A Corpus Driven Text Analysis of "Fake Social Media News": A Case Study of The Indian Chronicles

Wardah Azhar¹
Syed Kazim Shah²

¹Government College University, Faisalabad – Pakistan
²Government College University, Faisalabad – Pakistan

Abstract

Online media has become a huge strategy for the tremendous extent of information, it spread in all occupations, including publicizing, and news-projecting, and that is only the start. This use of change is a direct result of some novel features, for instance, adaptability, free speech, and insight. Thus, a layperson is always subjected to being exposed to illegitimate “news”. Hence, the need to explore this phenomenon from the textual and linguistic perspective arose. This paper aims to investigate how “fake news” is generated linguistically and which textual features are employed by the “fake news” to make them look legitimate. The data was extracted from the report “The Indian Chronicles” – which is a source of certified counterfeit news. A systemic functional language approach is applied, where for exploration of textual features; a small corpus of fake news is compiled and tagged and through the software UAM, textual features are extracted and identified. Results have displayed, that fake news operates via centralizing participants: inanimate authority material the most, the strategy of manipulation, persuasion was most recurring, and so was foregrounding of inanimate material authority and lastly application of UAM to apply SFL concludes that the socio-cultural perspective of war and hatred in general, is the cause of this generation of fake news.

Keywords: CDA, Fake news, SFL, Social media, UAM

Author’s E-mail: wardah.azhar@gcuf.edu.pk

Introduction

Online media has become a huge strategy for the tremendous extent of information, sharing of information, and correspondence in all occupations, including publicizing, news-projecting, and that is only the start (Zafarani, Abbasi and H. Liu 2014). This use of change is a direct result of some novel features, for instance, adaptability, free speech, and
insight. Nevertheless, the insignificant exertion, basic access, and quick dispersal of information of online media draw a huge crowd and engage the wide expansion of disinformation, and fake news i.e., news with deliberately false information. A customary way to deal with discourse analysis is to deal with language past a level of a sentence in the syntactic order; it transcendentally revolves around connections between language and the setting (Nelson, 1998; Paltridge, 2012). All the more critically, visual pictures, which are emotional, polysemous, and culturally loaded, are joined into online messages occasionally for the movement of affiliated implications (Leech, 1976; Wilkins, 2012).

For such an explanation, other than Discourse-Centered Online Ethnography (e.g., Androutsopoulos, 2008), Multimodal Discourse Analysis (MDA), which dissects implications passed on by methods for a blend of text and pictures, can unquestionably be utilized to contemplate online discourse (Kress, 2011). Fake news about current social issues or strategy-focused issues is an over-accelerated spread of misinformation through online media with tremendous speed (Mehrdad and Anna, 2017).

Stories that emerge from Fake news can hoodwink individuals in its vicious net since their sources are limited to chiefly utilizing names the web delivers like legitimate news associations and few others.

There are additionally situations where fake news is produced unintentionally, however, it may likewise confound and delude crowds similarly to the way intentional fake news operates (Corner, 2017). Numerous individuals burn-through news and are educated about current political what's more; all parties from web-based media stages and organizes as cited by (Allcott and Gentzkow, 2017) and hence, it can prove to be a complicated process to distinguish
whether online stores are dependable or not.

A general non-attendance perception of how the web capacities, have furthermore added an extension in the fake news or stories which are not truthful which now has the overburden and in the absence of comprehension of how websites function to produce and spread fake news? Both online media and customers can have a significant impact in extending the spread of such stories (Spohr, 2017). Either way, solitary clients and client social events are acting to counter the spread of fake news through electronic media. These people's social occasions and their exercises are the focus of this investigation with the goal being a more complete understanding of how customers can perceive and combat infringing information can be made.

Major social media channels such as Google and Facebook have declared new measures to counter fake news with the presentation and announcement of call equipment, and even conventional media associations such as the BBC and Channel 4 have established truth-checking destinations.

In this way, it is essential to analyze clients' demonstrations of control over spotting and controlling fake news via web media. The instruments and strategies they use to recognize a fake story, as well as how they connect to it can be used to gain valuable data on how customers might act online to counter fake news via web media.

**Statement of Problem and Importance of Study**

A scenario exists in Pakistan, where the social media handle of various politicians inclusive of representatives of both the government and the opposition provides the masses with a chance to directly approach and connect with these leaders.

For instance, internet user commitment over data, for example, news stories, including posting about
various things, remarking on, or suggesting the news via web-based media, bear certain decisions of the users to the news and could fill in as assistant data for information and news identification of any news affiliation with any political wing. Also, web-based media information is animated which makes it a sort of feeble oversight. Users in online media can be both aloof content buyers and dynamic substance makers, causing the nature of client-created substances to shift. Informal communities are likewise loud with the presence of vindictive clients, for example, spammers and bots.

A Global Digital Report has revealed a dynamic increase in online media customers in Pakistan by a significant percentage of 5.7% till January 2019. Pakistan's flexible web customers as a degree of the general population were recorded at 21% at end of January 2019, as per the report. Customers of online users with a passage of 22% are considered as powerful online media customers at 37 million with an invasion of 18% and versatile electronic media customers at 36 million with a passageway of 18%, as highlighted in the report.

This new (yet subtle) sort of information orders new computational examination moves towards consolidated social speculations and factual information mining strategies. Because of the idea of online media commitment, we term these signs as powerless social oversight.

Along these lines in this research, the endeavors of learning with frail social oversight to comprehend and recognize (dis)information and news by "The Indian Chronicles" of biased propaganda, creating an anti-Pakistan international political scenario via online media is explored. Specifically, it studies how provoking situations are created to inculcate hatred against Pakistan.
The Rationale of the Study

This research offers a gander at political explanatory contentions utilizing online media as a stage. There is still a considerable amount to be found out about the arrangement of explanatory contentions and how this is formed by web-based media specifically regarding the case of "The Indian Chronicles". Aristotle characterized the way of talking as the capacity "to see all accessible methods for influence" in each discourse circumstance. (Griffin, 1994) The call of Aristotle toward balance and manner of speaking is at the core of the reason and reasoning of this study.

Academic Contribution

As per one investigation by Stanford University, rudimentary, centre, and secondary school understudies are incredibly terrible at deciding truth from fiction (Turner, 2016). Considerably all the more stunning, the absolute most effortlessly confounded news articles introduced by the investigation conductors were genuine notices found on websites. Truth be told, the research goes as far as to an extent that it aims to express that youngsters are more influenced by "how high a wellspring of data shows up on [internet] indexed lists as opposed to by the wellspring of the data" (Turner, 2016). Thus, the commitments of this exposition are summed up as: Provide instances of genuine news, parody, and phony news for understudies to analyze and separate between real and fake news.

Aims and Objectives

This study aims to explore textual features of fake news identified in the report “The Indian Chronicles”. In such a manner, we dissect our study to aim at the exploration of the following:

- Recognize fake news by its style from standard news using textual features as a guideline.
- Identify fake news regardless of whether it is skeptical,
satirical, or parody by use of ideational features.

- Distinguish fake news by means of its style alone, using interpersonal features.

**Literature Review**

**Definitions of Terms Used**

- The following terms and definitions may aid readers in understanding the content of this study:
- Comment/Remark: On Facebook (social media), when a client makes a unique post, different clients may compose an underlying reaction to the post. This underlying reaction is known as a remark or a comment.
- News: Political information shared by official forums
- Fake News/Bogus News/Counterfeit News/Phony News: deliberate presentation of (typically) false or misleading claims as news.
- Post: Facebook (Social Media) or on any online forum, any client may make unique substance for independently publishing on their public or private Facebook Profile. This unique substance is known as a post.
- Public: Public-confronting or available by the overall population; instead of private, in which exceptional authorizations are required.
- Political Discourse: Conversation or exchange related to the administration of a country; comprehensive of strategy, lawmakers, recent developments, ideological groups, and so forth.
- Reply: On Facebook (social media), the reaction of any client straightforwardly to an underlying remark on a post. On Twitter (social media), the
underlying reaction of a client to a unique tweet (answers on Twitter may likewise remain as Tweets).

- Tweet: On Twitter (social media), a posting made on the web-based media site, Twitter by an individual.
- User/Client: An individual (human individual) utilizing an individual (personal) web-based media account (e.g., Facebook client, Twitter client, and so on).

Definition of Fake News

In the simplest form, the definition of fake news can be summed up as "Counterfeit news" which is purposefully and unmistakably hoaxed and could deceive its recipients.

The term "fake news" is not new. Contemporary conversations, especially media thoughts, appear to portray fake news as viral posts that depend on fake records and are deliberately made to appear like news reports. Like a news report by Allcott and Gentzkow (2017), articles that are purposely joking and that could defraud readers are presented as fake news. As (Albright, 2016) points out, fake news has received snappy verbalization, especially after the 2016 US elections.

A direct but precise definition of the term fake news was given by Leonhardt and Thompson (2017) in their article in the New York Times, in which they portray fake news as such a decree and deceive the uncovering that contains mindful deception, disseminated through strategies for standard news sources (their print or online interpretations) or online electronic media. Of course, a news report (Tandoc, E, Lim, Z and Ling, R 2018) explains more about this by doing a study of the different types of fake news out there. The rating quickly perceives six types of fake news: news parody, news falsification, creation, control, publication, and disclosure (2018, p.147).
Examples of Fake News

For example, an article was dispersed on November 11th, 2020, by the European Chronicles, the news included was maintained by a sentence, that it was "news from European affiliation", proposing that it is "real" – which surely is a cognizant exertion to spread fake news in a precise way. The report was named "Pakistan-where the speed of advancement of the graveyard is more prominent than the monetary turn of events" with a side extension of naming it as an appraisal, while supporting information of it showed up with being more like real information. This story was reposted on different events by means of electronic media and it was recognized as a component of the subject.

Types of Mis and Disinformation

It can be argued that seven distinct types of content can be labeled as fake or problematic in a broader spectrum of online media in accordance with the scale, one that loosely measures the intent to deceive. They are as:

1. Satire or Parody
2. Misleading Content
3. Imposter Content
4. Fabricated Content
5. False Connection
6. False Context
7. Manipulated Content
Figure 1 Illustrates the types as shown below:

Figure 1: Types of Misinformation

Using Eliot's original list four additional motivations have been identified for the creation of this type of content: Poor Journalism, Parody, to Provoke or 'Punk', Passion, Partisanship, Profit, Political Influence or Power, and Propaganda, also known as 4Ps. (Elliot, 2017). Figure 2, illustrates how these 4Ps work with the seven types of Fake News and is also called the Misinformation Matrix.

Figure 2: Misinformation Matrix
Background

Described as a free type, reports, which are stories not transiently sequenced, have unique primary and semantic characteristics just as orderly register factors (Rose, 2012, 2014). Being a sort of media discourse performing explicit socio-cultural capacities, the news is never the objective as far as language use is concerned, for language structures are setting reliant and ideational (Van Dijk, 1985).

Theoretical Basis

Social media and fake news

Undoubtedly, news sources have generated a lot of interest in creating and improving progressive thoughts in news, dispersing content due to its online characteristics, and expanding relationships through electronic media and the Internet, driving the use of electronic media (Ahn Y, Jeong, H, Han, S, Kwak, H and Moon, S., 2007), and thus the Internet media becomes one of the main stages of disseminating information (Xiang and Gretzel, 2010).

Who produces Fake News?

Fake news stories started with a couple of objections and are fully configured to print misleading and intentionally transmitted articles, for example, denverguardian.com - the design and layout of this website are intentionally catered to produce fake news. The names of these territories are reliably chosen to take those of certifiable news affiliations. Other crafty grumblings contain articles that might be translated as genuine when seen out of the setting, for instance, the blueprint of wtoe5news.com. Various regions, for example, www.endingthefed.com, print a combination of obvious articles, routinely with a partisan trend, close to some misleading articles. The
objections that give false news everywhere are fleeting; in addition, the only one that was huge at the high level towards the 2016 political election as of now does not exist.

**Why Fake News is produced?**

In this research, one cause of the production of fake news is stories for financial benefits and the other cause of motivation is philosophical; more inclined towards certain ideological agendas. Specifically, the latter is found to be the most probable cause of the production and generation of fake news as presented in *The Indian Chronicles*. Some fake news outlets are trying to attract new talent so they prefer to engage in the generation and spread fake news. For example, the Romanian who ran www.endingthefed.com explains that he started the website primarily to support Donald Trump's campaign (Townsend 2016). Various traditionalist fake news providers claim that they are recognized as leftists and expect to shame those who have the advantage by showing that they would honestly stream fake stories (Dewey 2016; Sydell 2016).

**Research Questions**

To detect how fake news on social media operates specifically with the underlying agenda of maligning Pakistan's reputation at the international forum and global level, we are faced with several challenges: thus, this investigation is aimed at discovering the following:

- Which textual characteristics of fake news most frequently recur to form a pattern to produce specific discourse to alter the image of Pakistan globally?
- Based on the textual features, how does the spread of fake news in the social setting of online media structures quantify syntactic features?

**Methodology**

**Design and Choice of Methods**

This examination consolidates quantitative investigation. The
quantitative investigation is led on to explore frequently recurring textual structures, which may be attributed to designing fake news specifically. UAM software is used to analyze the syntactic patterns along with textual features employed in the generation of fake news.

The objective of the examination is to decide how fake news generates false propaganda against Pakistan in an open political talk utilizing computer-mediated -intervened web-based media and how this commitment ponders the exploration inquiries above.

**Data Source**

The data is sourced from a report called “The Indian Chronicles” which was originally published in the year 2019 internationally. The report was compiled as a result of a fifteen year long sting operation in which fake news was not only identified but their sources of production, linkages, and affiliations were found to be rooted from India. The content of most of the fake news circled around maligning Pakistan globally and to tarnish Pakistan’s contributions against war on terror. The authenticity of the report is taken into consideration, and since it was presented as evidence by the Defense Ministry of Pakistan in the UN, and considered legit, thus this research takes it as it is.

**Size and Sampling of Data**

The report “The Indian Chronicles” consists of eighty pages in which fake news and its sources are identified. This research incorporates all the text written in the report, with the exception of the table of contents at the beginning and hyperlinks given at the bottom of each page (these hyperlinks are meant to track the fake post and its source plus what could have been the real post).

**Data Analysis**

Exploration of research question one, as recorded above, is concerned about how the substance—structure or form of conveyance is influenced by the media. These
inquiries serve to address the presence of logical influence; however, likewise, serve to break down whether the substance is molded by the media as Postman's (1992) innovative determinism proposes. Hence, a mini corpus was developed. For the development of the corpus, the report was sectioned into twenty-eight clippings; each clipping contains the fake news identified, its text, and its editorial by the report generators. As for research question two as mentioned above, textual features are then individually examined in their context with the help of software UAM – with which the text is then analyzed according to SFL and its three layers of Textual, Interpersonal and Ideational levels. This analysis brings insight into how the fake posts linguistically operate to appear more valid at their face value.

**Results and Discussions**

**Textual Analysis**

The following tables show the results extracted from analyzing the corpus, with the use of UAM software.

<table>
<thead>
<tr>
<th>LENGTH:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Number of segments:</td>
<td>3306</td>
</tr>
<tr>
<td>- Tokens in segments:</td>
<td>43721</td>
</tr>
<tr>
<td>- Words in segments:</td>
<td>38175</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TEXT COMPLEXITY:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Av. Word Length:</td>
<td>5.18</td>
</tr>
<tr>
<td>- Av. Segment Length:</td>
<td>11.55</td>
</tr>
<tr>
<td>- Min. Segment Length:</td>
<td>1</td>
</tr>
<tr>
<td>- Max. Segment Length:</td>
<td>55</td>
</tr>
</tbody>
</table>
This table clearly shows that generators of fake news centralize the use of inanimate material authority to exert the impact of their content, and to validate it at face value by attributing the fake content to some random authority rather than some person or any specific office/chair.

For a better understanding, a visual chart is drawn below:
These results show that there is uniformity among the three layers of systemic functional language, which proves that this fake news is carefully curated for the audience to make it seem like real news and to increase its believability on social media.

The following table describes the Textual layer of the feature theme of the whole corpus.

Table: Textual layer of the feature theme of the whole corpus

The numbers show that, the token, segments, and words of the segment include the length of the corpus and that it is substantial enough to deduce that features such as text complexity, lexical density, and subjectivity showcase the careful customization of fake news which
A Corpus Driven Text Analysis of "Fake Social Media News": A Case Study of the Indian Chronicles

derives that this production of fake news is intentional and is aimed at maligning a certain topic or misleading it recipients.

**Analysis of Discursive Practices used**

When compared with the checklist, it was found that the text of the fake news resided to opt the discursive strategy of manipulation, using the technique of foregrounding to enhance the impact of the fake news. For example, in the title of the fake news: *Pakistan summons Swiss ambassador over posters in Geneva*: here Pakistan is foregrounded to emphasize the actions of the country against another entity.

**Mechanism of Study**

The main task of this research so far is to establish a theoretical framework for analyzing the form of fake news that is systematically generated against Pakistan.

This investigation discovers legitimate logical influence created by the political inclination of "fake pages" on social media, the introduction and examination of information and news by posts. Aristotle expressed, "There are, at that point, these three methods for affecting influence… to reason intelligently, to comprehend the human character and goodness in their different structures, [and] to comprehend the feelings" (Aristotle, 1984, p. 25) Those considering and utilizing explanatory examination refer the enticing bids by their unique Greek words, logos (or consistent allure), ethos (moral allure), and sentiment (passionate allure). In a convincing contention, the three entreaties should cooperate to convey a viable contention (Barrett, et al., 2013).

Especially with research question one, for example, consequences of examination question one might be controlled by the
mathematical portrayal of occurrences of influence toward (or away from) arrangements, and individuals, separately. It employs the use of Corpus, where a small corpus is compiled using the news post excerpts enlisted in the certified report of "The Indian Chronicles". Through, these corpora, the frequency of the most used structures are extracted with the help of the software AntConc., additionally, a pattern associated with a particular affiliation of anti-Pakistan propaganda is derived.

**Ethical Considerations**

In respect of the privacy of social media users, the study uses only public content of the report The Indian Chronicles and does not divulge the name(s) of any users.

**Limitations of the Study**

To start with, the investigation zeroed in on the content identified as fake news spread on social media as identified in the report "The Indian Chronicles" inclusive of political news posts and tweets during a particular timeframe. Even though the information remains a sensible subjective marker of how people utilize web-based media, the Pakistani political atmosphere existing apart from everything else surely shapes that information. Likewise, the current polarization of Pakistani legislative issues may influence the information; that is, maybe if the examination were directed 2-4 years earlier or later, various other outcomes can be found.

**Conclusions**

Genuine worries have been raised about the concerns of "counterfeit news" in the discretionary cycles, most outstandingly the criticism of Pakistan by the Indian lobby to segregate Pakistan universally. This examination is a work to participate in contemplating "counterfeit news" by zeroing in on how "counterfeit news" has been characterized and verbalized by apparently ordinary citizens via virtual entertainment. Over the time of 10 years, the Report "The Indian
"Corpus Driven Text Analysis of "Fake Social Media News": A Case Study of the Indian Chronicles"

"Chronicles" is a compilation of schemed systematic generation of fake news propagated as anti-Pakistan discourse about “fake news”, which is increasingly characterized by identity-driven and affect-laden language from both the liberal and conservative sides. Moreover, similar people specifically enhance group messages to guarantee the ability to indicate misrepresentation and characteristic fault as per bunch interests. Discussing "counterfeit news" goes past the elite class technique and turns into a profoundly political practice in customary residents' internet-based discourses.

**Future Recommendations**

It is a recommendation for future researchers to include the comments and replies to the fake news post and analyze from a perspective of critical discourse analysis as well. Researchers may include data from fake news posts, along with the public response (in the form of online interactions under the original post) for a more comprehensive picture.

**References**


