

FeatureTeacher: Raising the Bar on Education

David Karapetyan

Introduction



Serves students ages
17 – 24 in **Bronx and
Brooklyn**

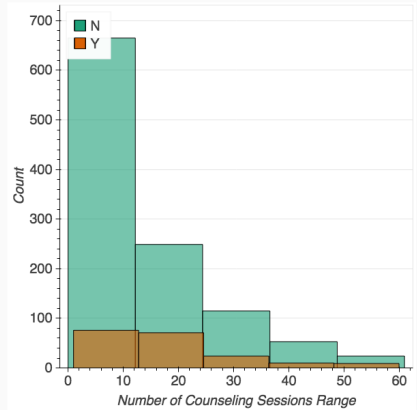
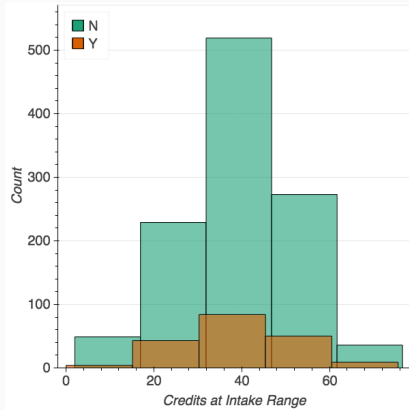
LifeLink program exists to
help students go to and
graduate from college

30 features per student

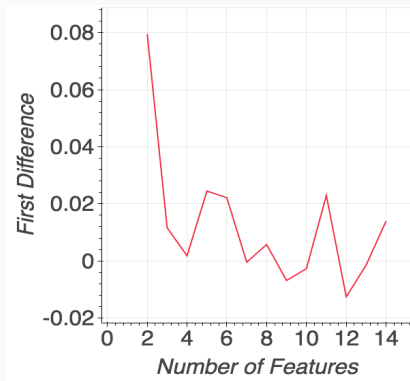
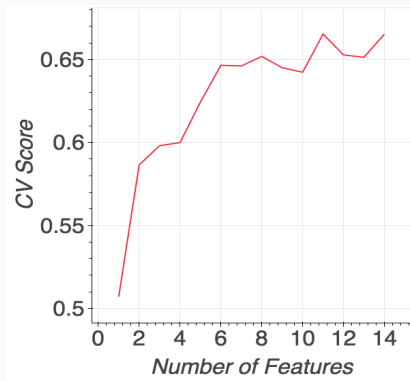
Which are most **correlated** with **LifeLink entry?**

How impactful is LifeLink on **graduation** rates?

Data Overview Using Histograms

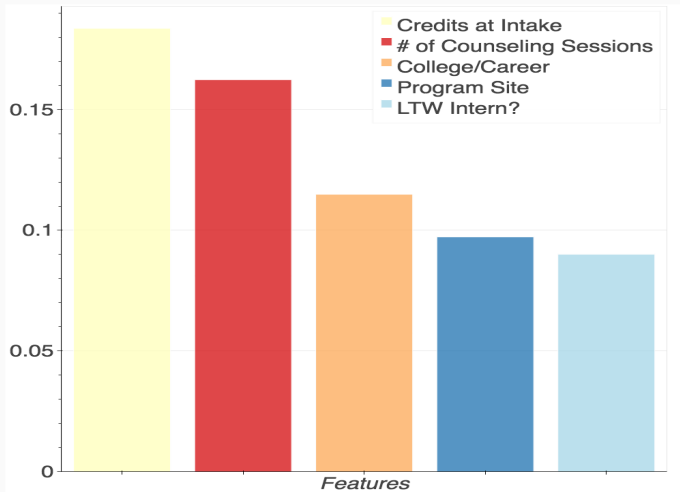


Algorithm (RFECV With Random Forests)



- 200 trees, with optimal parameters chosen via grid search
- Pruning at each step using 3-fold cross validation (with ROC-AUC as score)

Feature Importances



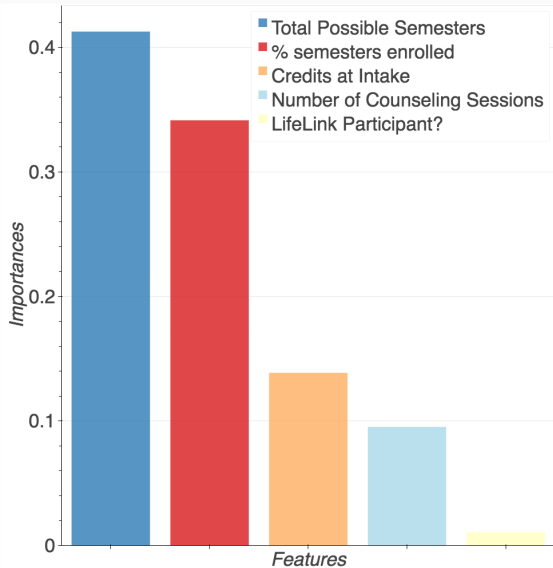
Conclusion

Using this information, we can take the top 2 features with respect to feature importances, i.e.

- Credits at Intake
- Number of Counseling Sessions

as the driving features (without significant degradation in model performance).

Does LifeLink Help Students Graduate?



Weight Balanced RF
Predicts Grad (Y/N)

recall: 37%

weighted: 96%

auc_roc : 90%

About Me



*City of Arts and Sciences
Valencia, Spain*

Before joining Insight, I was the head data scientist at Rudin Management.

I received my Ph.D in mathematics from the University of Notre Dame for work in analysis and partial differential equations.