

FAKE NEWS DETECTION

A SEMINAR REPORT

Submitted by

SUJOY DAS

Roll No : GAU-C-17/054

PRASTUTI KOCH

Roll Number: GAU-C-17/289

SHIVAM GUPTA

Roll Number : GAU-C-17/071

7th Semester

Department of Computer Science and Engineering

Under the guidance of

Mr. Dharani Kanta Roy

(Department of Computer Science and Engineering)

in partial fulfillment for the award of the degree
of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE AND ENGINEERING

CENTRAL INSTITUTE OF TECHNOLOGY, KOKRAJHAR



(Deemed to be University, MoE, Govt. of India)

April 2021

Certificate

This is to certify that this project report titled “Fake news detection” is the bonafide work of “Sujoy Das [Gau-c-17/054], Shivam Gupta [Gau-c-17/071], Prastuti Koch [Gau-c-17/289]”, who carried out the project under my supervision. Certified further, that to the best of my knowledge, the work reported herein doesn’t form any other project report or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

1 Abstract

Everyone depends on numerous online resources for news in our modern world, where the internet is ubiquitous. As the use of social media sites such as Facebook, Twitter, and others has grown, news has spread quickly among millions of users in a short period of time. The Internet is an incredible resource for news and information, but unfortunately not everything online is trustworthy. Fake news is any article or video containing untrue information disguised as a credible news source. While fake news is not unique to the Internet, it has recently become a big problem in today's digital world. In this project, we aim to perform binary classification of various news articles available online with the help of concepts pertaining to Artificial Intelligence, Natural Language Processing and Machine Learning. We want to give users the ability to identify news as false or actual, as well as verify the credibility of the website that published it.

2 Acknowledgements

We would like to express my deepest gratitude to our guide, Mr Dharani Kanta Roy for his valuable guidance, consistent encouragement, personal caring, timely help and providing me with an excellent atmosphere for doing research. All through the work, in spite of his busy schedule, he has extended cheerful and cordial support to us for completing this project work.

We express our heartfelt thanks to our Head of the Department, Dr Amitava Nag who has been actively involved and very influential from the start till the completion of our project.

We would also like to thank all teaching and non-teaching staff of the Computer Science and Engineering Department for their constant support and encouragement given to us. Last but not the least it is our great pleasure to acknowledge the wishes of friends and well wishers, both in academic and non-academic spheres.

Table of Contents

List of Figures

Introduction

Fake news has rapidly become a social issue, with people using it to spread misleading or rumoured information in order to influence the behaviour of the general public. During the 2016 U.S. Presidential Election, the proliferation of fake news revealed not only the risks of fake news's consequences, but also the difficulties of separating fake news from real news. While the word "fake news" is recent, it is not actually a new phenomenon. Fake news has been around since the 19th century, when one-sided, partisan newspapers first appeared and became popular. However, technological innovations and the distribution of news through multiple forms of media have amplified the propagation of false news today. As such, the effects of fake news have increased exponentially in the recent past and something must be done to prevent this from continuing in the future.

Therefore, our goal is to use machine learning to classify, at least as well as humans, more difficult discrepancies between real and fake news. There are two methods by which machines could attempt to solve the fake news problem better than humans. The first is that machines are better at detecting and keeping track of statistics than humans. Additionally, machines may be more efficient in surveying a knowledge base to find all relevant articles and answering based on those many different sources. Either of these methods could prove useful in detecting fake news. In this project, both these techniques have been used to predict fake news. We have used supervised learning in order to make a binary classifier that will predict fake news only within the source in question. Furthermore, to make the project dynamic in nature we have implemented the second technique of automatic surveying through the use of web page scraping and databases. The goal of this project is to find the effectiveness and limitations of language-based techniques for detection of fake news through the use of machine learning algorithms and computerised surveying. This type of solution is intended to be an end-to end solution for fake news detection. The project however, at times cannot predict fake news with complete certainty.

With that being said, the project is intended to be used as a tool and it is close to an end to end solution. This tool can be used for future applications and can be combined with other tools to provide much better results in fake news detection.