Insights for Surviving the Game of Thrones

Marcelo Tournier

Introduction

In the fictional realm of Westeros, there is a humongous struggle for power. In this setting, different groups are fighting to dominate over each other. Allegiances and treachery are part of the routine in the plot of Game of Thrones (GOT), added to a lot of story twists, who makes it even more difficult to predict what will happen next. This complex background makes survival a real challenge for all characters.

Objective & Methods

The objective of this study is to analyze the data from the world of GOT, bringing insights that could help its characters to improve their survivability chances throughout the narrative.

To identify insights, we studied data of 26 variables in 1946 characters, including information about names, titles, houses (family group and allegiances), cultures, popularity, among others. This data was used to build an exploratory data analysis, and further analytical modelling with focus on trying to predict the persons who will be alive (and who will die) in the story.

Results & Discussion

As **risk factors** (variables which are associated with reduced probabilities of being alive), we found that the more relevant ones for prediction modelling in our tests were the character's appearance in the first book, marriage, and higher popularity ratings.

As **protective factors**, we found that there is an association between the total number of books in which the character appeared and an improved chance of survival. Moreover, the character's appearance in fourth book of the series are tied to increased survivability. This combination could mean that characters who have a long plot, or appearance at mid-late story points could get a reduced probability of being killed. Belonging to a huge cultural group, and having no dead relations were found also as key variables related with improved survival chances.

As we have a lot of missing values in this dataset, we tried to use the number of absent information in our favor for predicting. Thus, we found that the total number of unknown (missing) values in each character profile is related with improved chances of being alive, together with the information if the character title is unknown. This could mean that maybe people who are outside the main plot for power struggle have a bigger probability of survival. We used all these factors to build a predictive model based on similarities between dead and alive character groups (KNN), based in all variables described above. Upon on this algorithm, we were able to correctly identify 82% of the survivors in tests with unused holdout data from GOT characters. Moreover, its high sensitivity ensures us that 95% of all characters the model predicted as alive or dead were correctly identified. An Area Under Curve of 0.802 means that the model predictions are robust and more reliable than a random guess.

Conclusion

For a new GOT character to improve its chances of survival (or watch out for a traitor eager to kill), we would recommend her/him to take special attention if they have any of the risk factors above: Consider though decisions, such as divorce (in Westeros, your significant other can be your worst enemy), or even appearing a bit less on the story spotlights. On the other hand, to protect herself/himself, it is important to care about the others of your group, so they wouldn't die (and so, they can care on your back), specially if you belong to a huge group. If their "mystery level" is high, we would recommend them to don't dig too much about their past – otherwise they could be buried by it. As final recommendation, this model could be implemented to help GOT fans to follow the story closer and speculate about the future of characters. Besides, this is an unusual way of trying to understand some of the patterns that the author George Martin used to build (and kill) the most beloved & hated personalities of Westeros.