C960 Task 2 Molina

May 23, 2024

Adam Silver, CTO

SNC Inc.

123 Easy Drive

St. Louis, MO

Dear Mr. Silver,

I am pleased to submit this project proposal for your consideration. As a Sports News Company™, it is understandably your goal to provide your customers with the most accurate and up to date sports news and stats. With AI and machine learning in the news, it would be a great time for your company to leverage the power that they hold.

The proposal outlines a machine-learning-based data product aimed at enhancing the analysis of basketball players' statistics for future performance predictions. For instance, if a player like T. Haliburton from the Indiana Pacers averages 18.5 points per game, the algorithm will provide a likelihood of the player scoring above or below his average in future games.

The clients will benefit from having the top-tier information regarding player performance, game outcomes and likelihoods, allowing SNC. Inc. to create better content for its viewers and have increased credibility as a news platform. The sportscasters will have their decision-making process improved by being given additional insight into future games/players.

The funding requirements will be $7000 for the initial upfront cost to cover the compensation for the chosen ML engineer. The ongoing costs for maintenance and upkeep of the infrastructure that the program runs on will be $2000 a year.

Sincerely,

Abraham Molina

Senior Director

Molina Software Solutions

Project Proposal: Predictive Analysis for Basketball Player Performance

**Summary of the Problem**

Sports News Company (SNC Inc.) seeks to enhance its analytical capabilities to provide deeper insights into basketball player performance. The current system lacks the advanced analytical tools needed to meet client demands for detailed performance predictions.

**Benefits of the Data Product**

The proposed data product will:

* **Improve Accuracy**: Utilize advanced machine learning algorithms for precise performance predictions.
* **Support Decision-Making**: Offer detailed insights to aid strategic decisions.
* **Increase Credibility**: Position SNC Inc. as a leader in sports analytics.

**Outline of the Data Product**

The data product will:

* Analyze historical player performance data.
* Predict future points scored utilizing feature analysis using stats like assists, steals, rebounds, etc.
* Feature an simple interface for result interpretation.

**Description of the Data**

The data product will use:

* **Historical Player Data**: Stats from past games, including shooting percentages, rebounds, assists, and points. In addition, the expected average of the players points will be added alongside the data so that the learning algorithm can have even more insight into why the player beat or missed their average
* **External Factors**: Data such as game location (home/away) and minutes played.

**Objectives and Hypotheses**

* **Objective**: Develop a model that accurately predicts if a player will score above or below their average points.
* **Hypothesis**: Advanced features and machine learning will improve prediction accuracy over traditional methods.

**Outline of the Project Methodology**

**Understanding -** First, I'll get a clear idea of what SNC Inc. wants. They need an algorithm to predict if a basketball player will score above or below their average points per game. This helps them in making better predictions and providing more insightful content for their viewers.

**Gathering and Cleaning -** Next, I'll gather all the data we can get on the player's stats. This includes points, assists, rebounds, minutes played, shooting percentages, and details about the games (like whether they were home or away), and I’ll clean this data to make sure there are no errors or missing values.

**Exploring the Data/Features –** I will create some sample visualizations so that I can see patterns and trends. Then, I will choose features to account for like minutes played, shooting percentages, etc.

**Building/Testing –** I will use a logistic regression for this project that will predict the probability of the player scoring above or below their average points. The features mentioned previously will serve as additional feature data that will affect the outcome of the test. For example, if player T. Haliburton is a consistently high scorer, but his rebounds + assists are climbing and they have a negative correlation, that can impact his predicted score.

**Evaluation/Delivery –** Once the model is trained, I will evaluate its performance by performing back-tests. Finally, I will deliver the algorithm to SNC. Inc. along with the associated documentation and reporting as necessary.

**Funding Requirements**

* **Initial Cost**: $7,000 for development.
* **Annual Maintenance**: $2,000.

**Impact on Stakeholders**

* **SNC Inc.**: Better decision-making and competitive advantage. Opportunity to lease access to key insights and analytics.
* **Players and Teams**: Insightful performance trends and improvement areas.
* **Fans and Analysts**: Better viewing experience with more accurate insights from sportscasters.

**Ethical and Legal Considerations**

* **Data Privacy**: Ensure data is anonymized and securely stored.
* **Transparency**: Clearly communicate how predictions are made.
* **Compliance**: Adhere to data protection regulations.

**Relevant Expertise**

Molina Software Solutions (MSS) has a proven track record in data science and machine learning. Our team has delivered successful projects across various industries, and we are confident in our ability to deliver this project effectively.

We look forward to working with SNC Inc. to develop this innovative data product and enhance our analytical capabilities.