ACKNOWLEDGEMENTS

We feel privileged to express our deepest sense of gratitude and sincere thanks to our guide **PROF. PANKAJ VANWARI** for his excellent guidance throughout the project work. His prompt and kind help directed us to the completion of dissertation work.

We would also thank the lab assistance of computer department for providing us with the requisite facilities to complete the project work.

Finally, we would like to acknowledge all our friends those who directly or indirectly helped us in the completion of our project.

TABLE OF CONTENTS

ABSTRACT

Fact checking is the process of attempting to verify or disprove assertions made in speech, print media or online content. The practice is essential for integrity in any area where claims are made, including government, journalism and business.
Misinformation may be intentional or simply a matter of lax fact checking or other human errors. Anyone who presents material as factual should ensure that it does not contain any false statements. Failing to do so can result in damage to a business' reputation and in more extreme cases can result in compliance issues and/or legal action. Automated fact checking systems are the need of the hour as large amount of fake news is being generated on a daily basis.

INTRODUCTION

Despite providing huge amounts of valuable information, the web is also a source of false claims in social media, other web-sources and even in news that quickly reach millions of users. This tends to mislead the people and makes them believe things which are not actually true. This issue is present in many domains, ranging from fake reviews on product websites, erroneous stock prices, manipulative statements about companies, celebrities, and politicians.

The researchers at the Ohio State University have found that fake news played a vital role in determining the results of the 2016 United States Presidential elections. The study suggests that fake news probably played a significant role in depressing Hillary Clinton's support on Election Day. A Princeton led study of fake news consumption during the 2016 campaign found that false articles made up 2.6 percent of all hard-news articles late in the 2016 campaign, with the stories most often reaching intense partisans who probably were not persuadable. A recent report by the Jumpshot Tech Blog showed that Facebook referrals accounted for 50% of the total traffic to fake news sites and 20% total traffic to reputable websites. Since as many as 62% of U.S. adults consume news on social media, being able to identify fake content in online sources is a pressing need. With the 2019 Indian General Election round the corner it is very essential to have accurate fake news checking systems in place in order to avoid a repeat of the presidential elections in the US. Detecting false claims and validating credible ones is a challenging task. Many times this turns out to be challenging for humans as well. This has resulted in the rise of many fake news detection websites like snopes.com, politifact.com and Buzzfeed which manually check each claim and provide information whether it is true or not. However these websites are usually constrained to a particular domain and require human expertise to verify the claims and provide any kind of generalization over other domains.

We propose a system to automate the task of detecting fake news which will not require any kind of human supervision. Our system aims at detecting the fake news without any bias towards a particular domain or personality.

AIM AND OBJECTIVES

The aim of this project is when given a claim it must identify whether it is true or not. Our system consults the web articles in order to determine the credibility of the claims. The main steps are:

1. Generating the data
2. Computing overlap score

3. Stance detection
4. Aggregating the scores to generate the output label

The objectives are:

- To automate the task of checking fake news.
- To improve the stance detection in case of implied news reporting.
- To reduce the noise in the training data by focussing only on the relevant parts of the news articles.
- Reducing the time required for training the classifier.

LITERATURE SURVEY

CredEye:

This research proposes an end to end system for claim checking. Initially a knowledge base lookup is performed to check if the credibility of the claim can be verified from the knowledge base itself. The claim is given as a query to Bing API and the top 30 articles are collected. Various regional filters are applied to the API. Based on the corpus obtained from the Bing API, classifiers are trained for identifying the stance of a news article towards a claim. The linguistic features are captured using a predefined set of lexicons. From each article a count of various verbs like assertives, factives, hedges, implicatives is maintained. Along with this a list of negative and positive words is maintained which is referred to while vectorising the claim. For data collection the snopes dataset is used which contains many claims which are labeled manually. While training the classifier it is assumed that all articles based on a certain claim will inherit the the claim’s label given in the snopes dataset. Along with detecting the stance the system also takes the credibility of the web sources reporting the article into consideration. A web-source is considered to be reliable if it contains articles that refute the false claims and support true claims. The vector obtained from credibility analysis of the web-sources is concatenated to the vector of linguistic features to give the final representation of data. The vector for each article forms the training data for the classifier. An aggregate of the credibility scores for each article is found to determine the final label of the claim.

ClaimBuster:

The main aim of this system is to identify the check-worthy claims in a news article. Each statement is categorised into three types:

- Non-Factual Sentence
- Unimportant Factual Sentence
- Check-worthy Factual Sentence

The data is taken from the US general election presidential debates. To label the data a survey was conducted where students were asked to manually classify each sentence into one of the three categories.

From the data generated supervised learning is used to classify the statements in the article as checking worthy and non check worthy. POS tagging, length and sentiments of the words are used as feature selection techniques.

Assertives by Hooper 1975:

Assertive verbs are those whose complement clauses assert a proposition. The truth of the proposition is not pre-supposed, but its level of certainty depends on the asserting verb. Whereas verbs of saying like say and state are usually neutral, point out and claim cast doubt on the certainty of the proposition

Hedges by Highland:

Hedges are used to used to reduce one’s commitment to the truth of a proposition, thus avoiding any bold predictions or statements. It is used to soften one’s stance towards a claim.

Factives by Kiparsky and Kiparsky:

Factive verbs presuppose the truth of their complement clause.

PROBLEM STATEMENT

The system must be capable of detecting the credibility of the claim given as an input. For every news article reporting on the claim, the stance of that article towards the claim must be identified.

The articles which are not reporting on the context of the claim must not be taken into consideration. A system must be able to understand the reporting type of the articles and capture the linguistic features present in them.

These features must be used to train a classifier to predict the label of a given claim.

SCOPE

The system we have proposed takes a claim from the user as an input and then checks whether it is true or not. Certain limitations of this system are:

1. Claim as a sentence:
Our system can only accept the claims as a single sentence. An entire news article or a paragraph can not be checked for veracity.

2. Factual statements:
The input must be a statement which contains a fact. Non-factual statements such as “Indian Cricket Team is the best cricket team in the world.” can not be assessed as it is an opinion and not a factual statement.
An example of factual statement is: “India is currently ranked 1st in the ICC Test rankings.”

3. Non-volatile claims (w.r.t Time):
Current work focuses only on claims which are stable with respect to time and do not keep changing continuously. Eg: “Crime rate is increasing.”

This claim can not be assessed by the system as it is too volatile to talk about. In terms of statistics the crime rate will keep changing for short intervals of time as well.

PROPOSED SYSTEM

Our system identifies the veracity of the claim by detecting the stance of the news articles towards that claim. We train a stance classifier using distant supervision to find the stance of any article towards a given claim. To improve upon the existing systems we are identifying the parts of the news articles which are more closely associated to a given claim in order to focus only on the important part of the information. This will eventually help in improving the signal to noise ratio. We are using the snopes dataset where claims are manually labeled to generate the corpus. It will then be divided into train and test data.

METHODOLOGY

The methodology of our system can be represented in the following six steps:

1. Extraction of news articles:

The claim is given as an input to the google search package. The top data from the top five articles is then dumped. This process is repeated for each claim in the dataset to generate a corpus.

2. Computing the overlap of the articles with the claims:

Each sentence in an article is passed through an overlap function that determines percentage overlap between the claim and the sentence. To overlap is computed by matching the unigram and bigrams of the claim and the target sentence.

3. Extracting the required snippets from each article:

A threshold on the percentage overlap is determined to identify which parts of the articles are to be considered for training. These parts are known as snippets.

4. Finding the linguistic features from the snippets:

Assertive and factive verbs, hedges, report verbs, subjective and bias words are used to capture the language style of an article. These are used to vectorise the data which is then passed on to the classifiers for training.

5. Training a classifier using distant supervision:

The vectorised data obtained from the previous step is used to train classifiers like logistic regressions, SVM and Random Forests. While training we assume that the unlabeled Web articles predominantly inherit the claim’s label thus making it distant supervision.

6. Generating the output:

The average of the credibility score of each article is calculated. Based on this average score the output label is determined.

ANALYSIS

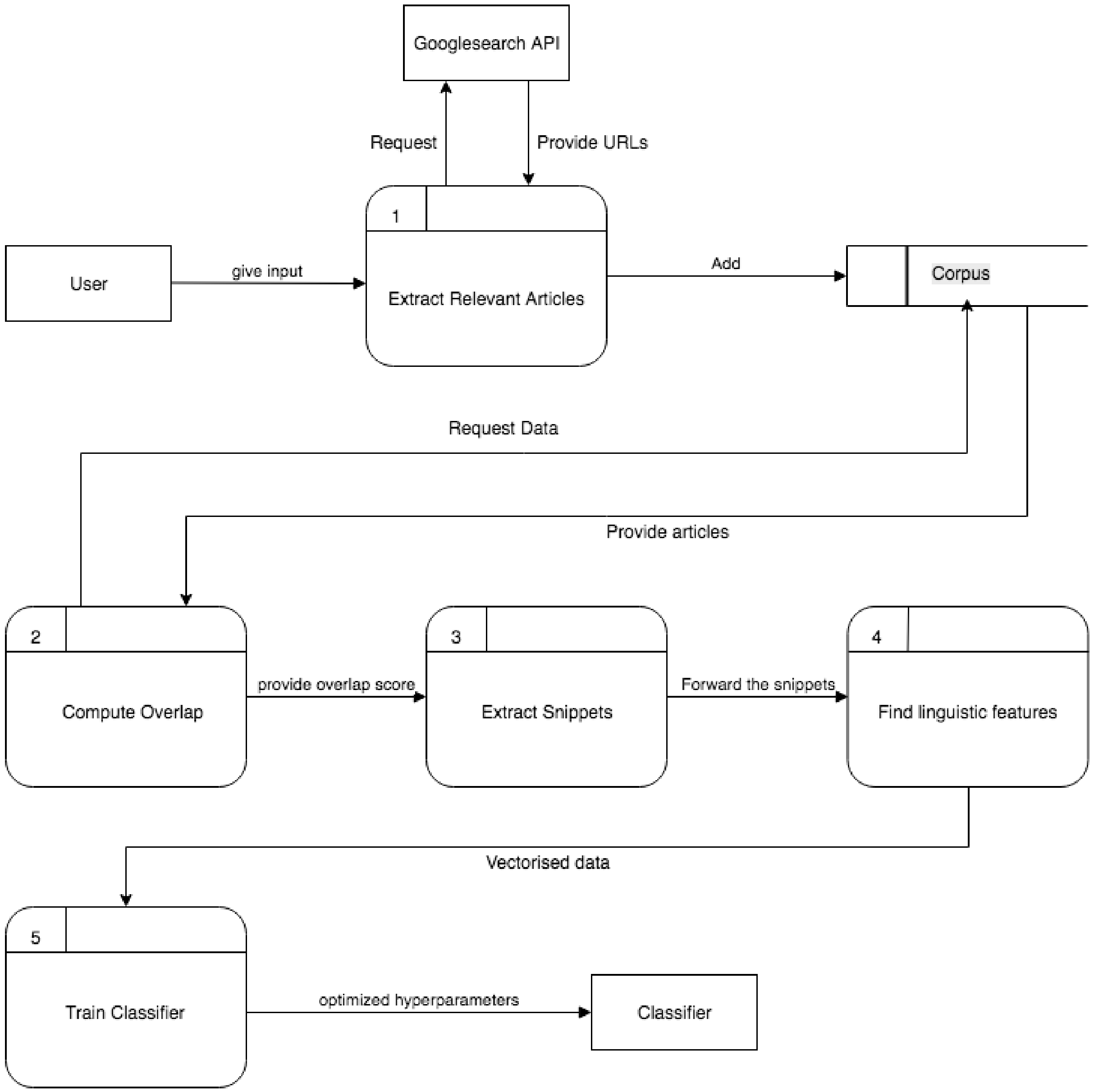
Feasibility analysis:

When the input is given in the form of text, it becomes much easier to find the articles that are directly reporting on that claim and have the same context as the context of the claim. In such cases a fully automated claim checking system can be developed as it is possible to detect the stance of each article towards these claims.

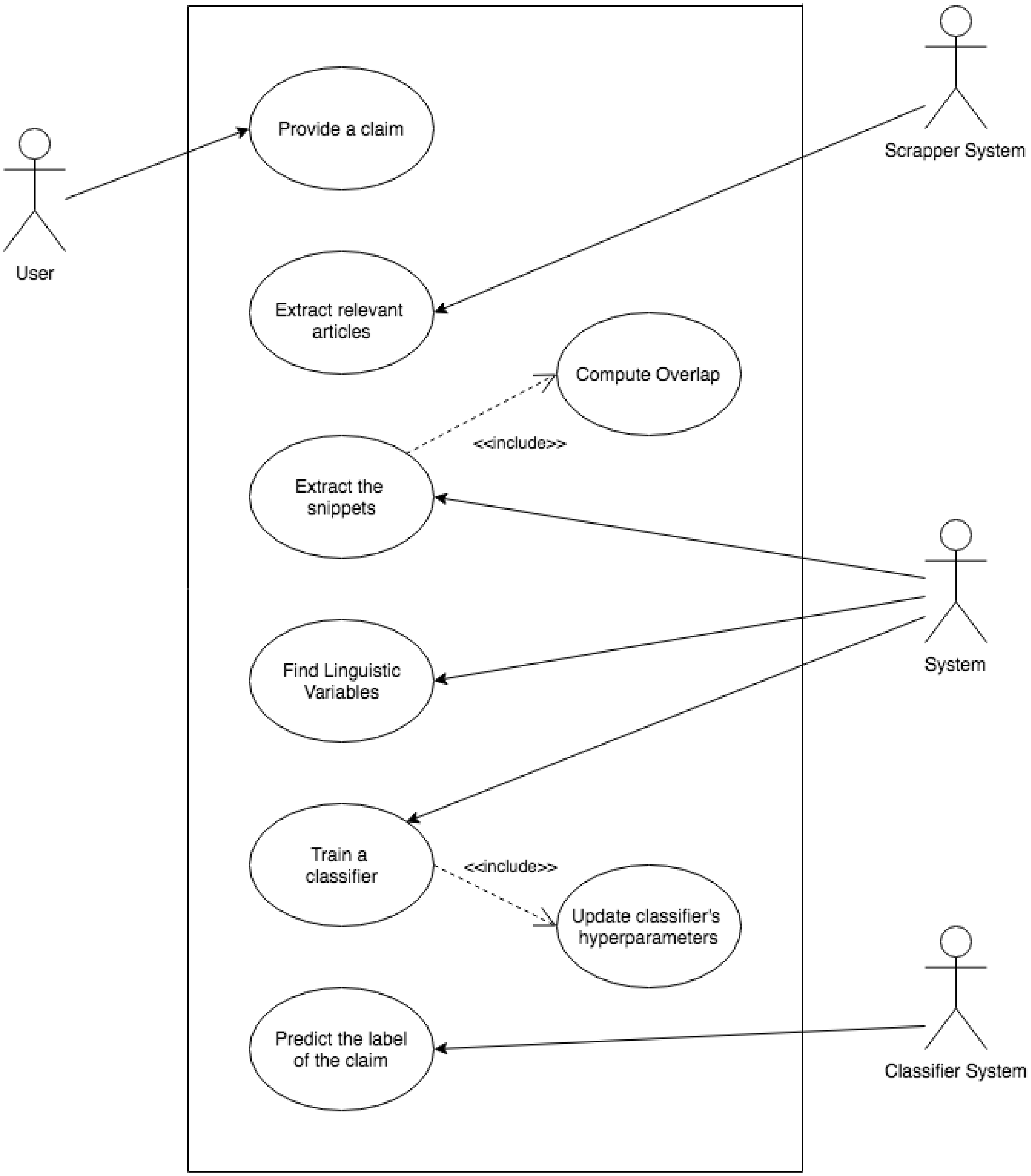
However when it comes to checking the story behind an image certain amount of human intervention is still required. For an image an automated system could make multiple interpretations and ultimately it comes down to a human to identify which one is the one in context. The system could provide the human expert with the various story leads associated with the image and then the expert can take it from there.

DESIGN

Data Flow Diagram:

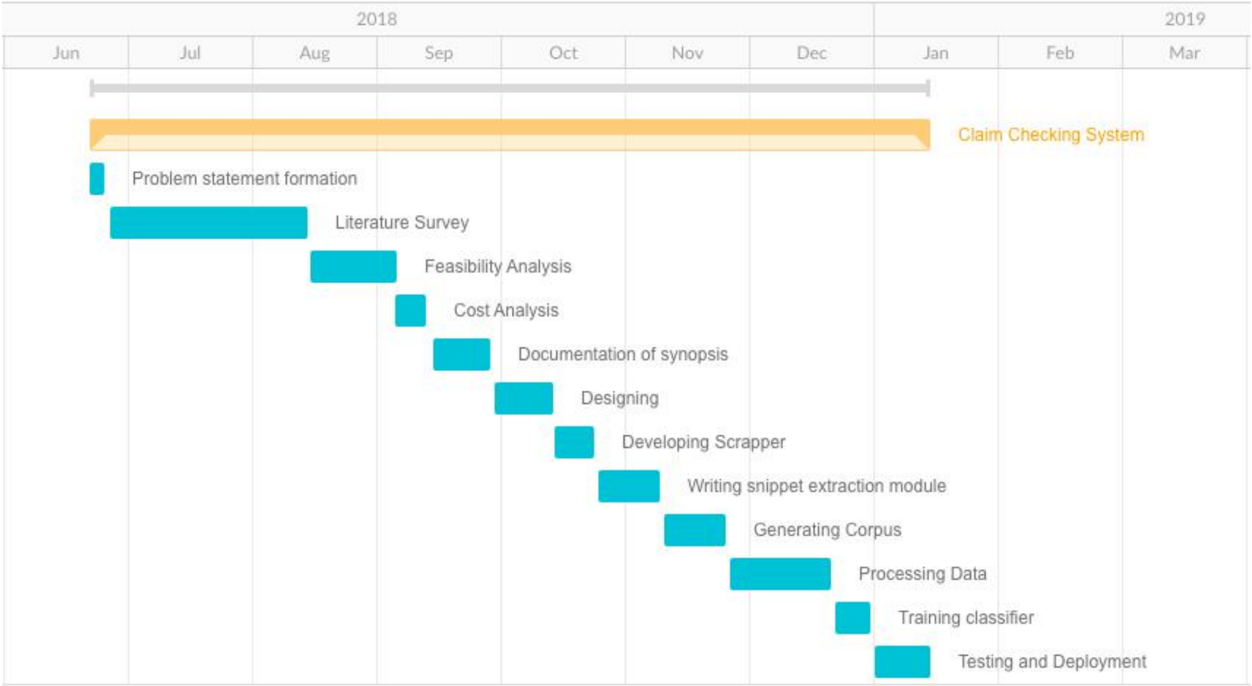


Use Case Diagram:



Timeline Chart

Sr.No	Task	Start Date	End Date	Duration
1	Problem Statement formation	21/6/2018	25/6/2018	4 days
2	Literature Survey	26/6/2018	14/8/2018	48 days
3	Feasibility Analysis	15/8/2018	5/9/2018	21 days
4	Cost Analysis	5/9/2018	12/9/2018	7 days
5	Documentation of Synopsis	14/9/2018	28/9/2018	14 days
6	Designing	29/9/2018	13/10/2018	21 days
7	Developing Scrapper	14/10/2018	24/10/2018	10 days
8	Writing snippet extraction module	25/10/2018	9/11/2018	15 days
9	Generating Corpus	10/11/2018	25/11/2018	15 days
10	Processing Data	26/11/2018	21/12/2018	25 days
11	Training classifier	22/12/2018	31/12/2018	9 days
12	Testing and Deployment	1/1/2019	15/1/2019	15 days



HARDWARE AND SOFTWARE REQUIREMENTS

REFERENCES

1. Popat, Kashyap and Mukherjee, Subhabrata, CredEye: A Credibility Lens for Analyzing and Explaining Misinformation in *Companion Proceedings of the The Web Conference 2018.*

2. Popat, Kashyap and Mukherjee, Subhabrata, Where the Truth Lies: Explaining the Credibility of Emerging Claims on the Web and Social Media in *Proceedings of the 26th International Conference on World Wide Web Companion.*

3. Mukherjee, Subhabrata and Weikum, Gerhard, Leveraging Joint Interactions for Credibility Analysis in News Communities in *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*

4. Popat, Kashyap and Mukherjee, Subhabrata, Credibility Assessment of Textual Claims on the Web in *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*

5. Marta Recasens and Cristian Danescu-Niculescu-Mizil and Daniel Jurafsky, Linguistic Models for Analyzing and Detecting Biased Language in *ACL*

6. Naeemul Hassan and Fatma Arslan and Chengkai Li and Mark Tremayne, Toward Automated Fact-Checking: Detecting Check-worthy Factual Claims by ClaimBuster in *KDD*