Using Machine Learning and Variable Importance to Support Program Development

Proposal for the Learning Analytics in Practice (LAP) Challenge 2024

Rishika Samala
rsamala@iu.edu

Dilip Francies
dfranci@iu.edu

Manali Shelar
mvhelar@iu.edu

Shiva Pendem
spendem@iu.edu

Joshua Quick
jdquick@iu.edu

Gina Deom
gdeom@iu.edu

Indiana University
Bloomington, Indiana, USA

Project Overview

Driven by the imperative to enhance student success and retention, leaders focused on student success are increasingly exploring statistical and machine learning-based insights. However, a significant challenge lies in bridging the gap between these complex insights and actionable knowledge for non-technical stakeholders. Current methods for analyzing variable importance, such as Shapley variable importance plots, permutation importance, tree-based feature importance and partial dependence plots, lack clear interpretability for non-experts. Additionally, accurately assessing a variable's overall contribution and its variability within the model remains elusive.

Key Challenges

Our team has identified three critical challenges at the intersection of machine learning, educational analytics, and decision support systems. These challenges will form the foundation of our LAP 2024 presentation and discussion:

Challenge 1: Interpretability for Non-Technical Stakeholders
How can we simplify the interpretation of machine learning insights for non-technical stakeholders, such as advisers, to facilitate informed decision-making?

Challenge 2: Variable Importance Methodologies
How can we develop methodologies for constructing variable importance metrics that accurately capture each variable's contribution to the model's performance?

Challenge 3: Model-Process Alignment
How do we ensure that machine learning models' insights and the modeled process align?

Proposed Approach

Our presentation will explore novel visualization techniques, hybrid variable importance metrics, and stakeholder-centered evaluation frameworks to address these challenges. We will demonstrate practical examples from our work with student success initiatives at Indiana University, showing how:

- Visual storytelling can transform complex machine learning outputs into intuitive insights for academic advisors
- Ensemble variable importance methods can provide more robust and reliable indicators of key factors affecting student outcomes
- Collaborative model validation can ensure that algorithms remain aligned with educational realities and stakeholder needs

---

Learning Analytics in Practice (LAP) Challenge 2024

June 19, 2024

https://www.solaresearch.org/events/lap/