Personalization and Contextualization of Large Language Models For Improving Early Forecasting of Student Performance

Ahatsham Hayat *M.R. Hasan*
Department of Electrical and Computer Engineering
University of Nebraska-Lincoln

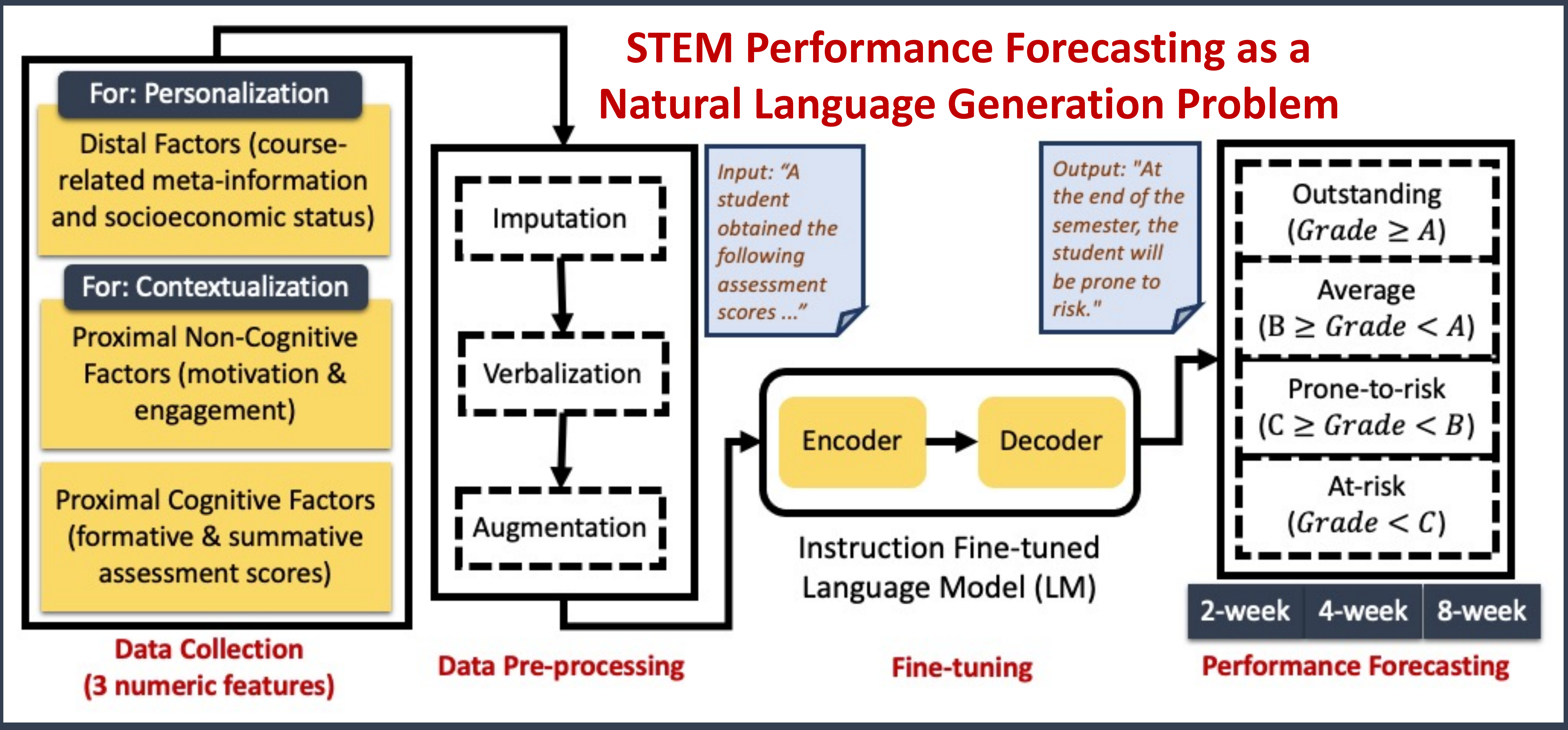
The 2023 Generative AI for Education (GAIED) Workshop
New Orleans, Louisiana
December 15, 2023

Key Ideas

- Leverage the general knowledge of pre-trained Language Models (LMs) for accurate early forecasting of college STEM performance with limited data (N=48).
- Contextualize and personalize LMs to improve performance forecasting.

Hypotheses

- **Personalization**: Distal factors correlate with academic trajectory, serving as a **useful prior for LMs**.
- **Contextualization**: Proximal non-cognitive factors being correlated with trajectory can capture subtle variations in performance.



Research Questions

- **RQ1**: Is the natural language generation approach more effective than numeric feature-based models in early forecasting of academic performance?
- **RQ2**: Do personalization and contextualization improve the LM's early forecasting efficacy?
- **RQ3**: How does the LM capacity (number of parameters) influence its forecasting performance?

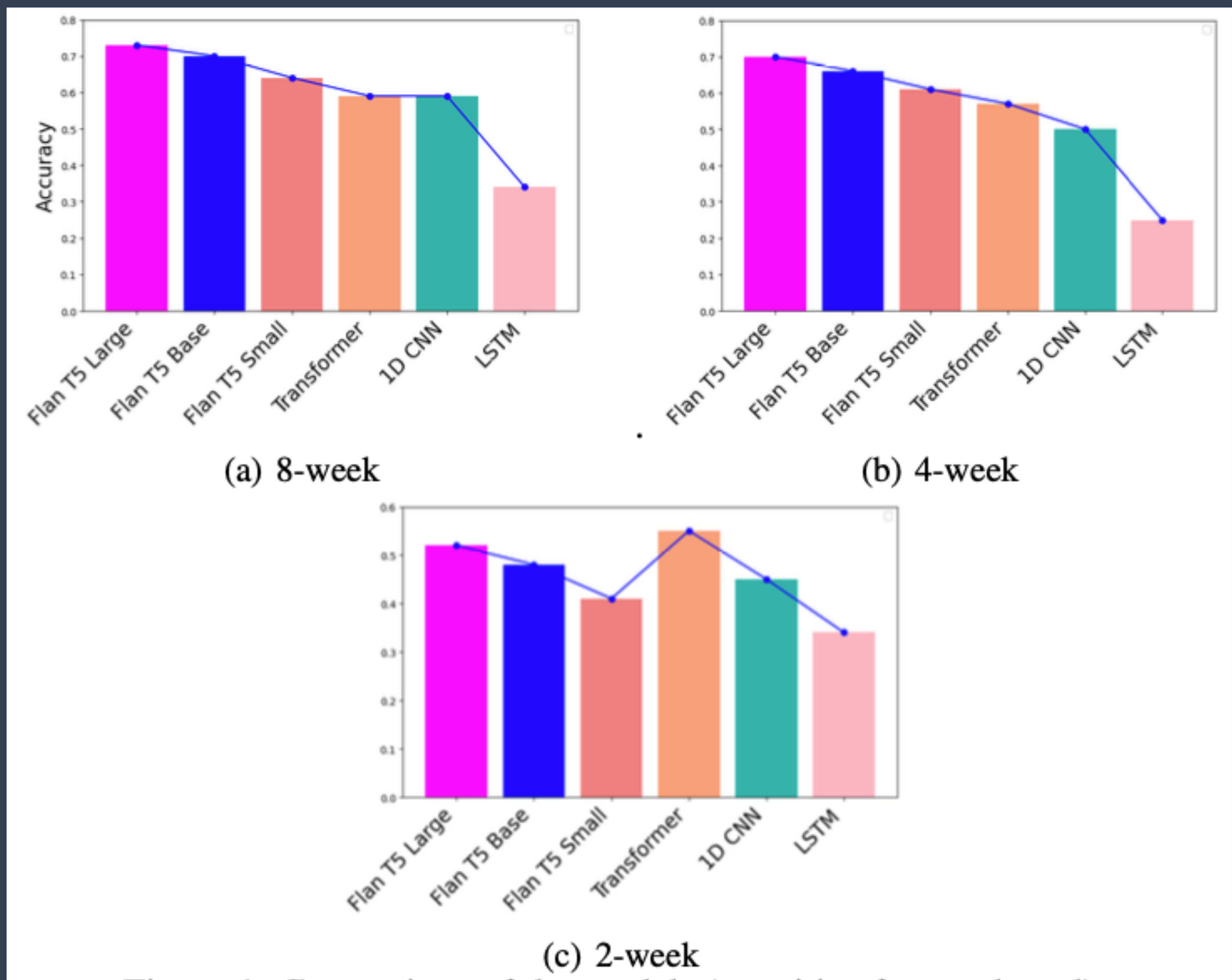


Figure 1: Comparison of the models (cognitive feature-based).

Table 1: Evaluation of the large LM (FLAN-T5-Large) fine-tuned with four combinations of the 3 feature types using the 8-week, 4-week, and 2-week datasets. The best results are in **bold**.
Legends: C=Cognitive, NC=Non-Cognitive, D=Distal, AR=At-Risk, PR=Prone-To-Risk, AV=Average, OU=Outstanding, P=Precision, R=Recall, F1=F1 Score, A=Accuracy

Features	Class	8-week				4-week				2-week			
		P	R	F1	A	P	R	F1	A	P	R	F1	A
C + NC + D	AR	0.78	1.00	0.88	**0.89**	1.00	1.00	1.00	**0.84**	0.64	1.00	0.78	**0.77**
	PR	0.89	0.80	0.84		0.89	0.80	0.84		1.00	0.50	0.67	
	AV	0.92	1.00	0.96		0.71	0.91	0.80		0.73	1.00	0.85	
	OU	0.93	0.81	0.87		0.86	0.75	0.80		0.85	0.69	0.76	
C + NC	AR	0.70	1.00	0.82	0.82	0.70	1.00	0.82	0.77	0.62	0.71	0.67	0.68
	PR	1.00	0.60	0.75		0.86	0.60	0.71		0.71	0.50	0.59	
	AV	0.73	1.00	0.85		0.69	1.00	0.81		0.62	0.91	0.74	
	OU	0.92	0.75	0.83		0.91	0.62	0.74		0.77	0.62	0.69	
C + D	AR	0.78	1.00	0.88	0.77	0.88	1.00	0.93	0.77	0.60	0.86	0.71	0.64
	PR	0.89	0.80	0.84		0.71	1.00	0.83		0.71	0.50	0.59	
	AV	0.67	0.73	0.70		0.69	0.82	0.75		0.70	0.64	0.67	
	OU	0.79	0.69	0.73		0.89	0.50	0.64		0.59	0.62	0.61	
C	AR	0.60	0.86	0.71	0.73	0.62	0.71	0.67	0.70	0.36	0.57	0.44	0.52
	PR	0.86	0.60	0.71		0.67	0.60	0.63		0.88	0.70	0.78	
	AV	0.60	0.82	0.69		0.67	0.91	0.77		0.54	0.64	0.58	
	OU	0.92	0.69	0.79		0.83	0.62	0.71		0.42	0.31	0.36	

Conclusion

- Personalization and contextualization significantly improve LM's forecasting capability.
- LM's general knowledge acts as an effective prior for learning nuanced patterns from the trajectory data.
- Larger LMs are the most performant.

Contact: aahatsham2@huskers.unl.edu
https://engineering.unl.edu/hasan/