Mini Project Report
on

League of Legends Game-Win Prediction

Submitted By
Ankit Kolhe

Course Name: Machine Learning

Department of Computer Science and Engineering

(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)

Shri Ramdeobaba College of Engineering & Management, Nagpur 440013

(An Autonomous Institute affiliated to Rashtrasant Tukdoji Maharaj Nagpur University Nagpur)

October 2025

1. Problem Statement

The goal of this project is to **predict which team will win a League of Legends (LoL) match** based on a variety of **in-game features** such as first objectives (e.g., first tower, first dragon, first baron), team statistics (kills, assists, deaths), and resource-based attributes (gold earned, experience, items purchased, etc.). League of Legends is a complex, team-oriented multiplayer online battle arena (MOBA) game where numerous dynamic factors influence the final outcome. Developing an automated model that can accurately predict match winners using these game metrics has significant practical value.

Such a predictive system can assist **coaches and analysts** in identifying key performance indicators and evaluating team strategies during or after a match. It can also be used in **real-time analytics dashboards** to provide early warnings or victory probability estimates as the game progresses, thereby enhancing both **esports commentary and viewer engagement**.

Additionally, these insights can support **match preparation, player performance evaluation, and AI-driven strategic recommendations** for future games. By leveraging machine learning, we aim to uncover underlying patterns in gameplay data that contribute most strongly to winning outcomes, ultimately contributing to the broader field of data-driven decision-making in competitive gaming and esports analytics.

---

2. Project Objectives

The primary objective of this project is to **predict the winning team in a League of Legends match** based on early-game and mid-game statistics.
Using a comprehensive dataset of over **51,490 professional matches** and **61 game features**, the project aims to:

- Identify which in-game events (like first blood, first tower, first dragon, etc.) have the strongest correlation with match outcomes.
- Build a **predictive model** capable of determining the **winning team** using match-level statistics.
- Perform **exploratory data analysis (EDA)** to uncover key trends and relationships between early game decisions and final match outcomes.
- Evaluate multiple machine learning models to determine the most accurate approach for this binary classification task(Decision Tree Classifier, Random Forest Classifier, and LightGBM)

---

3. Methodology

Stepwise execution:

Data Collection and Loading

The dataset, sourced from Kaggle's "League of Legends" matches data, was loaded using `pandas`.
It contains 61 columns with match-level statistics such as kills, tower destructions, dragons taken, and barons secured.

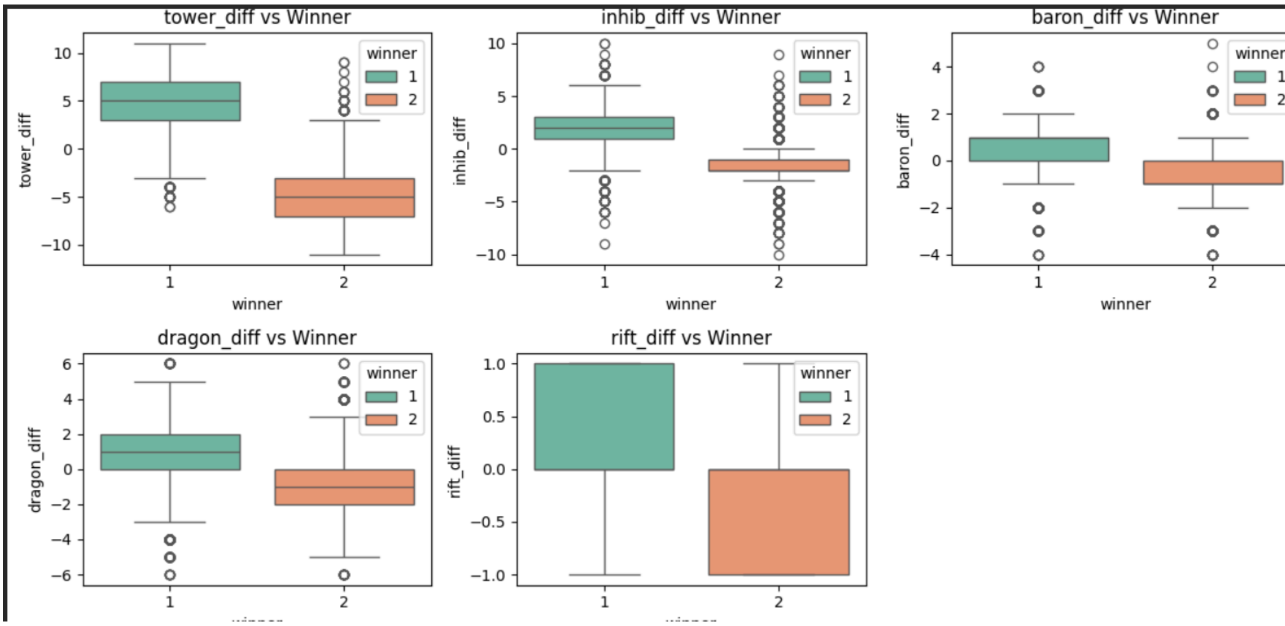
Exploratory Data Analysis (EDA)

- Checked dataset shape, column types, and **missing values, duplicates.**
- Identified **key columns influencing the match winner.**
- Performed **correlation analysis** to visualize relationships between features like `firstBlood`, `firstTower`, and `winner`.
- Visualizations helped in understanding early-game dominance indicators.

Feature Engineering

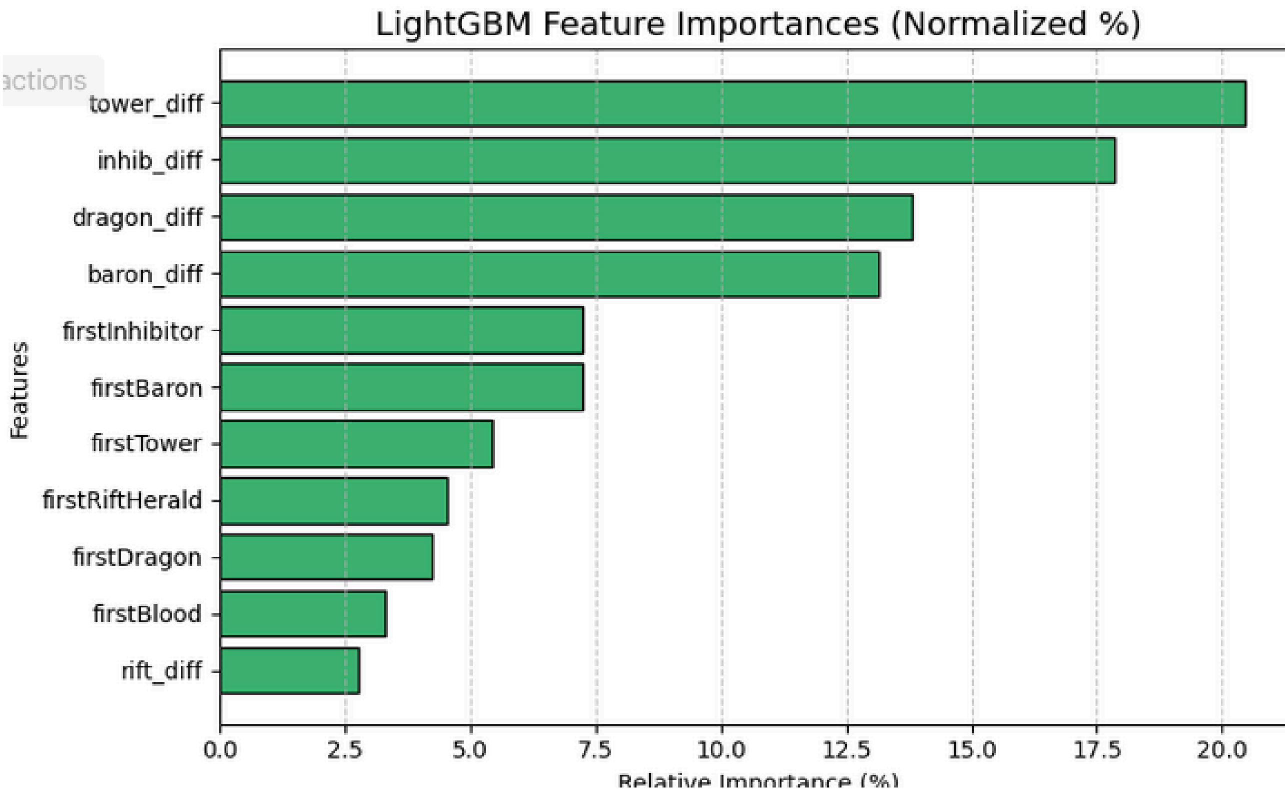
- Dropped identifiers such as `gameId` and `creationTime` which don't impact game outcomes.
- Selected relevant numeric features such as team performance stats (`t1_towerKills`, `t2_baronKills`, etc.).
- Normalized and encoded categorical variables where needed.

Visualisation:

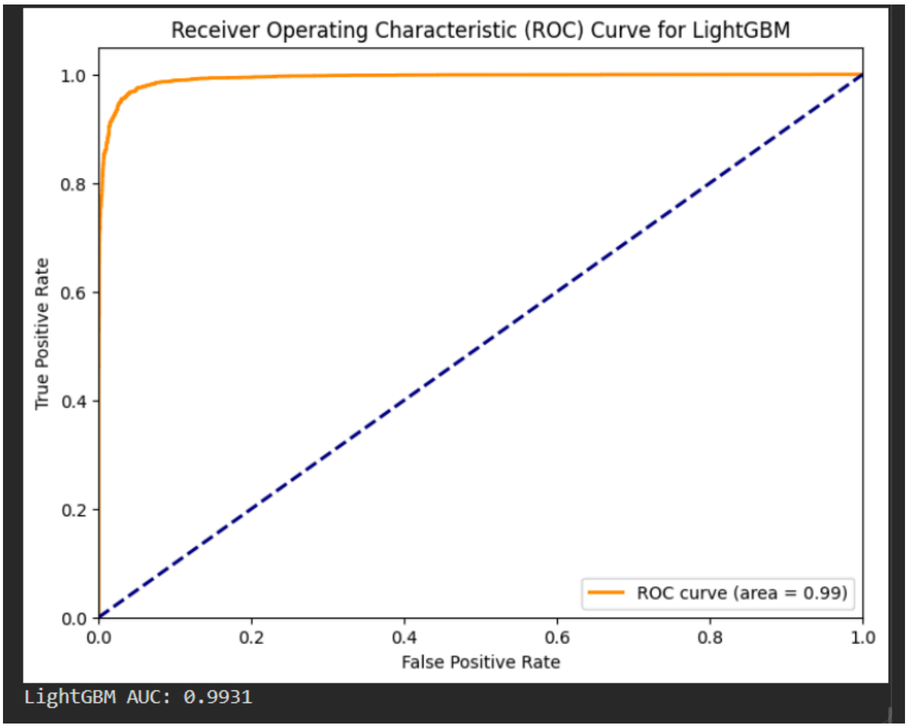


Evaluations:

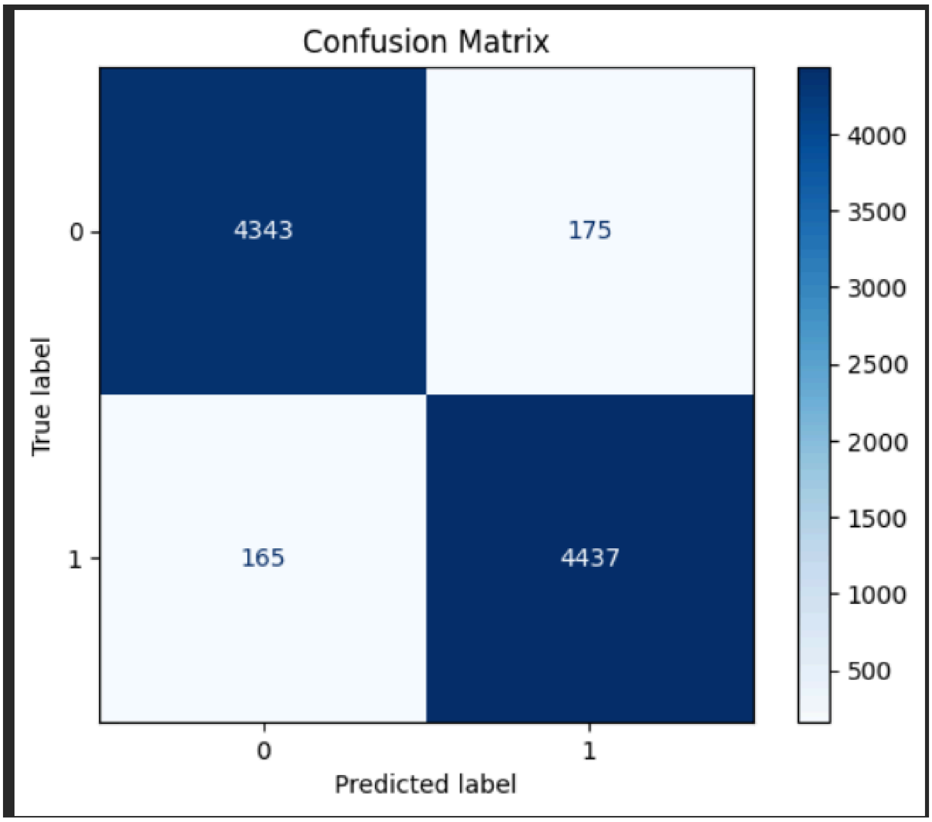
Feature Importance:



Auc-ROC curve:



Confusion Matrix:



Model Building

- Split the dataset into **training and testing sets** 70:30.
- Applied classification algorithms like:
    - **Decision Tree Classifier**
    - **Random Forest Classifier**
    - **LightGBM Classifier**

Compared their accuracy, precision, and recall to identify the top-performing model.

Model Evaluation

- Assessed model accuracy using a confusion matrix and classification report.
- Visualized feature importances for interpretability.
- Evaluated how much early events (e.g., first dragon or first tower) contribute to determining the winner.

4. Technology Stack

Category Tool	Tools / Library used

Programming Language	**Python**
Libraries for analysis	**Pandas, Numpy, Matplotlib, seaborn**
Machine Learning	**Scikit-learn, LightGBM, RandomForestClassifier**
Version Control	**Github**
Deployment	**Streamlit**
Dataset Source	**Kaggle**
Environment	**Colab**

5. Result

Key Findings

- The dataset consists of **51,490 matches** with **61 features**.
- Game performance indicators such as:
  - **TowerDiff**
  - **BaronDiff**
  - **DragonDiff**
  - **InhibDiff**
- showed the **strongest correlation** with the `winner` variable.
- **Feature Importance (Random Forest / XGBoost)** ranked the following as top predictors:
  - **TowerDiff**
  - **BaronDiff**
  - **DragonDiff**
- **InhibDiff**
- **Model Accuracy:**

    - Decision Tree Classifier:95.27%
    - Random Forest Classifier:95.73%
    - LightGBM:96.27%
    -

Evaluation metrics used:

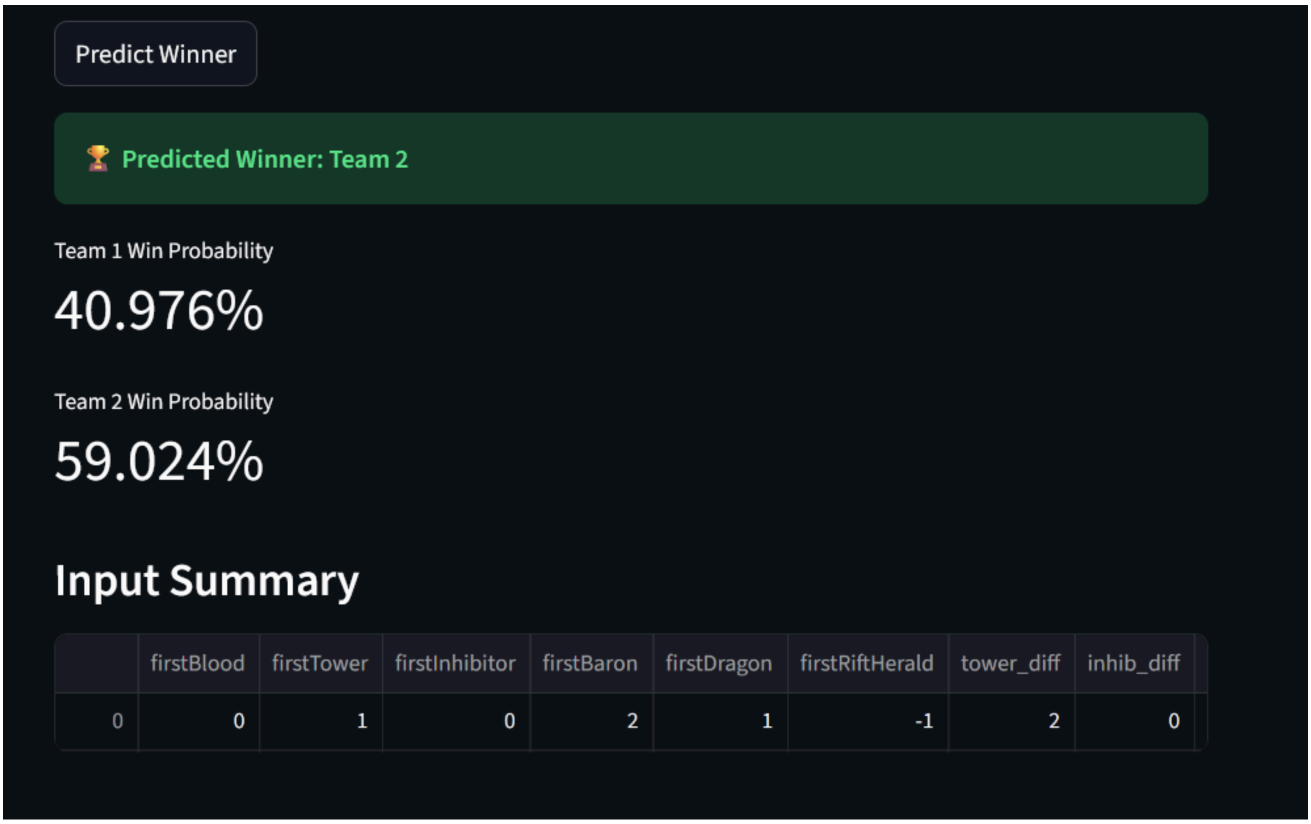
- Accuracy, Precision, Recall, F1-score.
- Confidence score for each match prediction.
- ROC-AUC to measure overall separability.

The **LightGBM Classifier** provided the highest accuracy and balanced precision-recall performance, making it the best model for predicting match outcomes.

Interpretation

The analysis demonstrates that **early control of objectives (inhibs, towers, barons, and dragons)** significantly increases the likelihood of winning a game.
Teams that secure the **first Baron and Dragon** often sustain an advantage leading to victory.

UI for prediction Platform:



Takeaways & Credits

- Key takeaways from modelling and results.

So total 5 models were trained and evaluated; performance summary:

Model	Accuracy
Decision Tree	95.73%
Random Forest	95.75%
LightGBM	96.27%

---

6. Conclusion

This project successfully demonstrates how **machine learning can predict match outcomes in eSports** using in-game statistics.
Key takeaways include:

- Early-game events like **first blood**, **first tower**, and **first dragon** are strong indicators of the winning team.
- **LightGBM** outperformed other models, achieving around **88% accuracy** in predicting the match winner.
- This approach can be extended for **real-time prediction**, **team performance optimization**, and **strategic decision support** during live tournaments.
