Classification of news spreading barriers



Abdul Sittar, Dunja Mladenić, Marko Grobelnik, Josef Stefan Institute, Slovenia

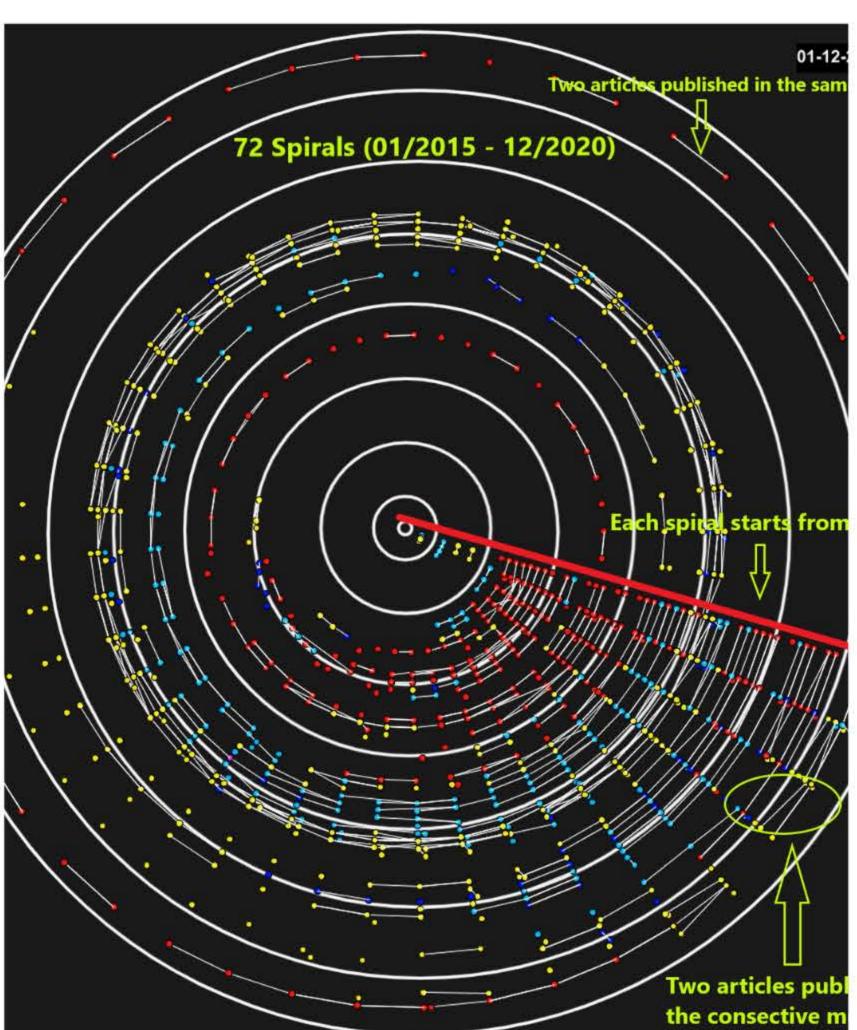
1 Introduction

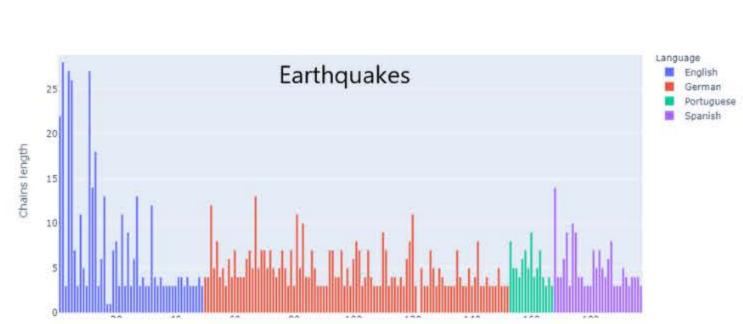
- News spreading is one of the most effective mechanisms for spreading information.
- Due to globalization, many events from different areas are internationally relevant such as COVID-19, and Brexit etc..
- News coverage directly catalogs the occurrence of specific events and indicates the local as well as global opinions of stakeholders.
- Representation of cross-lingual information about an event: to understand the entire story of current regional and international events that belong to diverse cultures.
- Determinants for news coverage are economic conditions and association between countries, cultural values (including publishing languages), geographical juxtaposition, and political alignment of news publishers.

2 Objectives

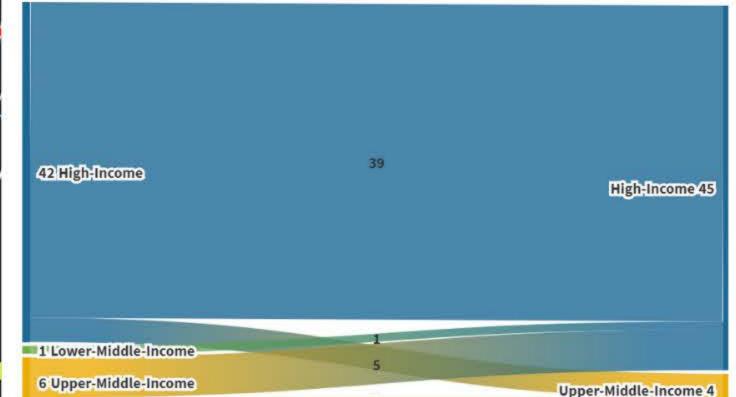
- Observing differences to news spreading on different events published by multiple publishers to understand what may influence the differences in the spreading patterns.
- Analyzing the news reporting differences and evolution of discussions across political and economic barriers using topic modelling.
- Developing a methodology to perform classification of news articles based on semantic knowledge including a wide range of common sense inferences and sentiments of news headlines.

3 Analysis





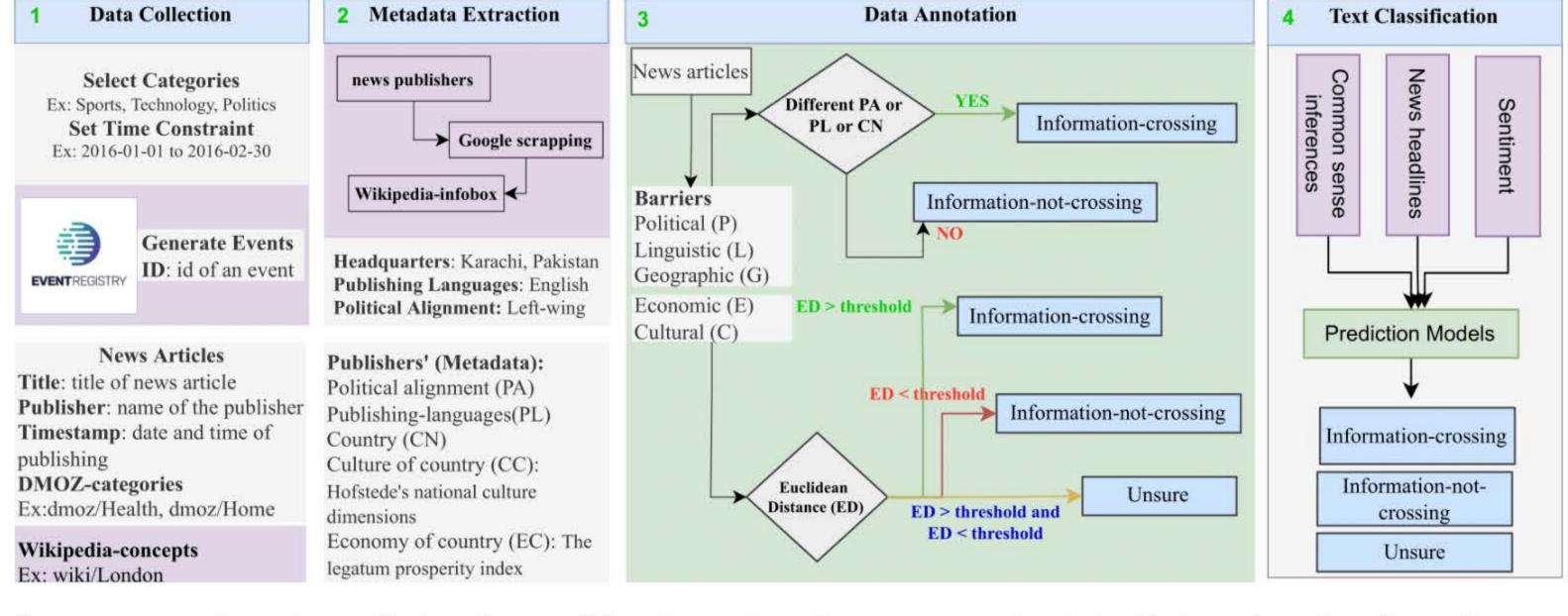
Overview of longest cascading chains in different languages related to events such as cascading chains in English, English and German for Earthquake, Global Warming and FIFA World Cup



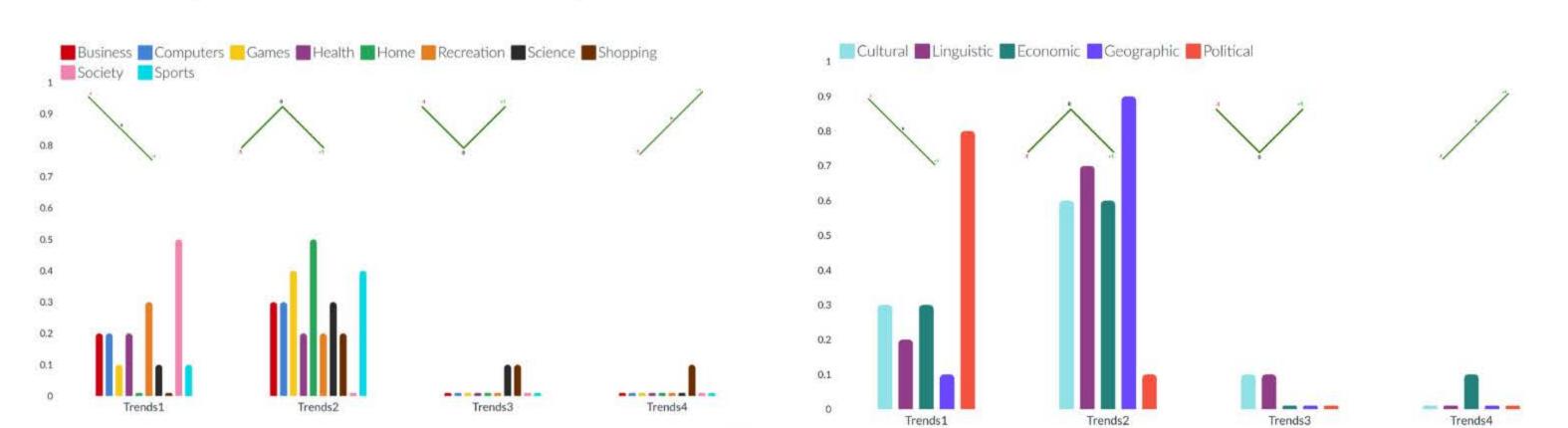
Visual depiction of multi-lingual temporal propagation for FIFA World Cup

Illustration of news spreading across different income levels

4 Approach



An approach to automatic barrier profiling based on the news meta-data. Data extraction from the Event Registry is the first step. Meta-data extraction through Google and Wikipedia scrapping is the second step. The third step is to label the news articles after calculating the euclidean distances. The classification with the classical machine learning models, deep learning, and transformer-based methods is performed in the last step.



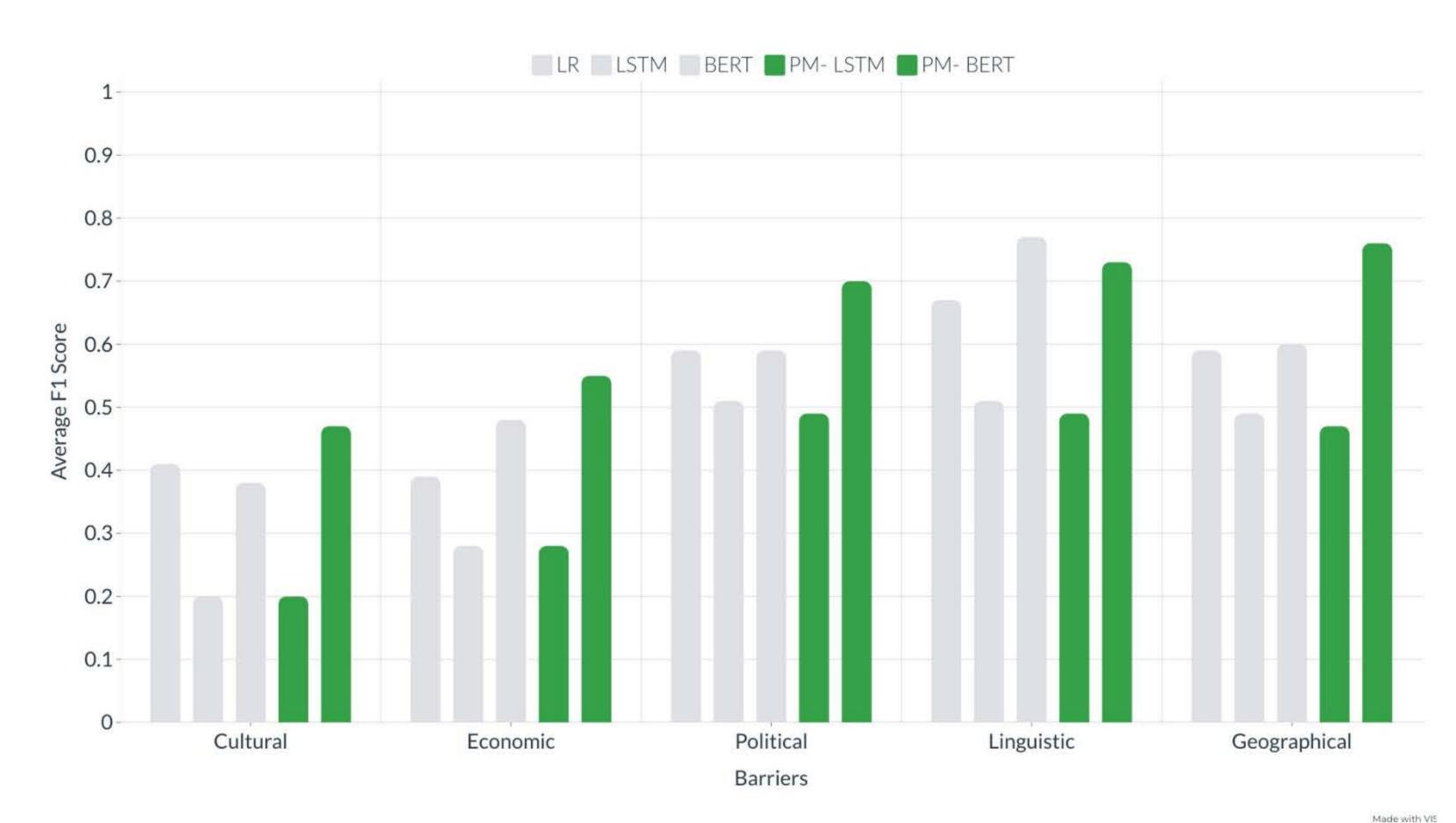
The bar charts present the distribution of different possible trends of sentiments across different categories and barriers (from left to right). The sentimental trends vary in four different types (see on the x-axis): trend1, and trend4 represent decrement and increment respectively in the percentage of news articles (see on the y-axis) with negative sentiment to neutral and then to positive: trend2, and trend3 represent decrement and increment respectively in the percentage of news articles with neutral sentiments than positive and negative sentiments



The network diagram presents an example of headlines with the common sense knowledge. Headline1 is "Uproar at Republican convention as anti-Trump delegates revolt", Headline2 is "Trump aims to show his softer side at Cleveland convention", and Head line 3 is "Protests turn violent outside Trump rally in New Mexico"

5 Results

- A novel methodology to analyze the news spreading barriers on different kinds of news events.
- A novel approach to enhance the topic modeling technique and understand political and economic differences in news reporting.
- An approach to barrier profiling by automatically annotating and classifying the news articles for the different barriers.



The green bars show the average F1-scores of all the ten categories for LSTM and BERT using commonsense-based semantic knowledge and sentiments. The gray bars show the average F1-scores of all the ten categories for LR, LSTM, and BERT using only headline text. The x-axis shows the groups of bars for all five barriers whereas the y-axis shows the average F1-score

6 Conclusions

Our findings suggest that:

- Events which may in some way involve political benefits are mostly published by those publishers which are not politically neutral.
- Countries with shorter time-zone differences and similar cultures tend to propagate news between each other
- Geographical size of a news publisher's country is directly proportional to the number of publishers and articles reporting on the same information
- The political alignment of a newspaper and the economic condition of a country influence news spreading
- Common sense-based semantic knowledge and sentiments of news headlines help to classifying the news-spreading barriers

References

- Sittar, Abdul, et al. "Political and Economic Patterns in COVID-19 News: From Lockdown to Vaccination." IEEE Access 10 (2022): 40036-40050.
- Sittar, Abdul, Dunja Mladenić, and Marko Grobelnik. "Analysis of information cascading and propagation barriers across distinctive news events." Journal of Intelligent Information Systems 58.1 (2022): 119-152.B
- Barrier Profiling in news spreading (Sittar, Abdul, Dunja Mladenić, and Marko Grobelnik)
- Cui, Yapeng, et al. "Modeling the dynamics of information dissemination under disaster." Physica A: Statistical Mechanics and its Applications 537 (2020): 122822
- Hong, Xudong, et al. "Cross-lingual event-centered news clustering based on elements semantic correlations of different news." *Multimedia Tools and Applications* 76.23 (2017): 25129-25143.
- Khosrowjerdi, Mahmood, Anneli Sundqvist, and Katriina Byström. "Cultural patterns of information source use: A global study of 47 countries." *Journal of the Association for Information Science and Technology* 71.6 (2020): 711-724.
- Kumar, Sanjay, et al. "Modeling information diffusion in online social networks using a modified forest-fire model." *Journal of intelligent information systems* 56.2 (2021): 355-377.