The effect of toxicity in game chats on match outcome in Dota 2

Abay Jumabayev

Central European University, Department of Economics

January 2022

Proposal

I test whether toxicity in game chats affects the outcome of a match. A message like "uninstall the game, freak" is considered toxic and, according to Richard et al. (2017), about every fifth player of online multiplayer games decides to quit playing because of such disruptive behavior.[1] To minimize the number of quitters and to maximize enjoyment of all players, the analysis of the effect of toxic behavior is useful for game companies. Proving that being toxic negatively affects players' performance allows game corporations to discourage players from toxic behavior and increase the popularity of the game.

I choose to analyze chat messages in the game called "Defense of the Ancients 2" (Dota 2) for three reasons. First, according to a survey mentioned in Ingersoll's report (2019), Dota 2 is the most toxic game among the 15 most popular games in the US analyzed.[2] The survey result showed that 79% of Dota 2 players experienced toxicity in-game at some time and 38% experience toxicity regularly. The second reason is that Valve Corporation, the developer of Dota 2, does not have a system that would discourage people from sending toxic messages. That means, that toxic players do not get punished for their behavior. The third reason is that Dota 2 is one of the most popular multiplayer online battle arena (MOBA) games having on average more than 400 thousand people playing the game.[3] Those three reasons contribute to a high frequency of toxic messages in game chats.

I found two papers that study the effect of toxicity in game chats on game success. Martens et al. (2015) found no strong correlation between the win rate of a player and the count of toxic words he/she sends to a chat.[4] The authors also found that the losing team starts using more toxic words in the late stage of a match, although the level of toxicity is consistent throughout the earlier periods. Traas (2017) continued the work of Martens et al. and focused on the impact of toxicity on the match outcome.[5] The author found that the team that sends toxic messages has fewer chances to win a match. My paper differs from the earlier works in three ways. First, I use newer and much larger data[6]: 50,000 matches from 2015 versus 13,000 matches from 2012 used by Martens et al. and Traas. Second, I will use Long short-term memory (LSTM), an artificial recurrent neural network (RNN) architecture, to analyze the level of toxicity in messages. This machine learning

---

[1] Richard, Tynes, and Kafai, "Diversifying Barbie and Mortal Kombat," 83.
[2] Ingersoll, "Free to Play?"
[3] "An Ongoing Analysis of Steam's Concurrent Players."
[4] Märtens et al., "Toxicity Detection in Multiplayer Online Games," 4.
[5] Traas, "The Impact of Toxic Behavior on Match Outcomes in DotA," 20.
[6] Dota 2 Matches

technique identifies the probability of message toxicity and performs better than labeling messages by the usage of profane words. Lastly, I use logistic regression to evaluate how toxicity contributes to the prediction of match outcomes. Logistic regression is a good choice given that the variable of interest, winning a match, is naturally dichotomous.

Conducting this research, I expect to find that toxicity affects match outcomes.

Bibliography

"An Ongoing Analysis of Steam's Concurrent Players." Accessed January 10, 2021. https://steamcharts.com/app/570.

"Dota 2 Matches." Accessed January 10, 2022. https://kaggle.com/devinanzelmo/dota-2-matches.

Ingersoll, Christina. "Free to Play? Hate, Harassment, and Positive Social Experiences in Online Games." ADL, July 2019. https://www.adl.org/media/13139/download.

Keras, Team. "Keras Documentation: LSTM Layer." Accessed January 10, 2021. https://keras.io/api/layers/recurrent_layers/lstm/.

Märtens, Marcus, Siqi Shen, Alexandru Iosup, and Fernando Kuipers. "Toxicity Detection in Multiplayer Online Games," 7. Zagreb, Croatia, 2015. https://doi.org/10.1109/NetGames.2015.7382991.

Richard, Gabriela, Brendesha M. Tynes, and Yasmin B. Kafai. "At the Intersections of Play." In *Diversifying Barbie and Mortal Kombat*. Pittsburgh: Carnegie Mellon: ETC Press, 2017. https://press.etc.cmu.edu/index.php/product/diversifying-barbie-and-mortal-kombat-intersectional-perspectives-and-inclusive-designs-in-gaming/.

Traas, Arjen. "The Impact of Toxic Behavior on Match Outcomes in DotA." Tilburg University, 2017. https://arno.uvt.nl/show.cgi?fid=145375#:~:text=We%20identified%20predictive%20variables%20of,chances%20of%20winning%20the%20game.

Varudandi, Shaunak. "Toxic Comment Classification Using LSTM and LSTM-CNN." towardsdatascience.com, January 26, 2021. https://towardsdatascience.com/toxic-comment-classification-using-lstm-and-lstm-cnn-db945d6b7986.