**ABSTRACT**

Intentionally deceptive content presented under the guise of legitimate journalism is a worldwide information accuracy and integrity problem that affects opinion forming, decision making, and voting patterns. Most so-called ‘fake news’ is initially distributed over social media conduits like Facebook and Twitter and later finds its way onto mainstream media platforms such as traditional television and radio news. The fake news stories that are initially seeded over social media platforms share key linguistic characteristics such as making excessive use of unsubstantiated hyperbole and non-attributed quoted content. In this paper, the results of a fake news identification study that documents the performance of a fake news classifier are presented. The Textblob, Natural Language, and SciPy Toolkits were used to develop a novel fake news detector that uses quoted attribution in a Bayesian machine learning system as a key feature to estimate the likelihood that a news article is fake. The resultant process precision is 63.333% effective at assessing the likelihood that an article with quotes is fake. This process is called influence mining and this novel technique is presented as a method that can be used to enable fake news and even propaganda detection. In this paper, the research process, technical analysis, technical linguistics work, and classifier performance and results are presented. The paper concludes with a discussion of how the current system will evolve into an influence mining system.

**EXISTING SYSTEM:**

Fake news has been shown to be detectable in several ways. Obviously, fact checking is one way to identify and debunk fake news; however, this is slow and difficult to automate. Batchelor [15] has suggested tasking libraries to help with this task. Automated detection, however, can occur at or near the speed of transmission, limiting the level of human involvement in certain areas of operations. Fake news has also been shown to differ from legitimate journalism in structural and other ways. Horne and Adali [16] note that fake and legitimate news differ in title length and the simplicity and repetitiveness of body text. Rubin,

**PROPOSED SYSTEM:**

Given the problem posed by fake news, a variety of approaches have been proposed to detect it automatically [19]. Chen, et al. [20] make the case for the need for automated detection, for speed and convenience, among other prospective needs that would be met. Unlike crowdsourcing [21] and using human employees for review, automation can result in nearinstant decisions and provides requisite scalability. Riedel, et al. [22], for example, proposed a headline stance based detection technique. Rashkin, et al. [23] use a language analysis-based approach, while Jin, et al. [24] propose a “hierarchical propagation” approach and Shu, et al. [25] use a datamining process.

The simple Python and Textblob system used to identify quotes in a paragraph worked well. Because the tool looks for the presence of double quotes inside strings, quote or content identification was simple. The tool identified 96% of the quotes in the training set. We assess that the tool was confused, in limited instances, by complex and malformed quotations.