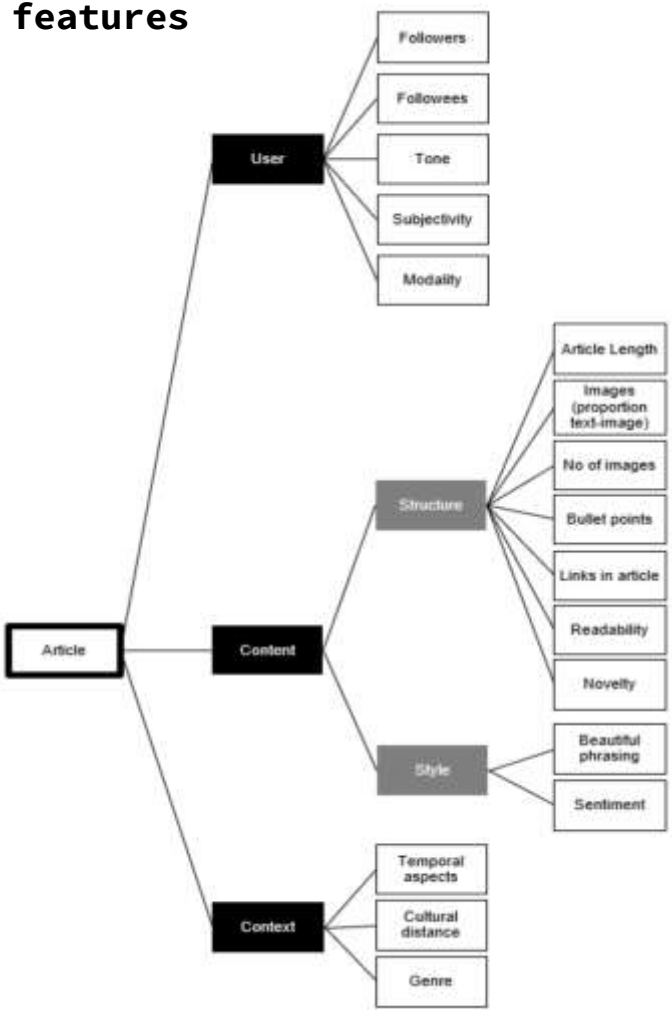
Feedback Matters! Predicting the Appreciation of Online Articles A Data-Driven Approach

Catherine Sotirakou | PhD student | University of Athens - Faculty of Media & Communication
Contact: @Cathrine_Sot| cathrinesot@gmail.com

Our features

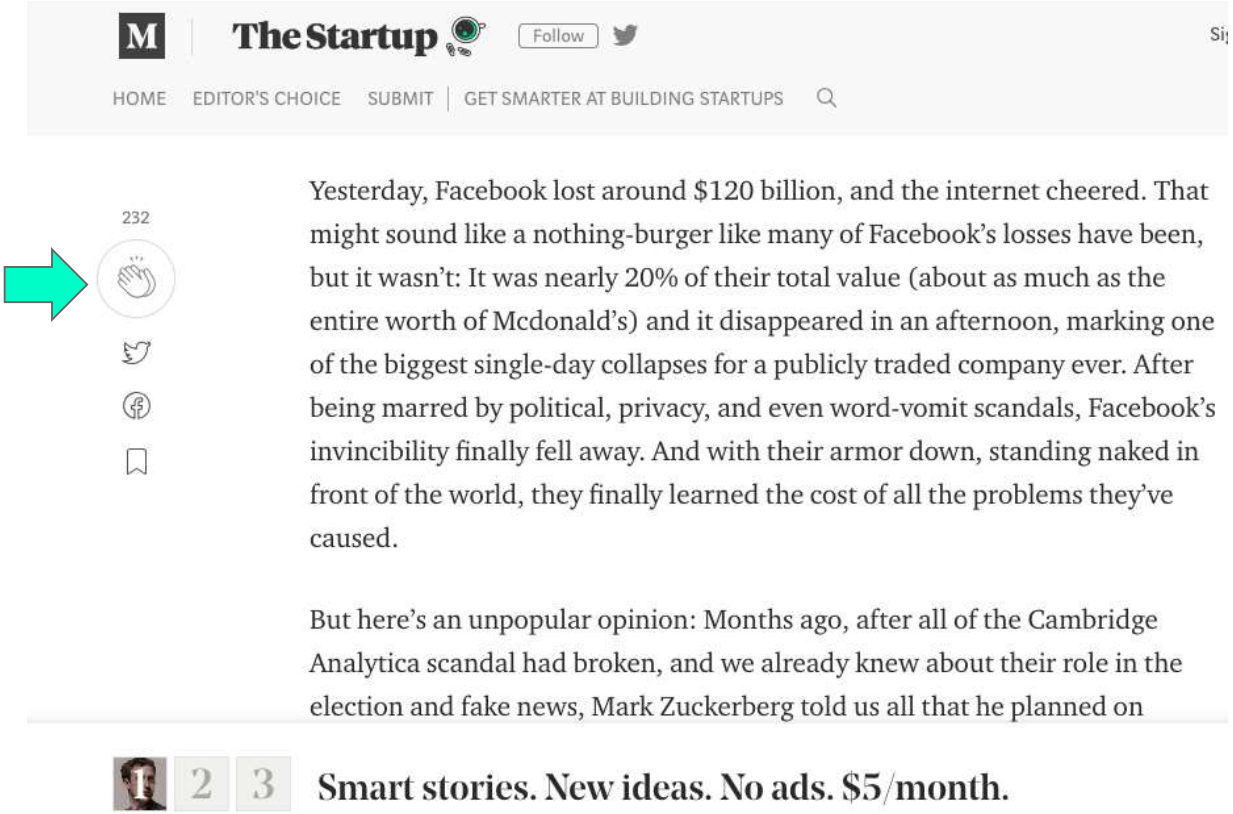


Predictive model

Data: We run experiments on a mixed-topic dataset of over 200K thousand articles published: Sept 2018-Sept 2019 on Medium.com and downloaded at the beginning of 2019. The articles had a large distribution of claps ranging from 84 to 157K claps

Dependent variable:

"Claps" (formerly called "Recommend") on Medium represent whether readers liked the story or not and would recommend it to other users on the platform (a user can clap more than once). In the world of Medium, the success of an article is measured regarding claps count, which is the number of times the readers have been clapped.



Method - random forest

The table presents the ordered list of the importance of the variables of the selected categories, such as Followers, Article Length, Number of images etc., starting from the most important variables that affect the acceptance of an article.

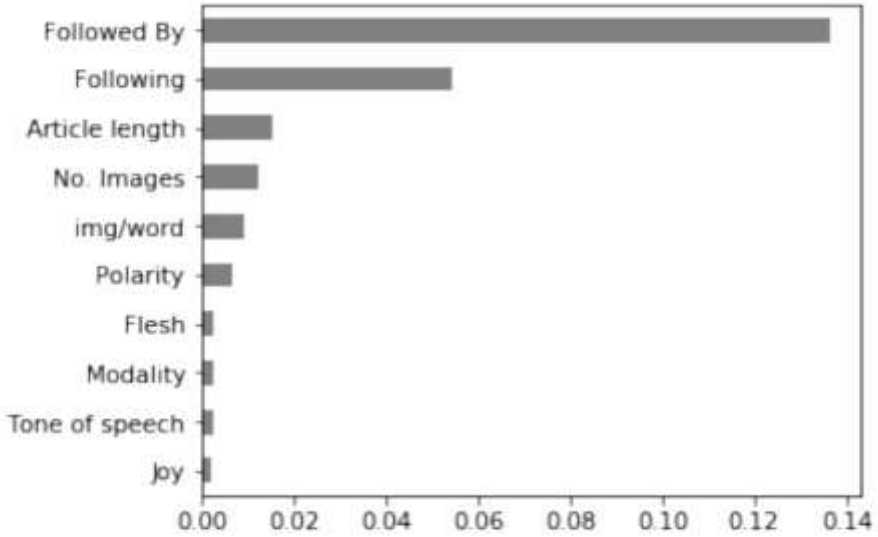


Figure 2: Features importance score for high claps bucket

The best score of all between the three different classifiers was generated by the XGBoost algorithm with F1-Score of 80%.

Thank you !
Pls give me some feedback!!!

Contact: @Cathrine_Sot|
cathrinesot@gmail.com|
aas2322@columbia.edu