Predicting MBA enrollment in Katz Business School

Introduction:

Katz Business School at the University of Pittsburgh seeks to increase MBA enrollment by identifying key factors influencing graduates' decisions to pursue an MBA. This study applies machine learning models to predict MBA pursuit and assess potential gender disparities.

Methods:

The dataset includes demographic, academic, and career-related factors. The target variable, decided_to_pursue_mba_yes, was analyzed using Logistic Regression and Random Forest Classifier.

Data Preparation

- Categorical variables were standardized and one-hot encoded.
- Numerical features were scaled.
- Interaction features were introduced based on feature importance, such as combining low salary with scholarship status.

Models

- Random Forest was used to determine feature importance.
- Logistic Regression (L1/L2 Regularization & Balanced Weights) was used for prediction.
- Performance was evaluated using accuracy, precision, recall, and F1-score.

Results:

- Logistic regression accuracy was 48.15%
- Key features encouraging MBA pursuit are networking, skill enhancement, and entrepreneurial interest.
- Key features discouraging MBA pursuit are Self-funding and certain post-MBA career paths.

Classification Report for Logistic Regression (Balanced Weights):

Class	Precision	Recall	F1-Score	Support
False (Not Pursuing MBA)	0.39	0.47	0.42	819
True (Pursuing MBA)	0.57	0.49	0.53	1181

MBA Pursuit Rate by Gender:

- Male: 58.80%
- Non-Male: 59.35%.

No gender bias noted in the model.

Discussion:

Some desired post MBA roles, such as finance manager and marketing director, appear to have negative correlation with MBA pursuit, which may not align with expectations. One possible explanation is hidden interactions between variables. For example, the model suggests finance managers are less likely to pursue an MBA, which might be due to entrepreneurs being more likely to enroll. Entrepreneurs are generally not interested in finance roles, the model may misinterpret this relationship, leading to misleading results. Similarly, self-funding may interact with career choices, further complicating predictions. These results suggest that MBA decisions depend on complex financial and career motivations beyond what the model captures.

Conclusion & Recommendations

- Networking and skill enhancement are major drivers of MBA pursuit
- Financial concerns (self-funding) and career aspirations may influence MBA decisions.
- Despite adjustments, model accuracy remains low (~48%), suggesting missing influential factors.
- Future improvements: Collect additional data, apply with alternative models such as neural networks to capture nonlinear relationships more effectively.
- Refine feature interactions.