KAZIRANGA UNIVERSITY
KNOWLEDGE & BEYOND

# League of Legends Data Analysis Project Synopsis


**Team members:**

- Veer Raj Singh - ET21BTHCS020
- Kuldeep Chamuah - ET21BTHCS071
- Sudarshan Biswakandali - ET21BTHCS056
- Debasish Phukan - ET21BTHCS022
- Nishant Prahar Saikia - ET21BTHCS078

## 1. Introduction

League of Legends (LoL) is one of the most popular multiplayer online battle arena (MOBA) games, where two teams compete to secure objectives and claim victory. Understanding the factors that influence match outcomes can provide valuable insights for players, analysts, and esports teams. This project focuses on analyzing match data using machine learning models to predict game results based on in-game statistics.

## 2. Objectives

To analyze key performance metrics that influence match outcomes.
To implement machine learning models for predicting game results.
To compare model performances and identify the most effective approach.
To provide data-driven insights into team strategies and gameplay patterns.

## 3. Methodology

**Data Collection & Preprocessing:** The dataset was loaded, cleaned, and encoded for machine learning models. Missing values were handled, and numerical features were standardized.

**Exploratory Data Analysis (EDA):** Visualizations and statistical analysis were used to identify patterns and relationships among key game attributes.

**Model Training & Evaluation:**

- Supervised learning models, including SVM, KNN, Decision Tree, and Random Forest, were trained.
- The dataset was split into 80% training and 20% testing to evaluate model performance.
- Accuracy, confusion matrices, and classification reports were used for assessment.

**Results Interpretation:** The most influential factors in match outcomes were identified.

## 4. Results

SVM and KNN achieved the highest accuracy, indicating that non-linear decision-making and proximity-based classification were effective.
Random Forest and Decision Tree performed similarly, capturing important data patterns but requiring further tuning.
Key winning factors included Gold Earned, First Blood, and First Tower, highlighting the importance of early-game advantages.
The Blue Side had a slightly higher win rate, suggesting potential map-based advantages.

## 5. Challenges & Limitations

Feature selection was challenging due to the variety of in-game factors.
The dataset did not account for real-time game dynamics or player decision-making.
Some models required hyperparameter tuning for optimal performance.

## 6. Conclusion & Future Work

The study demonstrated that machine learning models can effectively predict match outcomes based on in-game statistics. SVM and KNN were the most effective, while Random Forest and Decision Tree showed promise with further optimization. Future improvements can include real-time match data analysis, deep learning models, and additional game mechanics to enhance predictive accuracy.