A PROJECT ON COMMENT SENTIMENTAL ANALYSIS WITH N-GRAM METHODOLOGY



A Project submitted to the Islamic University in partial fulfillment of the requirements for the degree of B.Sc. in Computer Science & Engineering.

Supervised By

Joyassree Sen Assistant Professor Department of Computer Science & Engineering, Islamic University, Kushtia – 7003, Bangladesh

<u>Submitted By</u> M.M Mamuduzzaman

Roll No: 1414026

Reg No:1145

Session: 2014-2015

Department of Computer Science & Engineering

Islamic University, Kushtia – 7003, Bangladesh

Department of Computer Science & Engineering

CERTIFICATE OF THE SUPERVISOR

I am pleased to certify that **M.M Mamuduzzaman**, examination **Roll No.1414026**, has performed a project work title "**COMMENT SENTIMENTAL ANALYSIS WITH N-GRAM METHODOLOGY**

." under my supervision in academic year **2014-2015**, submitted in partial fulfillment of the requirements for the degree of Bachelor of Science. So far as I concern this an original project work that the carried out for fourth years in the department of Computer Science & Engineering, Islamic University, Kushtia, Bangladesh.

I strongly declare that this dissertation has not been copied from any other project or submitted to elsewhere prior submission to this department.

Signature
Date
Joyassree Sen
Project Supervisor
Assistant Professor
Department of Computer Science & Engineering
Islamic University, Kushtia, Bangladesh.

Dedicated To my parents

teacher

ACKNOWLEDGEMENTS

At first, I remember my almighty God for giving me the opportunity and strength to carry on and complete this project work successfully.

I express my sincere respect, thanks and gratitude to my project supervisor **Bappa Sarkar**, Assistant Professor, Department of Computer Science & Engineering, Islamic University, Kushtia, for his continuous supervision, encouragement and constant guidance, indefatigable care and keen interest at every steps of the work. Really his expanded helping hands enable me to complete the project. I am grateful to him.

I would like to express my intellectual debts of gratitude to the numerous authors, to all my classmates, well-wishers, elder brothers who give me confidence and instructions.

Finally, I would like to express my parents whose continued sacrifice, support and inspiration over the years were inevitable for the higher education of me and above all thanks to God.

Date: 3/15/2020 Author

ABSTRACT

Sentiment analysis has been studied in wide area of domain such as movie review, teaching review, product review, e-learning, hotel review and many more. Most scholars focused to quantitative data analysis. However, some studies have been done on qualitative data using sentiment analysis .Sentimental analysis is the interpretation and classification of emotions (positive, negative and neutral) within text data using text analysis techniques. Sentiment analysis allows businesses to identify customer sentiment toward products, brands or services in online conversations and feedback. Sentiment analysis models detect polarity within a text (e.g. a *positive* or *negative* opinion), whether it's a whole document, paragraph, sentence, or clause.

CONTENTS

Chapter 1

Introduction	
1.1 Purpose	07
1.2 Why sentiment analysis is needed	08
Chapter 2 Overall Description	
2.1 What are N-Grams	10
2.2 How many n-gram in a sentence	10
2.3 The code block for n-gram generation	11
2.4 N-gram data model	12
Chapter 3	
Specific requirements	
3.1 Use case report	13
3.2 Types of sentiment analysis	14
3.3 How does sentiment analysis work	14
Chapter 4	
Important code and figure	
4.1 The extraction of comment information and sentiment analysis	16
4.2 Code for positive feedback	17
4.3 Code for positive feedback	17
4.4 Percentage count	18
4.5 Show percentage	18
4.6 Technology overview	18
4.7 Project description	24
Chapter 5	
Design & development	
5.1 Snapshots	28
5.2 Features	37
Chapter 6	
Conclusion	38
References	38

Chapter -1 SOFTWARE REQUIRMENT SPECIFICATION

Introduction

1.1Purpose:

What is sentiment analysis? If we take your customer feedback as an example, sentiment analysis (a form of <u>text analytics</u>) measures the attitude of the customer towards the aspects of a service or product which they describe in text.

The task of sentiment analysis typically involves taking a piece of text, whether it's a sentence, a comment or an entire document and returning a "score" that measures how positive or negative the text is.

For example in customer feedback:

Customer Feedback Text	Sentiment
"This café is great, the staff are really friendly and the coffee is delicious"	Positive
"I would not recommend this café to anyone. Their coffee is terrible and is really expensive"	Negative

1.2 Why sentiment analysis is needed:

In today's environment where we're justifiably suffering from data overload (although this does not mean better or deeper insights), companies might have mountains of customer feedback collected; but for mere humans, it's still impossible to analyze it manually without any sort of error or bias.

Oftentimes, companies with the best intentions find themselves in an *insights vacuum*. You know you need insights to inform your decision making and you know that you're lacking them, but don't know *how* best to get them.

Sentiment analysis provides some answers into what the most important issues are, from the perspective of customers, at least. Because sentiment analysis can be automated, decisions can be made based on a significant amount of data rather than

plain intuition that isn't always right.

Imagine this scenario: you're the owner of a small delivery business and you receive about 20

responses to your email surveys every month. You could (and should), read these yourself

and perform your own analysis by hand.

Now, imagine receiving 30,000 responses per month (as many large enterprises do). Whoa,

that's more than a thousand responses that you would need to read and analyze each day.

This is why we need **Sentiment analysis**.

Sentiment analysis is useful for quickly gaining insights using large volumes of text data. In

addition to the customer feedback analysis use case, which we touched on above, here are

another two examples of where sentiment analysis can be useful.

Platform

Operating Systems: Microsoft Windows

Technologies Used:

Front End: HTML, CSS and Javascript

Web language: PHP

RDBMS(Back end): MySQL

Software Requirements:

☑ PHP 5.0

APACHE HTTP Server

Dreamweaver , FrontPage for FrontEnd Programming

Microsoft Windows or Linux

Hardware Requirements:

Intel Pentium IV processor or equivalent or higher

512 MB Ram or Higher

8

20 GB HDD or Higher

?

Technologies:

- PHP.
- MYSQL
- JAVASCRIPT
- HTML
- CSS

Overview:

The rest of this SRS is organized as follows:

Section 2 gives an overall description of the software. It gives what level of proficiency is expected of the user, some general constraints while making the software.

Section 3 gives specific requirements which the software is expected to deliver. Some performance requirements and constraints are also given and deal with other Non-Functional Requirements.

Section 4 deals with External Interface Requirements like Hardware and Software Interface.

Chapter -2

OVERALL DESCRIPTION

2.1 What are N-Grams:

N-grams of texts are extensively used in text mining and natural language processing tasks. They are basically a set of co-occurring words within a given window and when computing the n-grams you typically move one word forward (although you can move X words forward in more advanced scenarios). For example, for the sentence "The cow jumps over the moon". If N=2 (known as bigrams), then the ngrams would be:.

- the cow
- cow jumps
- jumps over
- over the
- the moon

2.2 N-Gram Function:

So you have 5 n-grams in this case. Notice that we moved from the->cow to cow->jumps to jumps->over, etc, essentially moving one word forward to generate the next bigram. If N=3, the n-grams would be:

- the cow jumps
- cow jumps over
- jumps over the
- over the moon

So you have 4 n-grams in this case. When N=1, this is referred to as **unigrams** and this is essentially the individual words in a sentence. When N=2, this is called **bigrams** and when N=3 this is called **trigrams**. When N>3 this is usually referred to as four grams or five grams and so on.

2.3 How many N-Grams in a sentence :

If X=Num of words in a given sentence K, the number of n-grams for sentence K would be::

$$Ngrams_K = X - (N-1)$$

2.4 The code block for N-gram generation:

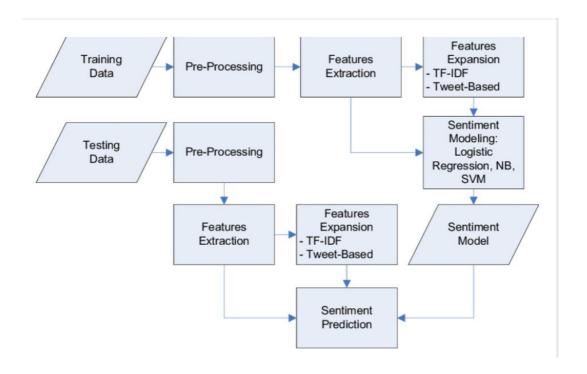
2.4 N-Gram data MODEL:



Chapter -3

Specific Requirements:

3.1 Use Case Reports:



3.2 Types of sentiment analysis:

Sentiment analysis assumes various forms, from models that focus on polarity (positive, negative, neutral) to those that detect feelings and emotions (angry, happy, sad, etc), or even models that identify intentions (e.g. *interested* v. *not interested*). Here are some of the most popular types of sentiment analysis:

Fine-grained Sentiment Analysis

If polarity precision is important to your business, you might consider expanding your polarity categories to include:

- Very positive
- Positive
- Neutral
- Negative
- Very negative

This is usually referred to as fine-grained sentiment analysis, and could be used to interpret 5-star ratings in a review, for example:

- Very Positive = 5 stars
- Very Negative = 1 star

Emotion detection

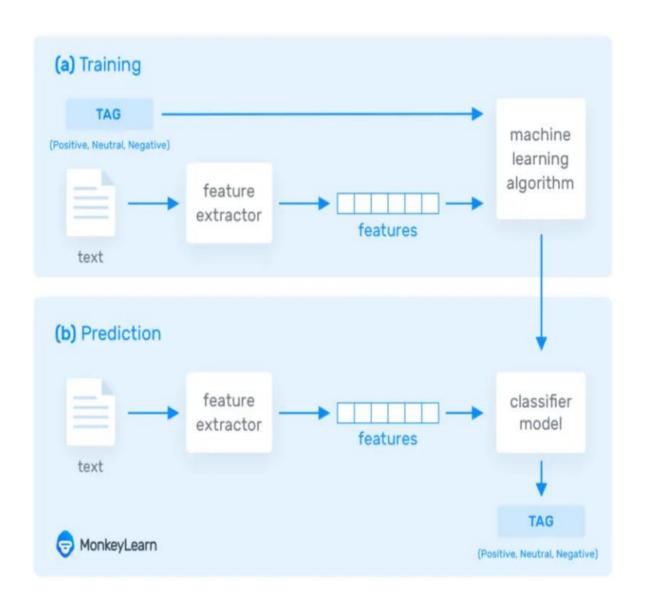
Emotion detection aims at detecting emotions, like happiness, frustration, anger, sadness, and so on. Many emotion detection systems use lexicons (i.e. lists of words and the emotions they convey) or complex machine learning algorithms.

One of the downsides of using lexicons is that people express emotions in different ways. Some words that typically express anger, like *bad* or *kill* (e.g. *your product is so bad* or *your customer support is killing me*) might also express happiness (e.g. *this is bad ass* or *you are killing it*).

Aspect-based Sentiment Analysis

Usually, when analyzing sentiments of texts, let's say product reviews, you'll want to know which particular aspects or features people are mentioning in a positive, neutral, or negative way. That's where <u>aspect-based sentiment analysis</u> can help, for example in this text: "The battery life of this camera is too short", an aspect-based classifier would be able to determine that the sentence expresses a negative opinion about the feature battery life.

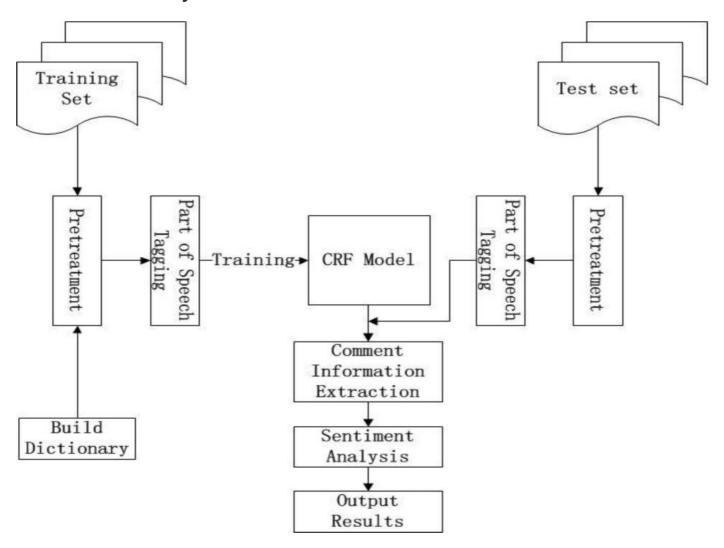
3.3 How does sentiment analysis work



Chapter-4

Important code and figure

4.1 The Extraction of Comment Information and Sentiment Analysis:



4.2 Code for positive feedback

```
$query="SELECT COUNT(table_name1.word) as good_word from table_name1,good where table_name1.word LIKE CONCAT('%',good.word
$'\kappa'';
$run=mysqli_query($connection,$query);
$result=mysqli_fetch_array($run);
 $positive=$result['good_word'];
  $query="SELECT COUNT(table_name2.word) as very_good_word from table_name2, very_good where table_name2.word LIKE CONCAT('
     %',very_good.word,'%') ";
 $run=mysqli_query($connection,$query);
 $result=mysqli_fetch_array($run);
  $positive_very=$result['very_good_word'];
 //for VERY_very_good postive feedback
| $query="SELECT COUNT(table_name3.word) as good_word from table_name3,very_very_good where table_name3.word LIKE CONCAT('
     %',very_very_good.word,'%') ";
$run=mysqli_query($connection,$query);
$result=mysqli_fetch_array($run);
  $positive_very_very=$result['good_word'];
   echo $positive=(($positive-$positive_very-$positive_very)+($positive_very-$positive_very_very)*1.25+($
       positive_very_very*1.5));
```

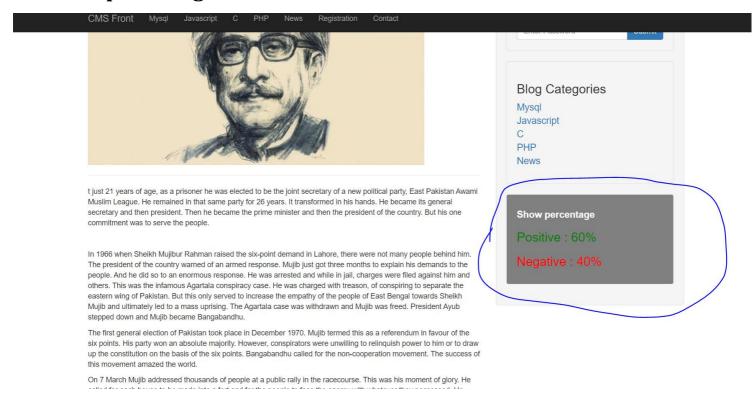
4.3 Code for negative feedback

```
$query="SELECT COUNT(table_name1.word) as bad_word from table_name1,bad where table_name1.word LIKE CONCAT('%',bad.word,
     '%') ";
$run=mysqli_query($connection,$query);
$result=mysqli_fetch_array($run);
$negative=$result['bad_word'];
//for very_bad negative feedback
$query="SELECT COUNT(table_name2.word) as bad_word from table_name2,very_bad where table_name2.word LIKE CONCAT('%',
    very_bad.word,'%') ";
$run=mysqli_query($connection,$query);
$result=mysqli_fetch_array($run);
$negative_very=$result['bad_word'];
$query="SELECT COUNT(table_name3.word) as bad_word from table_name3.very_very_bad where table_name3.word LIKE CONCAT('%'
     ,very_very_bad.word,'%') ";
$run=mysqli_query($connection,$query);
$result=mysqli_fetch_array($run);
$negative_very_very=$result['bad_word'];
echo "<br>";
echo $negative= (($negative-$negative_very-$negative_very)+($negative_very-$negative_very_very)*1.25+($
    negative_very_very*1.5));
```

4.4 Percentage count:

```
//percentage count
echo "<br>;
$positive_percentage= round((100*$positive)/($positive+$negative));
$negative_percentage=100-$positive_percentage;
echo "$positive_percentage %<br>";
echo "$negative_percentage %<br>";
```

4.5 Show percentage:



4.6 TECHNOLOGY OVERVIEW

The technology selected for implementing customer Information Management System is PHP/MYSQL.Apache is used as the HTTP server.The development was done in a 'windows' environment using adobe dreamweaver CS5.

PHP

PHP is a general-purpose scripting language that is especially suited to serverside web development where PHP generally runs on a web server.PHP code is embedded into the HTML source document. Any PHP code in a requested file is executed by the PHP runtime, usually to create dynamic web page content. It can also be used for command-line scripting and client-side GUI applications. PHP can be deployed on many web servers and operating systems, and can be used with many relational database management systems (RDBMS). It is available free of charge, and the PHP Group provides the complete source code for users to build, customize and extend for their own use.

MvSOL

MySQL is a relational database management system (RDBMS)^[1] that runs as a server providing multi-user access to a number of databases. MySQL is a popular choice of database for use in web applications and is an open source product. The process of setting up a MySQL database varies from host to host, however we will end up with a database name, a user name and a password. Before using our database, we must create a table. A table is a section of the database for storing related information. Ina table we will set up the different fields which will be used in that table. Creating a table in phpMyAdmin is simple, we just type the name, select the number of fields and click the 'go' button, we will then be taken to a setup screen where you must create the fields for the database. Another way of creating databases and tables in phpMyAdmin is by executing simple SQL statements. We have used this method in order to create our database and tables.

Apache

The Apache HTTP Server is a web server software notable for playing a key role in the initial growth of the World Wide Web. In 2009 it became the first web server software to surpass the 100 million web site milestone. Apache is developed and maintained by an open community of developers under the auspices of the Apache

Software Foundation. Since April 1996 Apache has been the most popular HTTP server software in use. As of November 2010 Apache served over 59.36% of all websites and over 66.56% of the first one million busiest websites.

XAMPP

XAMPP is a small and light Apache distribution containing the most common web development technologies in a single package. Its contents, small size, and portability make it the ideal tool for students developing and testing applications in PHP and MySQL. XAMPP is available as a free download in two specific packages: full and lite. While the full package download provides a wide array of development tools, XAMPP Lite contains the necessary technologies that meet the Ontario Skills Competition standards. The light version is a small package containing Apache HTTP Server, PHP, MySQL, phpMyAdmin, Openssl, and SQLite.

Obtaining and Installing XAMPP

As previously mentioned, XAMPP is a free package available for download and use for various web development tasks. All XAMPP packages and add-ons are distributed through the Apache Friends website at the address: http://www.apachefriends.org/. Once on the website, navigate and find the Windows version of XAMPP and download the self-extracting ZIP archive. After downloading the archive, run and extract its contents into the root path of a hard disk or USB drive. For example, the extract path for a local Windows installation would simply be C:\. If extracted properly we will notice a new xampp directory in the root of your installation disk. In order to test that everything has been installed correctly, first start the Apache HTTP Server by navigating to the xampp directory and clicking on the apache_start.bat batch file.

Next we will test if the server is running correctly by opening an internet browser and typing http://localhost/ into the address bar. If configured correctly, we will be presented with a screen similar to that of the one below.

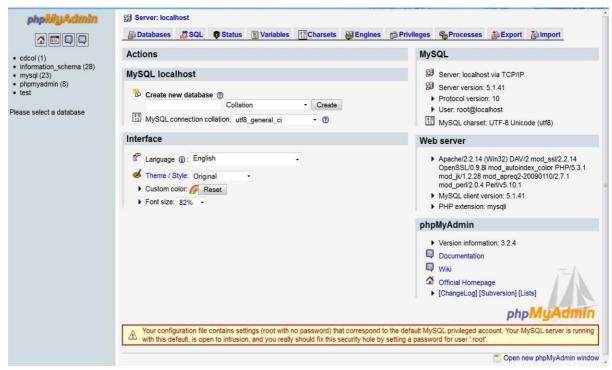


XAMPP splash screen.

In order to stop all Apache processes we do not close the running terminal application, but instead run another batch file in the xampplite directory called apache_stop.bat.

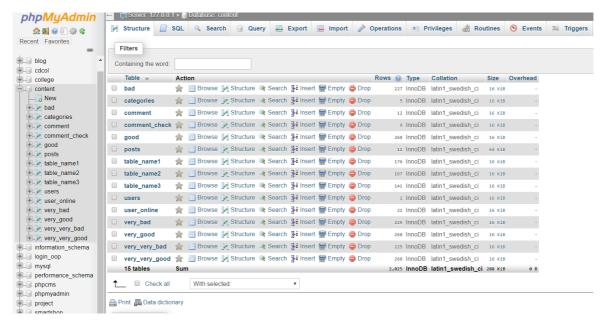
Creating a Database and Inserting Data

Now that we have run and tested Apache and PHP, the next step is running MySQL and creating a database and table which will hold information to be used by our website. In order to start MySQL, navigate to the xampp directory and run the mysql_start.bat batch file. The XAMPP package contains an application called phpMyAdmin which allows developers to administer and maintain MySQL databases. We will be using phpMyAdmin to create a database and table, and enter test data. Before testing phpMyAdmin, make sure that both Apache and MySQL are running by opening their respective batch files: apache_start.bat and mysql_start.bat. Along with Apache and MySQL running in the background, we type http://localhost/phpMyAdmin/ into our web browser. If successful we will be presented with a phpMyAdmin start page similar to the one shown below.



phpMyAdmin start page

The first step with phpMyAdmin running is creating a new database. We create a new database by directly executing SQL statements as shown below. The successful execution of the sql querry creates a database 'content with 15 tables in it. The tabels are bad, categories, comment, comment_check, good, posts, table_name1, table_name2, table_name3, users, user_online, very_bad, very_good, very_very_bad, very_very_good. We also inserted values in the table. The screenshot below shows the successful execution of the query thus creation of a database named customer.



Creation of database in mysql using phpMyadmin

Thus we have learned to create a database in MYSQL by executing sql statements. After creating the database and tables we are now ready to use them in our website "customer Information Management System".

4.7 PROJECT DESCRIPTION

Introduction

Achieving this objective is difficult using a manual system as the information is scattered, can be redundant and collecting relevant information may be very time consuming. All these problems are solved using this project

The directory structure of the project is as follows:

```
OPEN FILES
× functions.php
                                                                                                                                               bal $conn;
                                                                                                                              $array=explode(' ', $variable);
                                                                                                                             $\text{square} \text{square} \text{square} \\
$\text{square} \text{square} \\
$\text{square} \text{square} \\
$\text{square} \text{square} \\
$\text{square} \text{square} \\
$\text{if}(\text{square} \text{square} \\
$\text{square} \\
$\text{
   FOLDERS
    ▼ Em cms
      ▶ 🛅 admin
       ▶ 🛅 css
                                                                                                                                                                 $ng = '';
for($j = $n-1; $j >= 0; $j--) {
    $ng = $ng.$array[$i-$j]." ";
        ▶ m fonts
        ▶ ■ images
          ▼ includes
                     db.php
                     footer.php
                   functions.php
                                                                                                                                                                  $ngrams[] = $ng;
                     header.php login.php
                                                                                                                                 return $ngrams;
                     navigation.php
                     sidebar.php
                     widget.php
                                                                                                           function stored_value_in_database($variable)
          ▶ 🛅 js
               author_post.php
                                                                                                                                       lobal $connection;
or($i=1;$i<=3;$i++)</pre>
               🗅 category.php
              contact.php index.php
                                                                                                                                                $table="table_name".$i;
               index_old.php
                                                                                                                                               $array=getNgrams($variable,$i);
                                                                                                                                                                                 ($array as $value) {
lery="INSERT INTO $table(word) values('$value')";
               registration.php
                                                                                                                                                                $query="INSERT INTO $table(word) value
$run=mysqli_query($connection,$query);
               search.php
                                                                                                                                                  echo mysqli_error($connection);
                                                                                                            function empty_table()
                                                                                                                                 global $connection;
```

Description of root directory contents

- Images Directory: This directory contains the images uploaded by the user during registration process. Supported formats are the .jpg and .gif files.
- Resume Directory: This Directory Contains resumes of students uploaded during registration process of user. Files in this folder can be of .doc,.txt or .pdfformat.
- Admin_Edit_user_Info.php: Admin page for editing information of a user .The Authority can change details of a content in this page.Though facility of changing the image and resume are not yet provided but will be provided in future versions of the project.
 - Admin/includes/add_post.php :Add post from the admin panel.
 - Admin/includes/add_user.php: Add user from the admin panel.
 - Admin/includes/admin_footer.php :Admin page footer .
 - Admin/includes/admin_header.php :Admin page header .
 - ☑ Admin/includes/admin_navigation.php : Admin page navigation bar.
 - Admin/includes/edit_post.php :Edit the published post from the admin panel.
 - Admin/includes/edit_user.php: Edit the user information who are subscribe the site.
 - Admin/includes/update_categories.php: Update the category of the blog system
 - Admin/includes/view_all_comments.php: View all comments from the admin page.
 - Admin/includes/ view_all_post.php: View all posts from the admin page.
 - Admin/includes/ view_all_user.php: view all users from admin page.
 - **Footer.php**: Footer file for all pages.
 - **Header.php**: Header file for login page and homepage of the site.

- **categories.php** : Show all categories in H
- **author_post.php**: show the post by author.
- **Search.php**: Search page to search posts.

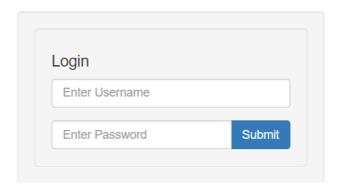
?

- **Search_Result.php** :Show the contact from .
- **post.php**: Show the single post of the website.
- **registration.php**: Open the registration form.
- **Includes/db.php**: Make connection to the database.
- includes/functions.php Add all the require function to keep the index page neat and clean.
- includes/login.php: Login Page for user/admin login. Appropriate message is displayed if the login is unsuccessful.
- includes/logout.php: logout from the website.
- **Includes/sidebar.php**: Show the sidebar in the homepage.
- **Includes/navigation.php**: Show the navigation bar in the homepage.
- **Style.css**: Stylesheet for the whole site design .
- **Validation.js**: Javascript validations used for validation of form values.

Various form entries are validated at the client side using this file onl

Description of database tables

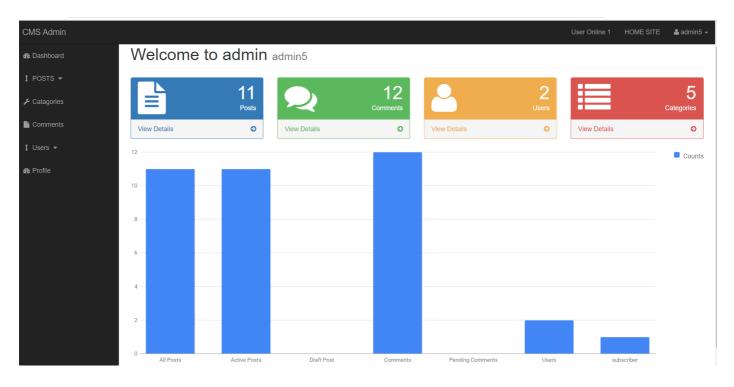
② admin_login:



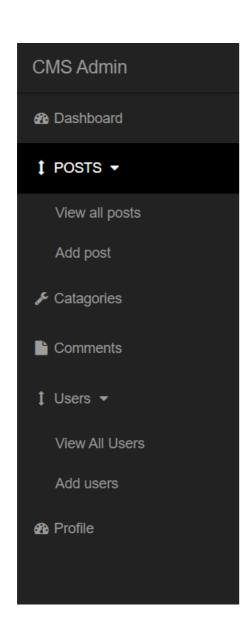
o user_name: Stores user name of Authority (s). o password : Stores password of the Authority (s). o last_login_date: Stores the last login date of the Authority (s).

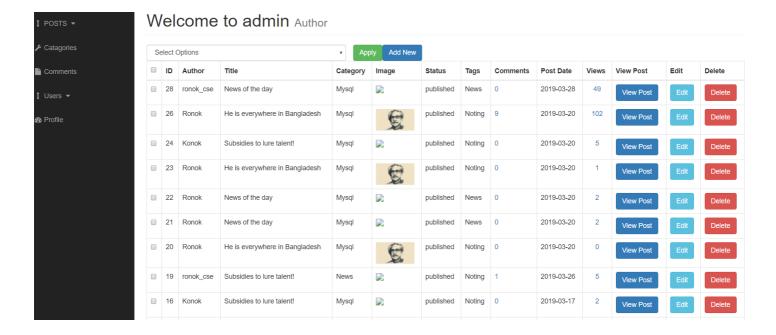
Chapter -5 Design & development

5.1 SNAPSHOTS

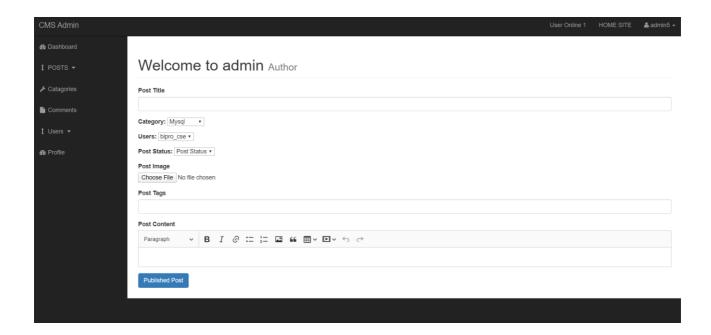


This is the admin dashboard.

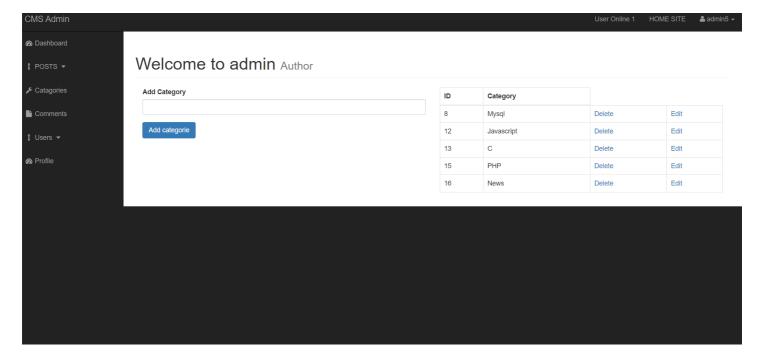




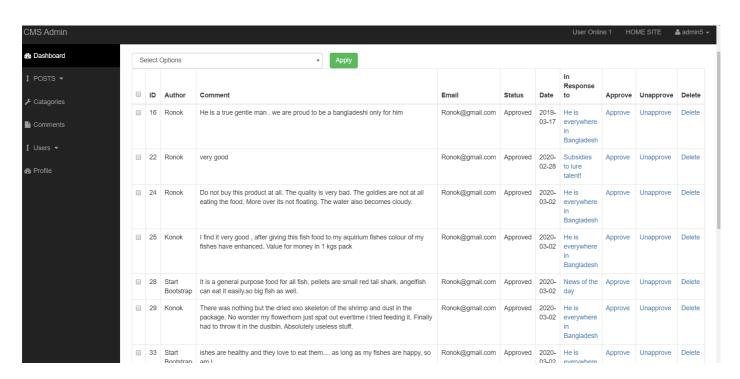
This is the admin view all post page.



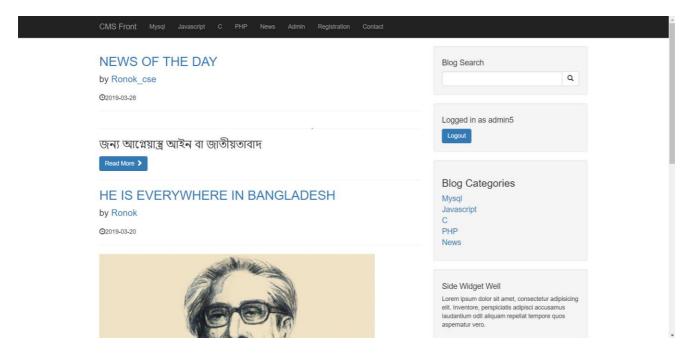
This is the admin add post page.



This is the admin add category page.



This is the addmin show all comments list page.



This is the Homepage of the Blog.

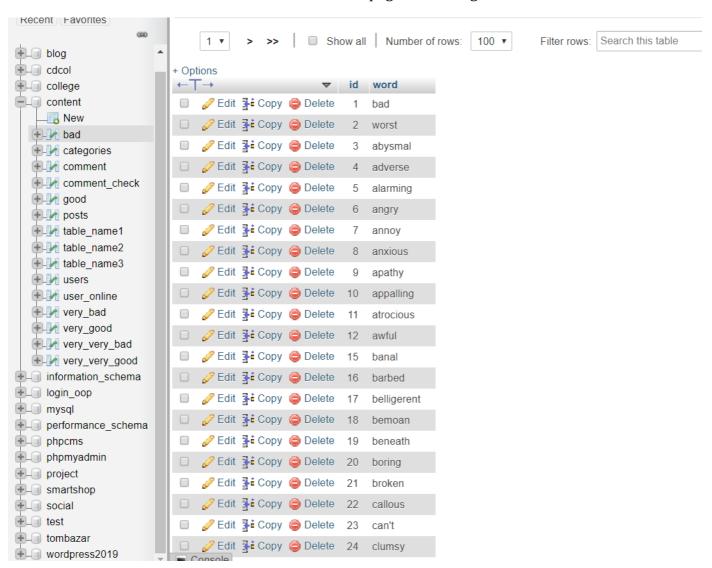


Table of all negative word.(bad)

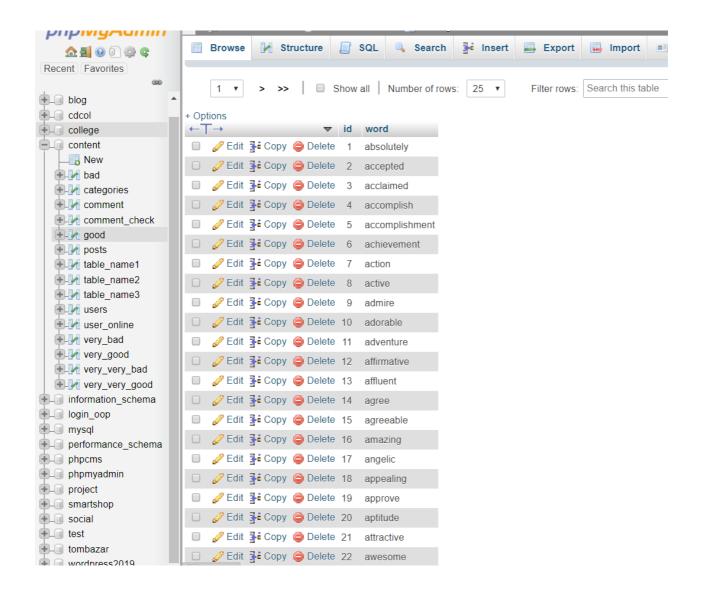


Table of all positive word.(good)

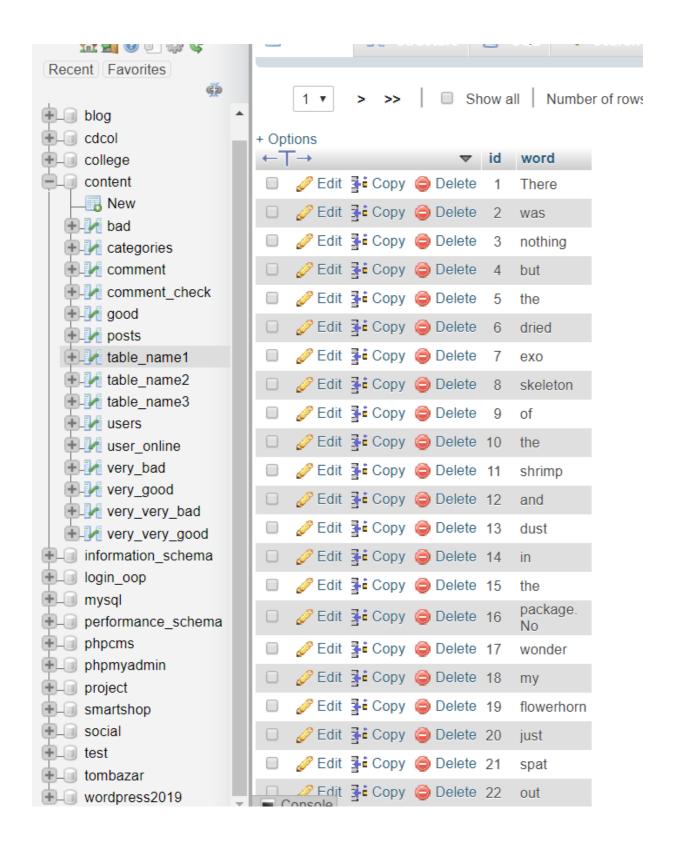


Table of uni-gram of the comment of the post (table_name1).

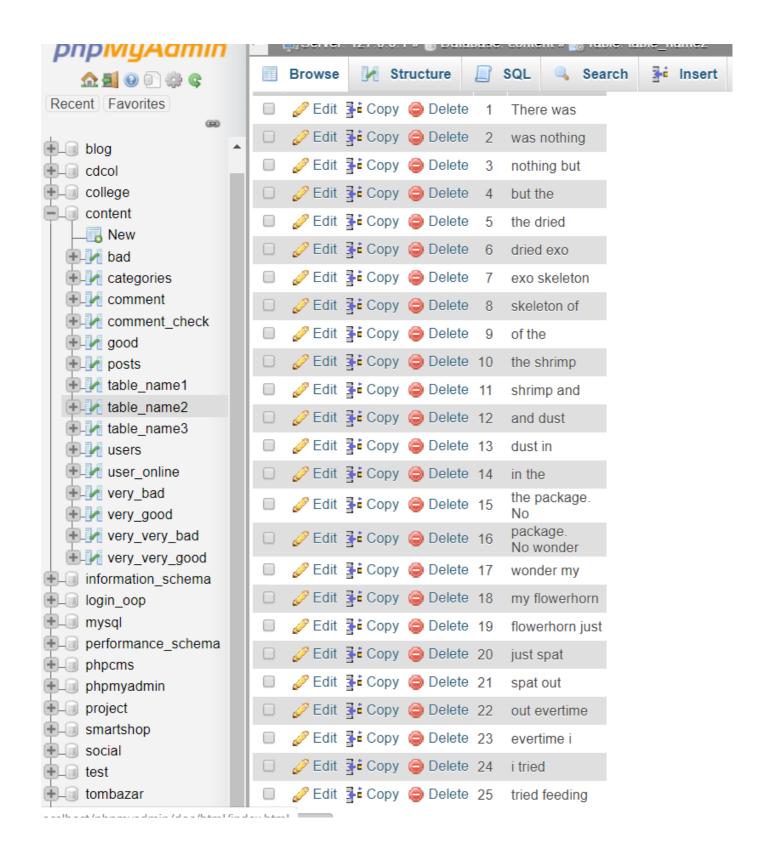


Table of bi-gram of the comment of the post (table_name2).

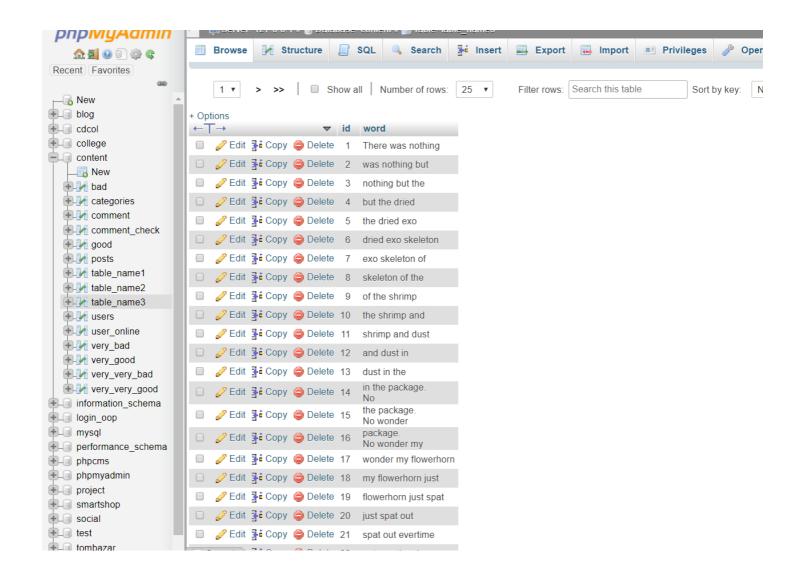


Table of tri-grapm of the comment of the post (table_name3).

5.2 Features

- © Comment sentimental analysis system can be enhanced to include some other functionality like newspaper, movie review, product review.
- Newspaper category can be added with this system.
- Social networking can also be added wherein user can identify review of such product.
- In ecommerce site it can be added.

Chapter-6

6.1 Conclusion

Undoubtedly Hotel management system can make the education system dynamic and advanced. If you wish to re-brand the school image and make it more customer and parent-friendly, then deploy an advanced hotel management system today.

6.2 References

- PHP book by Vasvani (TMH publications).
- Beginning PHP5 by WROX.
- www.google.com.
- www.wikipedia.com
- www.en. Wikipedia.org/wiki/network.com
- www.en. Wikipedia.org/wiki/network.CSS
- Head First PHP & MySQL by Lynn Beighley and Michael Morrison(O'Reilly)