

# Predict Community Interaction based on the Hyperlink Network of Reddit

Louis de COMMARMOND, Arthur COLL, Benjamin KARAOGLAN, Ismail KHALIL, Edouard VILAIN

# TABLE OF CONTENTS

01

---

Introduction & Motivation

02

Description of the Dataset

03

Related Work

04

---

Methodology

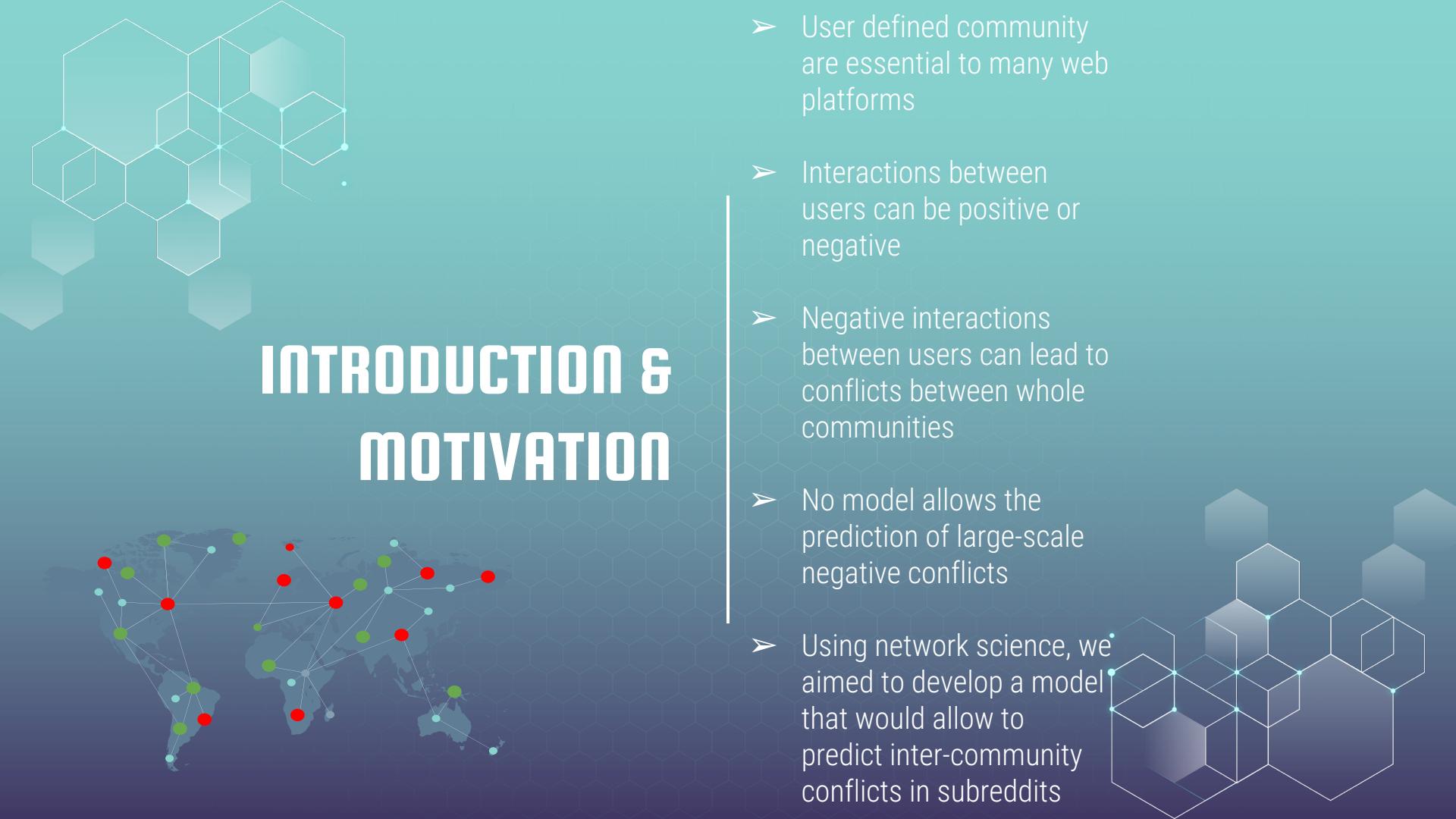
05

Evaluation

06

---

Conclusion

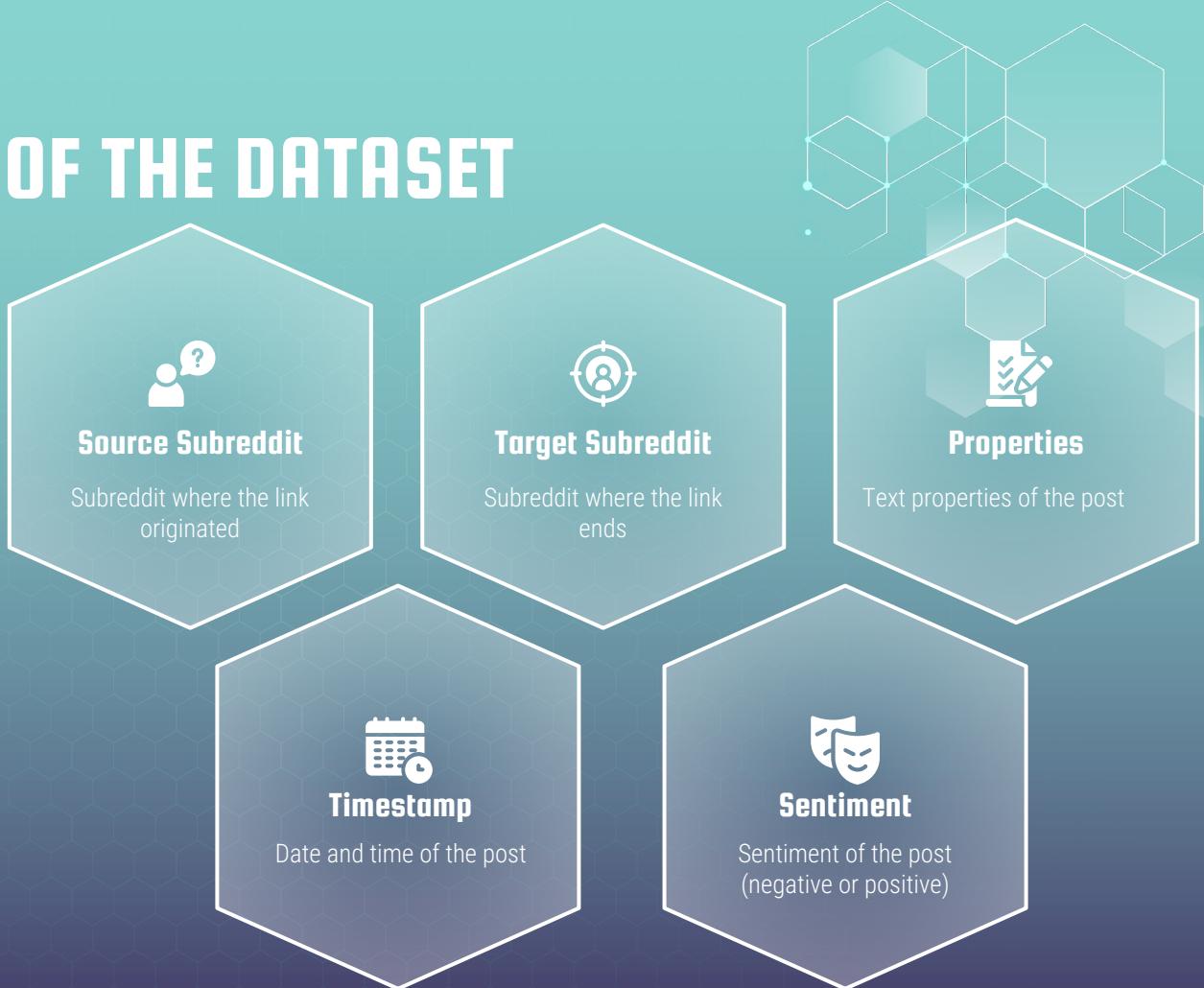


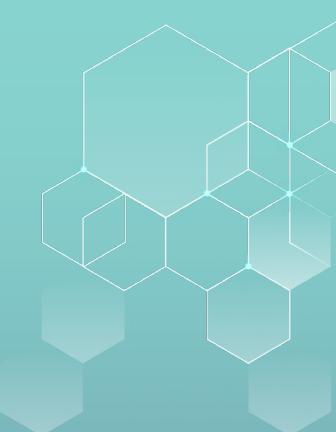
# INTRODUCTION & MOTIVATION

- User defined community are essential to many web platforms
- Interactions between users can be positive or negative
- Negative interactions between users can lead to conflicts between whole communities
- No model allows the prediction of large-scale negative conflicts
- Using network science, we aimed to develop a model that would allow to predict inter-community conflicts in subreddits

# DESCRIPTION OF THE DATASET

- Social Network dataset from Reddit
- Contains inter-subreddit post from 2014 to 2017
- 35 776 communities, 286 561 subreddits post
- Posts are classified as either positive (1) or negative (-1)





# Related Work

This dataset is the result of a research project conducted by Kumar et al. in 2018 entitled Community Interaction and Conflict on the Web. The main findings were:

- 1% of communities initiate 74% of all conflicts
- Conflict initiating nodes are rare and clustered together in certain communities.
- Less active users are the ones who participate the most into conflicts, even though active users are the ones who trigger conflicts
- Conflicts lead to “echo-chambers” where attackers mostly interact with attackers and defenders mostly interact with defenders
- Conflicts can have long term effects on the engagement of the defenders of the target community

# METHODOLOGY

- Using data from 2014 to 2016, predict subreddits interaction in 2017
- For each year, build features and make predictions using previous year data
- Use pagerank on 2017 obtained predictions to identify toxic communities



## Feature Engineering

Graph based features  
Node based features  
Null Values  
Normalization



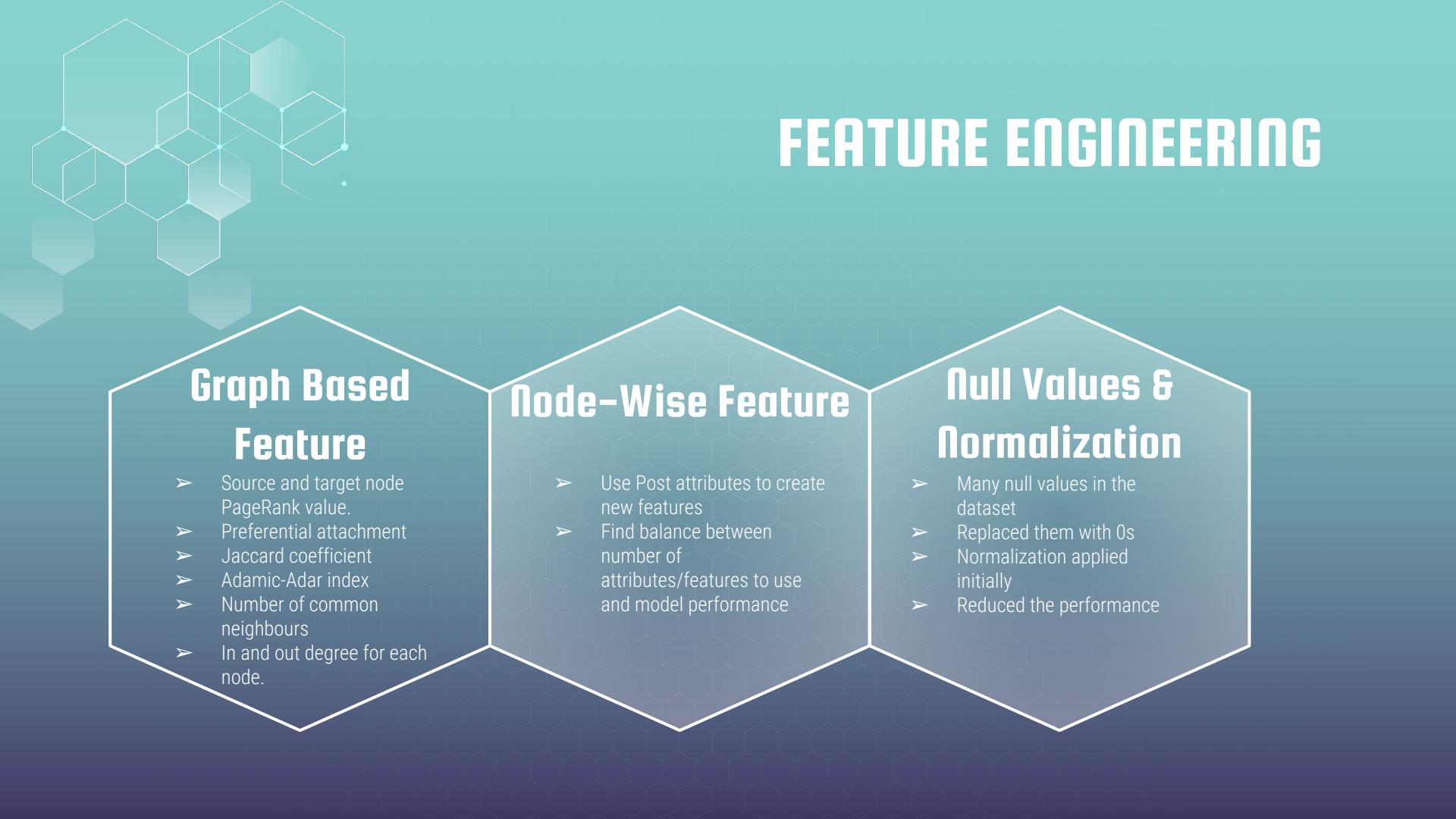
## Prediction Models

Catboost for regression  
Deep Learning MultiClass Regressor



## Automatic tool for monitoring

Pagerank for toxic  
subreddit  
identification



# FEATURE ENGINEERING

## Graph Based Feature

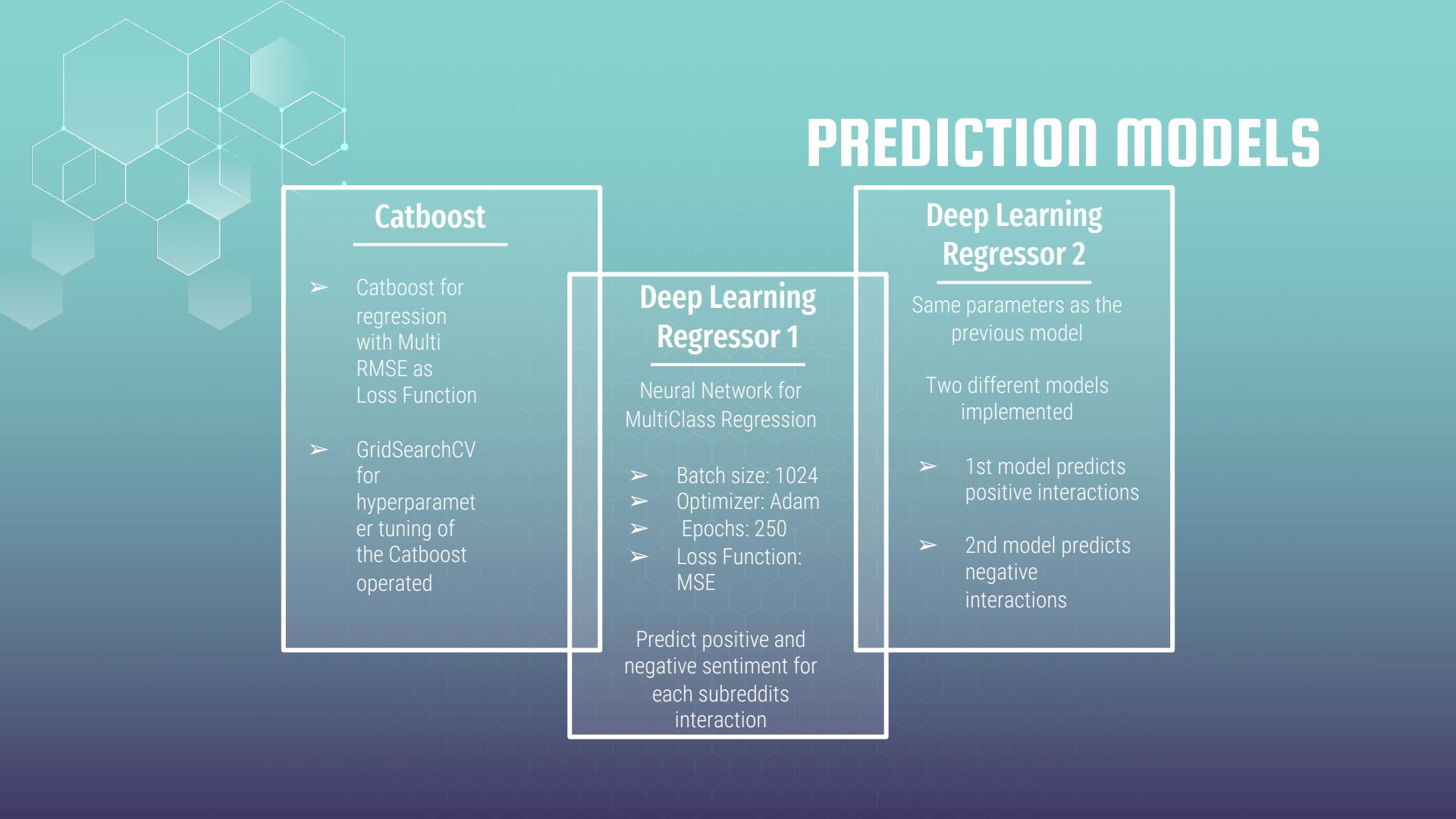
- Source and target node PageRank value.
- Preferential attachment
- Jaccard coefficient
- Adamic-Adar index
- Number of common neighbours
- In and out degree for each node.

## Node-Wise Feature

- Use Post attributes to create new features
- Find balance between number of attributes/features to use and model performance

## Null Values & Normalization

- Many null values in the dataset
- Replaced them with 0s
- Normalization applied initially
- Reduced the performance



# PREDICTION MODELS

## Catboost

- Catboost for regression with Multi RMSE as Loss Function
- GridSearchCV for hyperparameter tuning of the Catboost operated

## Deep Learning Regressor 1

Neural Network for MultiClass Regression

- Batch size: 1024
- Optimizer: Adam
- Epochs: 250
- Loss Function: MSE

Predict positive and negative sentiment for each subreddits interaction

## Deep Learning Regressor 2

Same parameters as the previous model

Two different models implemented

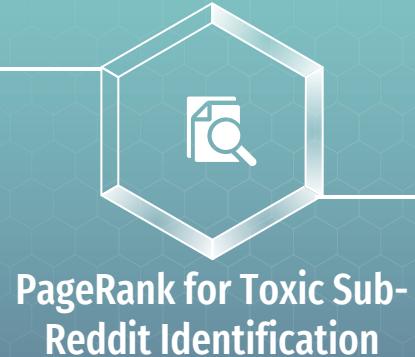
- 1st model predicts positive interactions
- 2nd model predicts negative interactions

# AUTOMATIC TOOL FOR MONITORING

Our predictions allows to create two weighted graphs:

- One for negative interactions
- One for positive interactions

We inverted the directions of each edges to know which subreddits have large number of negative interactions



We used a harmfulness coefficient in order to quantify to what extent will a subreddit be negative towards others in 2017.

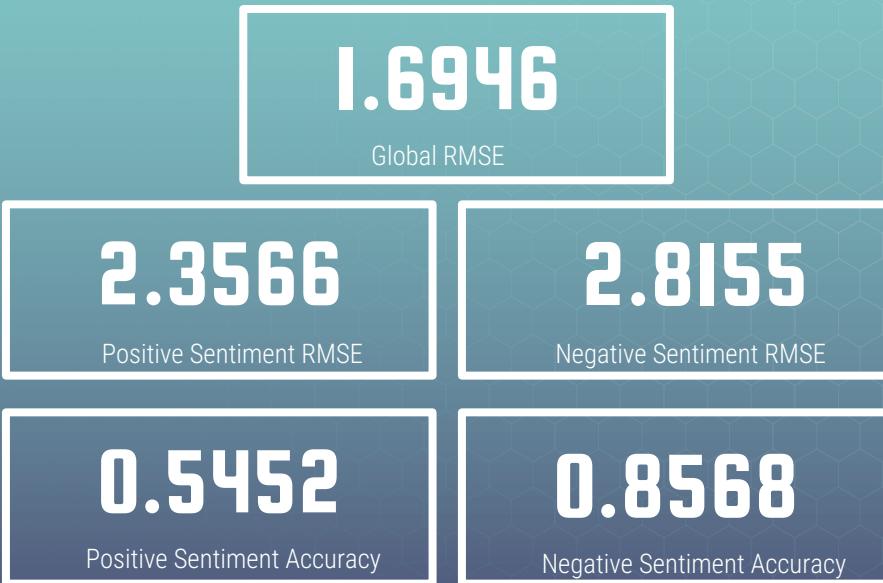
The coefficient is as follows:

$$harmfulness_i = \frac{NegativePageRank_i}{PositivePageRank_i + \epsilon}$$

$$\epsilon = \min_j PositivePageRank_j, PositivePageRank_j > 0$$

# EVALUATION

## Catboost



- We observed quite equally distributed predictions compared to ground truth values
- The model didn't accurately predict high values of negative interactions
- As 2017 data wasn't complete, ground truth value comparison isn't the most accurate estimator
- Lack of data in high values of negative interactions and 0s in positive interactions lead to less accuracy.



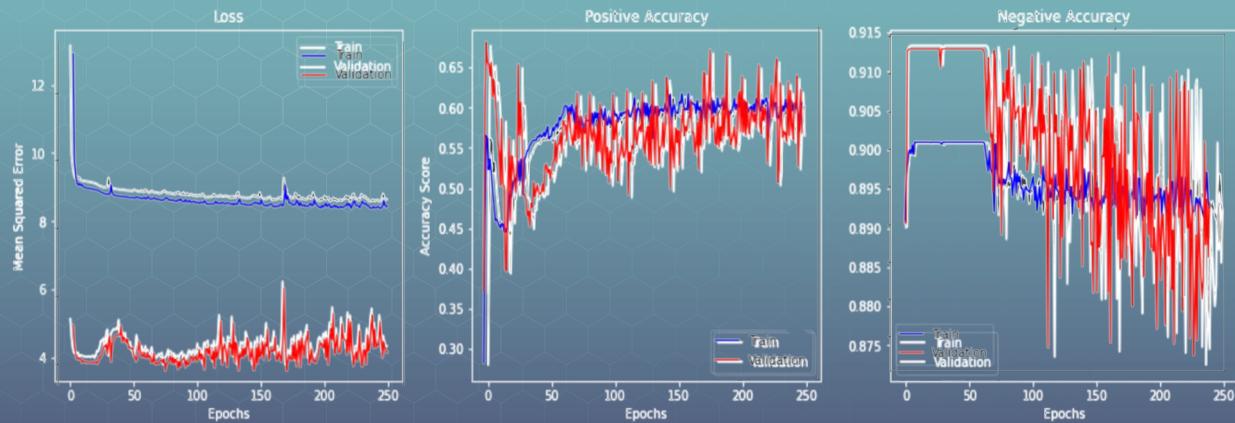
# EVALUATION

Model for prediction of both negative and positive interactions

Performance on the validation set oscillates and training is very noisy:

- Complexity of the task at hand
- Lack of data in the validation set (taken from 2017)

Two models may be a solution

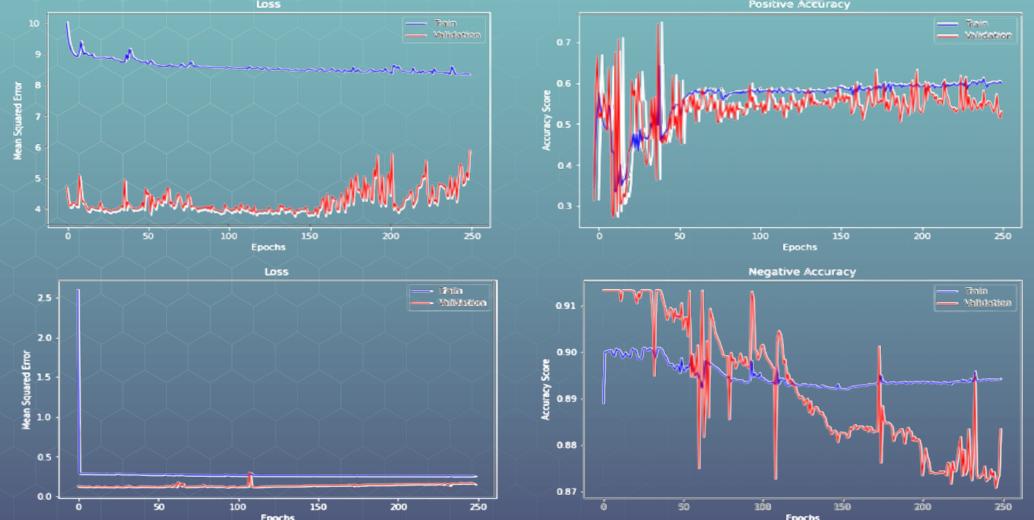


# EVALUATION

Models for prediction of first positive interactions only, then negative interactions only

Performances of the new model don't seem to improve the accuracy of the model

Thus, we decided to stick with the first Deep Learning model we built in order to go forward with the PageRank implementation



**Training curves for the positive interaction model (1st line) and the negative interaction model (2nd line)**

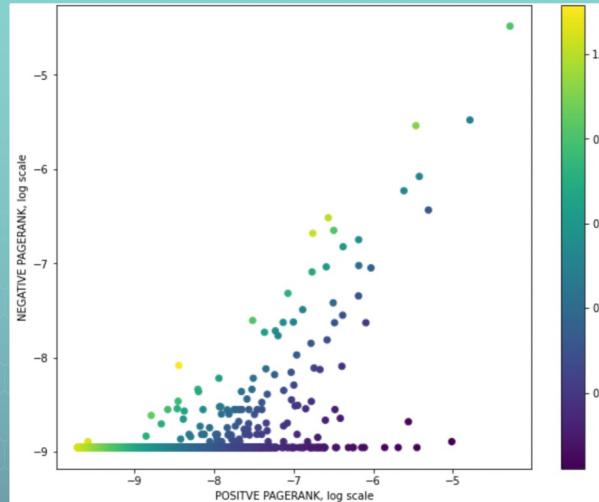
# EVALUATION

## PageRank for prediction of toxic subreddits

We observed that this tool allowed to discriminate well subreddits with a greater negative PageRank score than a positive one which was satisfactory

However, we also observed that many subreddits with high harmfulness were actually subreddits with not many interactions

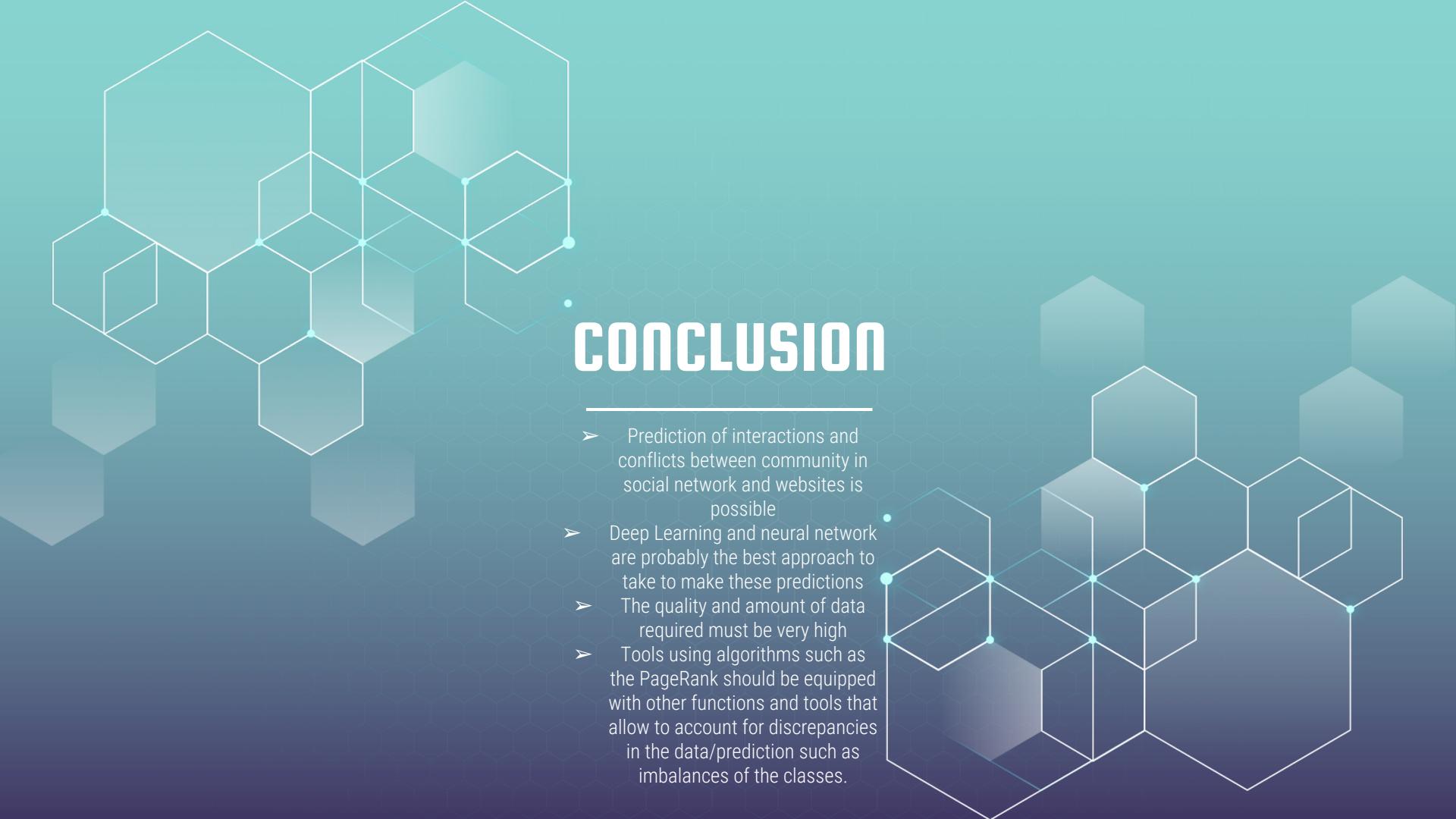
PageRank plot colorized by harmfulness



## Most and least harmfull subreddits of 2017

SUBREDDIT	harmfulness
explainlikeimfive	1.116096
academiceconomics	1.062438
rebbirl	1.062438
mongodb	1.062438
giftournament	1.062438
mustang	1.062438
ak47	1.062438
whatstheword	1.062438
fancyfollicles	1.062438
gorving	1.062438

SUBREDDIT	harmfulness
subredditoftoday	0.020503
rmsfhj	0.029686
nomorebamboozles	0.037442
tipofmypenis	0.043363
pokemonshuffle	0.044206
modhelp	0.056755
flyfishing	0.058581
place	0.064837
hiking	0.065370
askmen	0.068910



# CONCLUSION

---

- Prediction of interactions and conflicts between community in social network and websites is possible
- Deep Learning and neural network are probably the best approach to take to make these predictions
- The quality and amount of data required must be very high
- Tools using algorithms such as the PageRank should be equipped with other functions and tools that allow to account for discrepancies in the data/prediction such as imbalances of the classes.