A Project Report

on

Social Media Post Checker

Submitted in partial fulfilment of the
requirement for the award of the degree of

Bachelors of Technology in Computer Science and Engineering

Under The Supervision of
Dr. Ganga Sharma

Submitted By

NISHANT SINHA	19SCSE1010019
MAYANK SHEKHAR	19SCSE1260004

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY, GREATER NOIDA
INDIA
APRIL,2023

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING GALGOTIAS UNIVERSITY GREATER NOIDA

CANDIDATE'S DECLARATION

We hereby certify that the work which is being presented in the project, entitled **"Social Media Post Checker"** in partial fulfilment of the requirements for the award of the Bachelors of Technology in Computer Science and Engineering submitted in the School of Computing Science and Engineering of Galgotias University, Greater Noida, is an original work carried out during the period of March,2023 to May,2023 under the supervision of Dr. Ganga Sharma, of School of Computing Science and Engineering, Galgotias University, Greater Noida

The matter presented in the project has not been submitted by us for the award of any other degree of this or any other places.

Nishant Sinha 19SCSE1010019

Mayank Shekhar 19SCSE1260004

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Dr. Ganga Sharma

CERTIFICATE

The Final Project Viva-Vice examination of Mayank Shekhar 19SCSE1260004 and Nishant Sinha 19SCSE1010019 has been held on_____and his/her work is recommended for the award of -Bachelors of Technology in Computer Science and Engineering.

Signature of Examiner(s)   **Signature of Supervisor(s)**

Signature of Project Coordinator   **Signature of Dean**

Date: April,2023

Place: Greater Noida

ACKNOWLEDGEMENT

It gives immense pleasure in bringing out this synopsis of the project entitled

"Social Media Post Checker"

Firstly we would like to thank our guide **Dr. Ganga Sharma** who gave us her valuable suggestions and ideas when we were in need of them. she encouraged us to work on this project.

We are also grateful to our college for giving us the opportunity to work with them and providing us the necessary resources for the project

We would also thank to all of them who helped us to complete this project.

We are immensely grateful to all involved in this project as without their inspiration and valuable suggestion it would not have been possible to develop the project within the prescribed time.

With sincere thanks,

NISHANT SINHA 19SCSE1010019

MAYANK SHEKHAR 19SCSE1260004

ABSTRACT

As social media is not just a platform of communication but also holds a mass collection data which could be any image, article, news, memes etc... which is shared across the world. The problem exists here is that, if something is shared across the social media it, the users cannot find whether the message being conveyed through the post is authentic or not, most of the time it is found that the news spread across the social media is a fake or manipulated as recently seen after US 2020 elections, which lead riot at the capitol. Often fake news and manipulated media leads to several losses to an individual or a group.

In this project we will build a Social Media Post Checker App using different languages & Designing tools like Flutter, Dart, GenY motion emulator, Gimp which will help normal end user of social media websites to verify & authenticate the information of the post. This application will help in reducing spread of fake news and information which will be beneficial to the society.

Tools used for project include IDEs such as Android Studio and Microsoft's VS Code which are renowned in the programming world as one of the best IDE's for coding. Technology used for the deployment of the project includes Flutter which is Google's UI toolkit for building beautiful, natively compiled applications. Dart is the logical language used in flutter as programming language, Gimp for logo and icon design. Firebase for Machine Learning API for scanning image text in the social media post on the internet.

The resultant Social Media Post Checker application which is developed will be useful in finding out the source of information represented in the social media post on social media websites. Its User Interface is quite easy to use that even a person who is not fluent with technology can use it.

In the we would like to say our will help the end user and the society to verify the sources of the posts that are present on the internet and various social media websites like Facebook, Instagram, and Twitter. It will help the society to move forward in a positive way and will also help in reducing rumors and panic around in the society caused by such fake and morphed social media posts and will also the law enforcement and well as the social media company to catch or suspends those accounts who are spreading rumors on a regular basis.

Table of Contents

List of Figures

CHAPTER-1

INTRODUCTION

We all are living in the world of social media where we keep ourselves updated with several posts we share and see what others are sharing on it. As social media is not just a platform of communication but also holds a mass collection data which could be any image, article, news, memes etc... which is shared across the world.

The problem exists here is that, if something is shared across the social media it, the users cannot find whether the message being conveyed through the post is authentic or not in simple words if the post holds legitimate information or fake new, most of the time it is found that the news spread across the social media is a fake or manipulated as recently seen after US 2020 elections, which lead riot at the capitol. Often fake news and manipulated media leads to several losses to an individual or a group.

"India has more fake news than any other country in the world"

Business Insider India

Nearly 60% of Indians had seen fake news against the global average of 57%, said a recent survey.

Overall, India ranked seventh in Microsoft's 'Digital Civility' Index.

It found that the spread of online risks by family and friends in India have been increasing sharply by nine percentage points to 29%.

1.2 FORMULATION OF PROBLEM

Social Media is a double-edged sword. On the one hand, its low cost, easy access, and rapid dissemination of information lead people to seek out and consume information and sometimes even news and current affairs. It also enables the wide spread of "fake news", i.e., low quality news with intentionally false information or rumors without any background information attached to it. The extensive spread of fake news & rumors has the potential for extremely negative impacts on individuals and society. Therefore, fake news detection has recently become emerging research that is attracting tremendous attention. First, fake news is intentionally written to mislead readers to believe false information, which makes it difficult and nontrivial to detect based on news content.

So, to tackle such situations we users need to keep a check on any news or a rumor which is spread across the social media. So, we came up with the idea to develop a SOCIAL MEDIA POST CHECKER using machine learning algorithms. The algorithm must be able to detect fake information in a given scenario such as a social media post.

Users will be able to capture the social media post which they feel like is misleading or fake news and search the internet to verify the authenticity of the post's information.

1.3 MACHINE LEARNING:

Machine learning (ML) is the scientific study of algorithms and statistical models that computer systems use to perform a specific task without using explicit instructions, relying on patterns and inference instead. It is seen as a subset of artificial intelligence. Machine learning algorithms build a mathematical model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to perform the task. Machine learning is closely related to computational statistics, which focuses on making predictions using computers. The study of mathematical optimization delivers methods, theory, and application domains to the field of machine learning. "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E." This is Alan Turing's definition of machine learning. Deep learning is a class of machine learning algorithms that utilizes a hierarchical level of artificial neural networks to carry out

the process of machine learning. The artificial neural networks are built like the human brain, with neuron nodes connected like a web. While traditional programs build analysis with data in a linear way, the hierarchical function of deep learning systems enables machines to process data with a non-linear approach.

The word "deep" in "deep learning" refers to the number of layers through which the data is transformed. More precisely, deep learning systems have a substantial credit assignment path (CAP) depth. The CAP is the chain of transformations from input to output. CAPs describe potentially causal connections between input and output.

For a feed forward neural network, the depth of the CAPs is that of the network and is the number of hidden layers plus one (as the output layer is also parametrized). For recurrent neural networks, in which a signal may propagate through a layer more than once, the CAP depth is potentially unlimited.

Deep learning architectures such as deep neural networks, deep belief networks, recurrent neural networks and convolutional neural networks have been applied to fields including computer vision, speech recognition, natural language processing, audio recognition, social network filtering, machine translation, bioinformatics, drug design, medical image analysis, material inspection and board game programs, where they have produced results comparable to and in some cases superior to human experts.
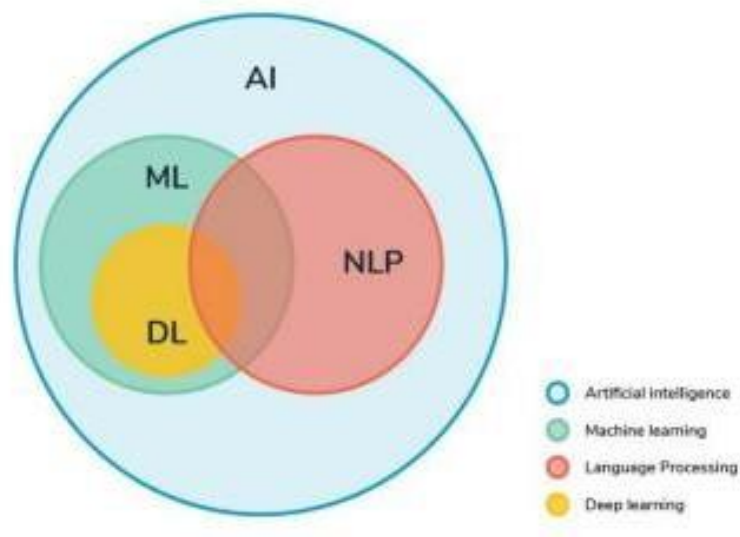
Fig. 1 : Graphical representation of relationship between various fields in artificial intelligence (source: devopedia.org)

1.3.1 NATURAL LANGUAGE PROCESSING

NLP is an area of computer science and artificial intelligence concerned with the interactions between computers and human (natural) languages, how to program computers to fruitfully process large amounts of natural language data. Natural language processing (NLP) is a subfield of linguistics, computer science, information engineering, and artificial intelligence concerned with the interactions between computers and human (natural) languages, how to program computers to process and analyse large amounts of natural language data.

1.3.1.1 STAGES IN NLP

1.3.1.1.1 LEXICAL ANALYSIS

Lexical Analysis involves identifying and the structure of words. Lexicon of a language means the collection of words and phrases in a language. Lexical analysis is dividing the whole chunk of txt into paragraphs, sentences, and words.

1.3.1.1.2 SYNTACTIC ANALYSIS (PARSING)

Syntactic Analysis involves analysis of words in the sentence for grammar and arranging words in a manner that shows the relationship among the words. The sentence such as "The school goes to boy" is rejected by English syntactic analyser.

1.3.1.1.3 SEMANTIC ANALYSIS

Semantic Analysis draws the exact meaning or the dictionary meaning from the text. The text is checked for meaningfulness. It is done by mapping syntactic structures and objects in the task domain. The semantic analyser disregards sentence such as "hot ice-cream"

1.3.1.1.4 DISCOURSE INTEGRATION

The meaning of any sentence depends upon the meaning of the sentence just before it. In addition, it also brings about the meaning of immediately succeeding sentence. So, in Discourse Integration gives the meaning based on all the sentences given before i.e... Consider the sentence "Water is flowing on the bank of the river" But bank has two meanings One Financial Institute and Two River of the bank here System must consider the second meaning.

1.3.1.1.5 PRAGMATIC ANALYSIS

During this, what was said is re-interpreted on what it meant. It involves deriving those aspects of language which require real world knowledge.

1.3.2 MOTIVATION OF WORK

The rise of fake news before and during the 2020 U.S. Presidential Election highlighted not only the dangers of the effects of fake news but also the challenges presented when attempting to separate fake news from real news. Fake news may be a relatively new term but it is not necessarily a new phenomenon. Fake news has technically been around at least since the appearance and popularity of one-sided, partisan newspapers in the 19th century. However, advances in technology and the spread of news through different types of media have increased the spread of fake 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. I have identified the three most prevalent motivations for writing fake news and chosen only one as the target for this project to narrow the search in a meaningful way. The first motivation for writing fake news, which dates to the 19th century one-sided party newspapers, is to influence public opinion. The second, which requires more recent advances in technology, is the use of fake headlines as clickbait to raise money.

As such, this paper will focus primarily on fake news as defined by politifact.com, "fabricated content that intentionally masquerades as news coverage of actual events." This definition excludes satire, which is intended to be humorous 8 and not deceptive to readers. Most satirical articles come from sources. Satire can already be classified, by machine learning techniques Therefore, our goal is to move beyond these achievements and use machine learning to classify, at least as well as humans, more difficult discrepancies between real and fake news.

The dangerous effects of fake news, as previously defined, are made clear by events in which a man attacked a pizzeria due to a widespread fake news article. This story along with analysis provide evidence that humans are not very good at detecting fake news, possibly not better than chance. As such, the question remains whether machines can do a better job.

1.4 TOOLS AND TECHNOLOGY USED:

1.4.1 FLUTTER

Flutter is a cross-platform UI toolkit that is designed to allow code reuse across operating systems such as iOS and Android, while also allowing applications to interface directly with underlying platform services. The goal is to enable developers to deliver high-performance apps that feel natural on different platforms, embracing differences where they exist while sharing as much code as possible.

During development, Flutter apps run in a VM that offers stateful hot reload of changes without needing a full recompile. For release, Flutter apps are compiled directly to machine code, whether Intel x64 or ARM instructions, or to JavaScript if targeting the web. The framework is open source, with a permissive BSD license, and has a thriving ecosystem of third-party packages that supplement the core library functionality.

1.4.2 DART

Dart is a programming language designed for client development, such as for the web and mobile apps. It is developed by Google and can also be used to build server and desktop applications. Dart is an object-oriented, class-based, garbage-collected language with C-style syntax. Dart can compile to either native code or JavaScript. It supports interfaces, mixing, abstract classes, reified generics, and type inference.

Fig.2: Logo of Dart programming language

The Dart SDK contains two Dart-to-JavaScript compilers. During development, darted supports quick refresh cycles. For the final version of an app, dart2js produces deployable JavaScript.

The first compiler to generate JavaScript from Dart code was Dart, but it was deprecated. The second Dart-to-JavaScript compiler was Frog. It was written in Dart, but never implemented the full semantics of the language. The third Dart-to-JavaScript compiler was dart2js. An evolution of earlier compilers, dart2js is written in Dart and intended to implement the full Dart language specification and semantics.

1.4.3 GIMP

GIMP is a cross-platform image editor available for GNU/Linux, macOS, Windows and more operating systems. It is free software; you can change its source code and distribute your changes. Whether you are a graphic designer, photographer, illustrator, or scientist, GIMP provides you with sophisticated tools to get your job done. You can further enhance your productivity with GIMP thanks to many customization options and 3rd party plug-ins.

1.4.4 FIREBASE ML VISION PLUG-IN FOR FLUTTER

A Flutter plug-in to use the capabilities of Firebase ML, which includes all of Firebase's cloud-based ML features, and ML Kit, a stand-alone library for on-device ML, which can be used with or without Firebase.
To use this plug-in, add firebase_ml_vision as a dependency in your pubspec.yaml file. You must also configure Firebase for each platform project: Android and iOS.

1.4.5 VISUAL STUDIO CODE

Visual Studio Code is a lightweight but powerful source code editor which runs on your desktop and is available for Windows, macOS and Linux. It comes with built-in support for JavaScript, Type Script and Node.js and has a rich ecosystem of extensions for other languages (such as C++, C#, Java, Python, PHP, Go) and runtimes (such as .NET and Unity).

CHAPTER-2

LITERATURE SURVEY

In the world of rapidly increasing technology, information sharing has become an easy task. There is no doubt that internet has made our lives easier and access to lots of information. This is an evolution in human history, but at the same time it unfocused the line between true media and maliciously forged media. Today anyone can publish content – credible or not – that can be consumed by the world wide web. Sadly, fake news and rumors accumulates a great deal of attention over the internet, especially on social media. People get deceived and do not think twice before circulating such mis-informative pieces to the world. This kind of news vanishes but not without doing the harm it intended to cause. The social media sites like Facebook, Twitter, WhatsApp play a major role in supplying these false news and rumors. Many scientists believe that counterfeited news issue may be addressed by means of machine learning and artificial intelligence.

Various models are used to provide an accuracy range of 60-75%. Which comprises of Naive Bayes classifier, Linguistic features based, Bounded decision tree model, SVM etc. The parameters that are taken in consideration do not yield high accuracy. The motive of this project is to increase the accuracy of detecting fake news more than the present results that are available. By fabricating this new model which will judge the counterfeit news articles based on certain criteria like spelling mistake, jumbled sentences, punctuation errors, words used.

So, to tackle such situations we users need to keep a check on any news or a rumour which is spread across the social media. So, we came up with the idea that we can build an application which will help users to check whether the news or information in the social media post is authentic and genuine or not.

Users will be able to capture the social media post which they feel like is misleading or fake news and search the internet to verify the authenticity of the post's information.

Among the many new sources of consumer information that have emerged in the last decade, social media data are among the most potent Anda game-changing for effective marketing research. Social media platforms offer a powerful opportunity to gain immediate access to the unfettered opinions of

consumers. Many companies are aware of the value of using social media data to gain marketing insights. But there is so much information out there. How can businesses tap this source to obtain deep, actionable insights?

A few excellent programs and services – some free and some commercial – have been developed for the analysis of social media data. Yet, the focus of most of these tools is to provide summary statistics of the data. Web analytics – for example, word counts, reach, word clouds, volume, sentiment analysis – can provide valuable, up-to-the-minute snapshots of Web content. Still, no algorithm is an adequate replacement for the in-depth analysis of consumer-generated feedback that can be conducted by a skilled analyst with a deep understanding of a brand and its challenges and opportunities.

Following are the steps for analysing Social media data:-

Step 1: Develop a problem definition and research objectives

For most research, developing focused research objectives is usually the most important step. What decisions will be made with this information? This guideline holds particularly true for social media analysis where a clear direction is needed to make sense of the copious amount of data. Limiting the focus to a defined topic and specific objectives will make the analysis more manageable.

Still, to take full advantage of social media data analysis, the research objectives should also allow for an element of discovery. The data may lead to unexpected places.

The following are examples of objectives that social media analysis is particularly suited to address: competitive analysis; product extensions; product strengths and weaknesses; new uses of products; and reactions to advertising and promotions.

Step 2: Identify key search terms

The identification of the proper key search terms is a crucial step to the successful analysis of social media data. The process is often an iterative one, with broader searches being followed by searches using combinations of terms or newly discovered synonyms or tangential phrases. Obvious terms to start a search include the product's brand name, competitors' brand names and the product class. More exploratory analyses might investigate activities, events and emotions related to a brand.

Step 3: Identify social media data sources

The identification of the most useful data sources is another important step to social media data analysis. Online aggregator tools, such as Tweet Deck and Scout Labs, can aid in this process. Still, sometimes these tools can miss some important types of social media platforms.

Depending on the research objectives, some types of social media sites that can provide consumer-generated data include the following:

- social network sites (e.g., Facebook),
- video-sharing sites (e.g., YouTube),
- photo-sharing sites (e.g., Flickr),
- product and service review sites (e.g., Yelp),
- Web-based communities (e.g., Chowhound),
- blogs (e.g., Gardenia), and
- microblogs (e.g., Twitter).

Finding the most current and germane sites is a moving target, since social media-oriented data sources ebb and flow in popularity. While this makes the task of identifying the best sites from which to gather data more difficult, it also means that new forms of exciting and relevant consumer-generated feedback are always emerging and can be uncovered with a bit of persistence.

Step 4: Organize data

Some of the most important consumer-generated data will not necessarily be in the form of text. Photos, videos, artwork, literature, and other forms of data might provide new insights into product feedback. As a result, organization of the data should be flexible and allow for diverse forms of media. Several commercial services (e.g., Hootsuite, Radian6) and software (e.g., NVivo) are available to assist in this process, as well as free online tools (e.g., Social Mention, Google Alerts). However, some analysts will prefer to replace or supplant these options with more of a do-it-yourself approach to organizing data to ensure versatility and comprehensiveness.

Analysts will also need to decide whether to view the data online, via hard copy or through a combination of paper and electronic sources when conducting the analysis, based on personal preferences and on to what extent the data analysis will involve collaboration among team members.

With the abundance of data available on the Web, and with all the twists and turns that can be encountered in the process of organizing data, it is important to know when to stop seeking new sources. The rule of thumb is that when a saturation point is reached – that is, when little new information is being acquired relative to the effort – it is time to end the searches.

Step 5: Analyse data

Once the social media data have been gathered and organized, the best practices for analysing social media data are the same as those used for traditional qualitative data. First the analysts should review the data thoroughly. As with all research, insightful analysis depends on a comprehensive knowledge and understanding of the data. Then the analysts should begin identifying key themes that emerge from the findings – beliefs, ideas, concepts, definitions, behaviours. The data should be coded according to themes, either by hand or via software

(e.g., NVivo) and then compared and integrated. To repeat: This step parallels content analysis of traditional types of qualitative data.

Step 6: Present findings

Following analysis of the data, the findings will be presented via oral and written presentation, using concrete examples and illustrations. Here is where social media data really stands out. Quotes can be presented from Twitter, reviews and blogs, just as verbatim quotes would be used to illustrate findings from focus groups and interviews. But consumer-generated social media data offer much more. Photos found online can illustrate exactly where, when, and how a consumer is using a product or service. Consumer-produced videos can demonstrate perceived advantages and disadvantages of products. Even textual quotes praising or criticizing products can be much more colourful when found online with the opinions offered spontaneously and not prompted by a moderator.

Step 7: Outline limitations

When using social media data, it is at least as, and probably even more, important than with other research methods to outline the limitations of the data. Explicitly stating the problems and gaps

encountered when gathering and analysing the data helps to provide a more complete understanding of the findings.

The following are some of the limitations that are most encountered with social media data:

- The online consumers are not necessarily demographically representative of the product's target consumers.
- Self-selection bias is inherent with social media data.
- Advocates and detractors can distort online conservations.
- The demographic and geographic information of the consumers is often not traceable.

Step 8: Strategies

As with all research, the final and most important step of the analysis is to use the finding to develop research-based, actionable recommendations related to the research objectives. Then, based on the project's results, the next stage of research should be planned.

Challenges and opportunities: -

Many of the basic steps used for the content analysis of text from structured data collection methods – such as interviews, focus groups, diaries, and managed online communities – can be generalized to social media data. However, social media data is different in several fundamental ways, representing both challenges and opportunities for analyses. It is useful to consider these differences.

- ➢ **Overwhelming amount of data.** Traditional interviews or focus groups offer a discrete amount of material to organize and present. Social media data, on the other hand, is available in abundance. Often much more social media data related to a topic exists than can be reasonably analysed. Analysts must place limits, by topics or time periods, on their search efforts.

- ➢ **Unrestricted comments.** With focus groups, interviews and even online communities, participants are responding to directed questions. The users of social media state whatever is on their minds. This represents a great opportunity to gain new understandings about

consumers' motives, needs, behaviours and emotions. It also means that the problem definitions and research objectives that researchers identify prior to analysis may miss the mark and require revision.

➢ **Much more noise.** Because social media data is not generally managed, many, if not most, of the comments that analysts sort through will be useless. For every insightful comment found, there are likely to be numerous useless posts, such as sales pitches ("My friend made $1,200 at home last month…"),

  empty comments ("So true." "What he said." "Yes."), and non-contextualized obscenities (no examples necessary).

➢ **Multiple languages.** Because social media is on the World Wide Web, relevant comments are frequently posted in multiple languages. As a result, depending on the objectives of the research, it may be beneficial to assemble a multilingual team for targeted projects.

➢ **Multiple forms.** Consumer-generated data found online can take many different forms. In addition to text, data might appear as videos, audios, photos, artwork, slideshows and other structures.

➢ **Lacks context.** Traditional qualitative methods allow quotes to be identified with specific individuals, providing key information such as gender, age, location, and income. It is much more difficult to ascribe demographics to social media quotes. Even if the information can be traced to a user profile, there are no assurances that the profile is factual.

CHAPTER-3

METHODOLOGY

3.1 PROPOSED SYSTEM

The proposed system when subjected to a scenario of a set of news articles, the new articles are categorized as true or fake by the existing data available. This prediction is done by using the relationship between the words used in the article with one another. The proposed system contains a Word2Vec model for finding the relationship between the words and with the obtained information of the existing relations, the new articles are categorized into fake and real news.
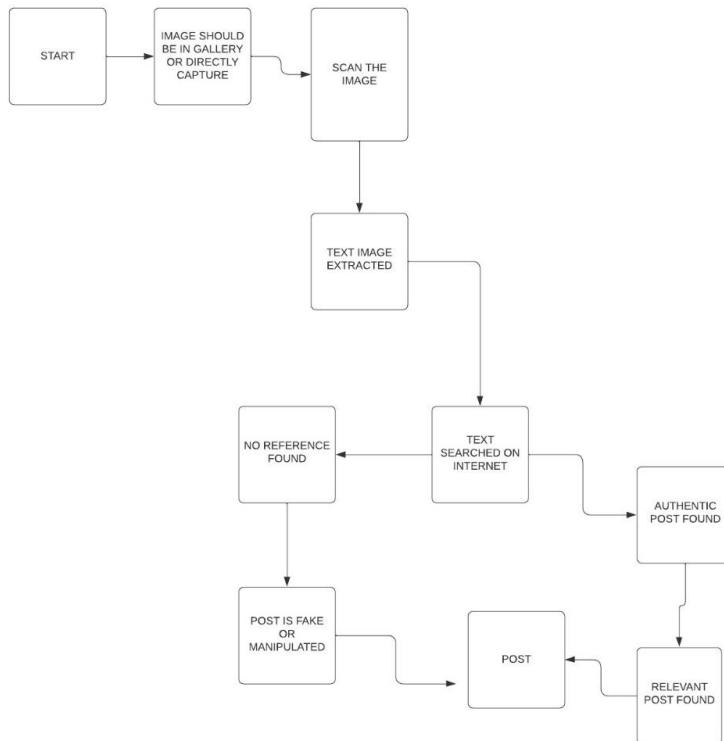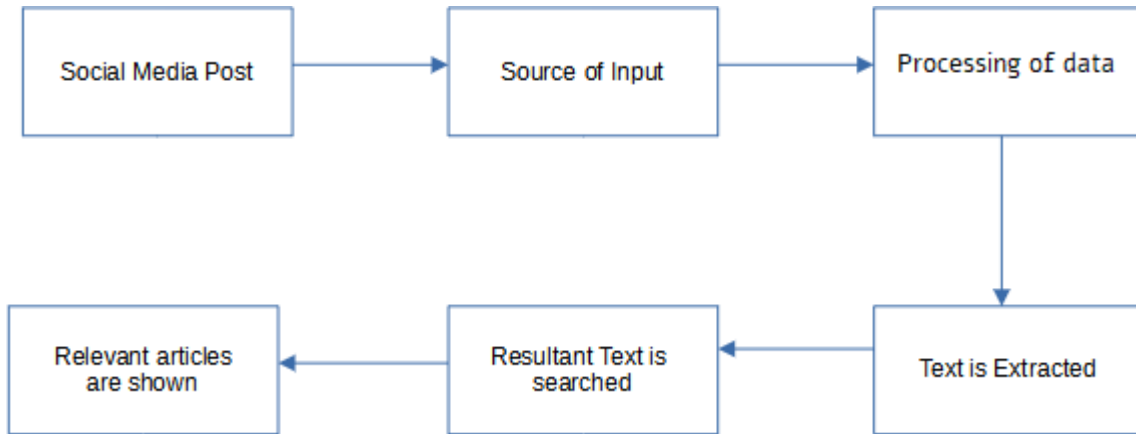
3.2 SYSTEM ARCHITECTURE

Fig no. 3 System Architecture

3.2 FLOW CHART

3.3 ALGORITHM FOR THE PROPOSED SYSTEM

Step 1: Start

Step 2: Input is collected from various sources like Facebook and Instagram.

Step 3: Processing of data is done.

Step 4: Resultant text is searched on the internet to check if there exists an article relevant to the post's information.

Step 5: Related articles to the articles are shown for user to read.

CHAPTER-4

CONCEPTS

4.1 PREPROCESSING:

In any Machine Learning process, Data Preprocessing is that step in which the data gets transformed, or encoded, to bring it to such a state that now the machine can easily parse it. In other words, the features of the data can now be easily interpreted by the algorithm.

In this fake news detection, preprocessing is the major thing that should be done. Firstly, as the data datasets is collected from various sources unnecessary information should be removed, converted to lower case, remove punctuation, symbols, stop words.

4.2 STEPS IN TEXT PRE-PROCESSING:

4.2.1 TEXT NORMALIZATION:

Text normalization is a process of transforming text into a single canonical form. Normalizing text before storing or processing it allows for separation of required data from the rest so that the system can send consistent data as an input to the other steps of the algorithm.

4.2.2 STOP WORD REMOVAL

A Stop Word is a commonly used word in any natural language such as "a, an, the, for, is, was, which, are, were, from, do, with, and, so, very, that, this, no, yourselves etc. ". These Stop Words will have a very high frequency and so these should be eliminated while calculating the term frequency so that the other important things are given priority. Stop word removal is such a Pre-processing step which removes these stop words and thereby helping in the further steps and reducing some processing time because the size of the document decreases tremendously.

4.3 GOOGLE'S MACHINE LEARNING KIT:

5.3.1 TEXT RECOGNITION

The ML Kit Text Recognition API can recognize text in any Latin-based character set. It can also be used to automate data-entry tasks such as processing credit cards, receipts, and business cards.

Key capabilities

- **Recognize text across Latin-based languages** Supports recognizing text using Latin script
- **Analyze structure of text** Supports detection of words/elements, lines, and paragraphs
- **Identify language of text** Identifies the language of the recognized text

- **Small application footprint** On Android, the API is offered as an un-bundled library through Google Play Services
- **Real-time recognition** Can recognize text in real-time on a wide range of devices.

5.3.2 TEXT STRUCTURE

The Text Recognizer segments text into blocks, lines, and elements. Roughly speaking:

- a Block is a contiguous set of text lines, such as a paragraph or column,
- a Line is a contiguous set of words on the same axis, and
- an Element is a contiguous set of alphanumeric characters ("word") on the same axis in most Latin languages, or a character in others

The image below highlights examples of each of these in descending order. The first highlighted block, in cyan, is a Block of text. The second set of highlighted blocks, in blue, are Lines of text. Finally, the third set of highlighted blocks, in dark blue, are Words.
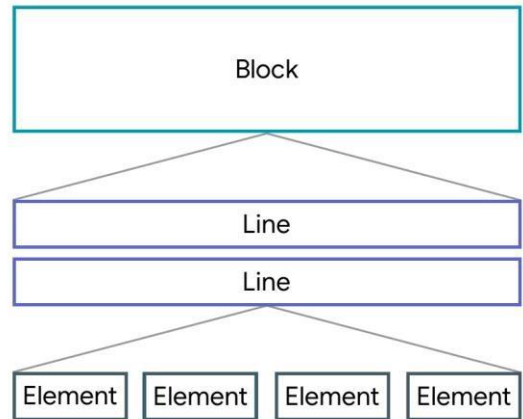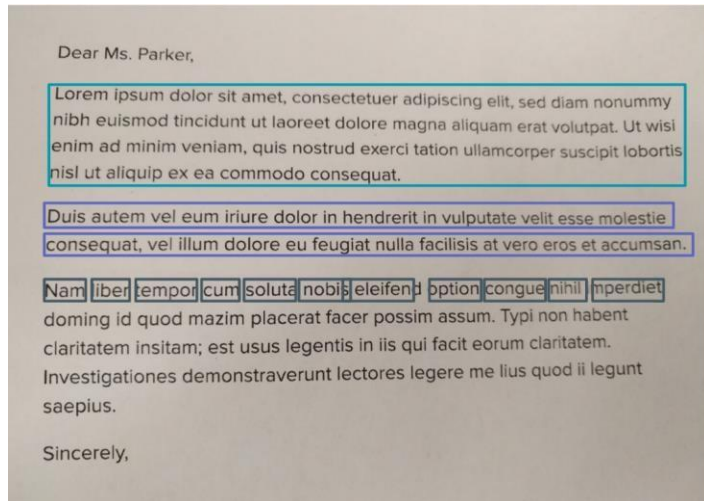
Fig No. 4 Example of Text Structure

For all detected blocks, lines and elements, the API returns the bounding boxes, corner points, recognized languages and recognized text.

Example Results

Example Results

Recognized Text	
Text	Wege der parlamentarischen Demokratie
Blocks	(1 block)

Block 0	
Text	Wege der parlamentarischen Demokratie
Frame	(117.0, 258.0, 190.0, 83.0)
Corner Points	(117, 270), (301.64, 258.49), (306.05, 329.36), (121.41, 340.86)
Recognized Language Code	de
Lines	(3 lines)

Line 0	
Text	Wege der
Frame	(167.0, 261.0, 91.0, 28.0)
Corner Points	(167, 267), (255.82, 261.46), (257.19, 283.42), (168.36, 288.95)
Recognized Language Code	de
Elements	(2 elements)

Element 0	
Text	Wege
Frame	(167.0, 263.0, 59.0, 26.0)
Corner Points	(167, 267), (223.88, 263.45), (225.25, 285.41), (168.36, 288.95)

CHAPTER-5

PROJECT DESIGN

4.1 System Requirements Specifications

System requirements are expressed in a software requirement document. The Software requirement specification (SRS) is the official statement of what is required of the system developers. This requirement document includes the requirements definition and the requirement specification. The software requirement document is not a design document. It should set out what the system should do without specifying how it should be done. The requirement set out in this document is complete and consistent. The software specification document satisfies the following: -

➢ It specifies the external system behaviours.

➢ It specifies constraints on the implementation.

➢ It is easy to change.

➢ It serves as reference tool for system maintainers.

➢ It record forethought about the life cycle of the system.

➢ It characterizes acceptable response to undesired events.

Functional Requirements:

The System must provide following functionalities—

➢ Function to import images from the user's device.

➢ Function to crop the image.

➢ Function to adjust image vertices.

➢ Function to scan text from image.

➢ Function to search that text on the internet.

➢ Function to copy that text to clipboard.

External Interface Requirements:

User Interface:

User of the system will be provided with the simple to use GUI.

Hardware Interface:

Hardware requirements for running this project are as follows:

Device: - Android Device with Android version 8.0 at least.

Ram: - 250MB

Space: - at least 100MB of storage

Connectivity: Internet Connectivity.

Software Interface:

Software required to make working of product is:

- Flutter
- Dart
- Flutter ML Vision API
- GIMP
- GenY motion Emulator
- VS Code
- Firebase account with Spark Plan (at least)

4.2 Conceptual Model

Activity Diagram

Activity diagrams fall under the category of behavioral diagrams in Unified Modeling Language. It is a high-level diagram used to visually represent the flow of control in a system. It has similarities with traditional flow charts. However, it is more powerful than a simple flow chart since it can represent various other concepts like concurrent activities, their joining, and so on.

Most people are familiar with a drawing package can design and create forms and most people with basic programming skills can double click on a control and enter some code. But does this approach lend itself to professional quality applications?

Activity diagrams, however, cannot depict the message passing among related objects. As such, it cannot be directly translated into code. These kinds of diagrams are suitable for confirming the logic to be implemented with the business users. These diagrams are typically used when the business logic is complex. In simple scenarios it can be avoided entirely.
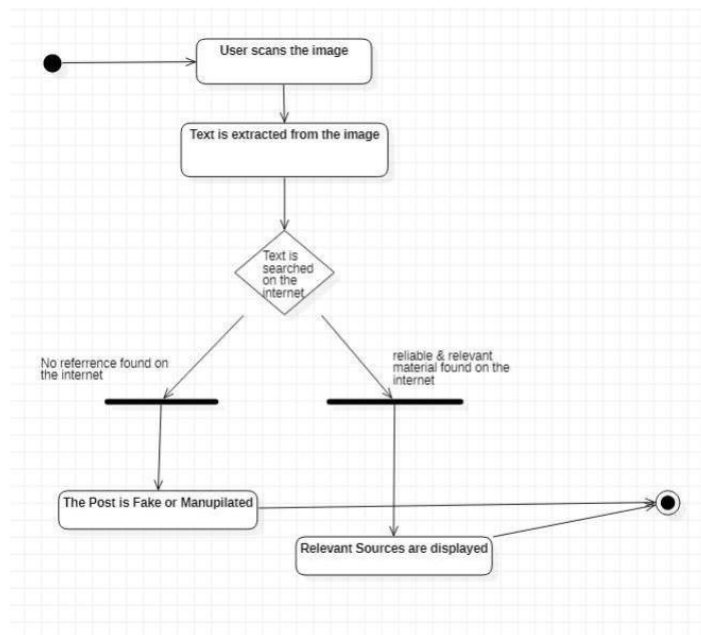
Fig No. 6 Activity Diagram

Use-case Diagram

Use-case diagram is a platform that can provide a common understanding for the end users, developers, and the domain experts. It is used to capture the basic functionality i.e., use cases, and the users of those available functionality, i.e., actors, from a given problem statement. In this experiment, we will learn how use cases and actors can be captured and how different use cases are related in a system.
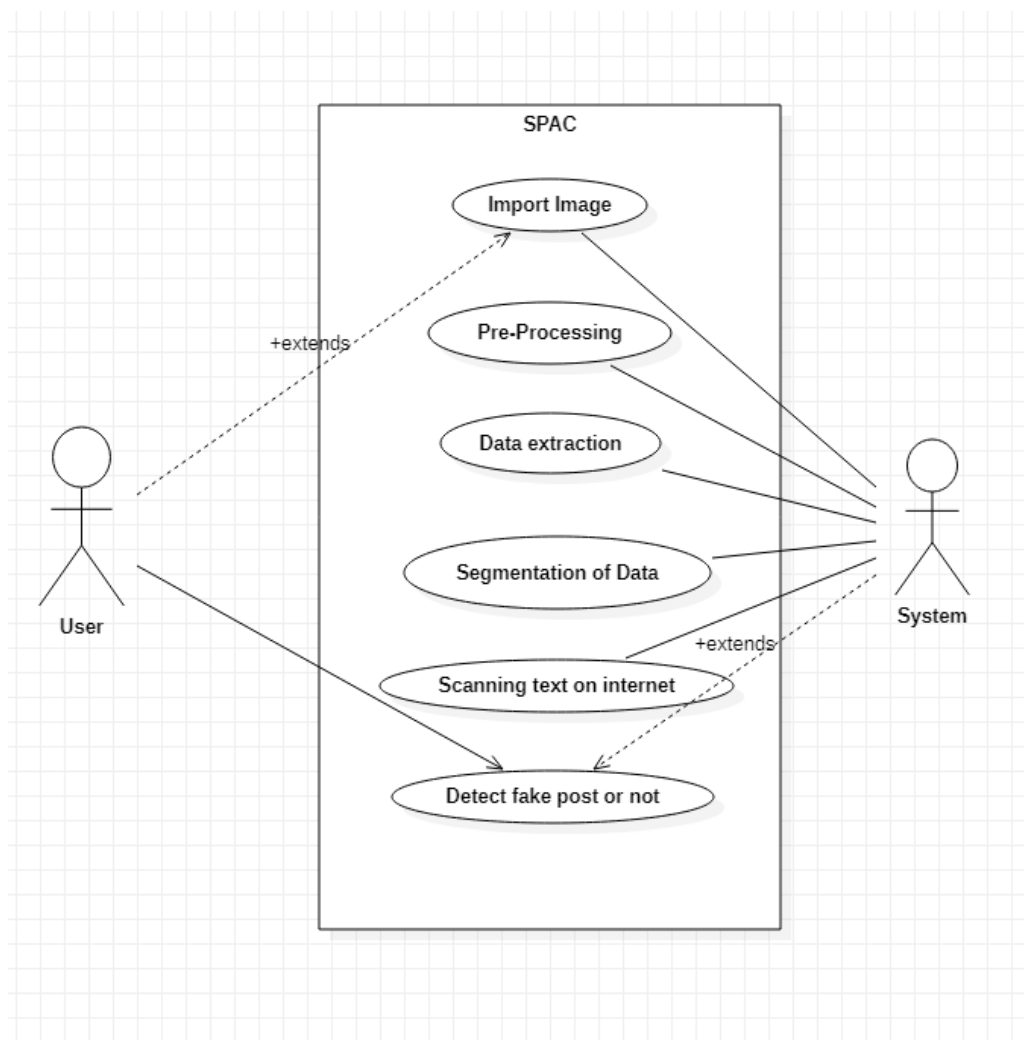
Fig No. 7 Use-Case Diagram

31

CHAPTER-6

WORKING OF PROJECT

System Designing

INTRODUCTION:

System design is the solution of a "how to approach to the creation of the new system. It is composed of several steps. It facilitates the understanding and provides the procedural details necessary for implementation of the system recommended in the feasibility study. Emphasis is given on translating the performance requirements into design specification. Design goes through logical and physical stages of development.

Logical design reviews the present physical system; prepares input and output specification; make editing; security and control specification; details the implementation plan, and prepare logical design walk through. The physical design maps out the details of the physical system; plans the system implementation plan and specifies hardware and software. System design translates the system requirement into the ways of the system as recommended in the feasibility study. Thus, the system design is the translation from user-oriented document to a programmer or a database personal oriented document. System design is a highly creative process that can be greatly facilitated by the following: -

- Strong Problem Definition

- Pictorial description of the Existing System

- Set of Requirements of the new system

Modules Description:

1. Gallery: Users can import images from gallery
2. Camera: Users can click images from the camera itself.
3. Clipboard: Users can copy text to clipboard from the post's image.
4. Search: Users can search for the information on the internet.

INPUT DESIGN

Very careful attention had to be given to input design, which is a major part of the overall system design. In order to make the data entry as easy, logical and error free as possible, specific standards had been followed. Validation checks, provided in the system prevented the user in entering incorrect, erroneous data. This made sure that, only valid data had been available for data processing. If valid data was entered, then meaningful error messages had been prompted to enter correct data. The interactive screen formats facilitate the entry of valid data.

INPUT DESIGN OBJECTIVES:

The numbers of clear objectives of input design are,

- To produce a cost-effective method of input

- To achieve the highest possible level of accuracy

- To ensure that the input is acceptable to and understand by the user staff.

OUTPUT DESIGN:

Output, as you probably know, generally refers to the results and information that are generated by the system. For many end-users, output is the main reason for developing the system and the basis on which they will evaluate the usefulness of the application. Most end users will not actually operate the information system or enter data through workstations, but they will use the output from the system.

When designing output, systems analysts must accomplish the following.

 Determine what information to present

 Decide whether to display, print, or "speak" the information and select

   the output medium.

 Arrange the presentation of information in an acceptable format.

 Decide how to distribute the output to intended recipients.

That alignment of information on a display or printed document is termed as

layout.

Accomplishing the general activities listed above will require specific decisions, such as whether to use preprinted forms when preparing reports and documents, how many lines to plan on a printed page, or whether to use graphics and colour.

The output design is specified on layout performs, sheets that describe the location characteristics, and format of the column headings and pagination. As we indicated at the beginning of this discussion, these elements are analogous to an architect's blue print that shows the location of each component.
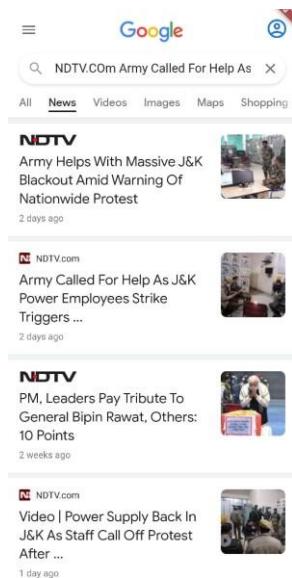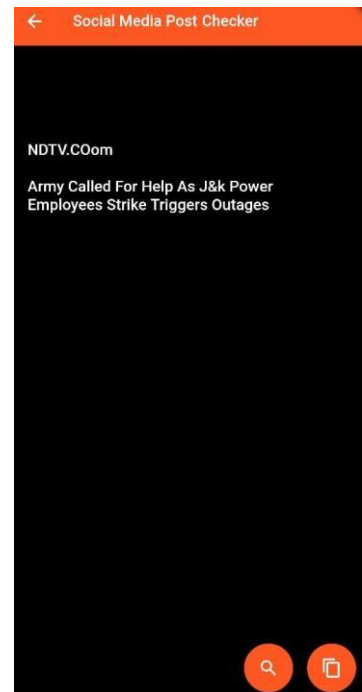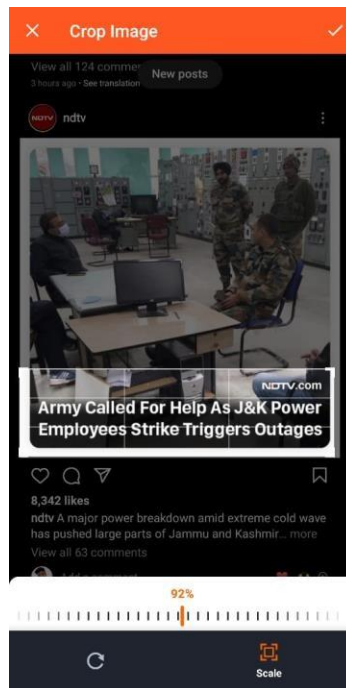
CHAPTER-7

RESULTS AND DISCUSSION

Screenshot of Application
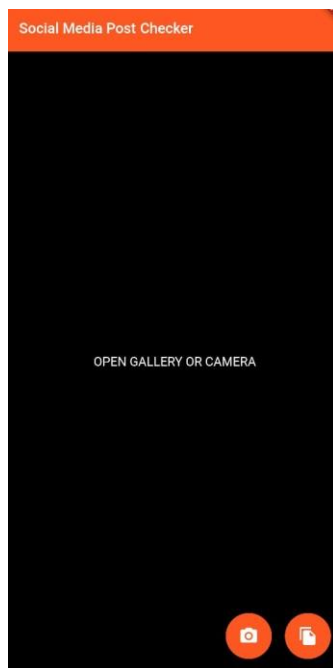
Opening page   **User crops the image from gallery Text is extracted**

Results are shown

CHAPTER-8

CONCLUSION

In the we would like to say our will help the end user and the society to verify the sources of the posts that are present on the internet and various social media websites like Facebook, Instagram, and Twitter. It will help the society to move forward in a positive way and will also help in reducing rumors and panic around in the society caused by such fake and morphed social media posts and will also the law enforcement and well as the social media company to catch or suspends those accounts who are spreading rumors on a regular basis. To implement & develop this we used Flutter as the main framework and Dart as the main programming language.

Google's Machine Learning Kit which is freely available helped us in implementing its true functions that is text recognition. We used GIMP as the main logo designing tool and for other artwork designing. We used VS Code as the main IDE for programming and writing our code. We used GenY motion Emulator as our main android emulator as it is lite weight and easy to use. We also used some native code from Java and Kotlin. XML was also used for designing the Splash screen. As we used basically all open-source software and APIs, it would be easier for others to implement these on their own system and hence make more modification for more refining in the future.

FUTURE SCOPE OF THE PROJECT

Software development is never –ending process and continues the life of the software as per the changing needs of the user from time to time. The project is no doubt has been developed keeping in mind easy modification and enhancement that may be required from time to time.
 However, there are many scopes to modify this software. As because due to shortage of time, we here become unable to include many things. We are trying to cover all their existing system for sales return records of the items but due to shortage of time we become unable to include many things. Due to lake of time, I here include none of them and a future scope one can develop these returns.

REFERENCES

- **https://www.edge-core.com/_upload/files/TEC_Social_Media_Login.pdf**

- **https://stackoverflow.com/questions/2661536/how-to-programmatically-take-a-screenshot-on-android**

- **https://www.androhub.com/android-floating-widget-like-facebook-messenger-chat-head/**

- **https://www.youtube.com/watch?v=eKxA3aa-uYI**

- **https://androidmonks.com/flutter-icons/**

- **https://www.businessinsider.in/india-has-more-fake-news-than-any-other-country-in-the-world-survey/articleshow/67868418.cms**

- **https://www.androhub.com/android-floating-widget-like-facebook-messenger-chat-head/**

- **S. Moghavvemi and M.H.N.B. MdNasir, 2016. SNS metrics analysis a study on fanpage interactive. Intl. J. Appl. Bus. Econ. Res., 14: 1405-1415.**

- **Factors influencing popularity of branded content in Facebook fan pages. Eur. Manage. J., 3**

- **https://www.quirks.com/articles/analyzing-the-content-of-social-media-data**