VISVESVARAYA TECHNOLOGICAL UNIVERSITY

Jnana Sangama, Belgaum-590018

A PROJECT REPORT (**15CSP85**)ON

SPAM MESSAGE FILTER USING NAIVE BAYES CLASSIFIER AND INTELLIGENT TEXT MODIFICATION METHOD

Submitted in Partial fulfillment of the Requirements for the Degree of

Bachelor of Engineering in Computer Science & Engineering

By

SAGAR S (1CR16CS142)

VISHWANATHA S (1CR16CS180)

VISHWAS TN (1CR16CS181)

PRABHUKUMAR GD (1CR17CS408)

Under the Guidance of,

DR.P.KAVITHA

Associate Professor, Dept. of CSE

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CMR INSTITUTE OF TECHNOLOGY

#132, AECS LAYOUT, IT PARK ROAD, KUNDALAHALLI, BANGALORE-560037

CMR INSTITUTE OF TECHNOLOGY

#132, AECS LAYOUT, IT PARK ROAD, KUNDALAHALLI,BANGALORE-560037

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

Certified that the project work entitled **SPAM MESSAGE FILTER USING NAIVE BAYES CLASSIFIER AND INTELLIGENT TEXT MODIFICATION METHOD** carried out by **SAGAR S**, USN **1CR16CS142**, **VISHWANATHA S**, USN **1CR16CS180**, **VISHWAS TN**, USN **1CR16CS181, PRABHUKUMAR GD**, USN **1CR17CS408,** bonafide students of CMR Institute of Technology, in partial fulfillment for the award of **Bachelor of Engineering** in Computer Science and Engineering of the Visveswaraiah Technological University, Belgaum during the year 2019-2020. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the Report deposited in the departmental library.

The project report has been approved as it satisfies the academic requirements in respect of Project work prescribed for the said Degree.

.

Dr.P.Kavitha	**Dr. Prem Kumar Ramesh**	**Dr. Sanjay Jain**
Associate Professor	**Professor & Head**	**Principal**
Dept. of CSE, CMRIT	**Dept. of CSE, CMRIT**	**CMRIT**

DECLARATION

We, the students of Computer Science and Engineering, CMR Institute of Technology, Bangalore declare that the work entitled **SPAM MESSAGE FILTER USING NAIVE BAYES CLASSIFIER AND INTELLIGENT TEXT MODIFICATION METHOD** has been successfully completed under the guidance of Dr.P.Kavitha, Computer Science and Engineering Department, CMR Institute of technology, Bangalore. This dissertation work is submitted in partial fulfillment of the requirements for the award of Degree of Bachelor of Engineering in Computer Science and Engineering during the academic year 2019 - 2020. Further the matter embodied in the project report has not been submitted previously by anybody for the award of any degree or diploma to any university.

Place: Bangalore

Date:

Team members:

SAGAR S (1CR16CS142) _____

VISHWANATHA S (1CR16CS180) _____

VISHWAS TN (1CR16CS181) _____

PRABHUKUMAR GD (1CR17CS408) _____

ABSTRACT

Spam emails have been a chronic issue in computer security. They are very costly economically and extremely dangerous for computers and networks. Despite of the emergence of social networks and other Internet based information exchange venues, dependence on email communication has increased over the years and this dependence has resulted in an urgent need to improve spam filters. Although many spam filters have been created to help prevent these spam emails from entering a user's inbox, there is a lack or research focusing on text modifications. Currently, Naive Bayes is one of the most popular methods of spam classification because of its simplicity and efficiency. Naive Bayes is also very accurate; however, it is unable to correctly classify emails when they contain leetspeak or diacritics. Thus, in this proposes, we implemented a novel algorithm for enhancing the accuracy of the Naive Bayes Spam Filter so that it can detect text modifications and correctly classify the email as spam or ham. Our Python algorithm combines semantic based, keyword based, and machine learning algorithms to increase the accuracy of Naive Bayes compared to Spam assassin by over two hundred percent.

ACKNOWLEDGEMENT

I take this opportunity to express my sincere gratitude and respect to **CMR Institute of Technology, Bengaluru** for providing me a platform to pursue my studies and carry out my final year project

I have a great pleasure in expressing my deep sense of gratitude to **Dr. Sanjay Jain,** Principal, CMRIT, Bangalore, for his constant encouragement.

I would like to thank **Dr. Prem Kumar Ramesh,** Professor and Head, Department of Computer Science and Engineering, CMRIT, Bangalore, who has been a constant support and encouragement throughout the course of this project.

I consider it a privilege and honor to express my sincere gratitude to my guide **Dr.P.Kavitha, Associate Professor,** Department of Computer Science and Engineering, for the valuable guidance throughout the tenure of this review.

I also extend my thanks to all the faculty of Computer Science and Engineering who directly or indirectly encouraged me.

Finally, I would like to thank my parents and friends for all their moral support they have given me during the completion of this work.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Spam refers to an email aimed manipulating an individual to whom it is aimed at or just randomly flooding the inbox. It is also called as junk mail and it floods Internet clients Inboxes. Today spam emails are of a variety of types ranging from ads to business promoting to doubtful products to some objectionable services. Therefore it is difficult to identify and classify an email as spam or non-spam.

Usenet also called as User Network is an email service that distributes group talks or emails aimed at a particular group of people associated with a certain service or product and are mostly informative but do crowd up the inbox of the user. The data that goes over the Internet is called Netnews" an accumulation of these data that is aimed at providing message about a specific topic is called a "newsgroup". People that read such news from these newsgroups are the prime target of Spammers. Spammers use these news groups for the promotion of certain unrelated ads or unrelated posts. Usenet spam robs clients of the utility of the newsgroups by promoting other unrelated posts.

1.1 Relevance of the Project

As the digitization of communication grows, electronic mail, or emails, has become increasingly popular; in 2016, an estimated 2.3 million people used email. In 2015, 205 billion emails were sent and received daily, which is expected to grow at an annual rate of 3% and reach over 246 billion by 2019. However, the growth in emails has also led to an unprecedented increase in the number of illegitimate mail, or spam - 49.7% of emails sent is spam - because current spam detection methods lack an accurate spam classifier. Spam is problematic not only because it often is the carrier of malware, but also because spam emails hoard network bandwidth, storage space, and computational power. Additionally, the commercial world has significant

interests in spam detection because spam causes loss of work productivity and financial loss. It is estimated that American firms and consumers lose 20 billion annually, even while sustained by the private firms' investment in anti-spam software. On the other hand, spam advertising earns 200 million per year. Although extensive work has been done on spam filter improvement over the years, many of the spam filters today have limited success because of the dynamic nature of spam. Spammers are constantly developing new techniques to bypass filters, some of which include word obfuscation and statistical poisoning. Although these two text classification issues are recognized, research today has largely neglected to provide a successful method to improve spam detection by counteracting word obstruction and Bayesian poisoning, and many common spam filters are unable to detect them.

1.2 Scope of the project

Naïve Bayes is very accurate; however, it is unable to correctly classify emails when they contain leetspeak or diacritics. Thus in this proposes, we implemented a novel algorithm for enhancing the accuracy of the Naive Bayes Spam Filter so that it can detect text modifications and correctly classify the email as spam or ham.

1.3 Problem statement

This project proposes "Spam Message Filtering using Intelligent Text Modification method". This project predicts spam messages more accurately than the existing projects. This project uses Naïve Bayes algorithm to modify the text.

CHAPTER 2

LITERATURE SURVEY

2.1 Paper 1

Email Spam Detection using integrated approach of Naïve Bayes and Particle Swarm Optimization[1]

- Naïve Bayes algorithm is a Bayes theorem based statistical machine learning based approach having properties of strong independence, probability distribution and ability to handle large datasets.
- In NB, probability distribution is evaluated from the frequency distribution of dataset.
- Particle Swarm Optimization (PSO) is swarm intelligence based concept derived in 1995 by Eberhart and Kennedy
- PSO work on the property of stochastic distribution and initially find the local search solution, then individual particle share their solution and global solution is obtained.
- NB having probability distribution property determines the possible class for the email content from the spam class or non-spam class on the basis of keywords present in the email textual data.
- PSO is used to further optimize the parameters of NB approach to improve the accuracy, search space and classification process.

2.2 Paper 2

Email Spam Classification by Support Vector Machine[2]

- This paper uses Support Vector Mechanism algorithm to identify spam emails.
- Descriptions as provided on Spam Assassin website for the dataset used in this paper.
- SVM is also considered as an important kernel methods, which is one of the most important areas in machine learning concepts.
- Smart Traffic Control System with Application of Image Processing Techniques
- In this work they have also compared Linear and Gaussian as two of the very popular kernel and employed them for the problem of email spam detection
- The two models have been proposed, trained and tested using popular and often used standard database.

2.3 Paper 3

Intelligent Model for Classification of SPAM and HAM[3]

- In this paper they have used machine learning and non machine learning approaches.
- Machine learning approaches like support vector mechanism, neural network etc. Non machine learning approaches like strong key word searching and whitelisting and blacklisting of words.
- The sets so formed are further used as training set and the classification set.
- The process is to use the first set as training set and the remaining N-1 sets as the sets to be classified.

- In the next iteration the second set is used as the training set and the remaining sets are sets to be classified. The process is repeated until all the sets are used as training sets.

- The emails are classified based on the spam percentage each mail gains.

CHAPTER 3

REQUIREMENTS SPECIFICATION

This chapter involves both the hardware and software requirements needed for the project and detailed explanation of the specifications.

3.1 Hardware requirements

- A PC with Windows/Linux OS
- Processor with 1.7-2.4gHz speed
- Minimum of 8gb RAM

3.2 Software specification

- Anaconda distribution package
- Python libraries

3.3 Software requirements

3.3.1 Anaconda distribution:

Anaconda is a free and open-source distribution of the Python and R programming languages for scientific computing (data science, machine learning applications, large-scale data processing, predictive analytics, etc.), that aims to simplify package management system and deployment. Package versions are managed by the package management system conda. The anaconda distribution includes data-science packages suitable for Windows, Linux and MacOS.

3.3.2 Python libraries:

For the computation and analysis we need certain python libraries which are used to perform analytics. Packages such as SKlearn, Numpy, pandas, Matplotlib, Flask framework, etc are needed.

SKlearn: It features various classification, regression and clustering algorithms including support vector machines, random forests, gradient boosting, k-means and DBSCAN, and is designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy.

NumPy: NumPy is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays. It **is** the fundamental package for scientific computing with Python.

Pandas: Pandas is one of the most widely used python libraries in data science. It provides high-performance, easy to use structures and data analysis tools. Unlike NumPy library which provides objects for multi-dimensional arrays, Pandas provides in-memory 2d table object called Data frame.

Matplotlib: matplotlib pyplot is a collection of command style functions that make matplotlib work like MATLAB. Each pilot function makes some change to a figure: e.g., creates a figure, creates a plotting area in a figure, plots some lines in a plotting area, decorates the plot with labels, etc

CHAPTER 4

SYSTEM ANALYSIS AND DESIGN

4.1 System Architecture

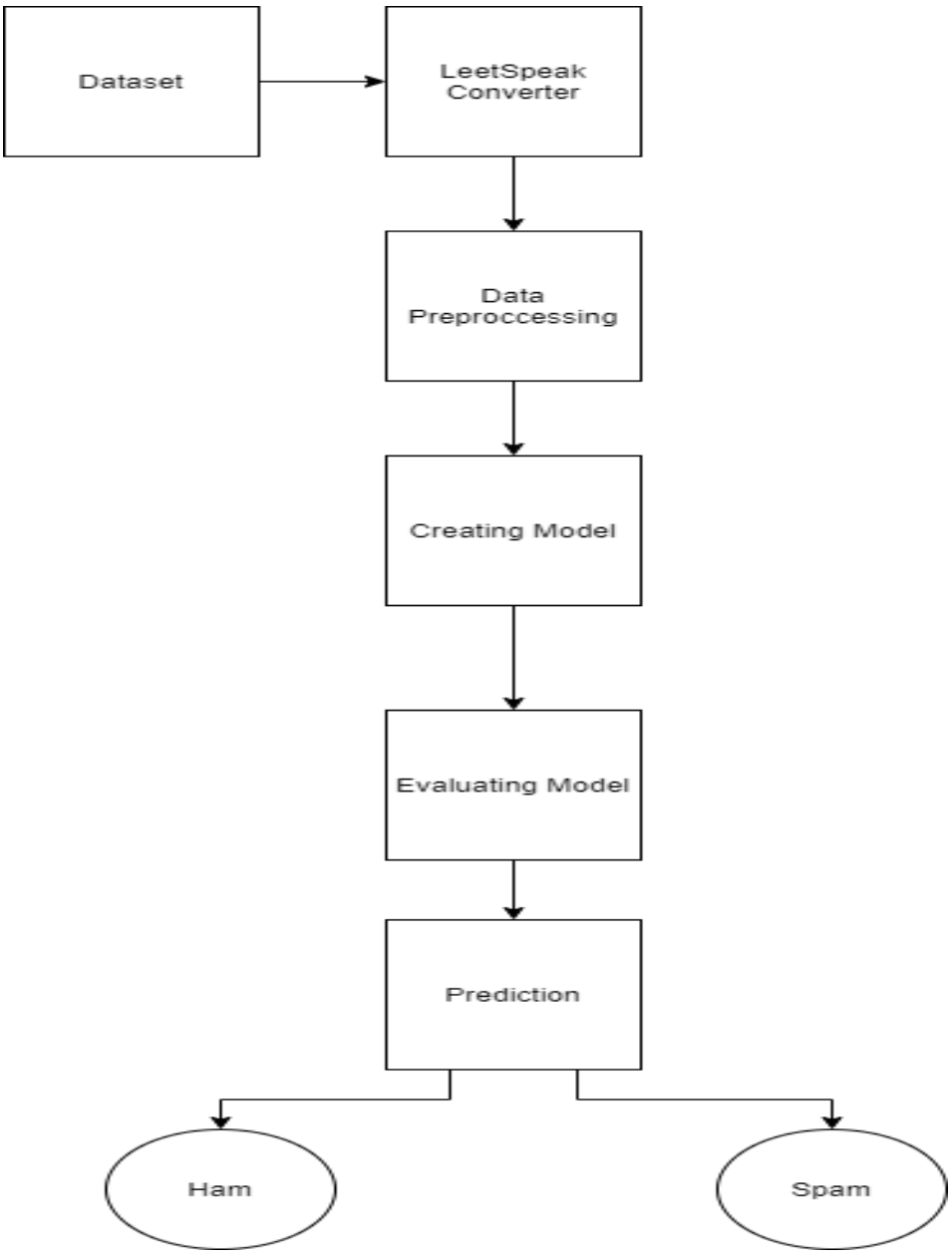

Fig 4.1 Proposed System

The above figure represents the Proposed System of the project. We first take the dataset then it is passed through the leetspeak converter which converts all the leetspeak characters into normal text. Then that data set is set for data pre processing where in which all the unwanted and stop words are removed in order to minimize the dataset as much as possible. Then a model is created to test the data which is evaluated. Finally the dataset is fed to this model which predicts the output whether the given text is spam or ham.

4.2 Dataset

```
v1,v2,,,
ham,"Go until jurong point, crazy.. Available only in bugis n great world la e buffet... Cine there got amore wat...",,,
ham,Ok lar... Joking wif u oni...,,,
spam,Fl2ee entrЧ 1n 2 @ wki_y com¶ +o w1n w1n Cu|> f!nal tkt5 21st Ma¥ 2005. Tex'|' Tex'|' 7`o 87121 '|'o recei\/e entr`/ que§tion T&C's app|_y 08452810075over18's,,,
ham,U dun say so early hor... U c already then say...,,,
ham,"Nah I don't think he goes to usf, he lives around here though",,,
spam,"FreeM$g He¥ '|'here darl!ng it's bëën 3 week's n0w an|) nΩ woЯd back! I'd l!ke zome f{_}n Чou u|* foЯ |t still? T|8 ok! X*X 5td chg5 '|'o send, ï¿½1.50 -|-o rcv",,,
ham,Even my brother is not like to speak with me. They treat me like aids patent.,,,
ham,As per your request 'Melle Melle (Oru Minnaminunginte Nurungu Vettam)' has been set as your callertune for all Callers. Press *9 to copy your friends Callertune,,,
spam,WINNER!! A§ /_\ valued ne~|~work cu5tomer λou ha\/e b33n $elected -|-o receiνea ï¿½900 pr|ze reward! T{} cla|m ca1_1_ 09061701461. Cla|m cod£ KL341. Val!d 12 hour$ only.,,,
spam,Ha|] `/our mob!le 11 monthz oи more? oи oи ent1tled `|`o Upda~|~e -|-o ~|~he late§t co1_our mobile5 w|th camel2a fol2 Free! Cai_i_ Th£ Mob!le Upda+e C0 C0 []n 08002986030,,,
ham,"I'm gonna be home soon and i don't want to talk about this stuff anymore tonight, k? I've cried enough today.",,,
spam,"[]n chancez -|-o w!n CASH! FЯom 100 7o 20,000 pound§ txt> CSH11 an|] $end ŧo 87575. Co$t 150p/day, 6days, 16+ T§andC§ appl_y Repi_y Repi_y 4 |nfo",,,
spam,"URGENT! Yo(_) have w{}n /_\ 1 w€€k w€€k member5hip !n ou|2 ï¿½100,000 Pr|ze Jackpot! Tx7` 7he word: 7he 7o No: 81010 T&C www.dbuk.net 7o 7o 4403LDNW1A7RW18",,,
ham,I've been searching for the right words to thank you for this breather. I promise i wont take your help for granted and will fulfil my promise. You have been wonderful and a blessing at all times.,,,
ham,I HAVE A DATE ON SUNDAY WITH WILL!!,,,
spam,"XXXMobileMovieClub: T{} uze Чour credit, cl!ck -|-he -|-he l!nk |n ~|~he nexŧ ~|~x~|~ me§§age oи cl!ck here>> http://wap.xxxmobilemovieclub.com?n=QJKGIGHJJGCBL",,,
ham,Oh k...i'm watching here:),,,
ham,Eh u remember how 2 spell his name... Yes i did. He v naughty make until i v wet.,,,
```

Fig 4.2 Dataset with leetspeak

The above screenshot consists of sample dataset that contains leetspeak words which needs to be removed.

```
v1,v2,,,
ham,"Go until jurong point, crazy.. Available only in bugis n great world la e buffet... Cine there got amore wat...",,,
ham,Ok lar... Joking wif u oni...,,,
spam,Free entry in 2 a wkly comp to win FA Cup final tkts 21st May 2005. Text FA to 87121 to receive entry question T&C's apply 08452810075over18's,,,
ham,U dun say so early hor... U c already then say...,,,
ham,"Nah I don't think he goes to usf, he lives around here though",,,
spam,"FreeMsg Hey there darling it's been 3 week's now and no word back! I'd like some fun you up for it still? Tb ok! XxX std chgs to send, £1.50 to rcv",,,
ham,Even my brother is not like to speak with me. They treat me like aids patent.,,,
ham,As per your request 'Melle Melle (Oru Minnaminunginte Nurungu Vettam)' has been set as your callertune for all Callers. Press *9 to copy your friends Callertune,,,
spam,WINNER!! As a valued network customer you have been selected to receivea £900 prize reward! To claim call 09061701461. Claim code KL341. Valid 12 hours only.,,,
spam,Had your mobile 11 months or more? U R entitled to Update to the latest colour mobiles with camera for Free! Call The Mobile Update Co FREE on 08002986030,,,
ham,"I'm gonna be home soon and i don't want to talk about this stuff anymore tonight, k? I've cried enough today.",,,
spam,"SIX chances to win CASH! From 100 to 20,000 pounds txt> CSH11 and send to 87575. Cost 150p/day, 6days, 16+ TsandCs apply Reply HL 4 info",,,
spam,"URGENT! You have won a 1 week FREE membership in our £100,000 Prize Jackpot! Txt the word: CLAIM to No: 81010 T&C www.dbuk.net LCCLTD POBOX 4403LDNW1A7RW18",,,
ham,I've been searching for the right words to thank you for this breather. I promise i wont take your help for granted and will fulfil my promise. You have been wonderful and a blessing at all times.,,,
ham,I HAVE A DATE ON SUNDAY WITH WILL!!,,,
spam,"XXXMobileMovieClub: To use your credit, click the WAP link in the next txt message or click here>> http://wap.
```

Fig 4.3 Dataset without leetspeak

The above screenshot consists of sample dataset where all the leetspeak words are removed and converted to the corresponding/similar words.

4.3 Design

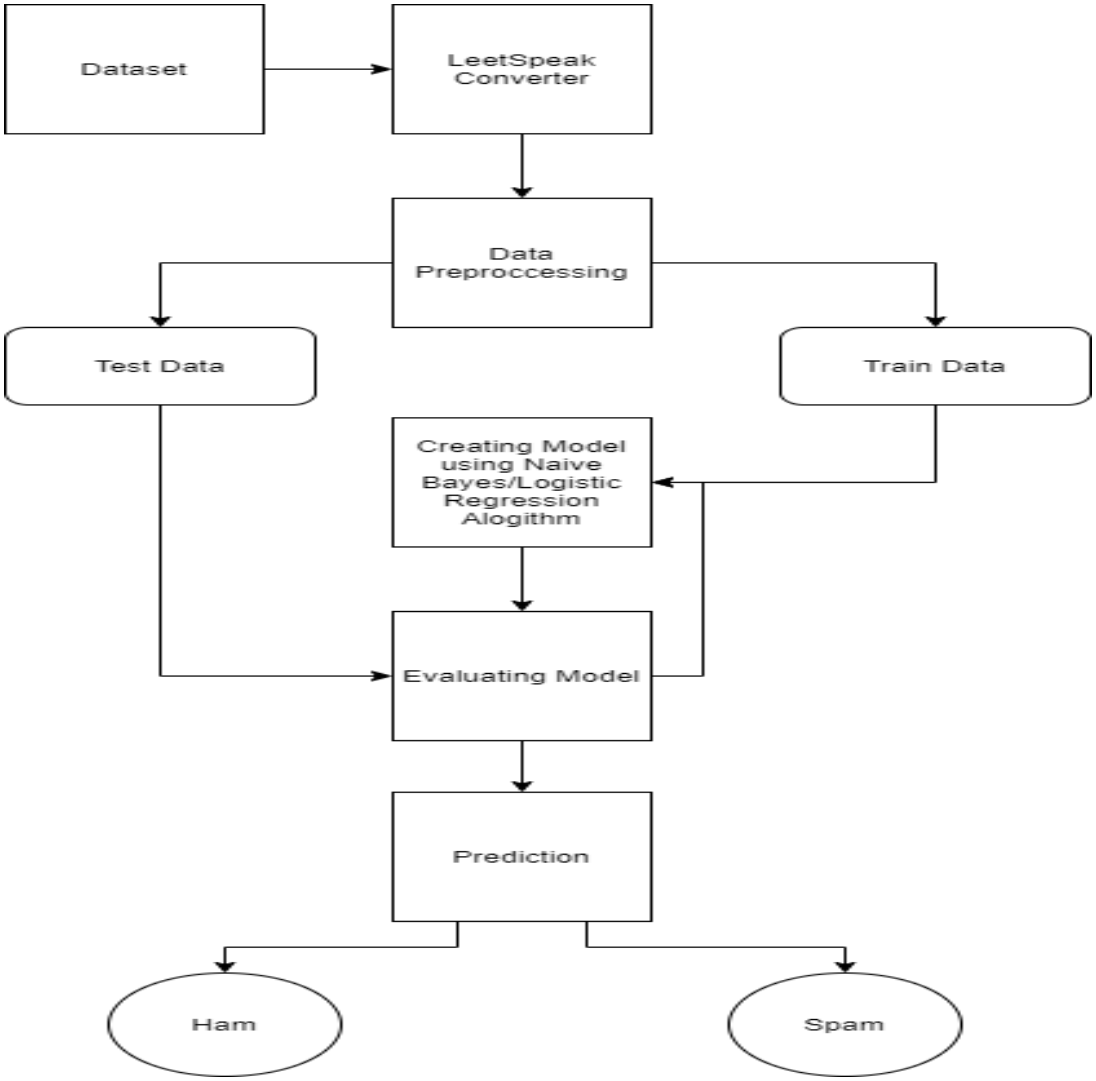

Fig 4.4 Flow chart

The dataset is sent to the leetspeak converter where all the leetspeak words are converted to corresponding words. The it is sent to data pre processor where all the stop and un wanted words are removed. The data is split into test data and train data. Each of these split data are sent to the model to evaluate their efficiency for the test and train data. Finally the out put is predicted for the test data.

CHAPTER 5

IMPLEMENTATION

5.1 Algorithm

Naïve Bayes Classifier applies to learning tasks where each instance x is described by conjunction of attribute values and where the target function f(x) can take on any value from some finite set V. A set of training examples of the target function is provided, and a new instance is presented, described by tuple of attribute values (a1,a2,…) . The learner is asked to predict the target value, or classification, for this new instance.

The Bayesian approach to classifying the new instance is to design the most portable target value, Vmap, given the attribute values(a1,a2,..) that describe the instance.

$$v_{MAP} = \underset{v_j \in V}{\operatorname{argmax}} P(v_j | a_1, a_2 \ldots a_n)$$

Use Bayesian theorem to rewrite this expression as

$$v_{MAP} = \underset{v_j \in V}{\operatorname{argmax}} \frac{P(a_1, a_2 \ldots a_n | v_j) P(v_j)}{P(a_1, a_2 \ldots a_n)}$$

$$= \underset{v_j \in V}{\operatorname{argmax}} P(a_1, a_2 \ldots a_n | v_j) P(v_j) \qquad \textbf{equ (1)}$$

The Naïve Bayes classifier is based on the assumption that the attribute values are conditionally independent given the target value. Mean, the assumption is that given the target value of the instance, the probability of observing the conjunction is just product of the probabilities for the individual attributes.

$$P(a_1, a_2 \ldots a_n | v_j) = \prod_i P(a_i | v_j)$$

Substituting this into equation 1

Naive Bayes classifier:

$$V_{NB} = \underset{v_j \in V}{\operatorname{argmax}} P(v_j) \prod_i P(a_i | v_j) \qquad \text{equ (2)}$$

Where, V_{NB} denotes the target value output by the naïve Bayes Classifier.

5.2 Pre processing the data

Remove punctuations. Remove

stop words :- Stop words like “and”, “the”, “of”, etc are very common in all English sentences and are not very meaningful in deciding spam or legitimate status, so these words have been removed from the emails.

```
def process_text(text):
    '''
    What will be covered:
    1. Remove punctuation
    2. Remove stopwords
    3. Return list of clean text words
    '''

    #1
    nopunc = [char for char in text if char not in string.punctuation]
    nopunc = ''.join(nopunc)

    #2
    clean_words = [word for word in nopunc.split() if word.lower() not in stopwords.words('english')]

    #3
    return clean_words
```

In pre processing we remove all the unwanted words and punctuations such as comma, full stops, extra spaces and other repeated words. This will help to reduce the data to be processed. This helps to process the data much faster than before due to small amount of data.

5.3 Creating Model and Training

- Naïve Bayes Classifier

```
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import MultinomialNB
messages_bow = CountVectorizer(analyzer=process_text).fit_transform(df['v2'])
X_train, X_test, y_train, y_test = train_test_split(messages_bow, df['v1'], test_size = 0.20, random_state = 0)
classifier = MultinomialNB()
classifier.fit(X_train, y_train)
```

Fig 5.3.1 Naive Bayes Model

Here we split the data into train and test data and then we convert our data into the desired matrix format. To do this we will be using Count Vectorizer(). There are two steps to consider here: Firstly, we have to fit our training data (X_train) into Count Vectorizer() and return the matrix. Secondly, we have to transform our testing data (X_test) to return the matrix. Note that X_train is our training data for the 'v2' column in our dataset and we will be using this to train our model. X_test is our testing data for the 'v2' column and this is the data we will be using(after transformation to a matrix) to make predictions. Import the Multinomial classifier and fit the training data into the classifier using fit(). Name your classifier 'classifier'. Now that our algorithm has been trained using the training data set we can now make some predictions on the test data stored in 'X_test' using predict().

- **Logistic Regression model**

```
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
messages_bow = CountVectorizer(analyzer=process_text).fit_transform(df['v2'])
message_train, message_test, spam_nospam_train, spam_nospam_test = train_test_split(messages_bow, df['v1'], test_size = 0.20,
Spam_model = LogisticRegression(solver='liblinear', penalty='l1')
Spam_model.fit(message_train, spam_nospam_train)
pred = Spam_model.predict(message_test)
```

Fig 5.3.2 Logistic regression Model

Here we split the data into train and test data and then we convert our data into the desired matrix format. To do this we will be using Count Vectorizer(). There are two steps to consider here: Firstly, we have to fit our training data (message train) into Count Vectorizer() and return the matrix. Secondly, we have to transform our testing data (message test) to return the matrix. Note that message train is our training data for the 'v2' column in our dataset and we will be using this to train our model. message test is our testing data for the 'v2' column and this is the data we will be using(after transformation to a matrix) to make predictions.

Import the Logistic Regression algorithm and fit the training data into the model using fit(). Name your model 'spam model'. Now that our algorithm has been trained using the training data set we can now make some predictions on the test data stored in 'message test' using predict().

CHAPTER 6

RESULTS AND DISCUSSION

Comparing Results for Naïve Bayes and logistic Regression model with 80-20 split ratio.

Naïve Bayes Model

Evaluating test data(20% of dataset)

	precision	recall	f1-score	support
ham	0.99	0.97	0.98	898
spam	0.84	0.97	0.90	152
micro avg	0.97	0.97	0.97	1050
macro avg	0.92	0.97	0.94	1050
weighted avg	0.97	0.97	0.97	1050

Confusion Matrix:
[[870 28]
[5 147]]

Accuracy: 0.9685714285714285

Logistic Regression Model

Evaluating test data(20% of dataset)

	precision	recall	f1-score	support
ham	0.97	1.00	0.98	888
spam	0.99	0.83	0.91	162
micro avg	0.97	0.97	0.97	1050
macro avg	0.98	0.92	0.95	1050
weighted avg	0.97	0.97	0.97	1050

Confusion Matrix:
[[887 1]
[27 135]]

Accuracy: 0.9733333333333334

6.1 SCREENSHOTS

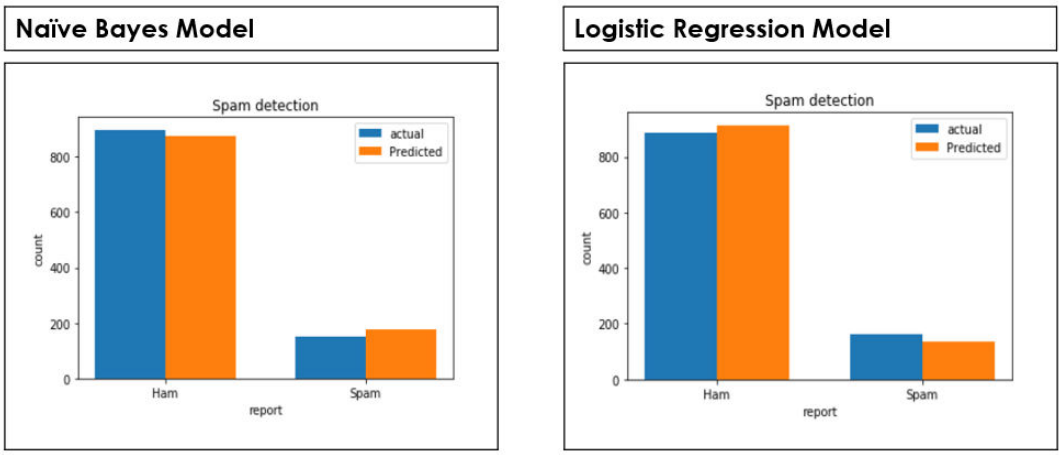

Fig 6.1 80-20 split ration comparison

The above screenshots shows the results of the output when the naïve bayes classifier and logistic regression models are tested for 80-20 split ratio that is the percentage of data provided while training the model and testing the model respectively.

CHAPTER 7

TESTING

- We tested the dataset and found out which is ham and spam indicated as spam and ham.
- We calculated the accuracy of the predicted output.
- We finally compared the output with the model created using logistic regression algorithm.

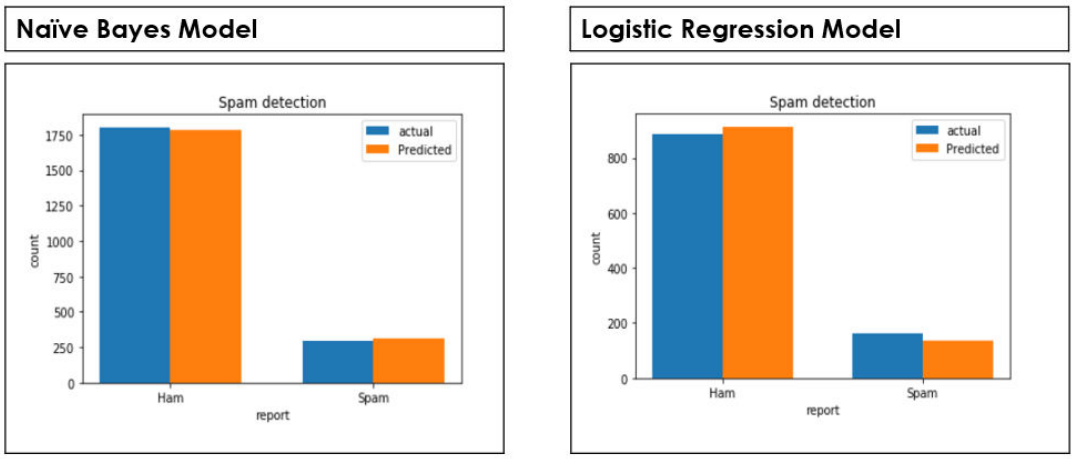

Fig 7.1 60-40 split ration comparison

The above screenshots shows the results of the output when the naïve bayes classifier and logistic regression models are tested for 60-40 split ratio that is the percentage of data provided while training the model and testing the model respectively.

Comparing Results for Naïve Bayes and logistic Regression model with 60-40 split ratio.

Naïve Bayes Model

Evaluating test data(40% of dataset)

	precision	recall	f1-score	support
ham	0.99	0.98	0.98	1804
spam	0.88	0.94	0.91	296
micro avg	0.97	0.97	0.97	2100
macro avg	0.94	0.96	0.95	2100
weighted avg	0.97	0.97	0.97	2100

Confusion Matrix:
[[1767 37]
[19 277]]

Accuracy: 0.9733333333333334

Logistic Regression Model

Evaluating test data(40% of dataset)

	precision	recall	f1-score	support
ham	0.97	1.00	0.98	888
spam	0.99	0.83	0.91	162
micro avg	0.97	0.97	0.97	1050
macro avg	0.98	0.92	0.95	1050
weighted avg	0.97	0.97	0.97	1050

Confusion Matrix:
[[887 1]
[27 135]]

Accuracy: 0.9733333333333334

CHAPTER 8

CONCLUSION AND FUTURE SCOPE

8.1 Conclusion

We proposed a novel algorithm for enhancing the accuracy of the Naive Bayes Spam Filter. The algorithm was implemented as an enhancement for Naive Bayes Classifier and also tested with logistic regression model. Naive Bayes has a very fast processing speed and allows for a small training set, hence is suitable for real-time spam filtering. We are also using Intelligent Text Modification method to identify messages containing leetspeak and diacritic. We are able to classify email as spam or ham. By creating an addition to Naive Bayes Classifier. We also found that our new addition helped improve ham classification due to the high recall and precision rates. We demonstrated that our algorithm consistently reduced the amount of spam emails misclassified as ham email.

8.2 Contribution

We have tried to introduce leetspeak characters into the dataset and then convert them to the normal text. We then predict the output of the dataset. This is our contribution to the project.

8.3 Future Scope

In the future we would like to create an API for the same and test that in real world environment. We will try to optimize this project for much larger amount of dataset. Since our addition successfully enhances the Naive Bayes spam filter, we will try to implement the addition onto other machine learning spam filters such as Vector Space Models, clustering, and artificial neural networks. Combining these other methods will allow the improvement of spam detection across many different systems to ultimately create a well developed spam detector for text modifications.

REFERENCES

[1] Email Spam Detection using integrated approach of Naïve Bayes and Particle Swarm Optimization

[2] Email Spam Classification by Support Vector Machine

[3] Intelligent Model for Classification of SPAM and HAM

[4] Team, Radicati. "Email Statistics Report, 2015-2019. The Radicati Group." (2015).

[5] Androutsopoulos I., J. Koutsias, K. V. Chandrinos, G. Paliouras,and C. D. Spyropoulos, "An evaluation of naive bayesian antispam filtering", In: *11th European Conference on Machine Learning,* pp.9-17, Barcelona, Spain, 2000.

[6]GitHub, Inc, "Spam Assassin," 21 April 2016. [Online]. Available: https://github.com/dmitrynogin/SpamAssassin.git. [Accessed 20 August 2017].