

---

# Random Forest Classification Model

— NBA Win-Loss Prediction —  
Seasons 2002 - 2016

---

# Project Team

- Dragan Bogatic
- Howard Edwards
- Jordan Gross
- Ryan Reisner

## Project goal

- Make accurate predictions about NBA game Win-Loss based on historical game statistics using Machine Learning Model

# Objectives

- Research available sources of historical NBA game data
- Download data in standardized format
- Clean the data
- Select model features
- Select prediction variable
- Transform data to format to fit machine learning model

## Objectives cont'd

- Select prediction model
- Train and test the model
- Evaluate prediction performance of the model
- Fine-tune model hyper-parameters to improve model prediction
- Evaluate results
- Conclusions

# Resources

- Data source: AWS Data Exchange / Stats Perform
- Data: NBA Seasons 2002 - 2016
- Data format: CSV
- Dataset type: detailed play by play stats by game played
- Dataset size: approx. 8 million rows by 326 columns

# Machine Learning Model

- Model selected: Balanced Random Forest Classifier
- Reduces overfitting
- Proven performer in classification prediction problems
- Works with categorical and continuous values
- Uses Ensemble Learning technique (many weak learners strong together)

# Model Features and Prediction

- Total features in raw dataset: 325
- Features selected for the base model:
  - HT\_POSS
  - GAME\_TIME
  - HT\_SCORE\_DIFF
  - HT\_ORBD
  - HT\_DRBD
  - HT\_TOV
- Prediction: OUTCOME (Win-Loss)



# Model Accuracy

**Balanced Accuracy Score : 0.8892515660382707**

# Confusion Matrix

	Predicted 0	Predicted 1
Actual 0	107355	13065
Actual 1	19194	150662

# Classification Report

	<b>f1-score</b>	<b>precision</b>	<b>recall</b>	<b>support</b>
<b>0.0</b>	0.869380	0.848328	0.891505	120420.000000
<b>1.0</b>	0.903295	0.920203	0.886998	169856.000000
<b>accuracy</b>	0.888868	0.888868	0.888868	0.888868
<b>macro avg</b>	0.886338	0.884265	0.889252	290276.000000
<b>weighted avg</b>	0.889226	0.890385	0.888868	290276.000000

# Most Important Features

