

Dave Han

DS 210

14 Dec 2024

Final Project Write Up

Source: <https://www.kaggle.com/datasets/justinas/nba-players-data>

Project Write-Up: Impact of Team Performance Metrics on NBA Success

1. Project Overview

This project analyzes the relationship between key team performance metrics and team success in the NBA. Using a dataset of 1000 rows, which includes aggregated team statistics such as average points, assists, and rebounds, the project builds a predictive model to estimate team success (measured as win percentage). Additionally, the project generates visualizations to highlight trends and relationships in the data.

The primary goals of the project are:

1. To identify key performance metrics that correlate with team success.
2. To visualize team performance trends and relationships between metrics.
3. To predict team success based on these metrics, providing actionable insights.

Steps to Run

1. Clone the Repository: Ensure the project folder structure matches the provided guideline:

project/

└─ src/

```

| | — main.rs      // Main program
| | — visualizations/ // Folder for visualization modules
| | | — mod.rs      // Module entry point
| | | — win_trend.rs // Visualization for Win Percentage Trend
| | | — scatter.rs   // Visualization for Points vs. Win Percentage
| — Cargo.toml      // Rust project configuration

```

2. Outputs: The program will:

- Save two visualizations in the project folder:
- `win_percentage_trend.png`: A trend line of win percentages.
- `points_vs_win_percentage.png`: A scatter plot showing the relationship between points and win percentage.
- Print predictions and evaluation metrics (e.g., Mean Squared Error) to the console.

3. Output Description

Visualizations

Win Percentage Trend (`win_percentage_trend.png`):

- A line chart showing how win percentages evolve across the dataset.
- Practical Use: Highlights overall trends in team success across the dataset.

Points vs. Win Percentage (`points_vs_win_percentage.png`):

- A scatter plot showing the relationship between team average points and win percentage.
- Practical Use: Reveals whether scoring more points correlates with higher success.

Multilinear Regression

- Predictions: The program predicts win percentages for each row based on team metrics.
Example: `[0.5636, 0.5237, 0.5621, ...]`
- Evaluation (Mean Squared Error): Measures the accuracy of predictions. A low MSE (e.g., `0.0218`) indicates a good fit for the model.

4. Interpretation of Results

Key Insights

- The regression model demonstrates that team average points, assists, and rebounds are strong predictors of win percentage.
- Visualizations corroborate these relationships, showing clear trends and positive correlations.

Practical Use

- For Teams: Insights from the model can guide training and recruitment strategies. For instance, if assists have a strong correlation with success, teams can prioritize playmaking.
- For Analysts: Provides a framework to analyze team performance and predict success rates.
- For Fans: Offers an accessible way to understand how specific metrics drive team outcomes.

Limitations:

- The model does not account for external factors like injuries, trades, or coaching changes.
- Predictions are limited to the metrics in the dataset and may not generalize to unseen conditions.

5. Future Improvements

- Expand Dataset: Add more features (e.g., shooting efficiency, turnovers) for a comprehensive analysis.
- Train-Test Split: Evaluate the model's performance on unseen data to assess generalization.

Significance: This project successfully establishes a data-driven framework for analyzing NBA team success, demonstrating clear relationships between key metrics and outcomes. It ultimately predicts each team's success for the future seasons.