



Manovaad: A Novel Approach to Event Oriented Corpus Creation Capturing Subjectivity and Focus

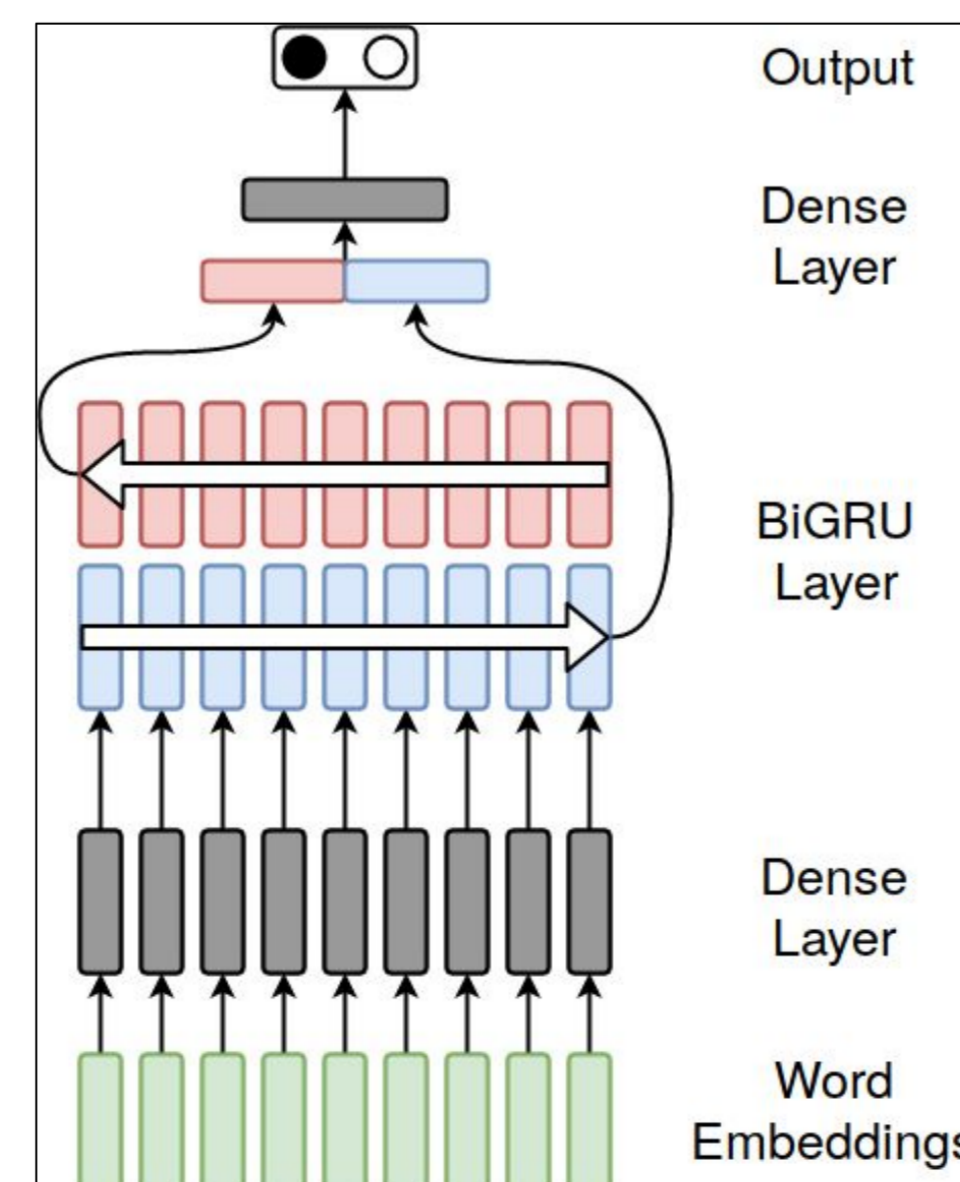
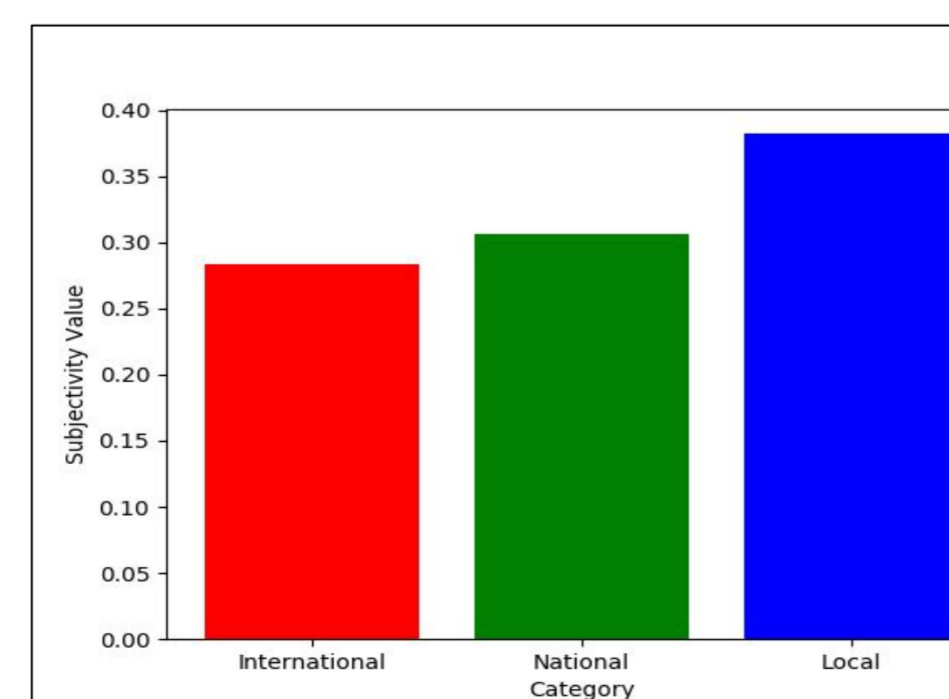
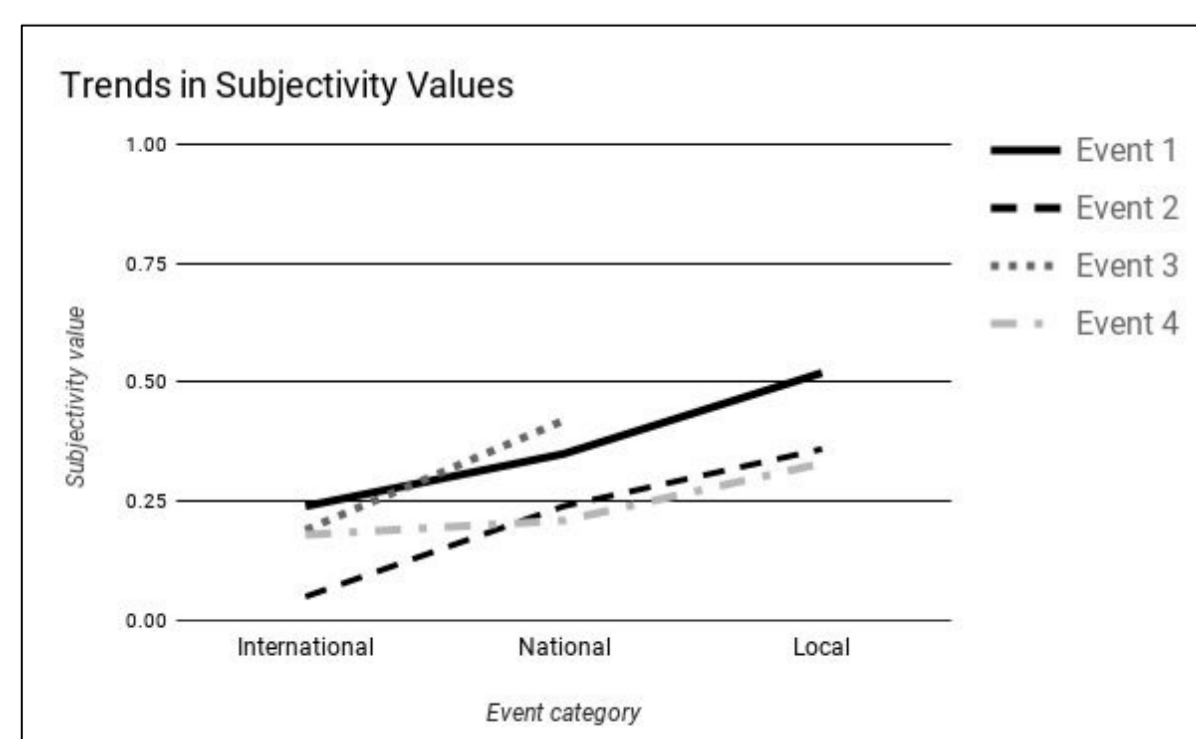
OVERVIEW

For every event across the world, there are conflicting opinions, arguments and disagreements, often reflected in print media and online social platforms. It is necessary to distinguish factual observations from personal judgements in news, as subjectivity in reporting can influence the audience's perception of reality. Several studies conducted on the different styles of reporting in journalism are essential in understanding phenomena such as media bias and multiple interpretations of the same event. This domain finds applications in fields such as Media Studies, Discourse Analysis, Information Extraction, Sentiment Analysis, and Opinion Mining.

We present an event corpus "Manovaad-v1.0" consisting of 1035 news articles corresponding to 65 events from 3 levels of newspapers viz., Local, National, and International levels. Using this novel format, we correlate the trends in the degree of subjectivity with the geographical closeness of reporting using a Bi-RNN model. We also analyse the role of background and focus in event reporting and capture the focus shift patterns within a global discourse structure for an event.

SUBJECTIVITY ANALYSIS

To analyse the trends in the degree of subjectivity of reporting at different levels, we implemented a Subjectivity Classifier using a Bidirectional Recurrent Neural Network. We observed subjectivity to be decreasing from local to international.



FOCUS SHIFTS

Focus: The most prominent information conveyed through a discourse element.

Unlike the usual method of studying focus independently without considering some global variables, we attempted to organise events and the corresponding articles at each level in a chronological order. For focus identification of each article, we used the Python NLTK and Gensim to generate the summary, considering the headline and the article. Then, we manually arrived at the focus for each article from its summary and the additional context provided from the previous article.

We observed maximum shift in focus in local articles, lesser in national, and no shift in International articles.

STRUCTURE BUILDING FRAMEWORK

Our observations are in agreement with the Structure Building Framework of discourse comprehension in which Gernsbacher (2013) proposed two main stages in creating a structure corresponding to each discourse.

Initial Stage: Creating a basic foundation for cognitive structure corresponding to an event

Mapping Stage: The readers further develop the cognitive structure representing the event by mapping new information onto the already existing foundation.

Time Period	Main Focus
9th - 12th Aug	Initiation of event, damage caused
12th - 20th Aug	Response of national level political parties
20th - 26th Aug	Solidarity and support from the country and other parts of the world

Time Period	Main Focus
11th - 13th Aug	Event initiation and description of the damage being caused
14th - 16th Aug	Rescue operations
17th - 19th Aug	Support from various state governments
20th - 27th Aug	Damage: deaths, casualties and financial loss
27th Aug - 1st Sep	Role of various political and other bodies
4th - 11th Sep	Damage caused to various industries and their specific support measures
12th Sep - 27th Oct	Post flood analysis - damage, causes, displacement of people, etc.
28th Oct - 7th Nov	Employment and Livelihood options for victims
8th - 24th Nov	Complaints and political accusations among various local and national parties
25th Nov - 4th Dec	Finances - loss, expenditure, compensation, etc.
5th Dec - 10th Jan	Rehabilitation measures