Dear Puzzle Editors in New York Times:

Thank you for providing us with the data. After 4 days of working, we construct some models, and get some information that may be helpful for you and your compony to make strategies for Wordle, here are some details of it.

First, we developed a model (ARIMA (4,1,2)) to make time series prediction, which will help you predict the number of users who will paly games in future. This model is robust by testing it with Random Forest method and true data in Feb,1,2023. Although Wordle was once a popular game and made a big hit in America, time series analysis showed that active game users gradually become stable, and it seems to fell slightly, which may not be a good news for your company, thus some new strategies are needed to be taken into consideration.

Then we took word attributes into account, and made the model above more specific, which can help you predicting associated percentages of each tries with specific dates and words using method of Random Forest Regression. We constructed quite a few attributes ourselves (word entropy, combination of letters, occurrence of single letters, the position of letters, word guessing correctness of Monte Carlo analysis…) to present factors that may influence the percentage distribution. this model was not as robust as the first one since it is more specific, but still have a high value for you to make future decisions, such as deciding to choose which letter to be the answer.

Taking the attributes mentioned above, we use K-means clustering, dividing all your words into 4 classes (Simple, Ordinary, Difficult, Rather Difficult) through training the data you give (the standards we use are mean & standard derivation). The general user’s score various with different classes. Combing the classification of words and some basic knowledge of psychology (such as Cwithfish Effect and Transfinite Effect), you could make a strategy to give words form different classes periodically to better get our users excited.

Finally, we build a model to simulate people’s word guessing process by taking word frequency and probability into account. Our model showed that people who report their grades on Twitter was just a small part of users who play the game (which can be seen from our essay), and they have a common characteristic that most of them are those who won the game. wining games can give them a passion for sharing their results, which is a good way of advertisement. Hence increasing the maximum try times may help increasing users.

That’s all the work we do, hoping our model will help you and your compony design better strategies for wordle, and all of us are willing to see Wordle develop better in the future.

Best wishes

One of your team

2.21.2023