NBA Salary Predictions

David Keyer

Background

- NBA Contracts and Player Salaries
  - For mainstays, very lucrative $\Rightarrow$ almost $10,000,000/year in 2017 among players in the league for five years
- TV Deal
  - Close to $25 Billion
  - How is that relevant to player contracts?
    - Rise in "salary cap"
- As time goes on...
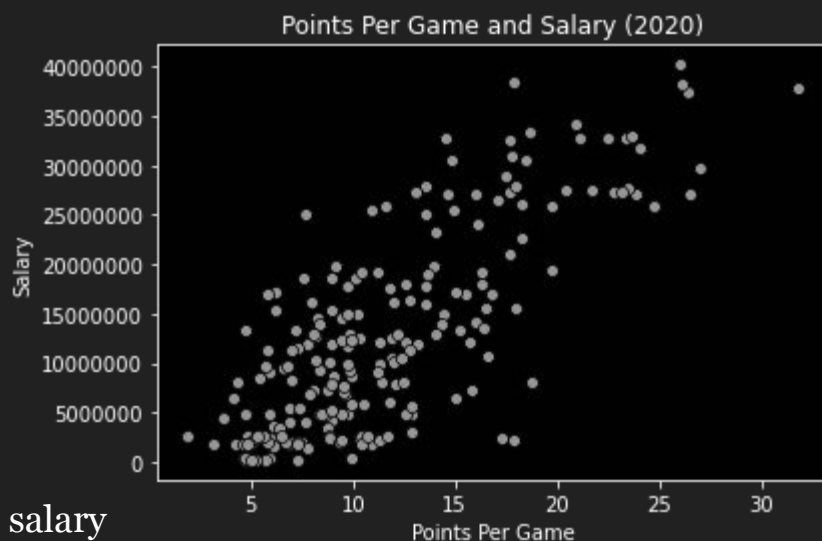 - Bigger player contracts

Approach

- Is it possible to accurately predict an NBA player's salary based on his statistics?
  - In-game statistics: Points Per Game, Rebounds Per Game, Assists Per Game, etc.
  - Other: Age
- TV Deal/shift in money ⇒ Something to consider
  - Monitor changes in model performance from year to year

Process

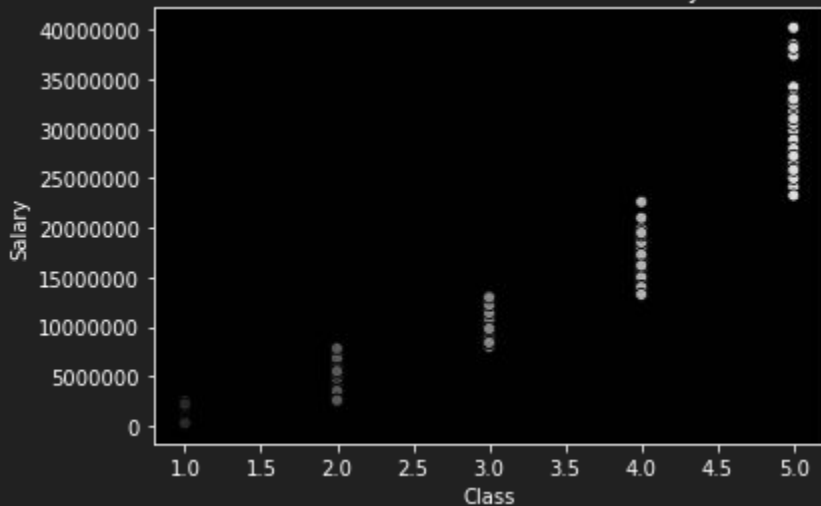Points Per Game and Salary (2020)

- Data collection, cleaning, and analysis
  - Dropping/filling nulls
  - Exploring trends
- Decisions
  - Merged statistics from previous five years with salary
  - Worked primarily with three datasets
    - 2017, 2018, 2020
- Originally started as a regression problem
  - Evolved into a classification problem
  - Large discrepancy in player salaries
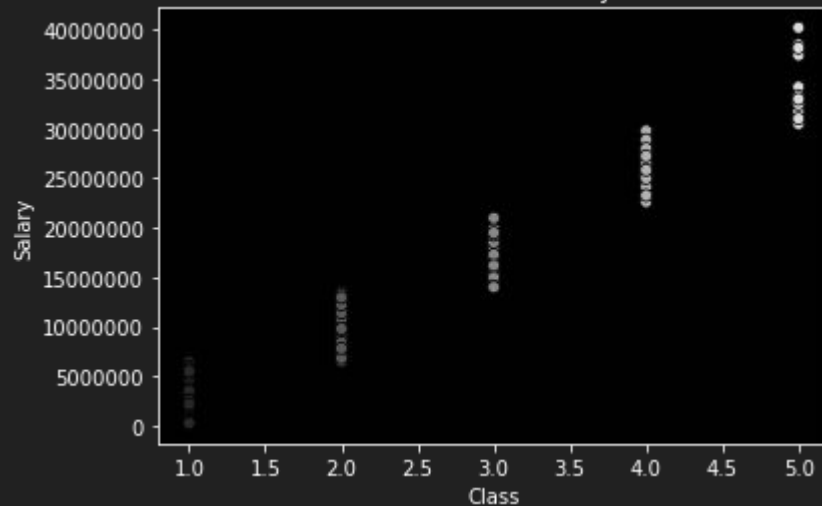 - Players not paid equally each year of contract

THE PROCESS

Modeling

- Datasets used
  - 2017 ⇒ Average player salary/year = $9.6 million
  - 2018 ⇒ Average player salary/year = $11.4 million
  - 2020 ⇒ Average player salary/year = $12.9 million

16-20 dataset classes set Manually

16-20 dataset classes set by Clusters

Modeling, continued

- Performance
- Accuracy Score
  - Best model(s)
  - 2017: Accuracy score: 50%
  - 2018: Accuracy score: 51%
  - 2020: Accuracy score: 60%
- Gridsearch
- Feature Importance
- Why did the model misclassify?
- What can we expect going forward?

Strengths and Weaknesses

Strengths:

- Increasing predictive power -- 60% improves significantly on baseline
- When misclassifying, real-world implications

Weaknesses:

- Amount of data
- Regression analysis difficult $\Rightarrow$ outliers, and how contracts are structured
- Could improve on 60%, that is the expectation

Recommendations/Conclusions

- If someone is in a contract year, can use the model to see if he should get more or less than current salary
- Likewise, front office management can consider it before bringing in a player
- Expectation: As money increases, accuracy should rise, be able to better predict which class a player is in
- Several models that improve on baseline

Next Steps

- See how models generalize to the future
- Uncertainty
  - How will the pandemic affect the season and salaries?
- If salaries keep rising, will it continue to be easier to predict salary class?
  - Timeseries

Questions?

- https://thunderousintentions.com/2020/09/15/okc-thunder-unlock-steven-adams-talents/
- https://www.essentiallysports.com/nba-news-portland-trail-blazers-damian-lillard-expresses-desire-to-have-skills-from-chicago-bulls-michael-jordan-los-angeles-lakers-lebron-james-and-brooklyn-nets-kevin-durant/
- https://clutchpoints.com/sixers-news-joel-embiid-sports-slides-with-the-process-written-on-them/
- https://bleacherreport.com/articles/2838515-report-clint-capela-traded-to-hawks-in-4-team-blockbuster
- https://6abc.com/espn-nba-tnt/338452/