Theran Meadows
Project 1: Master Chef
Milestone 2
June 23, 2024

Business Problem

The question/problem I am trying to answer is whether or not the winner of each season was the most skilled contestant of their respective opponents for that season to be crowned the winner. I would like to implement my own point system to confirm or deny my theory and provide a mock recommendation on how the FOX tv network could change how they decide who wins the Master chef reward.

I would also like to show a predictive classification model to see how the show performs between its own method and my own

Background/History

Currently the show has had a consistent number of contestants per season at 20. The show will then filter through one or two challenges per episode usually eliminating at least one person but sometimes two or more can also be eliminated. There may be partner challenges or group team challenges as well. The single or team winner of a challenger will then be given immunity from being eliminated and will therefore usually sit the following challenge out. The final 3 are then considered the finalists and complete through 3 final rounds cooking. The first round is the appetizer, the second is the entree and the third is the dessert. After the appetizer or the entree is usually when one of the finalists is eliminated. The winner is then chosen after both have completed the challenge

Data Explanation (Data Prep/Data Dictionary/etc)

The data is from the Master Chef Wiki Fan page that keep track of the contestants and the outcome of each challenge. It does keep track of the wins, tops, bottoms, and safes that a person landed in the current challenge but does not keep a running total. In my dashboard and predictive model, I used my own point system to keep a running total for every contestant. I used Excel and Sql to prepare the data and Python for the predictive model.

Methods

The target value for any of the predictive model tests was the winner variable. The purpose of the model was to see if the default Master Chef process was accurate or if another model should be used. A Decision tree was used and then a random forest was used to compare results

Analysis

Still ongoing

Conclusion

To be determined

Assumptions/ limitations

The data set, although complete as far as having every episode and contestant evaluated, the data set is only every measured with the highest episode count in a season at 33. That means that with any prediction, there are only 32 variables of data that can be used to see what the outcome of 33 would be.

Challenges

Because of the small data set, my model runs have run into over fitting a lot of times with accuracy being 100%. However, I have gotten 66% and 75% accuracy with different seasons and variables I throw in.

Future Uses/Additional Applications

This data set will be growing as even now there is a current season of Mast Chef airing. This set can always be added to but I think there won't be enough data fast enough to predict who would win.

Analyzing the quality of characteristics of contestants as well could also be used to predict who would win.

Recommendations

My recommendation would be to implement that point system into the show to see who has the most wins at the end of a season rather than reward those that happen to cook good enough to pass a challenge but not be in the top cooks.

Implementation Plan

This system would be used going forward for future seasons. It would have to be established to all contestants, judges and audience.

Ethical Assessment

There are no major ethical issues with the data set now as it is all public information. Nonetheless, going forward, judges in the show would have to not be biased towards any one or group of contestants. Telling someone to not be biased is impossible though. So no matter what, there will be an element of bias as long as judges have a say in who wins.

References

Wikimedia Foundation. (2024, June 13). *MasterChef (American TV series)*. Wikipedia. https://en.wikipedia.org/wiki/MasterChef_(American_TV_series)