

# Title: Building a Naïve Bayes Model

Subtitle: Predicting NBA Player Career Longevity Using Naïve Bayes

## ➤ ISSUE / PROBLEM

This project simulates the National Basketball Association (NBA) interests in identifying players who can endure the high-pressure environment and contribute to team success over an extended period. This challenge developed a predictive model that determines whether an NBA player will have a career lasting at least five years, using historical performance data.

## ➤ IMPACT

- **Strategic Decision-Making:** NBA teams can use the model's insights to make data-driven draft and player development decisions.
- **Talent Retention:** Improved player selection and management strategies could enhance long-term team performance and player retention.
- **Operational Efficiency:** Reducing the uncertainty around player career longevity could lead to better resource allocation for teams and coaching staff.

## ➤ RESPONSE

This project employs a **Naive Bayes classification model**, specifically the **Gaussian Naive Bayes algorithm**, to predict player career longevity. The model was chosen based on the assumption that the dataset features are normally distributed and continuous. The workflow involved **data preparation, feature engineering, model training, and evaluation** to ensure a robust predictive framework.

## ➤ KEY INSIGHTS

- The **Gaussian Naive Bayes model** was selected due to its suitability for continuous data despite the potential violation of strict normality assumptions.
- The **train-test split** resulted in a **75% training and 25% test dataset**, ensuring a balanced learning process.
- **Model evaluation** demonstrated a strong concentration of correct predictions in the confusion matrix, highlighting the model's predictive reliability.
- Feature selection and engineering played a crucial role in improving the model's accuracy, reinforcing the importance of **data preprocessing** in machine learning applications.

