Exploring StarCraft II

Siddharth Byale

Outline

- Introduction
- Dataset
- Exploration and Visualization
- Model
- Summary
- Future Work

Introduction

Why is this Important?

The primary objective is to develop a **predictive model for player rank** data from StarCraft II ranked games.

Understanding the key factors that influence a player's rank can **provide valuable insights** for players looking to improve their performance, coaches **developing training strategies**.

Dataset

Structure of Dataset

Size : Approx **3,400** rows

Features: **19 unique attributes**

↳ Age, Hours/Week, Actions/Min , Map Explored, Workers Created, Complex Abilities Used, etc.

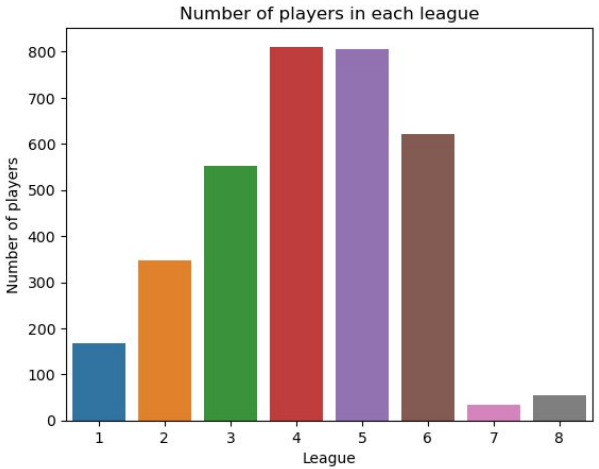
Player Rank: **1-8 League Index** (Bronze to Professional)

Initial Observation

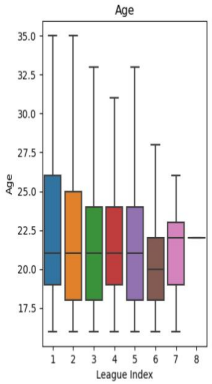
General **increasing trend** can be observed between leagues

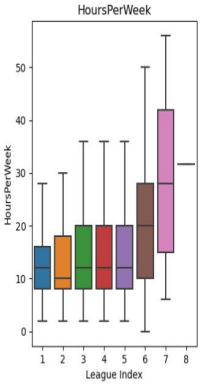
Age, Hours/Week and Total Hours **missing data** for Professional Players

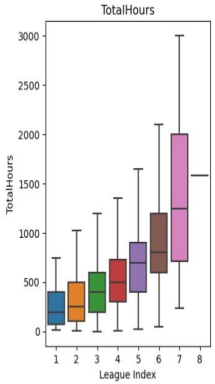
Lack of data on higher and lower ranks

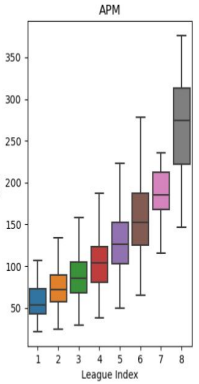


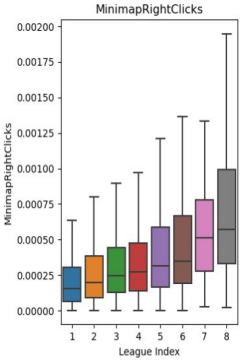
Exploration & Visualization

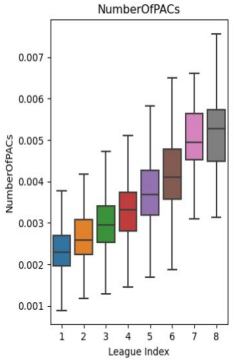


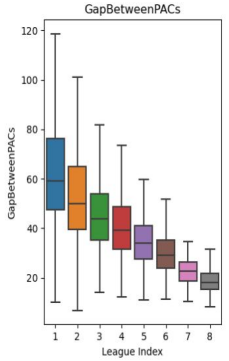


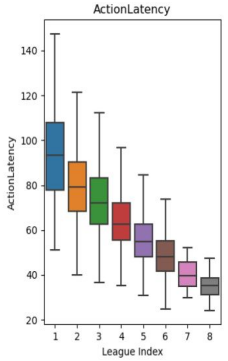


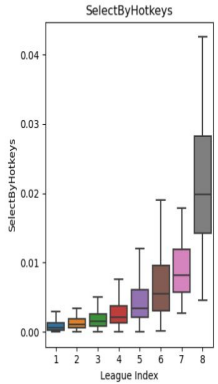


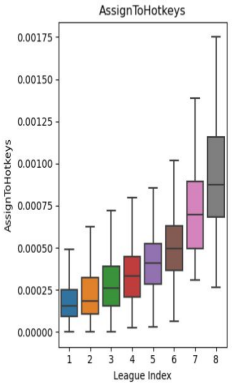


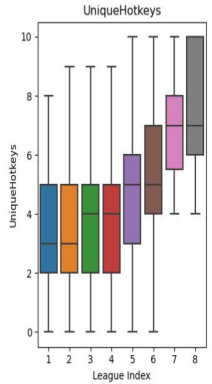


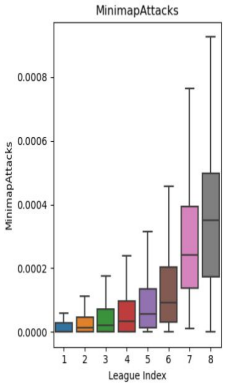


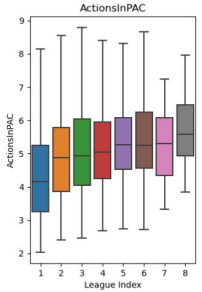


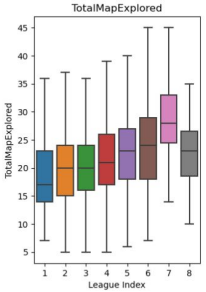


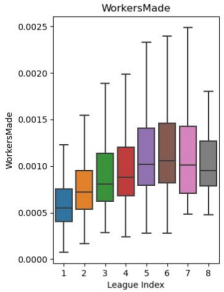


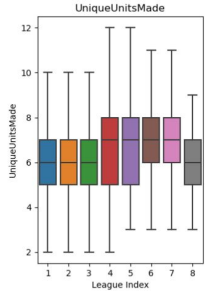


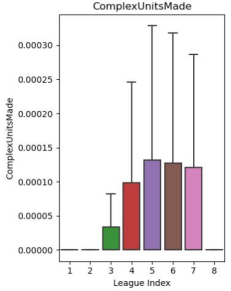


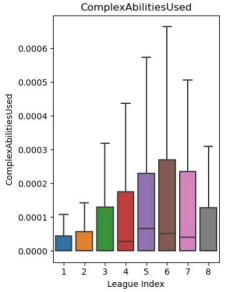




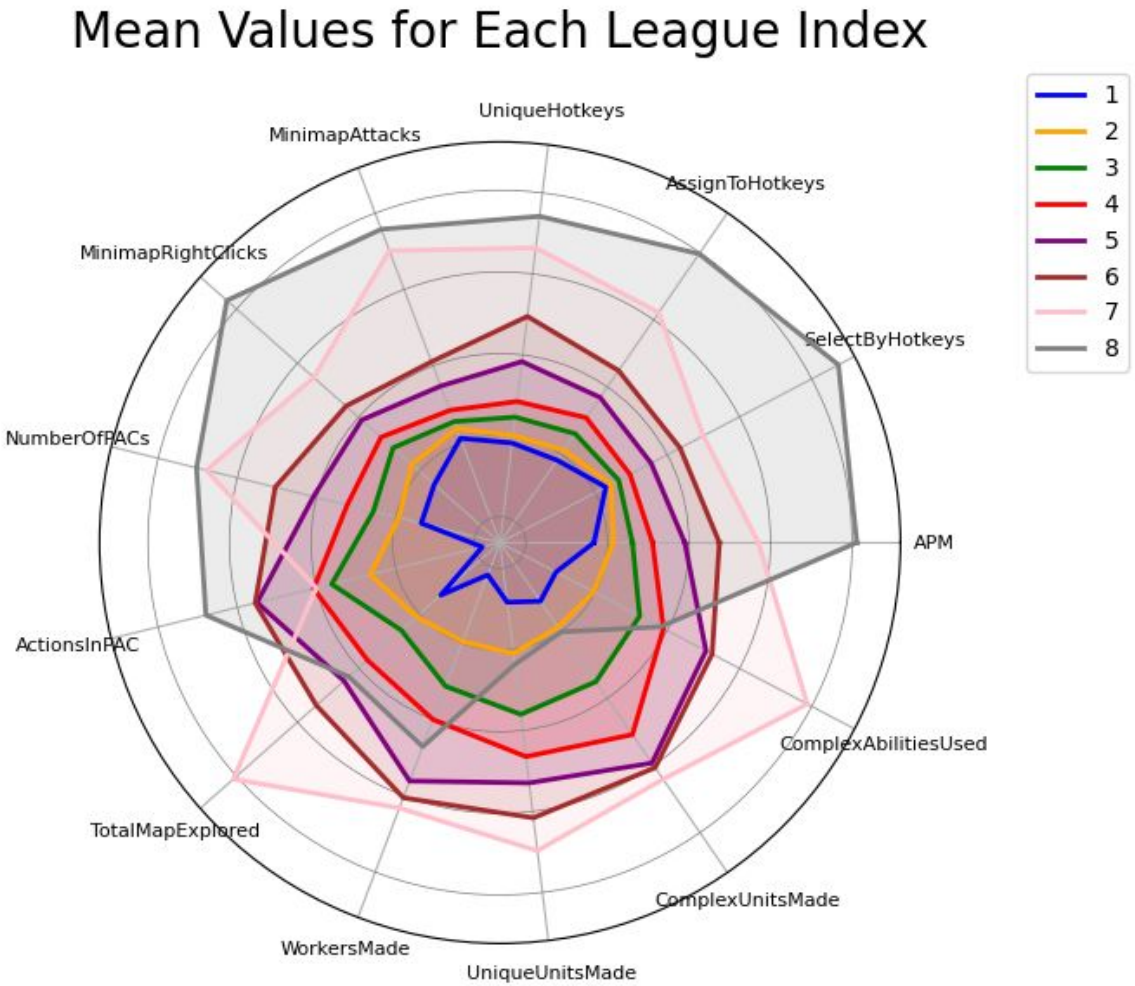






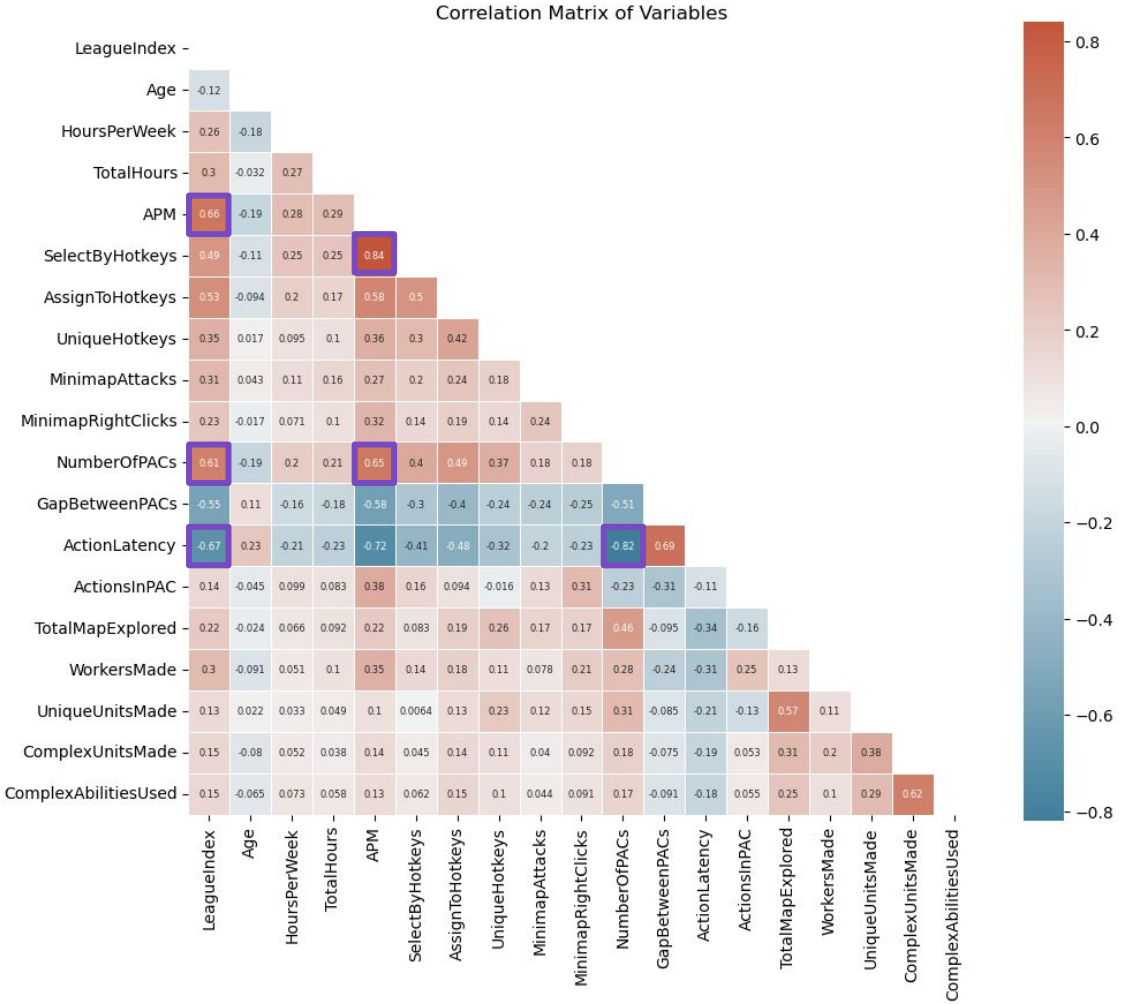


Feature Distribution by League Index



Mean Feature Variation for Each League

There is clear trend that how higher leagues have better averages in each features than lower skilled players



Variable Correlation

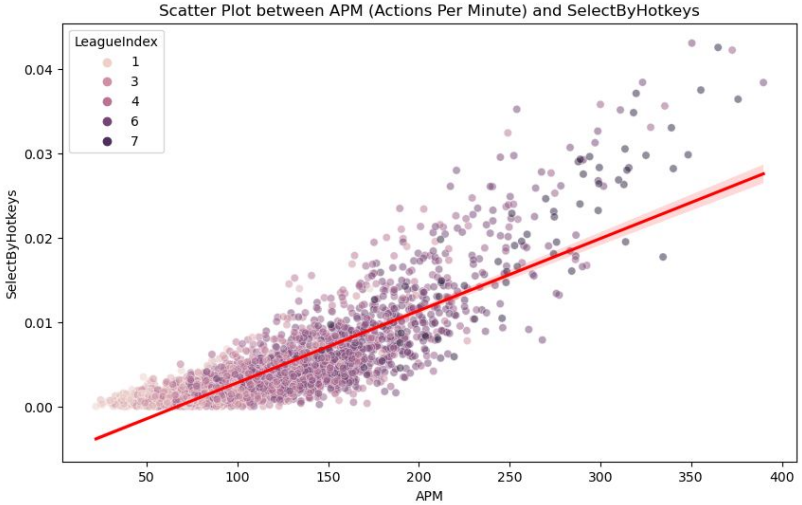
Positive Correlation

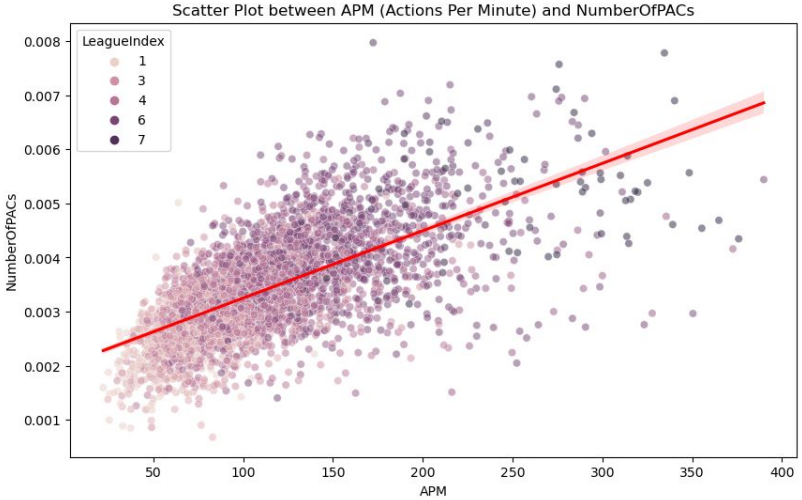
- League Index - APM, Number of PACs
- APM - SelectByHotkeys, Number of PACs

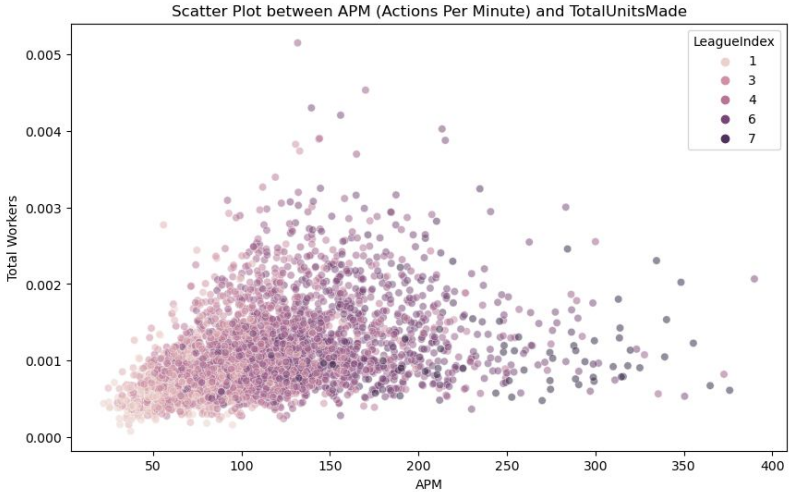
Negative Correlation

- League Index - ActionLatency
- ActionLatency - NumberofPACs

APM - Action/Min,
PACs - Measure multitasking
ActionLatency - Time Taken to perform an action
SelectByHotkeys- No. of units made using hotkeys







Analyzing Feature Relations

Positive Correlation

- APM & SelectByHotKeys
- APM & NumberofPACs

No Correlation

- APM & TotalUnitsMade

Model

Rank Prediction

Random Forest Classifier with Fine Tuning

Accuracy - 40.20%

Logistic Regression

Accuracy - 40.35%

XGBoost Model with Fine Tuning

Accuracy - **42.85%**

Summary

Need more data on low and high rank

All features were statistically significant for rank prediction

Model can be used to scout for Pro players & Strategy Development

Need new features for better Rank Prediction

Future Work

Need for More Data

More Detailed Gameplay Data

- Data on Player actions such as Race Selection, Average Unspent Resources, Time Spent Supply Capped, Resource Collection Rate

Data on Higher and Lower Ranks

Data at Different Timestamps about Player Action

Meta-Game Data

- Information about game state, Win/Loss data, types of units created

Thank You!