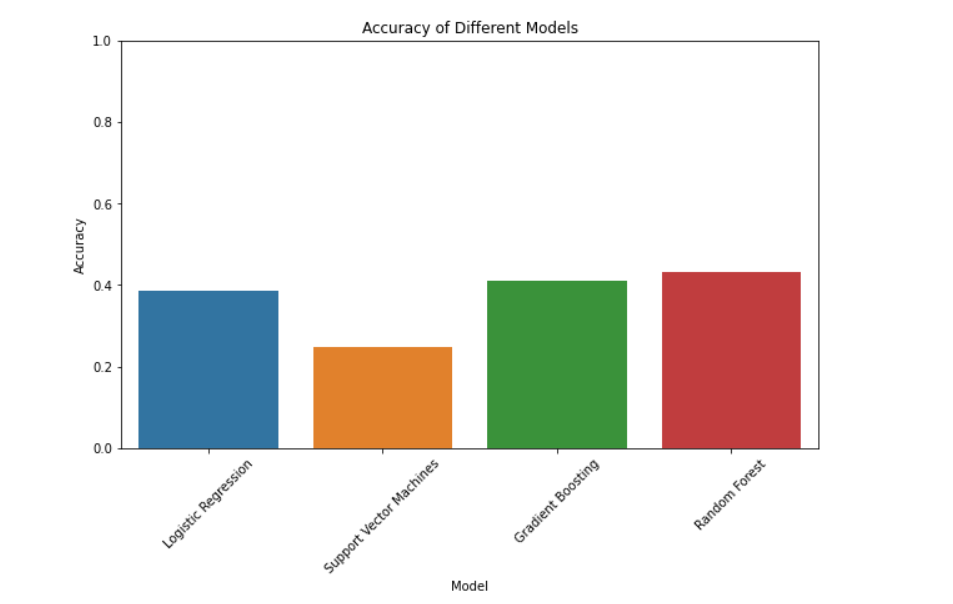
**Pranav Vishwanath Data Scientist Assessment**

Hypothetical: - after seeing your work, your stakeholders come to you and say that they

can collect more data, but want your guidance before starting. How would you advise

them based on your EDA and model results?

Answer:

Based on the analysis and model results, it is clear that the current dataset and models are facing challenges in accurately predicting the LeagueIndex, which represents a player's skill level in the StarCraft game. To address this, I would advise stakeholders to focus on collecting additional data that specifically captures a player's rank and the ranking system of the StarCraft game.

One important aspect to consider during data collection is capturing the player's rank, which serves as a crucial indicator of their skill level. This can be obtained by extracting information from the game's official ranking system. By incorporating this data into the existing dataset, we can provide the models with more precise and relevant information about each player's skill level.

Furthermore, it is essential to gather data related to the ranking system of the StarCraft game itself. This includes information about the different tiers or divisions within the ranking system, the criteria for advancing or demoting ranks, and any additional factors that impact a player's progression. By incorporating these details into the dataset, the models can better understand the intricacies of the game's ranking system and its influence on a player's skill level.

Additionally, collecting data on specific gameplay attributes and strategies used by players at different ranks can also be valuable. This could include information such as preferred race, playstyle preferences, average game duration, resource management, and decision-making patterns. By including these features, the models can capture the nuances of different player profiles and make more accurate predictions.

Lastly, it is important to ensure that the collected data is representative and diverse, encompassing players from various skill levels and regions. This will help to avoid biases and ensure that the models generalize well to different player populations.

In conclusion, to improve the accuracy of predicting the LeagueIndex in StarCraft, I recommend focusing on capturing player ranks, incorporating details about the game's ranking system, and gathering specific gameplay attributes. By considering these technical aspects of data collection and enriching the dataset with relevant information, we can enhance the models' understanding of player skill levels and improve the accuracy of the predictions.