Education
- STEM Jobs will increase by 10.8% between 2022 – 2032 (Bureau of Labor Statistics)

Potential of Artificial Intelligence (AI)
- AI may effectively identify at-risk students early on

Purpose
- Traditional systems perform basic forecasting of cognitive data and early forecasting is nearly impossible with two weeks of data, for example.
- We aim to address this challenge by integrating additional background information.

Materials & Methods

Comprehensive dataset of college students' academic journeys over the Fall 2023 and Spring 2024 semesters

Experiential Data
- Qualitative measurements of non-cognitive attributes such as student motivation and engagement

Non-Experiential Data
- Quantitative analysis of cognitive attributes such as formative & summative assessment scores

All data was converted into text form and inputted into the FLAN-T5 language model for fine-tuning.

Dataset	Week	Stand. Dev	Mean
C	2	0.120291	0.642
C	3	0	0.73
C	4	0.026833	0.718
C+D	2	0.038341	0.842
C+D	3	0	0.87
C+D	4	0.031305	0.744
C+NC	2	0.058566	0.614
C+NC	3	0.032863	0.694
C+NC	4	0.032863	0.706
NC	2	0.038341	0.558
NC	3	0.026833	0.718
NC	4	0	0.6
NC+C+D	2	0.031305	0.814
NC+C+D	3	0.076811	0.64
NC+C+D	4	0.031305	0.786
NC+D	2	0.026833	0.718
NC+D	3	0.071204	0.652
NC+D	4	0	0.73

Figure 1

	precision	recall	f1-score	support
At the end of the semester, the student will be vulnerable.	0.00	0.00	0.00	4
At the end of the semester, the student will exhibit proficient performance.	0.73	1.00	0.85	11
accuracy			0.73	15
macro avg	0.37	0.50	0.42	15
weighted avg	0.54	0.73	0.62	15

Figure 2

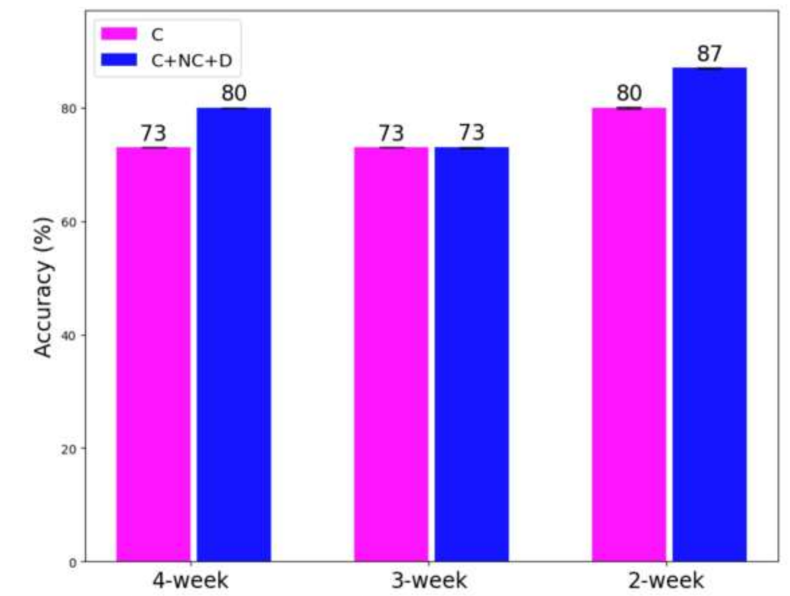

Figure 3

Results & Discussion

Figure 1
Experimented with 3 types of features: Cognitive factors (formative and summative scores), Distal factors (academic meta-information and socioeconomic info), and Non-Cognitive factors (answers from survey questions about engagement in the course).
Took the standard deviations of all the accuracies & calculated the mean.

Figure 2
Datasets were then run through Non LMs (Transformer, 1D CNN, and LSTM). Cognitive data only. Proficient is the majority class in our dataset, so even if they are predicting proficient all the time accuracy is 73%. Accuracy is worse when using only cognitive data.

Figure 3
LM had higher accuracy when predicting student performance when all 3 features were included

Conclusions

Promising results when non-cognitive and distal data is included. There is potential for early forecasting, but there needs to be more data and rigorous testing.

Future Work: Data collection & testing will continue Fall 2024 – Spring 2025

Acknowledgements

Dr. Mohammad R. Hasan, UNL Faculty Mentor
Ahatsham Hayat, UNL Graduate Student Mentor

MCNAIR SCHOLARS PROGRAM