CHAT REVIEWS

Harischandra K.P.I.E

IT14093760

B.Sc (Special Hons) Degree of Information technology

Department of Information Technology

Sri Lanka Institute of Information Technology

October 2017

CHAT REVIEWS

Harischandra K.P.I.E

IT14093760

Dissertation submitted in partial fulfillment of the requirements for the

B.Sc (Special Hons) Degree of Information technology

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

October 2017

CHAT REVIEWS

Member	ID Number	Signature
Harischandra K.P.I.E	IT14093760	

Supervisor: Dr. Dharshana Kasthurirathna

.....

B.Sc (Special Hons) Degree of Information technology

Sri Lanka Institute of Information Technology

Sri Lanka

October 2017

DECLARATION

We declare that this is our own work and this project proposal does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

ABSTRACT

The use of Internet chat applications has benefited many different segments of society. It also creates opportunities for criminal enterprise, terrorism, and espionage. We present a study of a real-world application of chat analysis which will analyze chat messages in four ways such as topic detection, Emotion Extraction, Evaluate healthy and Personal information sharing analysis. Also analyzing chat traffic has important applications for both the military and the civilian world. Here on this document, it compares the results of an unsupervised learning approach with those of a supervised classification approach with regards to chat review application. The paper also discusses some of the specific challenges presented by this chat review application. Unsupervised learning techniques such as clustering are very popular for analyzing text for topic identification as well as emotion extraction. These techniques have several attractive features, the most significant being that they do not require labeled training examples. This however is also a disadvantage under some circumstances. Therefor meantime we do this research we will discover more and more technologies required for analyzing chat massages based on four different categories such as topic detection, Emotion Extraction, evaluate healthy and Personal information sharing analysis. With use of this chat analysis application user will be able identify the chatting partner in analytical way. And system will keep an analytical review for each chat session user interacted. Also system will be capable of showing its analytical data in a user friendly manner (in a graphical way).

Key Words : chat , system , analyze , classification , application , review , sharing , information

ACKNOWLEDGEMENT

I would like to express my special thanks of gratitude to our project supervisor, Dr. Dharshana Kasthurirathna for his patient guidance, enthusiastic encouragement and useful critique of our research work. He give us his willingness and time so generously has been very much appreciated. Especially mentioning the moral support and the continuous guidance by providing important feedback enabled us to complete our work successfully. And also for his support that he provided to us by giving more information and patiently guiding us in the correct direction, for giving us the courage to carry out the project successfully I would also like to thank our lecturer-in-charge, Mr. Jayantha Amararachchi for constantly providing us with the guidance and assistance to carry out the research successfully. Also, I would like to thank my research group members as well as all the other friends who supported me and encouraged me throughout this project. Finally, I would like to thank my parents for their support and encouragement throughout my studies.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AI	Artificial intelligence
VS	Visual studio
NU	Neural Networks
SLIIT	Sri Lanka Institute of Information Technology
IM	Instant Messaging
AI	Artificial intelligence
Database	A set of data or information monitored by the system.
SRS	A document which contains description of a software system to be developed and described all the functional and non-functional requirements.
NE	Named Entity
HE	Honey Encryption

Table 1: Abbreviations

1.0. Introduction

This document provides all of the requirements for the Chat Reviews. All parts are developed individually and finally combined all together. This document will describe how to do the work flow and final outcome to the industry. This will help to identify trustworthiness of a chat session and identify trust users.

1.1. Problem to be addressed

In the fast moving world we meet different types of people. There are people with different ages and life styles. The main problem of trust worthiness of a people is a challenge for social media users. Personal information is used by many businesses for legitimate communication. However this is not always the case and some personal information can be misused by criminals or inappropriately by marketers. The important thing is true of the information. There have more social medias now and people connect with it. People need some trust about information. Many online services require users to provide some personal information in order to use their service. Prior to providing personal information, you should think about what can be done with your personal information and assess whether you are still happy to pass on these details. In addition to inappropriate or illegal use of information, disclosing personal information online can impact your digital reputation. In this case we find solution for detect trust person by analyzing messages. There can have more issues of personal information share in online such that Spam, scams, identity theft and fraud are just some of the more serious issues that you might face if you chat with another chatting partner.

We find a better solution to give better result for trust worth of people. By analyzing previous history of partner and give output to client. It helps to get some idea about unknown chatting partner. In this case this basically goes on identify personal information in the chat history. In this study, the term refers to any information that could be used to know the identity and build a profile of an individual. For this study, this term refers to web sites that allow members to construct a profile to articulate their relationship to other users in a way that is visible. In this

case lower trust will broke the relationship between chatting partners. So, idly this will make high in relationships. For this research we mainly give a solution for detect spam, and fake users by analyzing personal information of a message.

1.2. background context

Real Time Chat Review is a research type project, which has been proposed under the course, Comprehensive Design Analysis Projects in Sri Lanka Institute of Information Technology (SLIIT). This system will especially helpful for social media users to share their ideas with the world on their own way and trust other people. Researchers found out the difficult part is identify relevant algorithms and procedure to do that. The main aim of this project is to develop social chat tool that can be used effectively every users. This provides a user friendly environment to detect and achieve our target. We mainly deal with algorithms to develop this [1] . We use open NLP libraries and other open source libraries to gain more efficient to our system application. We develop models by using algorithms, and we create models to detect information. By using sample information we create model. When we implement model its hard to find sample data because there is no sample data according to personal information sharing. So I need to create custom dataset and model to develop this as well. The following image will show one data sentence of dataset.

<START:person>Johny<END> and<START:person>Ricky<END> are <START:relation>brothers<END>.

Figure 1

The system detect sentences gives lots of information. There have personal information and other information as well. People have some basic understand about personal information. So that is a plus point to the web site. Site analysis give what they ask from other user and then analysis the details. More personal information share is better for make better trust worthiness. When we

analysis we consider more factors. Such that sentence , tokens and some relevant informations as well.

1.3. Research gap

Real Time Chat Review is a research type project, it will cover trustworthiness of social media. As previously explained, currently there are no proper systems or tools to provide the facility. We cover the areas that we explain early to detect chatting partner. In early case there have no proper solution to detect personal information o chat messages. So then I try to give better solution for this as well. There can have profile analysis but I try to give analysis by using messages. Also there are no systems available including the detect personal information of a chat messages, detect spam message session, evaluation session and history session as one helpful package.

The focal point of our project is to provide an interactive environment to analysis any time they wish. When time and chat history change there output may be change according to given inputs. Inputs given as chat messages. No need to worry about that the system automatically get input chat history and give outputs. It is a package of help full fore components of personal information sharing, emotion sharing, evaluate healthy and topic detection.

The team going to achieve this goal by providing, display graphs according to the given chatting partner in the social media. To develop our system uniquely, we have come up with a better solution with our own algorithms which are facilitated by the open NLP. This research is basically depends on model training using algorithm and data science. We use open NLP libraries to get more accuracy as well. Social media network get the messages and save it to database. That data will do some analysis and identify relevant personal information on that given messages.

The main objective of our project is to provide an interactive environment to analysis trustworthiness of a social media chatting partner. The full system gives better output of personal chat history.

To develop our system uniquely, we have come up with a better solution with our own algorithms, different libraries, different kinds of new technologies and features. We can ensure that there is no similar product which is currently using as Chat Reviews.

1.4. Research questions

The project consists of various kinds of different and unique features. Therefore considerable amount of challenges had to face to the project team while developing the system and problems arises for user while using it.

From user perspective

- **Lack of existing social media analysis tools**

In order to do referencing there are only few helpful social media analysis websites that use messages but they not use personal information in a message to analysis trustworthiness of a chatting partner. There are some profile analysis to get less understand that the information is accurate or not. Through those websites users can only learn and evaluate up to basic level of person. This system can make people who are trust by chat with them. Majority of the people cannot afford to buy such a technology even they really needed.

- **Currently there is no personal information that they share.**

Even if there are some chat conversation they have its enough. Not only personal information the whole system consider more factors. Some of that are topic, health , spam , emotion as well.

- **Is it a requirement to use any special tool or third party software?**

There no need to implement such software. This is complete system. This is a web application. No need much computer resources. This will work as other web applications.

From developer perspective

- **What are the best algorithms for this research?**

There are many algorithms available for the machine learning. Among those algorithms, there are methods which match with our requirement. Most of the algorithms are not given the best answer for personal information sharing, and available detection techniques are not accurate. We need to find how to develop custom model and train data to get accurate results.

- **Accuracy in run time analysis.**

The accuracy of the given input stream is high because the system get input as chat history. So there have available messages in history always. The algorithms must use minimum running time and exhibit maximum efficiency. Practical session of the Chat Review system totally depends on this machine learning. User can freely and easily view partner and analysis.

- **Most of systems not support.**

This is a challenge for us. Therefore we develop web application. It run on web. If it is stand alone application we need to develop it for more platforms. In this situation we get exact solution for it and develop to cover most of platforms.

- **How do we find dataset?**

This is a challenge for us. There have some dataset to other modules. But the main problem is hard to find personal information dataset. We need to make dataset and train our model. The project team has overcome from these challenges in order to develop a better product.

2.0 Literature and Methodology

2.1. Addressing the literature

2.1.1 Research attempts related to Chat messages analysis

Up to now several number of research Attempts have been taken regarding to some personal information analysis. Some of the identified research projects are stated below.

Topic detection and categorization

Topic Detection (TD) is one of the issue in the field of Natural Language Processing (NLP) from the beginning [2]. It becomes a hot spot and many classic models and methods have been made currently, as it faces the information retrieval and extraction detecting on unknown topics and track exist ones. The paper reviews the development status of TD research. The definitions of Topic, Story, Event and Activity are stated. TDT corpus and primary tasks in current research are introduced. The integral framework and key technologies including topic models, correlation calculation and clustering and classification, are proposed. The status of development of topic detection and tracking is concluded in the end. This system only compatible for topics categorization but we need to consider more factors then we need to reject this system but we get some idea about to detect topic detection as well.

Named Entity Reorganization

Named Entity Reorganization (NER) tools and frameworks implement a broad spectrum of approaches, which can be subdivided into three main categories: dictionary-based, rule-based and machine learning approaches [4]. The first systems for NER implemented dictionary-based approaches, which relied on a list of named entities (NEs) and tried to identify text in sentence Following work then showed that these approaches did not perform well for NER tasks such as recognizing proper names [3]. Thus, rule-based approaches were introduced. These approaches rely on hand-crafted rules to recognize NEs. Most rule-based approaches combine dictionary and rule-based algorithms to extend the list of known entities. Nowadays, hand-crafted rules for recognizing NEs are usually implemented when no training examples are available for the language to process [5]. When training examples are available, the methods of choice are borrowed from supervised machine learning. Approaches such as Hidden Markov Models [6], Maximum Entropy Models [7] and Conditional Random Fields [8] have been applied to the NER task. Due to scarcity of large training corpora as necessitated by supervised machine learning approaches, the semi-supervised and unsupervised machine learning paradigms have also been used for extracting NER from text. In [44], a system was presented that combines with stacking and voting classifiers which were trained with several languages, for language-independent NER. [4] gives an exhaustive overview of approaches for the NER task. In this system the check only name just like person names. So we need more information related person then the system not meet our requirement. So then we have to introduce new application.

Identify fake and spam messages

In this development tool discuss about Detecting Misinformation During Natural Disasters on Social Media, we summarize some of the previous research work related to spam and fake accounts or contents on social media. They studied how to identify spammers or spam messages on social media. They study and prove how to detect that. [9] analyzed how shortened URLs on Twitter have been used to link malware and spam links. McCord and Chuah [10] used machine learning approach to detect social media chat and other spammers.

2.1.2 Existing Applications

Honey Chatting

There have been many efforts to strengthen security of Instant Messaging (IM) system. One of the typical technologies is the conventional message encryption using a secret or private key. However, the key is fundamentally vulnerable to a brute-force attack, causing to acquire the original message. In this respect, a countermeasure was suggested as the way to generating plausible-looking but fake plaintexts, which is called Honey Encryption (HE). In this paper, we present a HE-based statistical scheme and design a Honey Chatting application, which is robust to eavesdropping. Besides, we verify the effectiveness of the Honey Chatting by comparing the entropy of decrypted messages through experiments.

Group Wise

Group wise is used to real-time chat on desktop or mobile devices without the risk. The feature of that are Secure instant message delivery with Micro Focus eDirectory™ user authentication and SSL message encryption, A contact list that displays user information based on the information available in the Micro Focus eDirectory, Fully-secured iOS and Android mobile applications, Support for Windows, OS X, and Linux platforms. There have message encryption and topic detection facilities in this tool.

intelligent diagnosis system

This system mainly develop for topic detection purpose. Topic detection module detects the topic associated with messages. There are defined topic classes along with the training data associated with each topic class. For topic analysis in IM Analysis, we aim to identify chat sessions limited only to a number of important topics. Therefore, we have adopted supervised topic detection

approaches based on Naïve **Bayes Theorem** [11], and **Bernoulli document model** [12] which have been demonstrated with good performance for classifying text.

Topic detection module use training dataset collected from social networks based on related topics and classification algorithm (supervised learning) [13] to classify message. Text classifiers often don't use any kind of deep representation about language. This is an extremely simple representation. It only knows which words are included in the message (and how many times each word occurs), and throws away the word order. Then it gives output to user with topics.

Compare existing applications with the proposed CHAT REVIEWS web application

	Intelligent Diagnosis System	Honey Chatting	Group Wize	Proposed System
Keyword based Searching				
Topic Identification	✓		✓	✓
Emotion Extraction			✓	✓
Message Encryption		✓		✓
Detect Personal Information				✓
Evaluate Healthy				✓

Figure 2: Compare existing applications with the proposed Chat Reviews

2.1.3 Our Approach

This is proposed to develop as a solution for a potential way of integrating and finding trust worthiness of Social media like Facebook. Because today there will be more fake account holders in social media network. Mention will be made throughout this document of selected probable features of future releases. The purpose of this is to guide developers in selecting a design that will be able to accommodate the full-scale application.

Other than that the system extracted data from the social media will be categorized automatically. The information can be general knowledge based information regarding the customer requests which is provided from above mentioned social media channels.

The final outcome of the project is identify the trustworthiness of social media account holder. It will analyze user inputs and provide a service using its built in knowledge. The Objective is to design a flexible intelligent, efficient and real time system and it's a kind of a cost effective, time saving way to a profitable business with new technologies.

Now a days there have lots of social media sites and there have lots of fake account holders. We propose the system to identify real account holders and trustworthiness. We can reduce fraud activities, then we can make secure site for users. This proposed method will come up with successful chat in social media chat sessions.

2.2 Methodology

2.2.1 Work Breakdown Structure

Bellow diagram shows the four main components of the Chat Reviews and the sub functions which listed under the main components.

Figure 3: Work Breakdown Structure

This is the overall system diagram of chat reviews. As the above diagram, system contains four components. Each component has sub functions to facilitate the user to calculate trustworthiness.

2.2.2 System Overview Diagram

Bellow diagram shows the quick overview of the system and how it happen when message is come to the system.

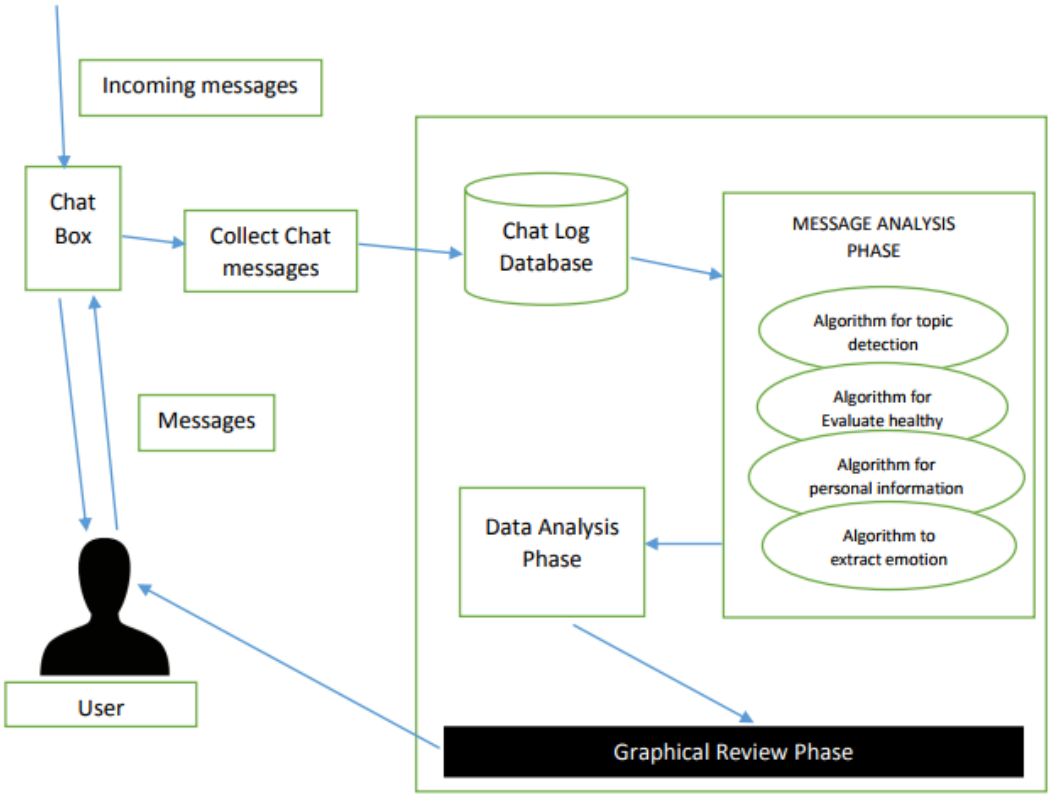

Figure 4 : System Overview Diagram

2.2.3 Use Case Diagram

The use case diagram describes the externally visible behavior of personal information sharing and it also describes the relationships among the user and system. Therefore it is easy for developers to identify the requirements.

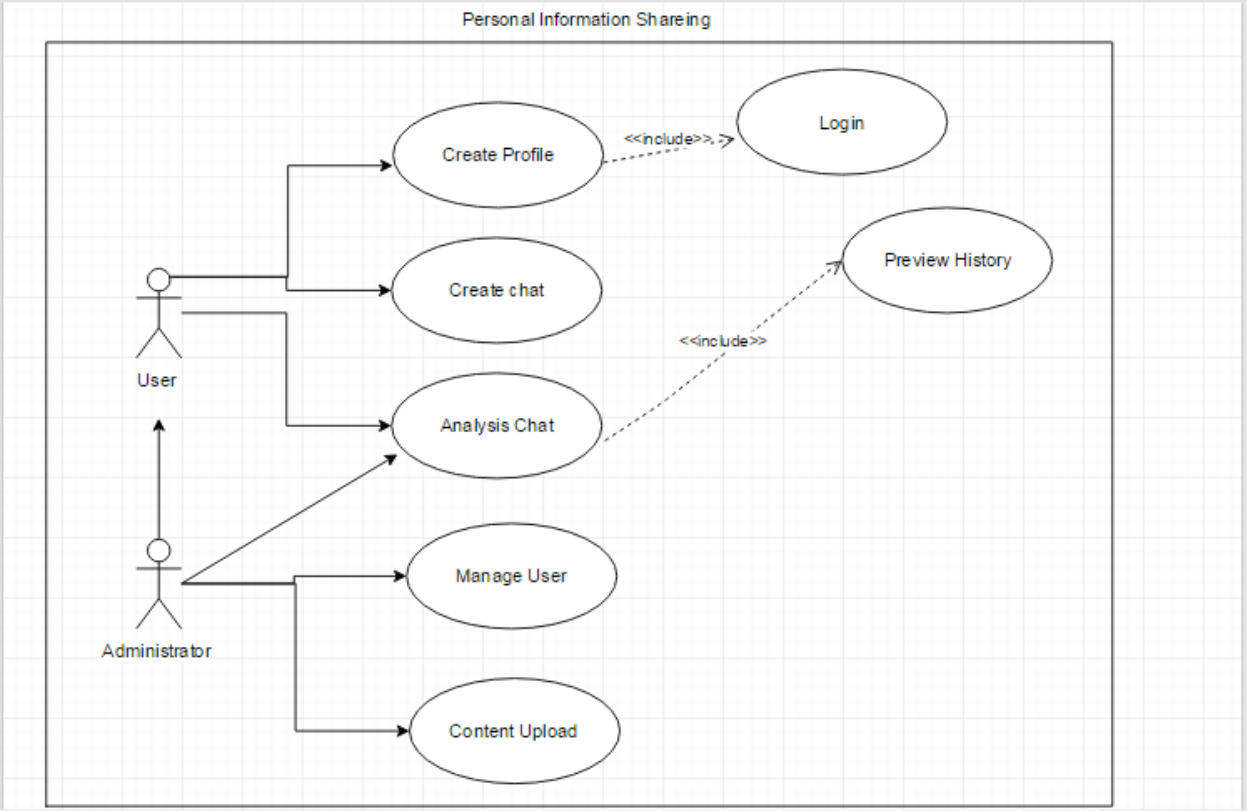

Figure 5 : Use case Diagram

This use-case diagram contains all the action that can be performed by the normal/deaf user. Accesses the system considers as a precondition since user needs to access the system before perform any action. And also how the personal information function going on.

2.2.4 Class Diagram

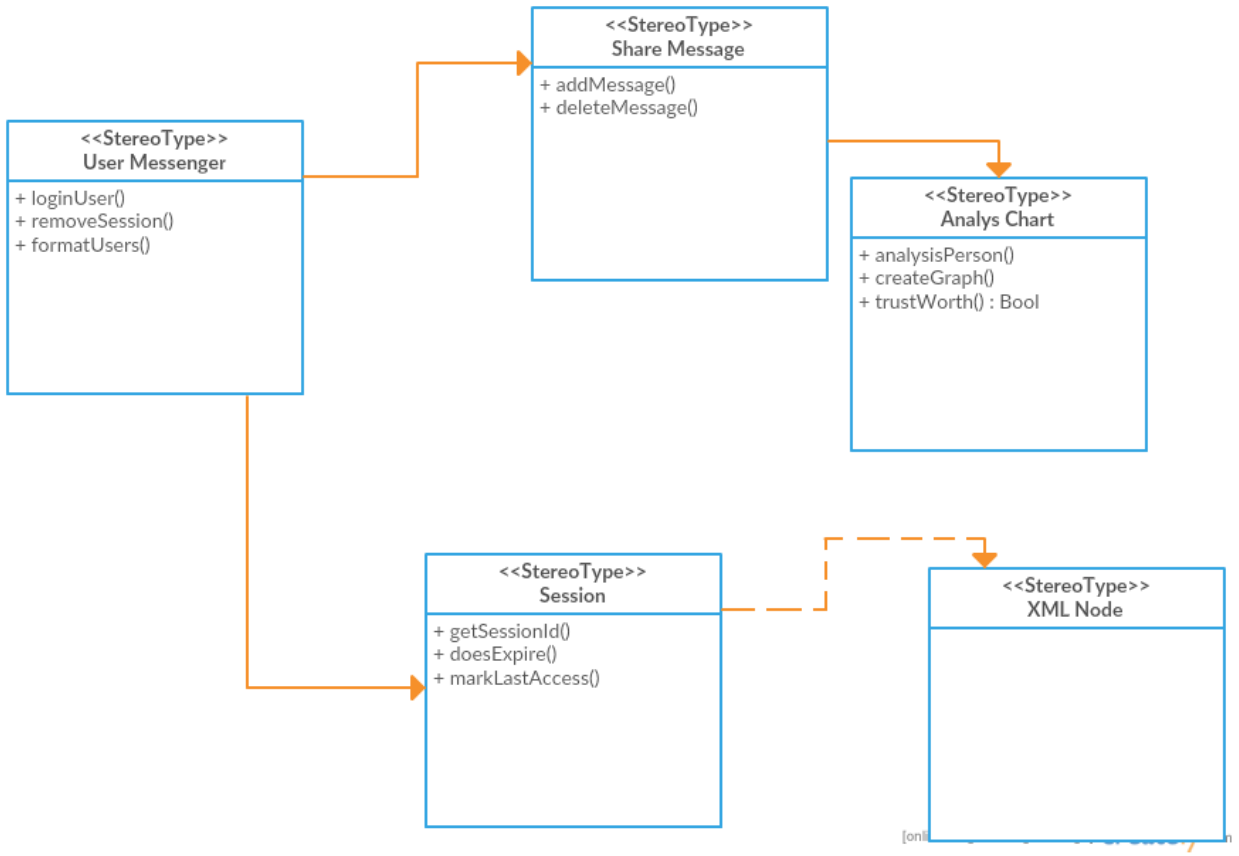

Figure 6 : Class Diagram

2.2.5. Requirement Gathering

We develop our system by mainly targeting Social media users. So this system should be user friendly, efficient and also should have more accurate text recognition system and also system must be reliable because it's also works as a background application. So that it should work for specifications of the all platforms of web. So we need to identify and declare how we are going to do and what are the necessary resources for that. Also we have to find out how the existing systems carried out and what the faults of those systems are. Throughout a good research we identified the suitable methodologies to develop a solution for the current problems. Requirements could be gathered by, Studying on researches about machine learning, neural

network and data mining to come up with a new best solution. Referring online materials such as tutorials, research papers, web articles, programming code and e-books related to do the optimization. Go through books and research articles regarding develop.

2.2.6 Overall Design

Real Time Chat Message Analysis (Chat Reviews) is mainly focused on giving an opportunity to social media users to identify partners. Now a days there have lots of social media sites and there have lots of fake account holders. We propose the system to identify real account holders and trustworthiness. We can reduce fraud activities, then we can make secure site for users. This proposed method will come up with successful chat in social media chat sessions.

Benefits

- Identify true account holders.
- Users can trust each other.
- Risk of the system will goes down.
- More reliable among users.
- Make the data to more secure.

Description of the features and the techniques used in Chat Reviews

Login interface

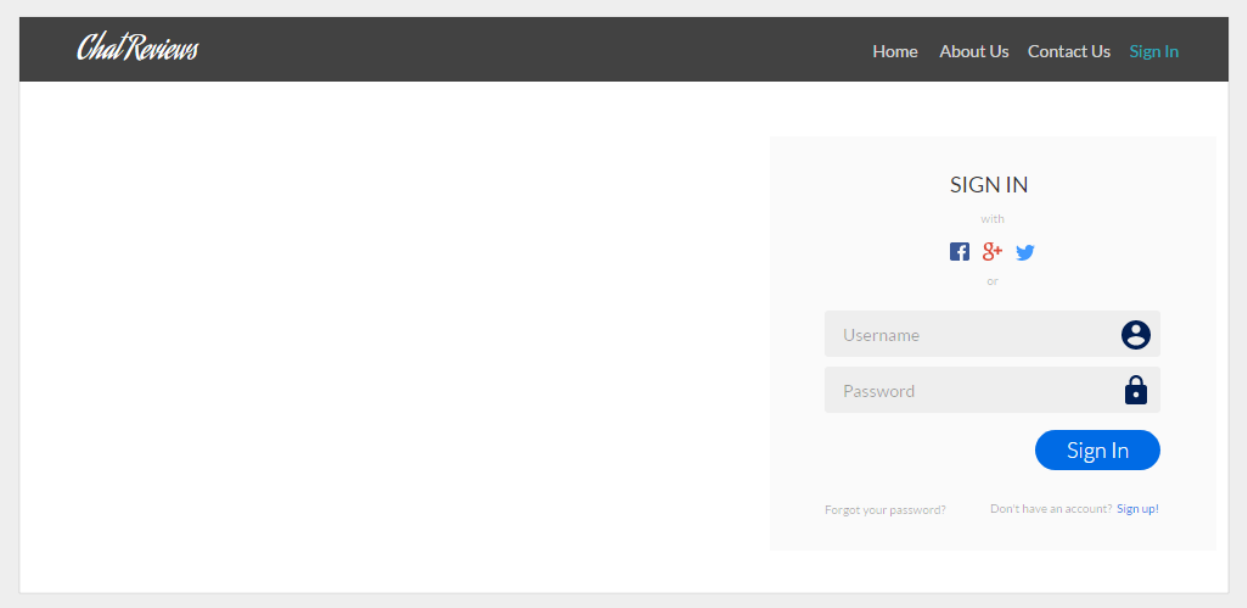

Figure 7: Login Page

The above screen is displayed as soon as the user selects the start button on the initial interface of the application. This interface provides user the facility of selecting the user mode he or she prefers. Administration mode is for the maintainers to do the modification of the database. For all the users who willing to do analysis should log to in to the system. After successfully proceed with this interface, the user of the system is directed to the Main Menu Interface.

Main Menu Interface

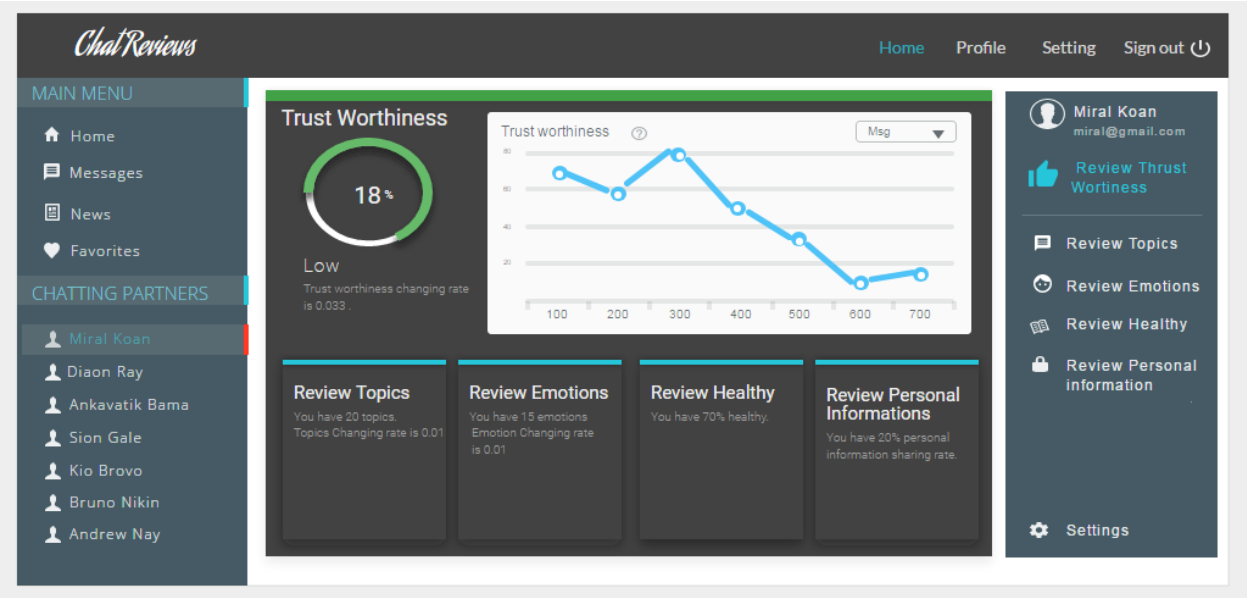

Figure 8: Main Menu / Dashboard

The above screen is displayed as soon as the user logged in to the system. This interface provides the navigation buttons for the main four components of the web app. Each of these buttons will directs to the named functions on the buttons.

Personal Information Sharing

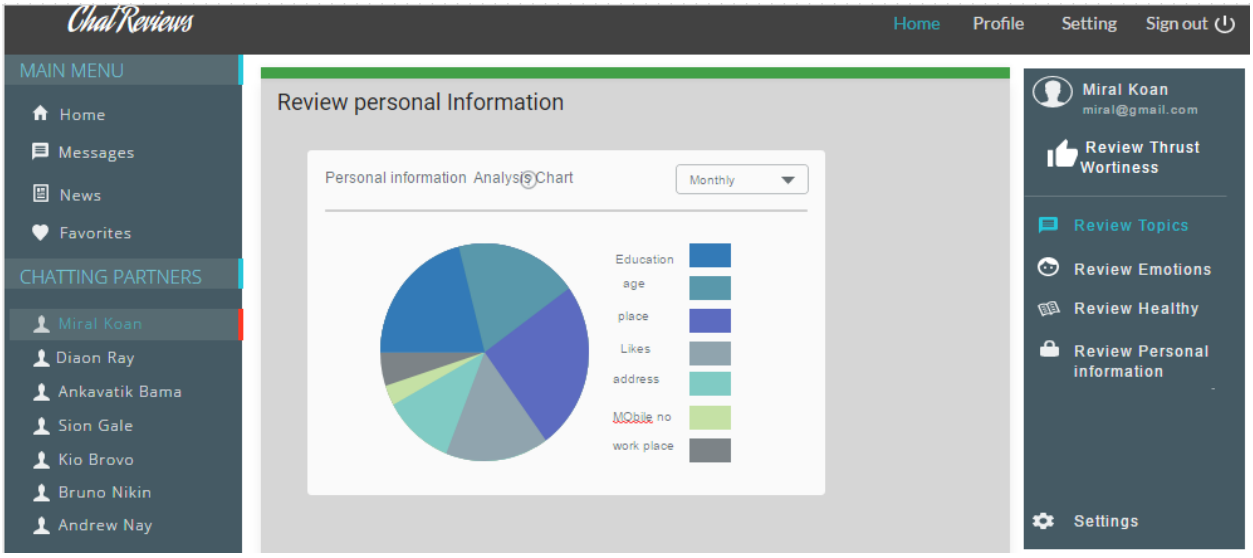

Figure 9: Personal Information Sharing

The above screen is displayed analysis details of personal information sharing . This interface provides the navigation buttons for the main menu as well. So then can analysis other analysis as well. Each of these buttons will directs to the named functions on the buttons. The system displays graphical details then the user can easily get idea of analysis.

This application contains GUI as the front end, Message and other resources database file as the backend and a layer to connect database and the GUI as the logic layer.

Three tier architecture can be further explained as follows.

1. **Data tier** :- Database access is define in this layer. This contain static classes and it act as an interface between the text database and the business tier.

2. **Business Tier** :- This layer define the structure of the storing data. This is the core module of this application.

3. **Presentation Tier** :- This tier consists of the GUIs that have direct access to the business tier and has nothing to do with the data tier. The code is placed in the Deaf Dictionary Form.

2.2.7 Data Training

We create models to train data to our use. We use algorithm to do this and find better solutions to train data. We have to create data for train because there have no such data for personal information. We need to create more than 15000 data for training. The model is always train data to give a better output from the model.

```
Indexing events using cutoff of 1

    Computing event counts...  done. 1392 events
    Indexing...  done.
Collecting events... Done indexing.
Incorporating indexed data for training...
done.
    Number of Event Tokens: 1392
        Number of Outcomes: 3
      Number of Predicates: 9268
Computing model parameters...
Performing 70 iterations.
  1:  . (1358/1392) 0.9755747126436781
  2:  . (1387/1392) 0.9964080459770115
  3:  . (1390/1392) 0.9985632183908046
  4:  . (1392/1392) 1.0
  5:  . (1392/1392) 1.0
  6:  . (1392/1392) 1.0
  7:  . (1392/1392) 1.0
Stopping: change in training set accuracy less than 1.0E-5
Stats: (1392/1392) 1.0
...done.
Compressed 9268 parameters to 428
4 outcome patterns
Finding types in the test sentence..
person : Alisa Fernandes      [probability=0.6643846020606172]
```

Figure 10: Sample Data Train Output

2.2.8 Testing and Implementation

Software testing can be defined as the verification of the program's behavior on a finite number of test cases, selected appropriately from infinite executions. Simply software testing is finding out of bugs by executing a program. Software testing is not a single activity that is a process which takes place throughout the SDLC. We can reduce the bugs being introduced to the program by planning test cases early in the SDLC . Software testing is one of the most important areas since we all make mistakes. Some are minor mistakes and some are not. Checking each and every part thoroughly is a must before handing over to the customer. It's considered as a best practice to hire someone else to check the bugs since he/she may spot the points that the developers have missed. The errors that have been made by developer can be identified by software testing. Since customer is the most important role in the SDLC it is very important to test the product and check whether it is up to the customer's requirement. Maintaining the quality of the product is very important since that builds the confident of the customer. To maintain an effective performance of the system testing is very much important. Maintaining cost will be very much higher than the development cost, so it is advised to minimize the failures since correcting those failures will be very expensive.

In a project performance, reliability, security, usability, availability, portability, quality, platform compatibility, efficiency, stability, supportability should be tested well before it's delivered to the customer.

Hardware interfaces

This system does not need any external hardware interfaces.

Software interfaces

- Net Beans / Java Use to build the system interfaces and backend of the system.
- MS SQL Server 2010 Use to handle the database functionalities of the system.

Communication interfaces

This system need to analysis real time chat. Therefore it need to connect to internet to access chat history. So then it need to have a modem to have access internet. Private or public wireless connection internet access is also possible to access internet.

Memory constraints

- web space will be needed for store server components.
- Pentium IV 3.0GHz processor.
- The system should be capable of running with a minimum of 1GB of RAM. However higher the amount of memory the better.

Operations

People who can access internet privileged for following operations.

- Chat with person.
- Analysis the trustworthiness of other person that we need to communicate.
- Analysis personal information of the person chat with you.
- In order to perform any task in chat system first of all users have to login to the system.
- By clicking on the tabs or the links users can navigate through the pages.
- Profiles should create automatically when the user send a request or log to the web application.
- When the customer sends the request system should able to categorize it based on the Natural language processing.
- All the data in the database should backup for eliminate the risk of data losses.

Site adaptation requirements

There has to be a chat history because the information and their results are really important and that information must available in all the time. Therefore system needs a history in running. Server has to be stored data for data protection.

User interfaces must be in English since the user can be of any nationality. provide validity of the chat clearly. We use English as the only language to respond the chat and analysis process. For use this product need to have some more English knowledge.

2.3. Research Findings

- There were applications and systems for Chat message analysis before the proposed chat reviews system come to the market. But those applications are mostly areas of chat. There has not been any application or system to give all the features such as personal information rate like that.
- Though prevailing systems are only focused on message encryption and topic detections, the proposed system gives all the outputs which are mentioned above at once and more. That is one of the main research finding in proposed system.
- On time acting is a common feature that all the prevailing applications and systems have. But they can't analysis more areas in one system.
- Though there are some applications which are capable of analysis, our system will be the first application which is going to analysis 4 different areas in one application.

3.0 RESULTS & DISCUSSION

3.1 Testing

The purpose of testing is to identify and correct bugs in the developed project. Testing is vital to the success of the system. Software testing can be stated as the process of validating and verifying system. There are several levels of testing. They are Unit testing, System testing and Acceptance Testing. In our system, testing will supports to check whether product goals are achieved or not. Some of product goals are,

- meets the requirements that guided its design and development
- works as expected
- can be implemented with the same characteristics
- And satisfies the needs of user requirements.

Unit testing

Test each unit (method) individually. Unit testing needs to be carried out for each part divisions individually. It means on receiving system design documents, the work is divided in modules/units and actual coding is started. The benefits of unit testing are improving the design without breaking, test-first reduces the cost of bugs, make better designs etc.

System Testing

Test the overall functionality of the system. System testing is performed on the entire system. System testing tests not only the design, but also the behavior of the user. In system testing, includes different types of tests. There are load testing, performance testing, stress testing etc.

Acceptance Testing

The system is given to the client and gets their approval. If user accepts with the system and satisfies as per their demands and requirements, the process is completed.

Test Case ID	1
Test Case	Getting the chat history
Test Input Data	Previous chat history
Expected Output	Store messages in database
Actual Output	Store messages in database
Pass or Fail	Pass

Table 2: Test case 1

Test Case ID	2
Test Case	Retrieve chat from database
Test Input Data	Chat messages
Expected Output	Message collection
Actual Output	Message collection
Pass or Fail	Pass

Table 3: Test case 1

Test Case ID	3
Test Case	Put messages to analysis
Test Input Data	Message collection retrieve from database
Expected Output	Graph of personal information share
Actual Output	Graph of personal information share
Pass or Fail	Pass

Table 4: Test case 1

Test Case ID	4
Test Case	Analysis trustworthiness
Test Input Data	Overall output of all 4 module
Expected Output	Trustworthiness graph
Actual Output	Trustworthiness graph
Pass or Fail	Pass

Table 5: Test case 1

3.2 Quality Expectations

The proposed system is focused on supporting following aspects within it. These are the software quality related objectives and the value of the final product is based on these objectives. At the end of this report we provide our evaluation on this project based on the level to which we could satisfy these qualities.

Reliability

The reliability is an attribute of any mobile-related component that consistently performs according to its specifications. This is the most important attribute of a software since the entire system process depends on that. The proposed application will be implemented to provide a reliable and efficient service to the users thus there will be a low rate of any hardware or software failures.

Availability

Chat Reviews system should be acting for use 24 x 7 and should support concurrent access. Because users are use online social networks any time.

Maintainability

Since the application is built as separate modules, maintaining each module can be done individually. The updates for the system can be easily released through our system website, so that the users can simply update their applications. Since the knowledge base of the application is stored in a centralized location, the maintainability will be much more easy and simple.

3.3 Evidence (Results)

We test our system using the Face Book to testing and getting final result. Also after the implemented the system we tested our system to real time scenarios as well with manual messages input. We testing our system to the functional wise to test cases and we get result implementing our system as a reliable system. beginning after the using tests we identify what are the drawbacks and bugs have in our system and then implementing and fix them successfully. For detecting high spam messages and incomplete messages we develop module and use library. But after done several testing we analyze real case scenarios. we After the end we test our whole system with the test scenarios and real world scenarios as well, finally we come up with a successfully working chat review system. With test cases system identify topics, healthy of messages , emotions and personal information successfully.

3.4 Discussion

- When users chat it store to database for analysis purpose.
- To improve the user friendliness, we add attractive colors, fonts and widgets .
- Limitations - Currently the application is working only for face book. Because we get messages from face book chat sessions.
- Use eye catching buttons to easy of use.
- Graphical output for easy identification.
- The systems will not use pre requirement because it run on web.

There are two types of users involve with the system.

Administrator: Administrator is the person who has the full authority of the system. our application and any Facebook user will be the main users of this subsystem. They are the people with highly experienced in the field of business environment in Information technology and

social networking field. Administrator mainly assigns agents for the system, Display dashboard results, add knowledge articles to the system, do the customer request categorization whenever NLP unable to do this task by pattern

Users : people who are interact with chat through social media. This is easy to analysis other user details by having less knowledge.

User type	Description
Account Holder	Person who is directly involve with the chat.
Administrator	Person who is doing experiments and analyzing for efficient usage.

Table 6: User characteristics

4.0 CONCLUSION

This system is a methodological analysis of social media services like Facebook for a chat system with the use of techniques and technologies like Natural Language Processing (NLP) techniques to enable the organization to engage with customers in order to provide mutually beneficial value in a trusted and transparent business environment. There are also many social chat systems in today's business but the problem is there are lots of impacts and they are not correctly use focusing on business to gain more benefits through social media by overcoming the difficulties of traditional chats. Social network analysis emerged as an important research topic in sociology decades ago.

And system is working for every platform. Because of that our proposed "Chat Reviews" is need to be the best system in the market.

5.0 References

[1] https://opennlp.apache.org/

[2] Topic Detection and tracking. Retrieved 23 March 2017, from Topic Detection and tracking IEEE: http://ieeexplore.ieee.org/

[3]Sampson, G.: How fully does a machine-usable dictionary cover english text. Literary and Linguistic Computing 4(1) (1989)

[4]Nadeau, D.: Semi-supervised Named Entity Recognition: Learning to Recognize 100 Entity Types with Little Supervision. PhD thesis, Ottawa, Ont., Canada, Canada, AAINR49385 (2007)

[5]Nadeau, D., Sekine, S.: A survey of named entity recognition and classification. Linguisticae Investigationes 30(1), 3–26 (2007)

[6]Zhou, G., Su, J.: Named entity recognition using an hmm-based chunk tagger. In: Proceedings of ACL, pp. 473–480 (2002)

[7]Curran, J.R., Clark, S.: Language independent ner using a maximum entropy tagger. In: Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003, vol. 4, pp. 164–167 (2003)

[8]Finkel, J.R., Grenager, T., Manning, C.: Incorporating non-local information into information extraction systems by gibbs sampling. In: ACL, pp. 363–370 (2005)

[9] C. Grier, K. Thomas, V. Paxson, and M. Zhang, "@ spam: the underground on 140 characters or less," in CCS, 2010.

[10] M. McCord and M. Chuah, "Spam detection on twitter using traditional classifiers," in Autonomic and trusted computing. Springer, 2011, pp. 175–186.

[11]K. Tzeras and S. Hartman, "Automatic indexing based on Bayesian inference networks". In: 16th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'93), pp. 22-34, 1993.

[12]Mattshomepage.com, 'Bernoulli Naive Bayes Classifier', 2014. [Online]. Available: http://mattshomepage.com/articles/2016/Jun/07/bernoulli_nb/. [Accessed: 20- July- 2017].

[13]Wikipedia.org, 'Supervied learning overview', 2014. [Online]. Available: https://en.wikipedia.org/wiki/Supervised_learning. [Accessed: 19- July- 2017].