



Understanding the Impact of Geographical Bias on News Sentiment A Case Study on London and Rio Olympics

Swati, Dunja Mladenić

Jožef Stefan Institute and Jožef Stefan International Postgraduate School, Ljubljana, Slovenia

Overview

Introduction

Related Work

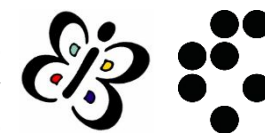
Graphical Abstract

Dataset

Materials and Methods

Results and Analysis

Future Direction



Introduction

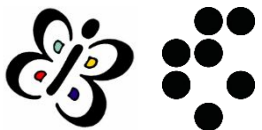
Is geographical bias a factor affecting news sentiment?

Case Study: Olympic legacies of Rio and London



Related Work

- **Existing work:**
 - News sentiment analysis
 - Geographical bias analysis in terms of news coverage
- **Proposed work:**
 - Investigate the impact of geographical bias on news sentiments



Graphical Abstract

News article on London Olympic Legacy

Headline: Five Years On, London's Olympic Real Estate Legacy Is A Clear Winner
Category: Business

Summary: The Olympics have become famously bad for real estate...

Sentiment: 0.067

Headline: Rio 2016's venues abandoned and vandalised in worrying legacy
Category: Business

Summary: Rio de Janeiro pulled off last year's Olympics, keeping crime...

Sentiment: -0.372

News articles on Rio Olympic Legacy

Headline: Rio 2016: An Olympics legacy of despair and wreckage
Category: Business

Summary: When any country hosts a major sporting event, the organisers and those...

Sentiment: -0.435

Headline: Rio de Janeiro suffering painful post-Olympic hangover
Category: Sports

Summary: The Brazilian metropolis of Rio de Janeiro was the epicenter of the...

Sentiment: -0.388

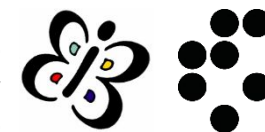
Article Similarity: 0.509, **Sentiment Dissimilarity:** 0.439

Article Similarity: 0.632, **Sentiment Dissimilarity:** 0.502

Article Similarity: 0.503, **Sentiment Dissimilarity:** 0.455

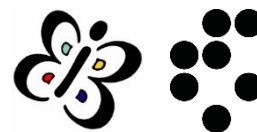
Dataset

- **Datasource:** EventRegistry
- **Language:** English
- **Duration:** January 2017 to December 2020
- **Fields:** Summary, Category, Sentiment
- **Type:** News Articles about London and Rio Olympic Legacy



Dataset

- **Fields:**
 - **Summary:** 290 to 6,553 words
 - **Sentiment:** -1 to 1
 - **Category:** business, politics, technology, environment, health, sports, arts-and-entertainment
- **Article Count:**
 - **London:** 8,690
 - **Rio:** 5,120



Materials and Methods

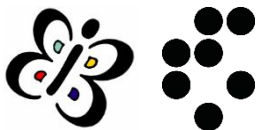
- **Methodology:**

- **Primary task:** to compute the average difference in sentiment scores between similar news articles about the Olympic legacies in Rio and London.
- **Secondary task:** to assess the primary task with respect to news categories.



Materials and Methods

- **Average Sentiment Dissimilarity**
- **Article Similarity:**
 - **Embedding:** Sentence-BERT (SBERT)
 - **Similarity score:** Cosine similarity
 - **Range:** -1 to 1
 - **Similar article:** score > 0.5

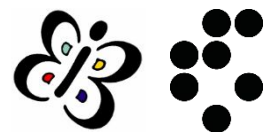


Results and Analysis

- Possible article pairs: 44,492,800
- Similar pairs: 375,008
- Average sentiment dissimilarity: 0.171
- Percentage of reversed sentiment polarity: 27%

	Positive	Negative
Positive	69	15
Negative	11	4

Confusion matrix to show the percentage of similar article pairs with respect to their sentiment polarity.

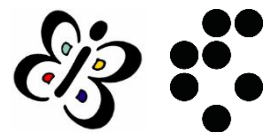


Results and Analysis

Impact of news categories:

News category	Average Sentiment Difference
Sports	0.19
Business	0.20
Politics	0.18
Health	0.16
Environment	0.22
Technology	0.14
Arts and Entertainment	0.19

Distribution of average sentiment difference across news categories for similar article pairs with identical category.



Results and Analysis

Impact of news categories:

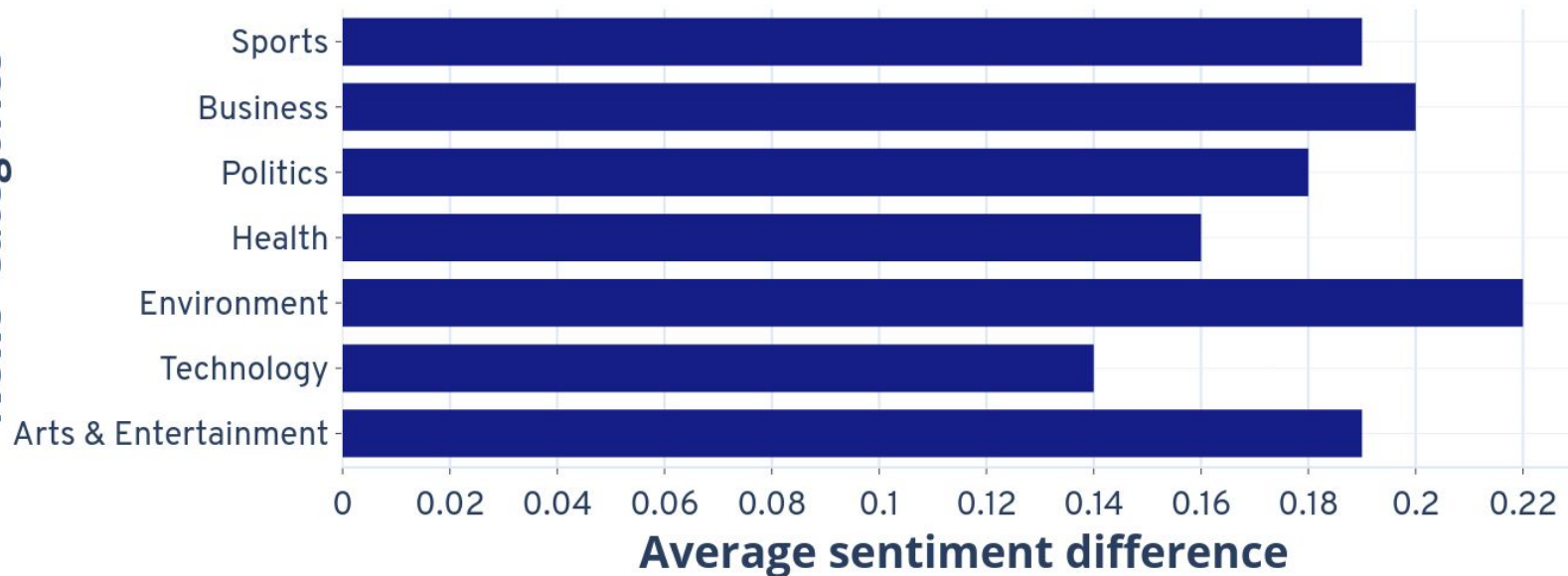
	Sports		Business		Politics		Environment		Health		Technology		Arts & Entertainment	
	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg
Pos	77	10	62	28	42	18	55	18	29	12	87	4	59	16
Neg	11	2	7	4	23	16	14	12	12	46	1	0	7	18

Category-wise confusion matrix to show the percentage of similar article pairs with respect to their sentiment polarity.

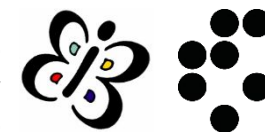
Results and Analysis

Impact of news categories:

News Categories

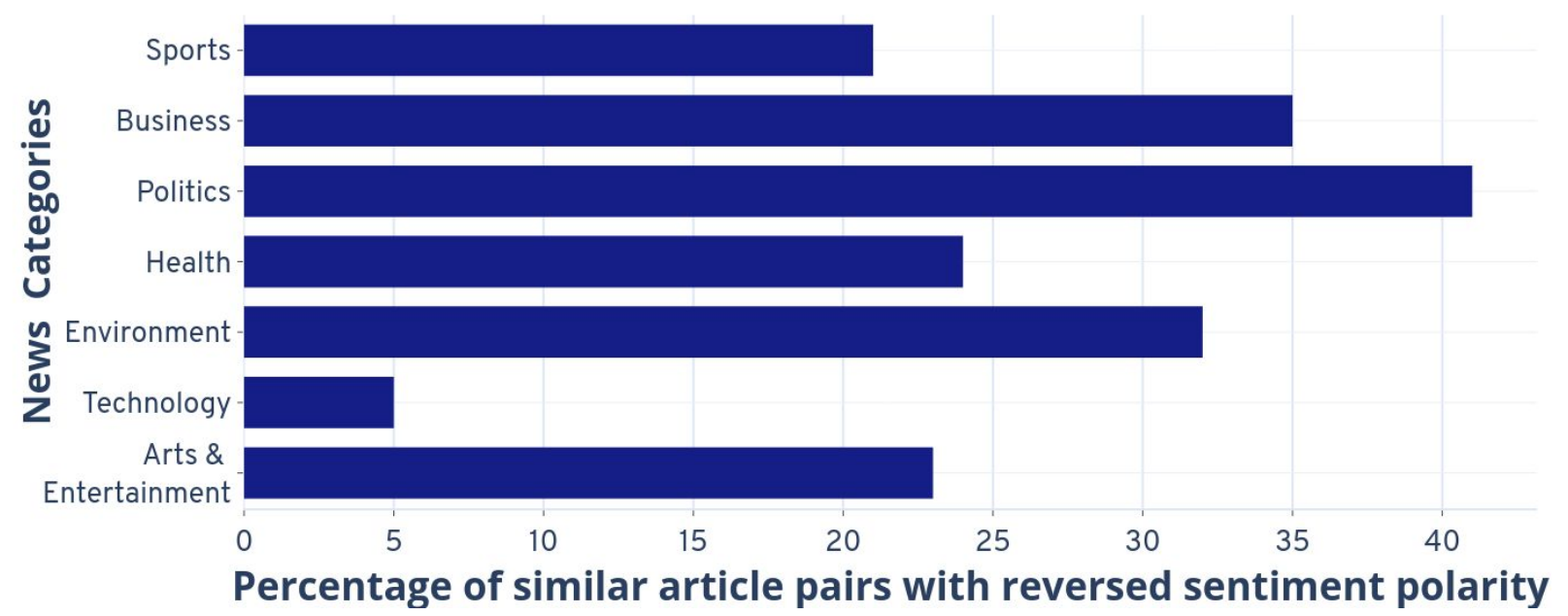


Distribution of average sentiment differences across categories for similar articles in the same category.



Results and Analysis

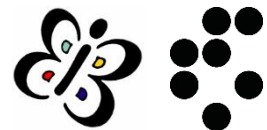
Impact of news categories:



Effect of category on sentiment polarity.

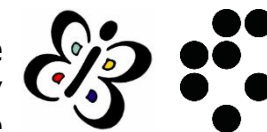
Future Direction

- Identify the most frequently mentioned topics in the Olympic legacy corpus.
- See how they affect the news sentiment of articles about different geographical locations.
- Explore the role of cultures and languages in this bias analysis.
- Discover the adjectives used to describe the negative and positive legacies of Rio and London.



Acknowledgements

This work was supported by the Slovenian Research Agency and the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 812997





Any Questions

