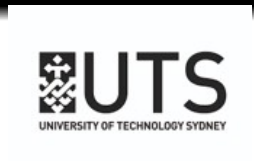


Political Stance Prediction Based on Online Information Diffusion

Pilli Yogesh Kumar

Laboratory Hubert Curien - University of Sydney

Advisors:- Christine Largeron Andrei-Marian Rizoiu

7th July 2020

Overview

State of Art

Myers, Seth A., Chenguang Zhu, and Jure Leskovec. (2012) "Information diffusion and external influence in networks."

Information Diffusion: Spreading of information from node to node along the edges of a social network

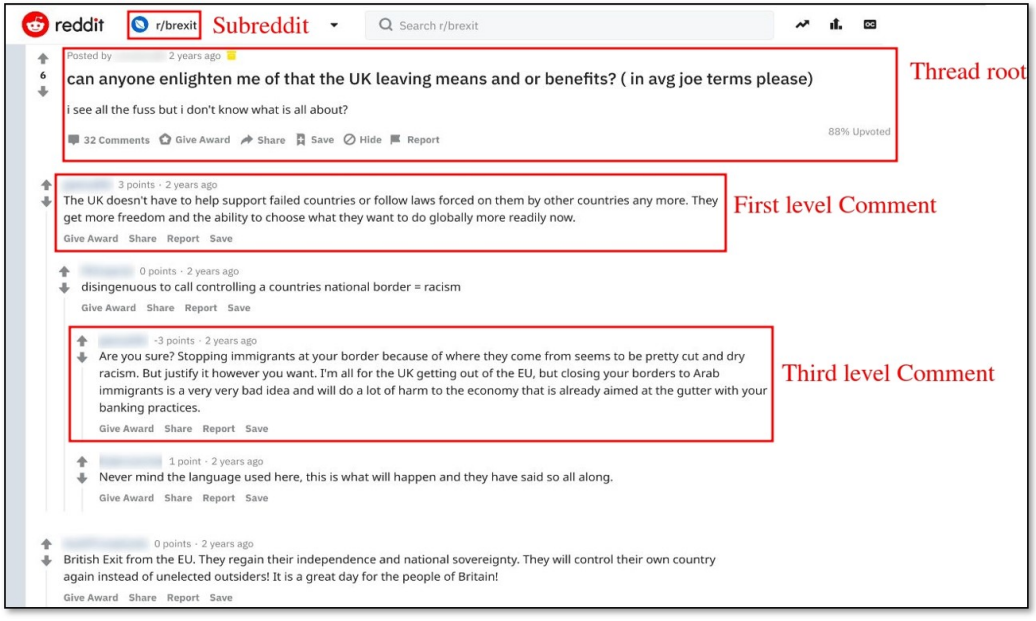
Shuai, X., Ding, Y., Busemeyer, J., Chen, S., Sun, Y., Tang, J. (IJSWIS , 2012) - "Modeling indirect influence on twitter

Social Influence: Process in which an individual's thoughts, feelings or actions are affected by other people

Benoit, K. and Matsuo, A. (2018) - "Network analysis of Brexit discussion on social media."

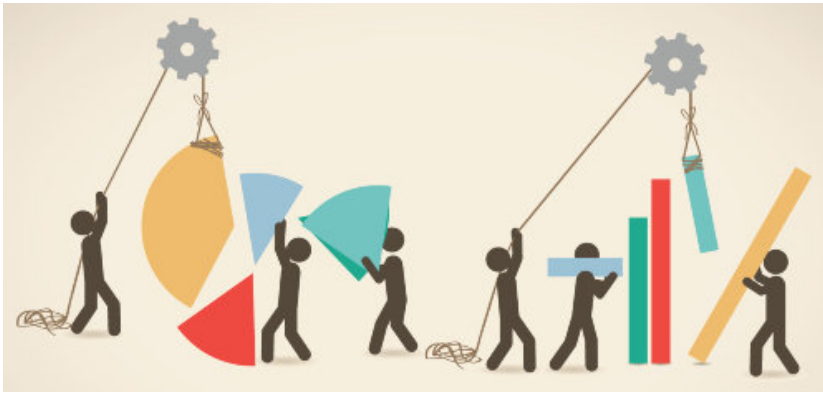
Politial Stance Classifier: Labels users Stance as AgainsBrexit, ProBrexit or Neutral.

Data



▸ Link

"https://files.pushshift.io/reddit/subreddits/brexit"



- **Time Frame:** Nov'2015 - Apr'2019
- **Users:** 14,362 users
- **Threads:** 21,725 threads
- **Comments:** 207,894

Social Network Analysis

Mark, N. P. (2003). "Culture and competition: Homophily and distancing explanations for cultural niches". American Sociological Review. 68 (3): 319–345.

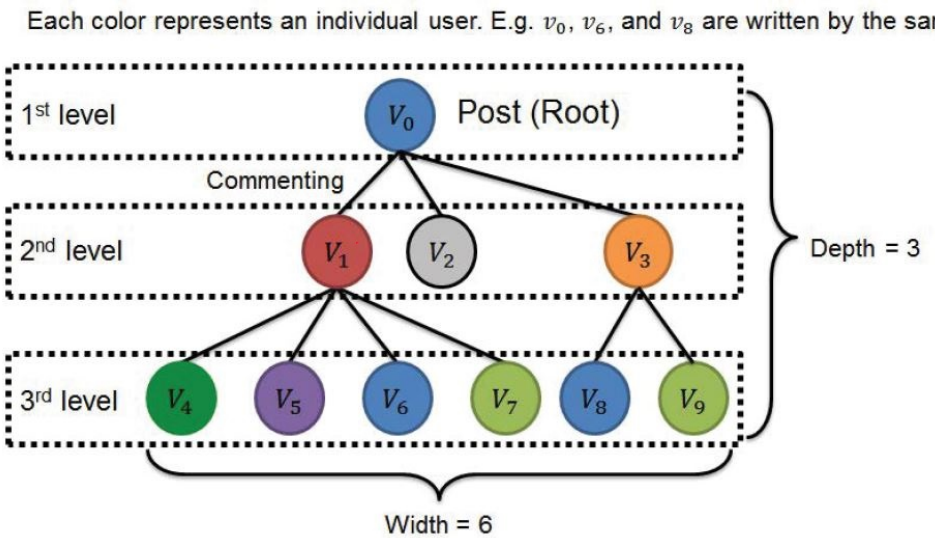
Homophily:-It is tendency of people to form friendships with others with similar characteristics

Crandall et al Anagnostopoulos et al. (KDD'08) Bonchi WI2011 - Influence Propagation in Social Networks: A Data Mining Perspective

Social Influence:- It is tendency of people may modify their behaviors to bring them more closely into alignment with the behaviors of their friends.

Social Network Analysis

Illustration of Comment Tree



- We define Comment Tree as an directed Tree, T = (V,E), where V is the set of all messages, which includes the original post(root) and follow-up comments in the thread, and E is the set of edges, each of which connects two messages that are linked by comment

Social Network Analysis

Notion of Edge Homogeneity

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H.E., Quattrociocchi, W.: The spreading of misinformation online. Proceedings of the National Academy of Sciences 113(3), 554–559 (2016)

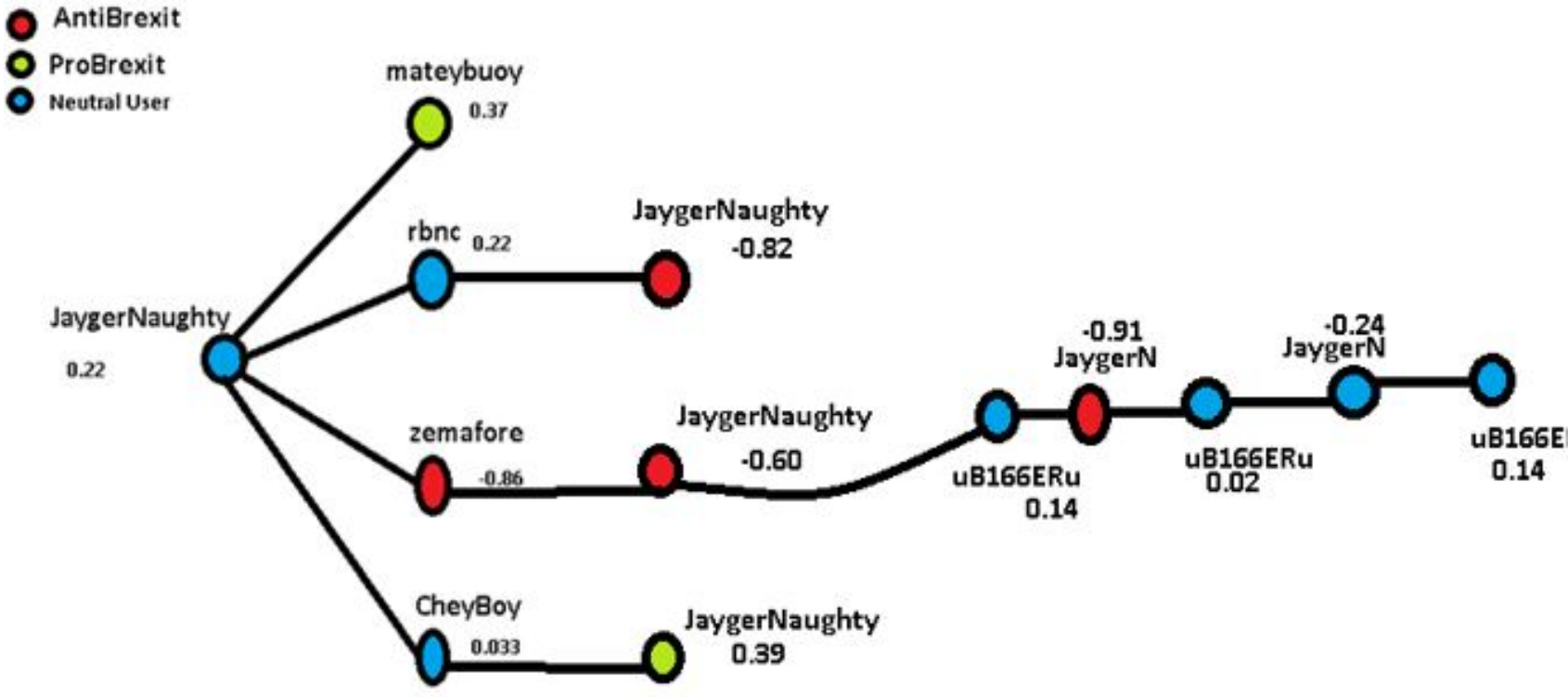
We define user polarization $\sigma = 2p - 1$, where p is probability, which lies in [0,1] and hence σ lies in between [-1,1]

From user polarization we define edge homogeneity , for any edge e_{ij} between the nodes i and j

$$\sigma_{ij} = \sigma_i \sigma_j \tag{1}$$

Social Network Analysis

Notion of Source Node and Target Node on smallest thread



Social Network Analysis

Case1

Scenario 1 - Fixing the Source Node to Neutral Stance and all the Target nodes which is attracted to Source Nodes

Case1

Scenario 2 - Fixing the Source Node to AgainstBrexit Stance and all the Target nodes which is attracted to Source Nodes

Case1

Scenario 3 - Fixing the Source Node to ProBrexit Stance and all the Target nodes which is attracted to Source Nodes

Social Network Analysis

Case2

Scenario 1 - Fixing the Target Node to Neutral Stance and all the Source nodes which is attracted by target Nodes

Case2

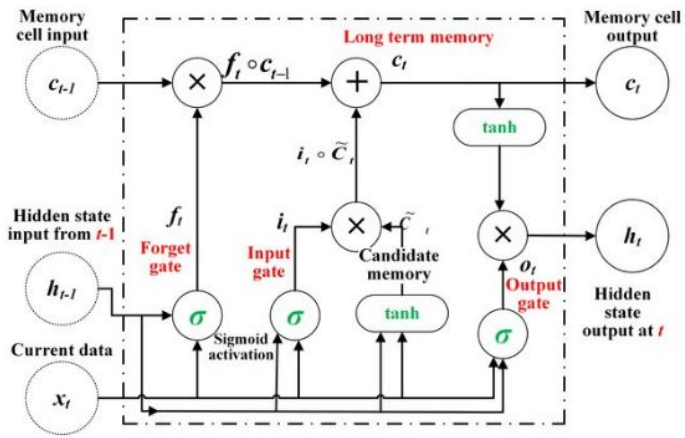
Scenario 2 - Fixing the Target Node to AgainstBrexit Stance and all the Source nodes which is attracted by target Nodes

Case2

Scenario 3 - Fixing the Target Node to ProBrexit Stance and all the Source nodes which is attracted to target Nodes

Political Stance Predictor

- LSTM is Variant of Recurrent Neural Network which introduces number of special and internal gates
- Internal gates help with the problem of learning relationships between long and short sequences in data
- **PRO:** Introduces many more internal parameters which must be learned - **Flexible**
- **CON:** Due to introducing many more internal parameters, It take time to learn - **Time Consuming**

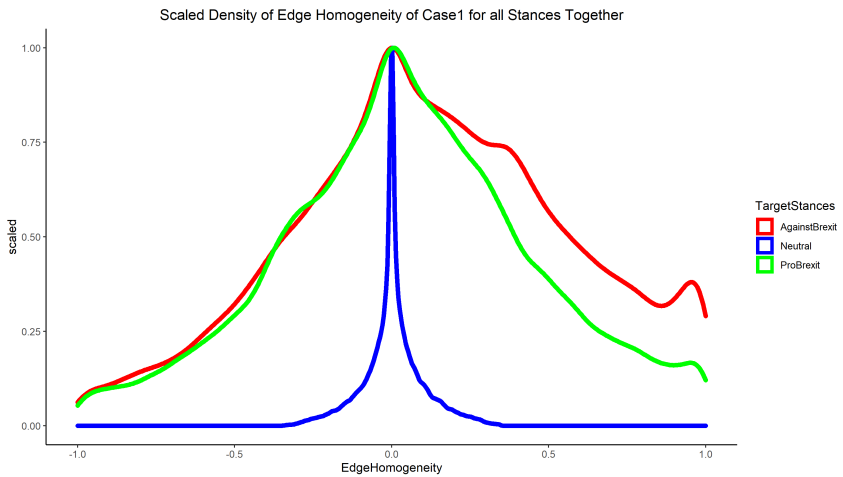


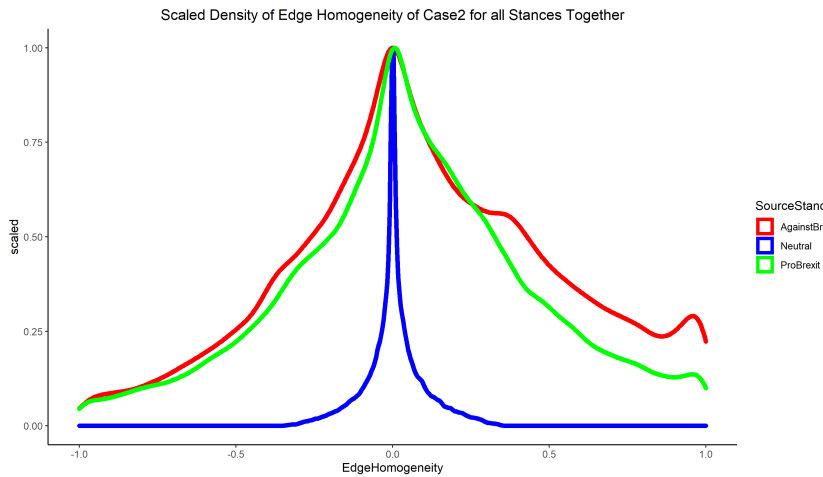
Political Stance Predictor

- **Intuition:**
  - We learn when to retain a state, or when to forget it.
  - Parameters are constantly updated as new data arrive
- **LSTM without Textual Features**
- **LSTM with textual features**
- **Performance metrics is measures with F1 Score, Accuracy, Recall and Precision**

Results

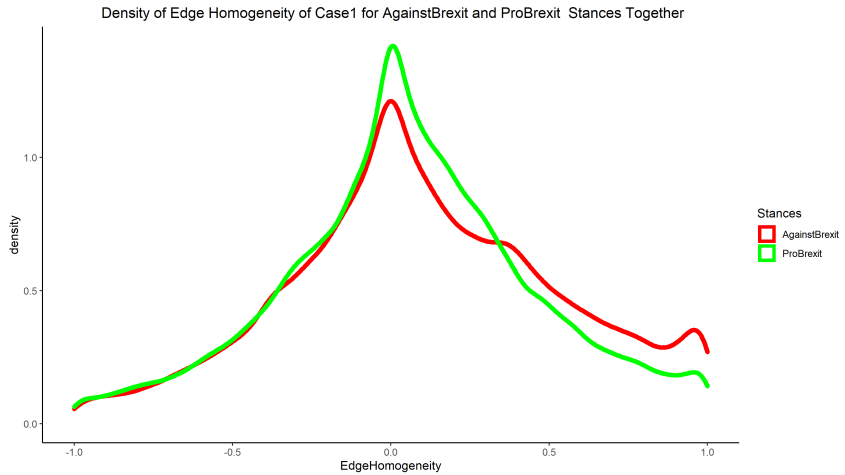
Probability Density Function of Edge Homogeneity of Case1 and Case2

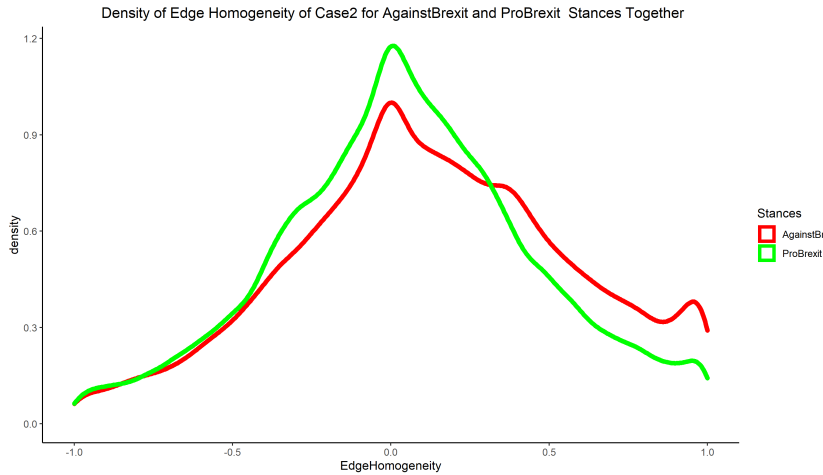




Results

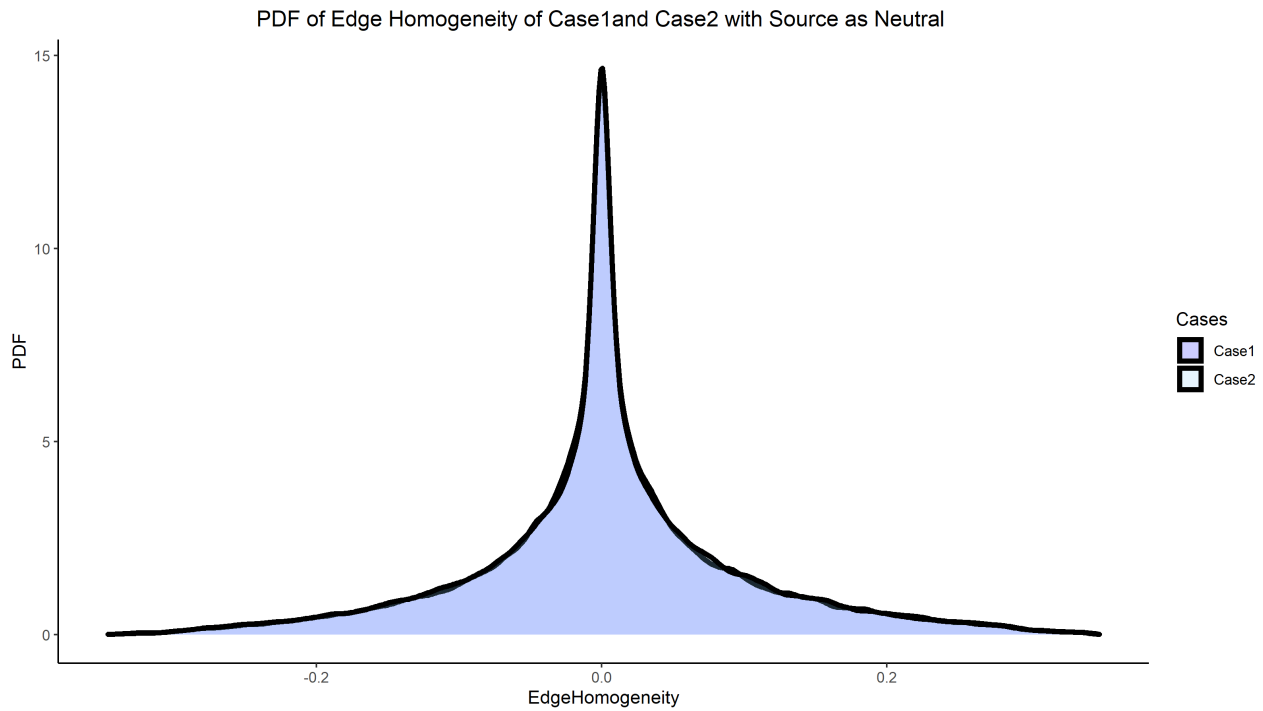
PDF of Edge Homogeneity of Case1 and Case2 without Neutral Stance





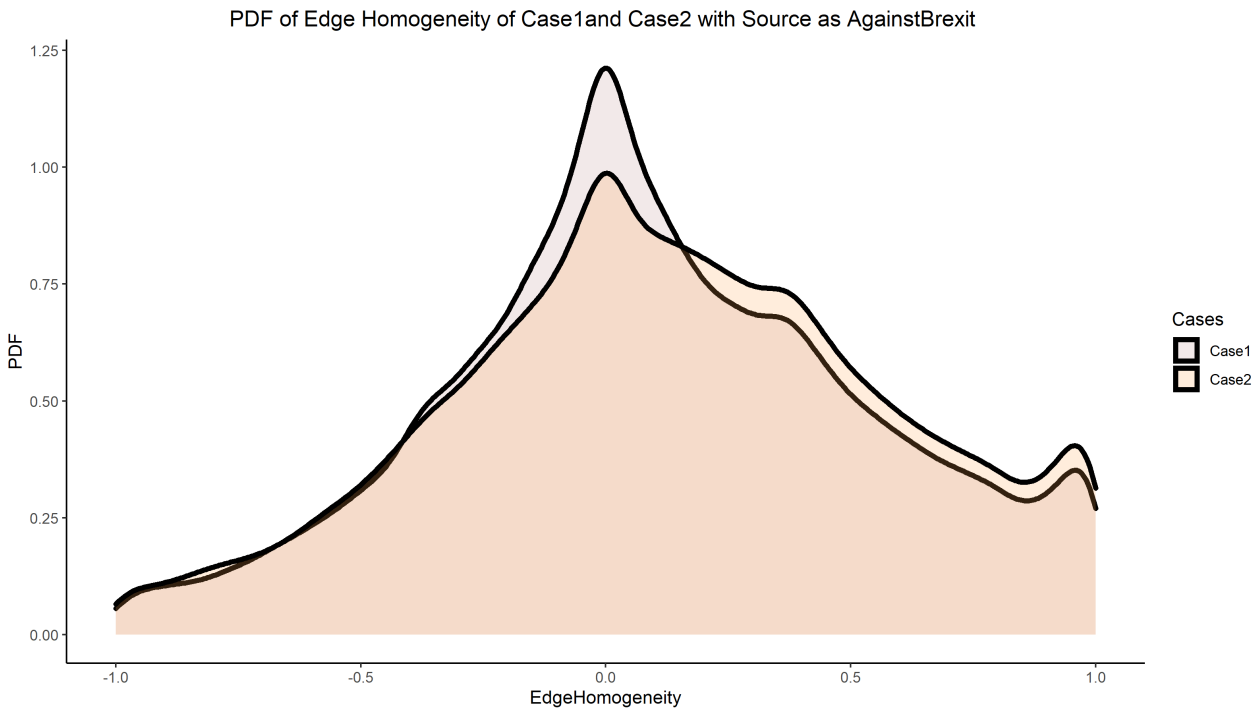
Results

Probability Density Function of Edge Homogeneity for Case1 and Case2 For Neutral Stance



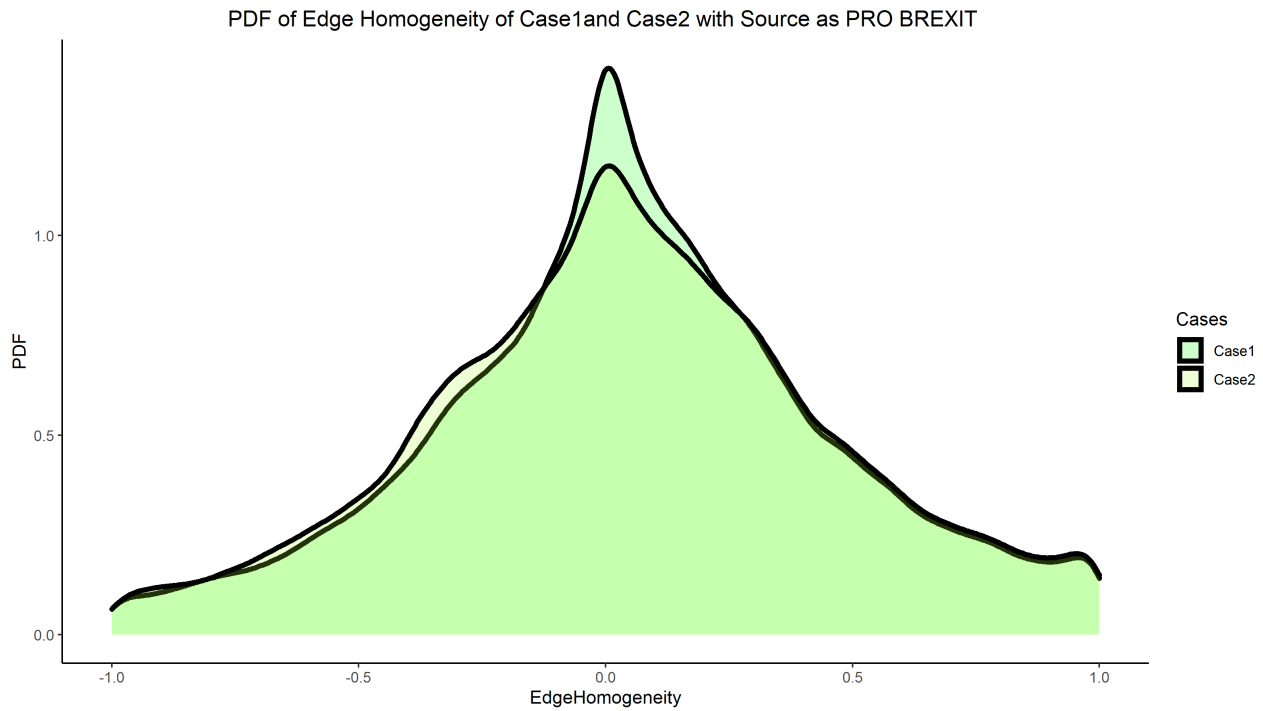
Results

Probability Density Function of Edge Homogeneity for Case1 and Case2 For Against Stance



Results

Probability Density Function of Edge Homogeneity for Case1 and Case2 For ProBrexit Stance



Results

Performance Evaluation of the Political Stance Predictor

Table: Confusion Matrix of Actual Stance and Predicted Stance from LSTM model

Actual —Predicted	**Against**	**Neutral**	**Brexit**
Against	3790	5544	3354
Neutral	5817	8174	4818
Brexit	3857	5238	3210

Results

Table: Performance Metrics for LSTM Model without Textual Features used for predicting user's stance based on their submitted posts.

Set	Accuracy	Precision	Recall	F1-Score
Test	0.350	0.3496	0.3496	0.3496

Table: Performance Metrics for LSTM Model with Textual Features used for predicting user's stance based on their submitted posts.

Set	Accuracy	Precision	Recall	F1-Score
Test	0.430	0.4236	0.4216	0.4236

Conclusion

- Edge Homogeneity on Source Node and Target Node around Brexit on reddit dataset
- Prediction of future political stance based on different features defined using the structure online diffusion
- Apply the Model on Dividing the data on different time periods
- Improve the model with adding additional hyper parameters

Thank You