42

Enhancing Student Success:
Predicting Early Dropout and At-Risk Students Using Advanced Time Series and Meta Modeling Approaches

Bala Gopinath Kuchibatla,
Sanketkumar Patel,
Kyung Jin Kwak,
Abhi Subramaniyam Kamuju,
Rajashekar Allam

*COLLEGE OF COMPUTING,
ILLINOIS INSTITUTE OF TECHNOLOGY*

BACKGROUND: Student dropouts and at-risk students' lower graduation rates and waste resources. Our study aims to proactively identify and support these students to improve retention and success.

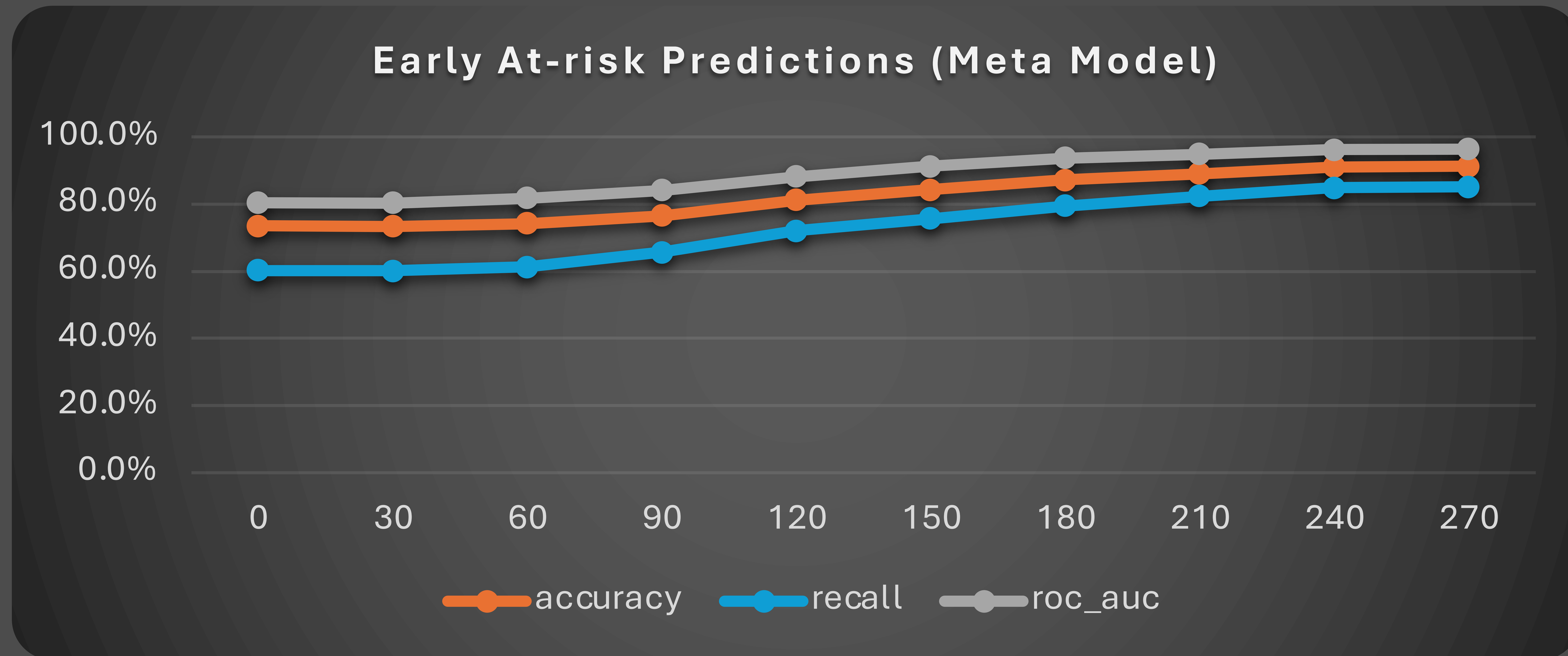
METHOD:
Data: Student demographics, behavioral interactions, assessment results - **OULAD**
Target: Students across various academic levels and institutions.
Models: Time Series Forest, Meta Model, Decision Trees, Random Forest, XGBoost, LightGBM.
Process: Implemented algorithms to predict early dropouts and at-risk students, using evaluation metrics like Recall, AUC, and Accuracy.

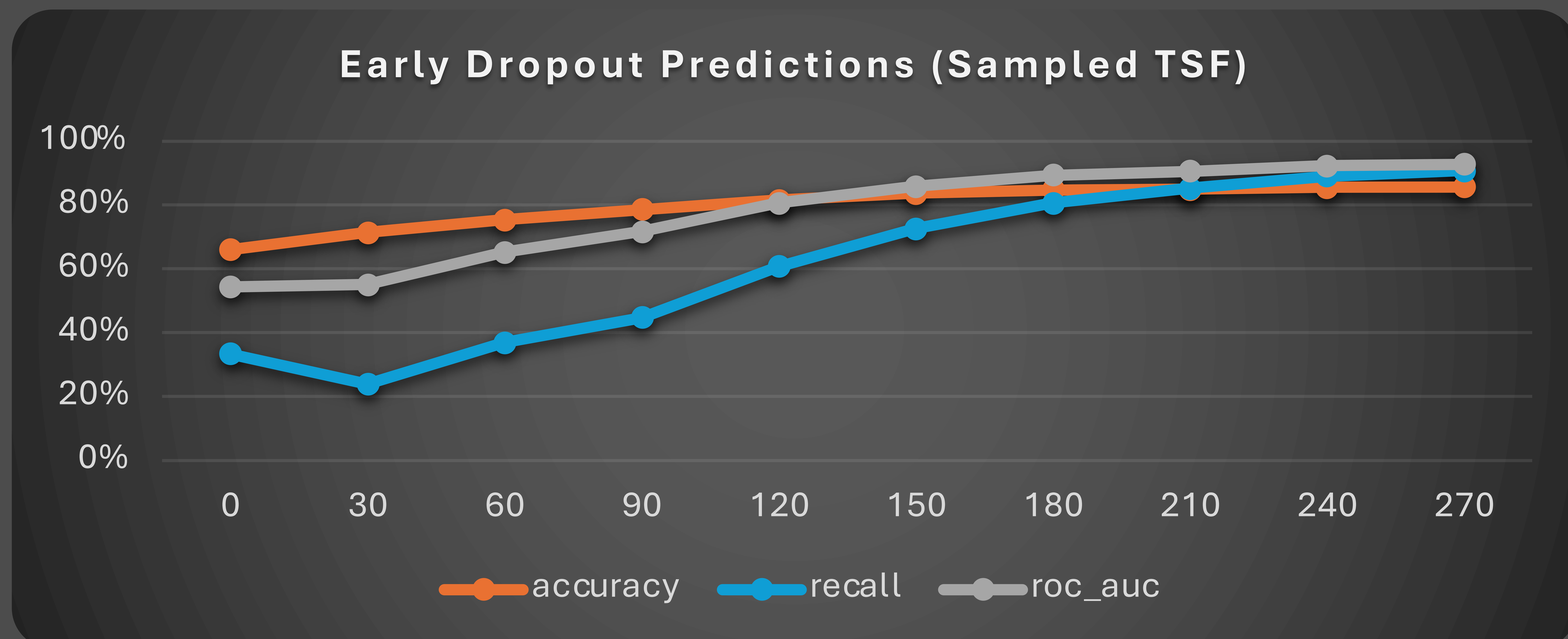
RESULTS:
Dropout Modeling:
Sampled TSF shows recall rising from 33.4% to 90.9% and AUC reaching 92.8%. It outperforms the Meta Model and traditional ML models.

At-Risk Modeling:
Meta Model leads with recall improving from 60.2% to 85.1% and AUC reaching 96.3%. It surpasses all other models, including traditional ML models.

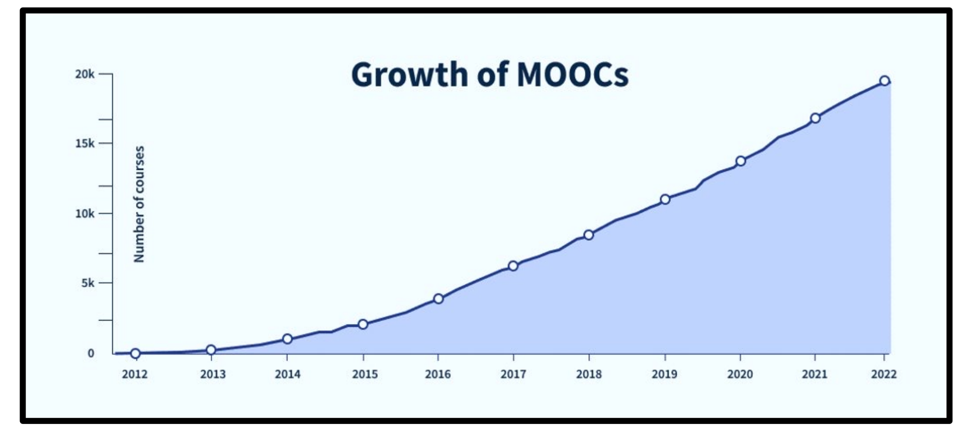
CONCLUSION:
Sampled TSF model helps schools proactively prevent student dropouts, while the Meta Model enables targeted support for students at risk of academic failure, enhancing overall educational outcomes.

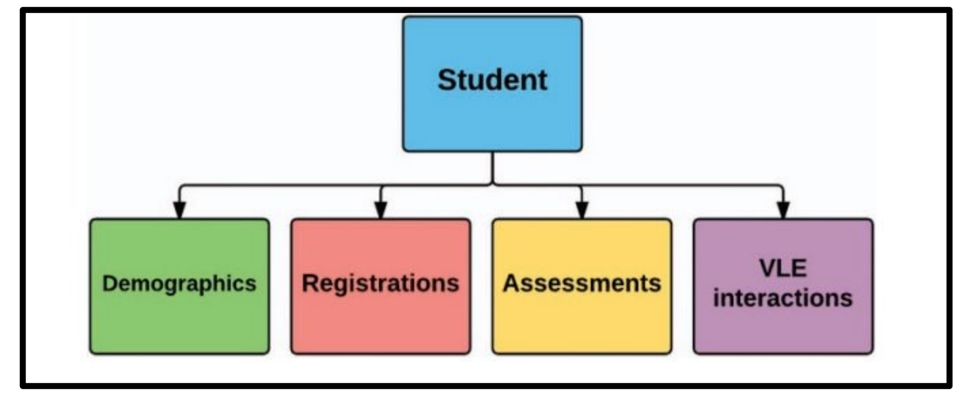
Time Series Forest (TSF) and Meta Model stood out as best performers compared to Traditional machine learning models like DT, LightGBM, RF, and XGBoost in predicting student dropouts and at-risk students.

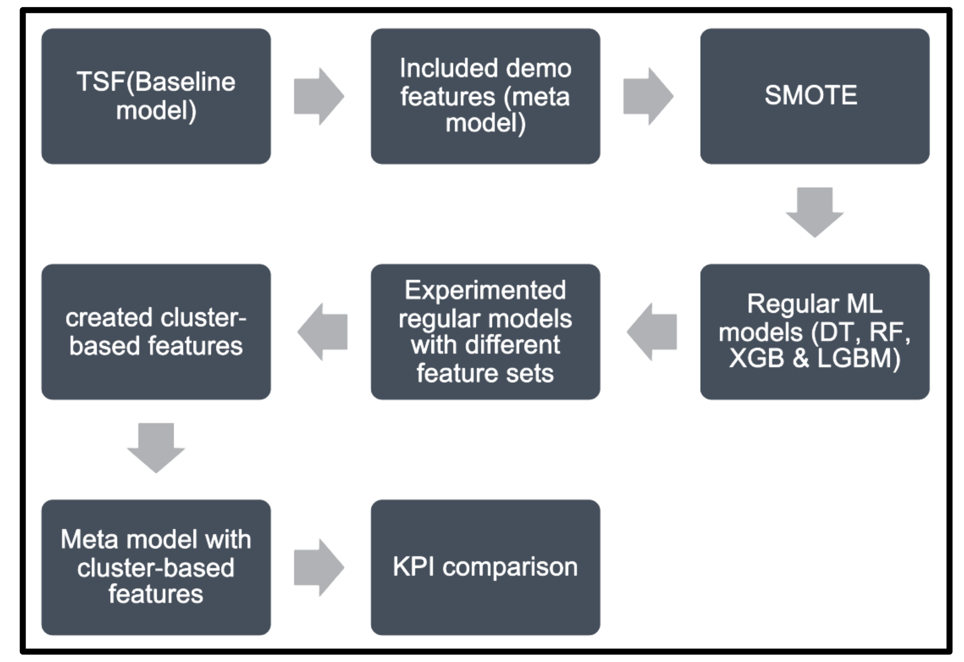




Acknowledgement:
This work is part of the Data Science Practicum, guided by Prof. Yong Zheng, whose insights and mentorship were invaluable to the project's success. We extend our gratitude for his continued support and expertise.







Dropout Modeling AUC analysis

Cuts	Baseline TSF	Sampled TSF	Meta Model	DT	LightGBM	RF	XGBoost
0	54.9%	54.4%	69.6%	53.0%	57.4%	56.6%	56.7%
30	63.5%	55.1%	69.6%	52.1%	60.1%	56.7%	59.1%
60	74.3%	65.3%	71.6%	52.8%	64.6%	59.4%	62.2%
90	78.8%	71.6%	75.1%	54.3%	65.4%	59.6%	63.5%
120	84.5%	80.7%	82.2%	54.5%	66.7%	59.7%	63.4%
150	88.0%	85.9%	86.7%	55.1%	66.8%	60.3%	63.3%
180	90.0%	89.3%	89.8%	56.3%	66.3%	61.5%	63.6%
210	91.1%	90.7%	91.0%	55.3%	65.1%	60.7%	62.6%
240	92.1%	92.4%	92.5%	55.6%	67.2%	61.9%	63.8%
270	92.2%	92.8%	92.9%	55.4%	69.1%	61.9%	65.8%

At-risk Modeling AUC analysis

Cuts	Baseline TSF	Sampled TSF	Meta Model	DT	LightGBM	RF	XGBoost
0	60.3%	56.1%	80.4%	54.6%	61.0%	62.2%	59.0%
30	69.4%	60.7%	80.4%	56.3%	66.1%	62.1%	64.7%
60	77.6%	69.3%	81.7%	58.3%	73.1%	67.1%	71.1%
90	81.5%	74.7%	84.1%	59.8%	74.7%	69.3%	73.3%
120	87.0%	82.2%	88.2%	61.4%	76.0%	71.4%	74.9%
150	90.4%	86.9%	91.2%	62.7%	76.8%	73.3%	75.7%
180	92.6%	90.2%	93.7%	63.2%	77.1%	73.8%	76.2%
210	94.1%	92.0%	94.9%	63.5%	76.3%	73.8%	76.4%
240	95.3%	93.9%	96.2%	63.8%	77.5%	74.2%	76.7%
270	100.0%	94.1%	96.3%	62.8%	77.1%	74.5%	76.0%

The Open University

ILLINOIS TECH
College of Computing