Making predictive modeling of NBA scores via

fastAI-bets.com

co-founder : Dyas Utomo and Sarah Kessler

business incubator : the Erdos Institute
inspired from Susquehanna Investment Group

Team 2

https://github.com/sarahjaynekessler/ErdosBootCampSIGproject

Business opportunity

Facts: The total amount of points scored during any NBA game can be attributed, somewhat, to a series of factors that are known in advance such as the starting lineup. This suggests that we can predict the probability of final score.

Opportunity: Create a machine learning classifier that predicts whether the total score of a NBA game is over or under a given threshold value.

Business value: NBA is one of favorite sports in America and world wide with average viewer of about 1 million people per season in the last 3 years. Fanatic fans love to make bets on their favorite team. Even for only 10% of market share with $1 fee per prediction, we make $ 120 million in revenue per regular season.

What fastAI-bets does?

VS

=

Features we used:

- Team name
- Starting players
- Bench players
- Referee names
- Dates in the season
- Time in the day
- Travel distance
- Temperature
- Attendance
- Recent win/lose
- Recent offensive performance
- Recent defensive performance

~~Overtime or not?~~

How fastAI-bets generates its prediction?

Web scraping

BASKETBALL REFERENCE

WEATHER UNDERGROUND

OpenCage

Feature engineering

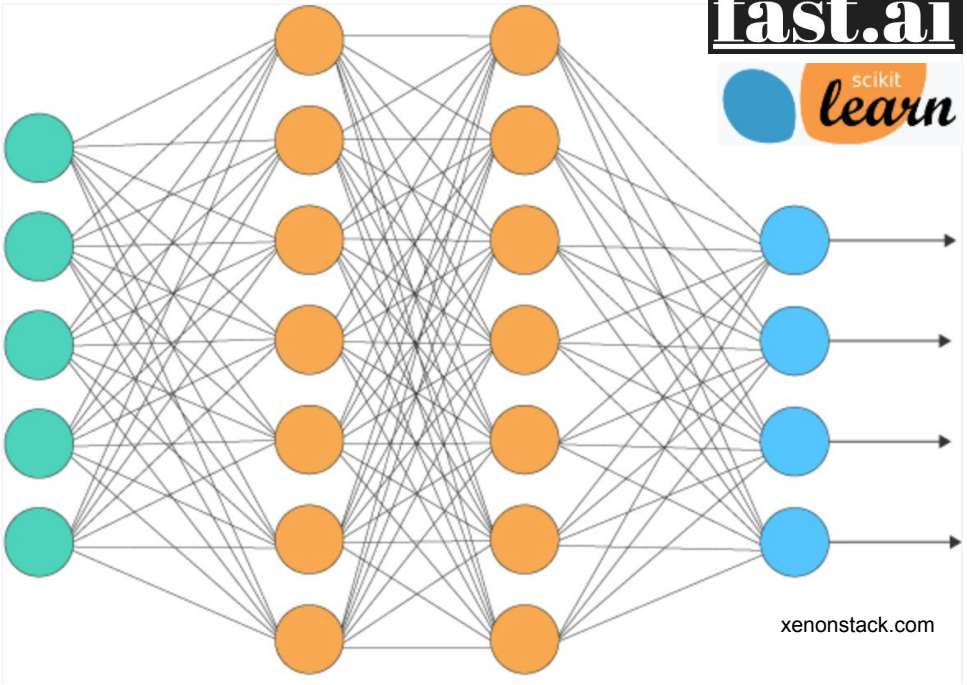
1. Time series analysis for past performance
2. One-hot-encoding for teams and referees
3. Expected score based on player stats
4. Give meaning to dates and times

LASSO, Random Forest, Neural Network

H$_2$O.ai

fast.ai

scikit learn



How accurate is fastAI-bets prediction?

	LASSO regression	Random Forest	Ensemble method	Neural Network
Accuracy	7%		73%	78%
AUC	--	0.83	0.93	0.84
Advantages	Feature selection		Fast, Best AUC	Fast, Simple, Best Accuracy
Drawbacks	Didn't work		Lower accuracy than fastai	AUC is lower than Ensemble Method

Next Steps

1. Expanding data to past year seasons (more rows)
2. Expanding more features (more columns)
3. Rerun the machine learning for different value of total scores (i.e. continuous prediction)
4. Exploring various architecture of neural network (how deep the layer is, number of perceptron per layer, different activation functions, etc.)
5. Seamless interaction via app: given the data just before the tip off, the machine gives probability of total score > threshold.
6. Business expansion to other sports: NFL, MLB, and MLS.

Thank You!

TEAM 2:

fastAI-bets.com

https://github.com/sarahjaynekessler/ErdosBootCampSIGproject

Dyas Utomo
dyas.utomo@gmail.com

Ohio State University

Sarah Kessler
sarahjaynekessler@gmail.com

Ohio State University