



Machine Learning and the NFL Draft

Luke Newman

Introduction



- Use machine learning to predict college players with good probability of being drafted



ED OLIVER

HOUSTON

DT

Prospect Info

COLLEGE

Houston

HOMETOWN

TX

CLASS

Junior

HEIGHT

6' 2"

WEIGHT

287 lbs

ARMS

31 3/4"

HANDS

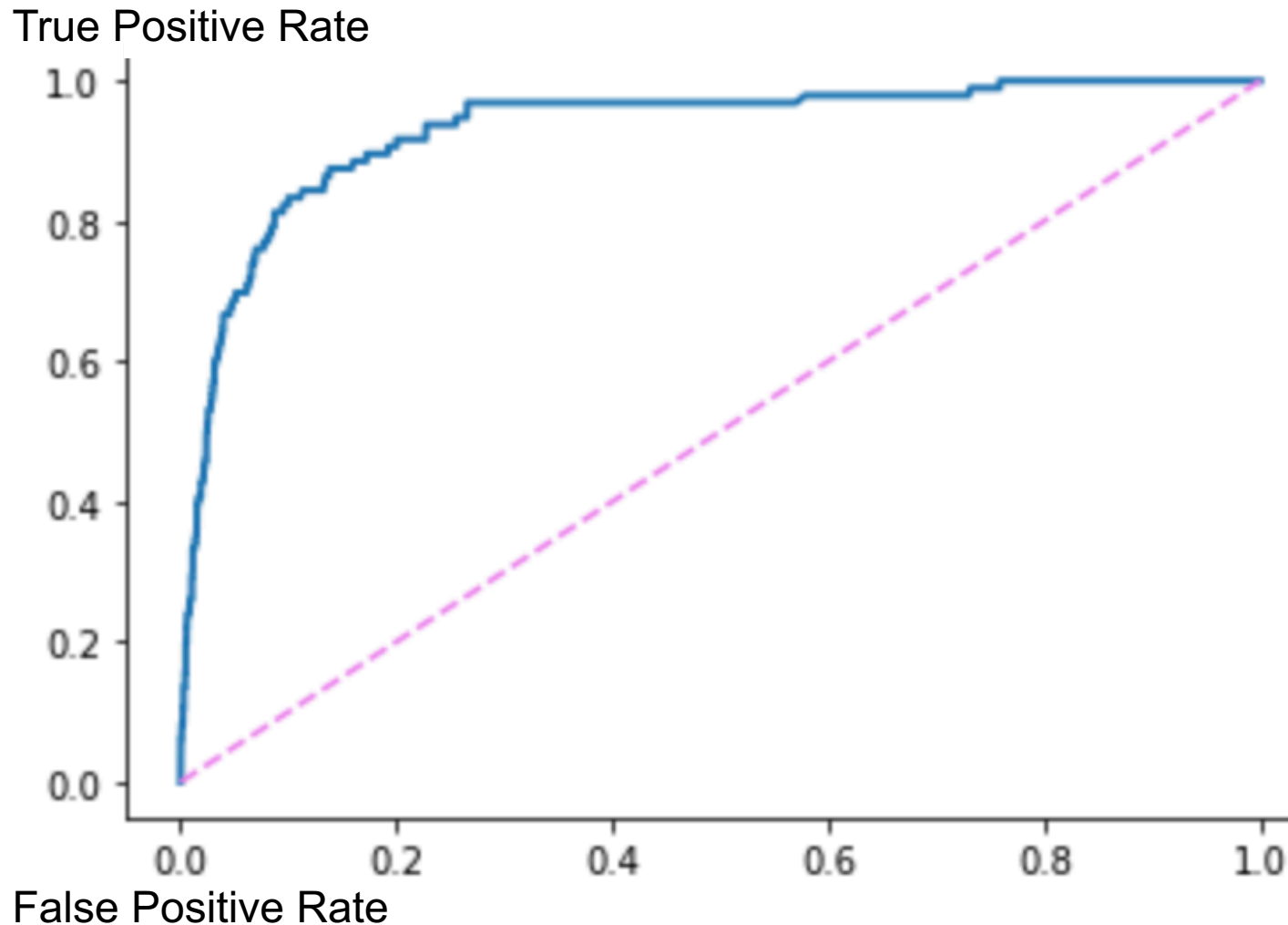
9 1/4"

Methodology



- Webscraped data
 - Cfbstats.com
 - Pro-Football-Reference.com
- Used ROC AUC score to measure performance of models
- Pandas, numpy, sklearn, imblearn, xgboost

ROC Curve



XGBClassifier	
ROC AUC Score	0.935

Confusion Matrix

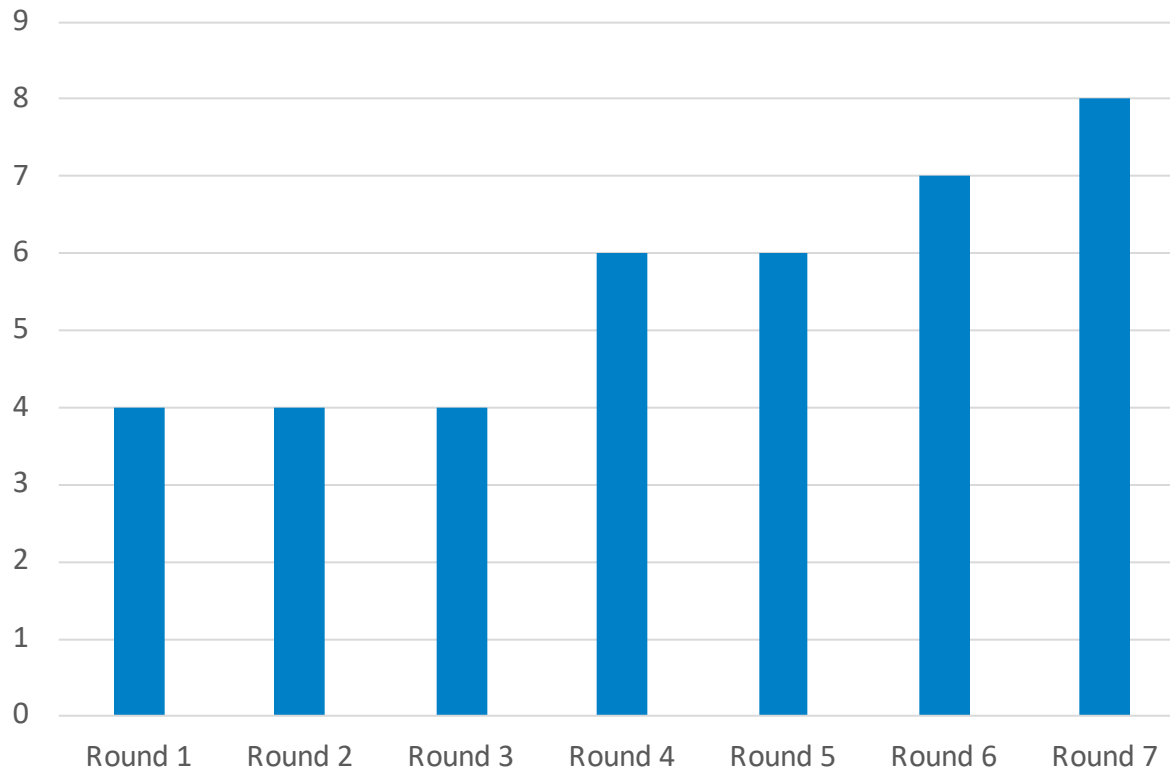


	Predicted Undrafted	Predicted Drafted
Actual Undrafted	3607	219
Actual Drafted	39	85

False Negatives



Drafted Players



- Drafted players who were predicted Undrafted
- Typically higher round draft picks

False Positives



Jr	43
Sr	174

- Undrafted players who were predicted drafted
- Potential late round draft picks
- Young future prospects

Conclusions



- Machine learning can help recruiters narrow down search
- Late round draft picks can be very crucial
- Tom Brady was a 6th round draft pick

Further Work



- Gather more data including new features
 - College
 - Conference
 - Wins
- Build models based on a given position



Questions?

Appendix



- Hyper parameter tuning can take a long time but is very important for models that don't come with good default settings. Using RandomizedSearchCV and GridSearchCV boosted performance
- Dealing with imbalanced data can be arduous. Using oversampling techniques help a lot but require a lot more model testing