Web-Based Student Opinion Mining System Using Sentiment Analysis

Olaniyi Abiodun Ayeni[1, 2],
[1]Department of Cyber Security, Federal University of Technology, Akure, Nigeria
Email: oaayeni@futa.edu.ng

Akinkuotu Mercy
[2]Department of Computer Science, Federal University of Technology, Akure
Email: akinkuotucsc134976@futa.edu.ng

Thompson A.F[1]
[1]Department of Cyber Security, Federal University of Technology, Akure, Nigeria
Email: afthompson@futa.edu.ng

Mogaji A.S[2]
[2]School of Computing, Federal University of Technology, Akure , Nigeria
Email: alabadepoola@yahoo.com

Abstract: Collecting feedback from a few students after the exams has been the norm in educational institutions. Forms are given to students to assess the course the lecturer has taught. The main purpose of developing student opinion mining system is to create a faster and easier method of collecting feedback from student, and also give lecturers and school administrators an easier way of analysing the feedback collected from students. The significance of this application is that it is less expensive and present a more confidential way of getting students opinion. The major tools used in developing this application are Python, Scikit learn, Textblob, Pandas and SQLite.. Django provides an in-built server that allows the application to run on the localhost.. In this project dataset gotten from online feedback form distributed to students was used for the sentiment analysi ,Chi-square was used for feature selection and the support vector machine algorithm was used for sentiment classification. The application will help the university administrators and lecturers to identify the strengths and weaknesses of the lecturer based on the textual evaluation made by the students.

Index Terms: Opinion Mining, Sentiment analysis, Students, Feedback

1. Introduction

Opinion Mining which is sometimes referred to as sentiment analysis, is the use of natural language processing, text analysis, computational linguistics and biometrics to systematically identify,extract quantify and study affective states and subjective information. In other words, sentiment analysis is an application of natural language processing,computer linguistcs and text analysis that identifies and process text in order to extract specific information from it. Sentiment analysis seeks to identify the view-points of an underlying text (Bo Pang and Lillian Lee 2004 [3]). There are different approaches to create sentiment analysis models : Lexicon-based, machine learning and hybrid approach (Altrabsheh Nabeela 2016[5]).Data mining is the application of data mining techniques to identify and address problems in education. (Altrabsheh, Nabeela, Gaber, M. and Cocea, Mihaela 2013[1]). In educational data mining the data mined could include student performance, lecturer's feedbacks on their students' performance, student feedbacks/opinion on their lecturers and courses e.t.c. The Paper (Altrabsheh Nabeela, et al 2013[1]) studied the use of data mining to improve education and address problems in education by monitoring Student Performance, by combining the naive bayes and support vector machine classifier to analyze student Feedback in real time, but had the limitation of getting real-time getting feedback from student requires the use of SMS which is expensive or clickers which gives limited feedback.

(Altrabsheh Nabeela 2016[5]) studied different methods for sentiment analysis and identified the best preprocessing level, features and machine learning techniques, and also explores the detection of emotion in student feedback. But was limited to small amount of dataset in the educational domain. This research proposes a less

expensive and a more confidential way of getting students opinion by using online forms which is cheaper compare to SMS and Introduces a level of animosity which will give them opportunity to give less bias opinions and feedbacks. The research objectives are to design a web-based system for Sentiment Analysis using Machine Learning and to implement the design.

2. Review of Related Works

Five different works were reviewed below. The strength and weakness of this works were highlighted, the essence of which is to develop an improved version of the application.

(a) Student Feedback Mining System Using Sentiment Analysis (R.Menaha Et Al 2017)

In this paper a student feedback mining system (SFMS) was developed, the system which text analytics and sentiment analysis approach to provide instructors a quantified and deeper analysis of the qualitative feedback from students that will improve the students learning experience.

Objectives:

a) To develop an efficient approach for providing qualitative feedback for the instructor that enriches the students learning.
b) To classify the comments of students using sentiment classifier and apply the visualization techniques to represent the views of students.

Methodology:

In this paper three methods were adopted to extract the keywords from the students' feedback document they include: Tokenization, Stop word removal, Clustering, Classification and Sentiment Analysis. Tokenization breaks up a sequence of strings into pieces such as words, keywords, phrases, symbols and other elements called tokens (words, phrases etc.). Clustering is the process of making a group of abstract objects into classes of similar objects. The feedbacks are collected from the students for a single course for easy evaluation and to improve student's learning file, after feedbacks have been collected, they are tokenized and stop words are removed, then topics are extracted for clustering

Limitation:

Semantic similarity for clustering the student feedback was not considered in the paper. The project did not employ the use of different visualization methods.

(b) Feedback Analysis On The Performance Evaluation Of Pedagogue Using Opinion Mining (Raj Kumar 2017).

Individual Views about various entities are the expressions of that perticular individual feelings, reactions about that entity. The research area of analyzing views contained in texts is generally known as opinion mining. An opinion mining system helps the customers in choosing the right product, helps the administrations of an organisation to choose a proper working plan, and also help political parties to have public opinion about their regime or what needs improvement in the opposition party government so they can use it to their advantage. Categorizing view or public opinion scientifically by opinion mining tools saves time and money significally than directing time consuming reviews or market research.

Motivation:

The Motivation of the research is to builds a system with explicit access rights to get student feedback and performance analysis on it,so as to facilitate the institute management to be responsive of their stake holder's viewpoints.

Methodology:

In the research work an opinion mining tool called RAPMINDER was utilized, RAPIDMINER is the free software used for data mining software that is used for data research, machine learning, deep knowledge, text excavating, and logical analytics.
The system as three main module which are: Admin module, Student Evaluation, Module Report Generation Module If all the metrics are fulfilled, then the teacher has attained highest percentage with excellent class label. If some of the metrics are fulfilled then the class teacher attained middle percentage with good, average and satisfactory class labels. If all the metrics are not fulfilled then the class teacher attained need to improve class labels. The report module generates result for each user depending on the access privillegde of the particular user.

(c) Learning Sentiment From Students' Feedback For Real-Time Interventions In Classrooms (Nabeela Altrabsheh Et Al2014 [8])

This paper studies different methods that could be used for learning sentiment from students'feedbacks. It utilized four techniques namely Naive Bayes, Complement Naive Bayes (CNB), Maximum Entropy and Support Vector Machine (SVM) on real-time students' feedback to identify sentiments.

Motivation:

In Education knowledge about students' thoughts, feelings can be used to address problems which affect student engagement and participation. Students' feedback can help the lecturers understand their students learning behavior and thus improve teaching their mode and methods of teaching. Taking feedbacks can highlight different issues student have. Feedbacks are usually collected at the end of the course when students don't understand a part of the course the lecturers get to know after the course but taking the feedback in real time can address that issue. Collecting real-time feedback cannot be used to their full advantage without support for the analysis of the collected data. There is a need to present the feedbacks from to student to lecturer in a presentable and understandable format.

Objectives:

a) To assess the ability of several machine learning techniques to learn sentiment from
students' textual feedback.
b) To create a system that will automatically analyze students' feedback in real-time and present them to the lecturer.
c) To allow lecturers to have an overall summary of the students Opinion.
d) To improve the quality of teaching and communication between lecturers and student.

Methodology:

Four major steps were taken during the course of the experiment: Collection of data, preprocessing of data, feature selection and applying machine learning techniques.

Data is collected at real-time from feedback gotten from Student Response Systems (SRS) which could be Clickers, mobile phones and social media. The total amount of data collected is 1036 instances.

Limitations:

Naive Bayes does not work well with uneven class sets. Maximum Entropy is not very realistic in many practical problems, as real datasets contain random errors or noises which create a less clean dataset. The effectiveness of the SVM can be affected by the kernel.

(d) Online Student Feedback Analysis System With Sentiment Analysis (Divyansh Shrivastava Et Al 2017):

The Online Student Feedback Analysis System is a web based system which collects the feedback from every individual student and provides an automatic generation of a collective feedback which has been taken by the students. The system was developed to provide the feedback in an easy and quick manner to any particular department in a college or an educational institute.

Objectives:

a) To provide a management information analysis system for educational institutes to manage student feedback data.
b) To create a faster and a stress free way to analyze students feedbacks.

System Implementation:

The key features of the system are Functionalities and Database. The core functionalities of the system include:

1. ADMINISTRATOR:

i. Can insert /update/delete new student
ii. Generate analysis
iii. View students who haven't given feedback

2. STUDENT:

i. Can fill the objective questionnaire by filling the marks out of a constant value
ii. Can give the comments/compliments/reviews about the respective faculties, course structure, subject topics, contents etc.
iii. Can verify the identity and edit his/her personal profile

Limitation:

The design of the screen requires further enhancement. Current forms in the system requires enhancement so as to retrieve the feedback details even in a better way

(e) Sentiment Analysis Using Support Vector Machine (Nurulhuda Z And Ali Selamat 2014)[7]

This work present an experiment using machine learning open source data mining software tool while there is no single tool or technique that always achieves the best result. However, some achieve better results more often than the other.

Objectives:

The aim of this experiment is to improve SVM on benchmark datasets by Pang Corpus and Taboada Corpus.

Methodology:

The framework consists of pre-processing, feature extraction, feature selection and classification stages. Dataset and pre-processing: The experiment used two dataset the Pang Corpus and Taboada Corpus

Classification Method Selection:

Support Vector Machine (SVM) was chosen for the classification in the experiments due to the fact that SVM works well for text classification, and has the potential to handle large features. The text categorization effectiveness is usually measured using the F1, accuracy, and AUC (Area under the (ROC) Curve).

Limitation:

It can be concluded that other n-grams and bi-grams models has a lower performance compared to the unigram model for both of the datasets. The experiment did not consider the embedded method of feature selection.

2.1 Research Base for the Work

The Paper (Altrabsheh Nabeela, et al 2013[1]) studied the use of data mining to improve education and address problems in education by monitoring Student Performance, by combining the naive bayes and support vector machine classifier to analyze student Feedback in real time, but had the limitation of getting real-time getting feedback from student requires the use of SMS which is expensive or clickers which gives limited feedback. (Altrabsheh Nabeela 2016[5]) studied different methods for sentiment analysis and identified the best preprocessing level, features and machine learning techniques, and also explores the detection of emotion in student feedback. But was limited to small amount of dataset in the educational domain. This research proposes a less expensive and a more confidential way of getting students opinion. By using online forms which is cheaper compare to SMS and Introduces a level of animosity which will give them opportunity to give less bias opinions and feedbacks. It can be concluded that other n-grams and bi-grams models has a lower performance compared to the unigram model for both of the datasets. The experiment did not consider the embedded method of feature selection

2.2 Objectives

a) Design a web-based machine learning system for Sentiment Analysis of student's opinion
b) Implement the design in (a)

3. System Analysis and Design

This chapter describes the system design, input, output and processing requirements of the proposed system and the hierarchical design of the proposed system. It also describes the system architecture, flow chart, use case diagram, which were used to design the system to function properly.
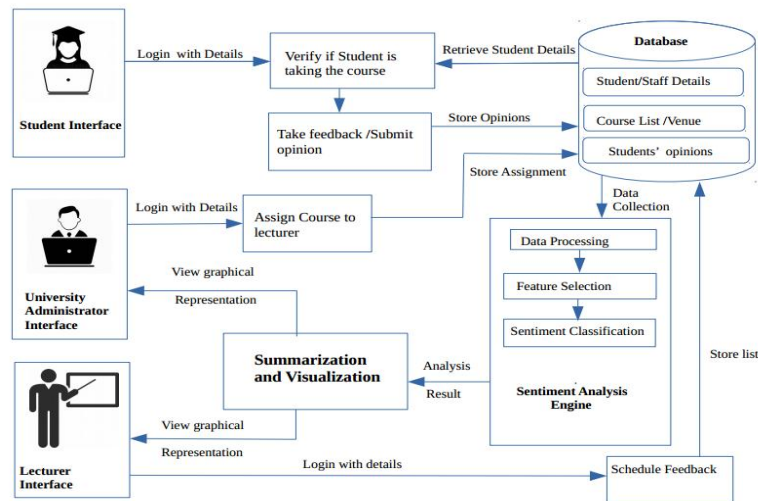
Fig. 1. The System Architecture (adapted from Francis F. and Benilda Eleonor V. 2016)

The system architecture of this project consists of five components;
i. The University Administrator
ii. The Lecturer
iii. The Student
iv. Database
v. Sentiment Analysis Engine

3.1 The University Administrator Interface

The University administrator includes the Head of Departments, Department officers,School officers, Sub-deans of Schools,Dean of Schools, Vice-Chancellor. The university administrator has the functionalities which includes; adding all the courses offered in the university, adding department according to the school they belong to, updating the lecturer assigned to course in situations whereby the course lecturer is changed. The university administrator has the permission to view the summary and charts of feedback for all lecturers and courses in the university. They can also view all the details of students and lecturers registered on the application.

3.2 The Lecturer Interface

The Lecturer is required to register, before he/she can use the application, after registration the university administrator add the lecturer to all the courses he/she is teaching for the semester or session. The lecturer can view all the courses that have been assigned to him/her by the university administrator, he/she can schedule feedback for the students. He/She can also view the summary and charts of feedback of the students, for all the courses he/she is teaching.

3.3 The Student Interface

The student is also required to register, before using the application, the student can add all the courses he/she is offering. The student is only allowed to take feedback for the courses he/she has registered for.

3.4 The Database

The database stores all the users (university administrator lecturer, students) details, it stores course and department information, it also stores the feedback that is submitted by the students which is later inputted into the sentiment analysis engine. It stores the result of the analysis from the sentiment engine.

3.5 The Sentiment Analysis Engine

This module is where sentiment analysis is done using the support vector machine with the radial basis function kernel. The feedback from the database is fed into the engine as its datasets, the polarity of the feedbacks are then classified. The output is then displayed as charts to the user
In the sentiment analysis engine three processes takes places in it, the processes includes:
i. Data Preprocessing
ii. Feature Selection
iii. Sentiment Classification

4. Data Preprocessing

In this project, the data used are students' feedbacks, gotten through google forms from students of the Federal University of Technology,Akure who have taken or are currently taking CSC102(Introduction to computing) and CSP 210(Introduction to Agricultural Practicals).

4.1 Flow Chart of the Application

The application was developed through several steps as shown in the figure below. The steps comprises of all the design and development efforts that was carried out. The following sections describe each step in details
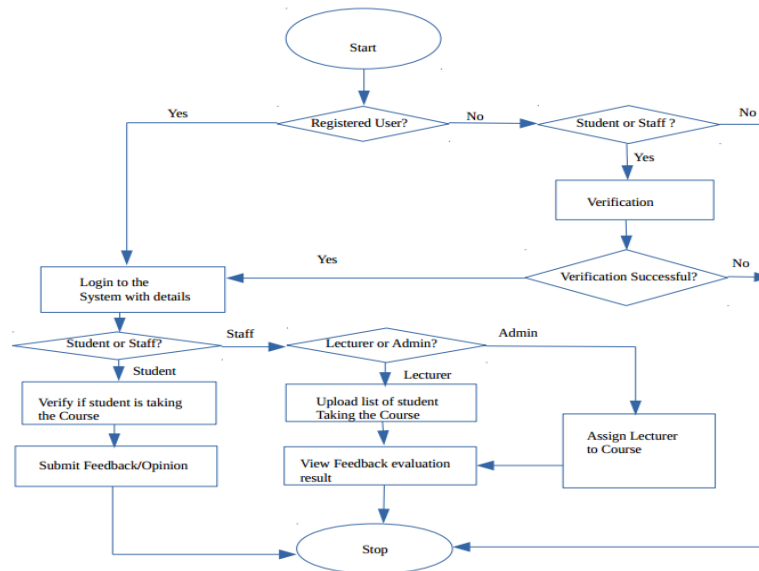
Fig.2. Flowchart of the system

5. System Implementation

This chapter introduces the full documentation of the designed web-based student opinion mining system. The hardware and software specifications of the system used in developing the application are stated in table 1.

Table 1. Specification of the system developed

Specifications category	Specifications Details	
Hardware Specifications	Processor Type	Intel Core i3-2350M CPU
	Processor Speed	2.30 Ghz
	Primary Memory	3.8 gb RAM
	Secondary Memory	310.8 GB HDD
Software Specifications	OS type	64bits
	OS	Ubuntu 16.04 LTS
	Web Browser	Moxilla Firefox Quantum 61.0.1,Google Chrome Official build 69.0.3497.81
	Web Server	Django 2.0.3 server
	Database Server	Sqlite3, DB Browser for Sqlite
Other Specifications	Internet Access	

Table1shows the minimum specification under which the application will work. In other words, the system to be use must have the specification in table 1 as the minimum before it can be used.

5.1 The Home Page

The home page (fig.3) is the first page that will be seen once the web-application is initialized from the local host at http://127.0.0.1:8000/. On the Home there is a Login Button(for existing users) and Create Account(for new users) button
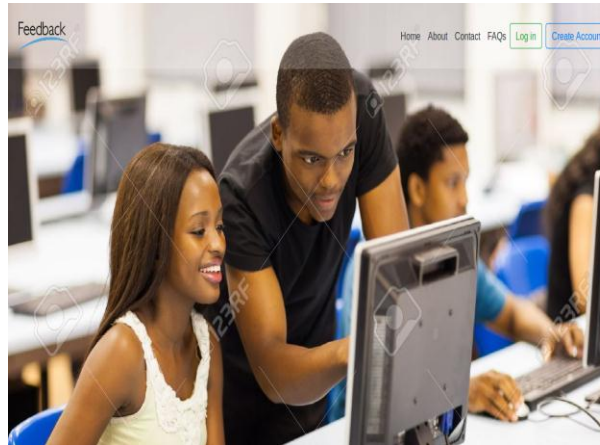
Fig. 3. Home page of students' opinion mining system

The Login Page :

This page allows existing users (Students, Lecturers, University Administrators) to access the application. Before anyone can use the application, the user details has to be authenticated
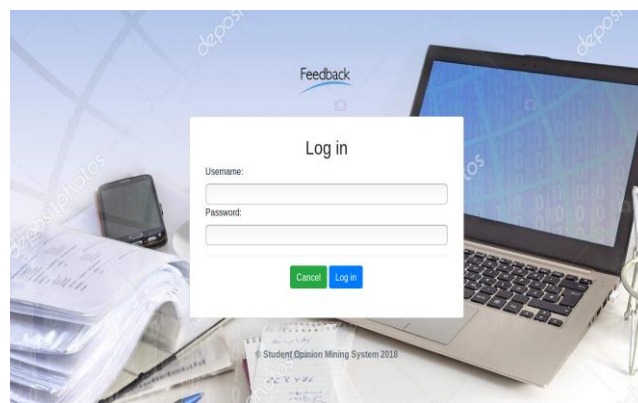

Fig. 4. Login page

5.2 The Registration Page

The three possible users (Students, Lecturers, University Administrators) of this application require slightly different details for them to user the application. Students are required to provide their matric number for unique identification, while on the other hand Lecturers and University Administrators provides their staff id. Thus the registration form for each user is different
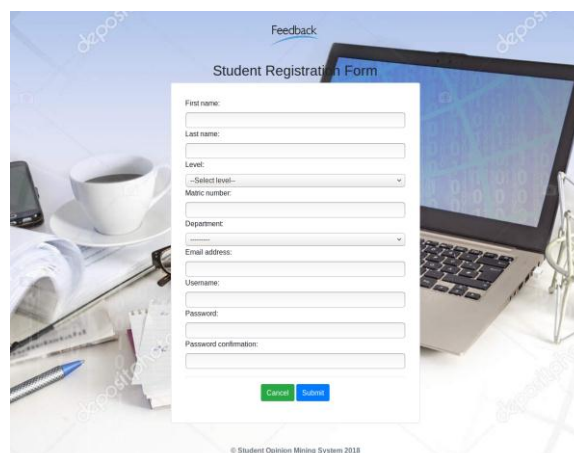
Fig.5. Student's registration page
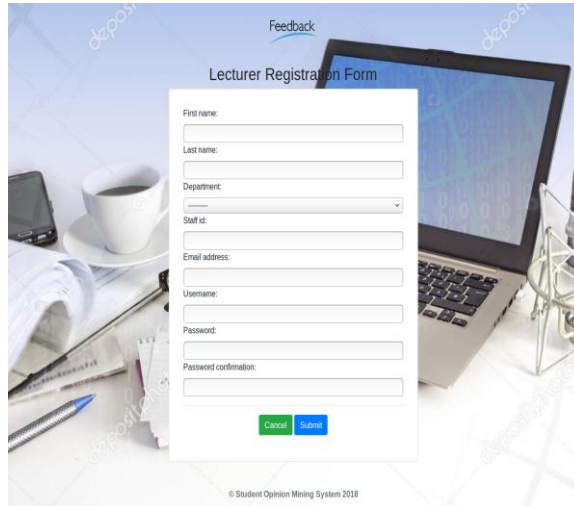
Fig. 6. Lecturer's registration page

5.3 The Dashboard Page

When a user, logs in into the app. A unique dashboard is displayed depending on the user(students,lecturer,university adminstrator).For Lecturers a general summary of all courses taken by the lecturer is display.For university administrators All the list of courses and lecturer is displayed on a table
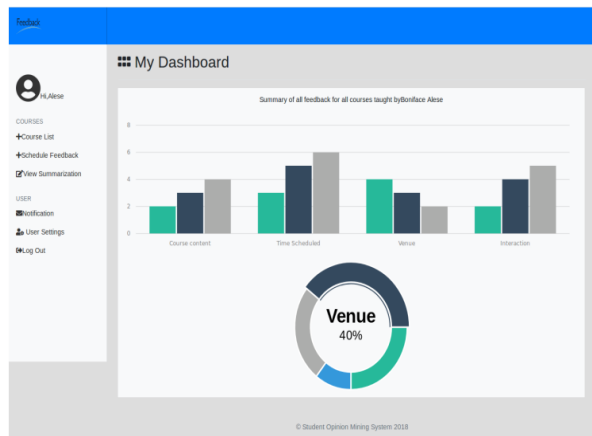
Fig.7. Dashboard

Fig.8. Administrator dashboard
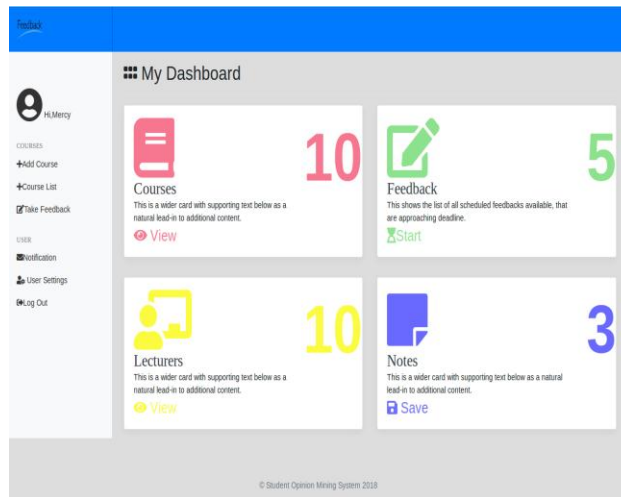
Fig.9. Student dashboard

5.4 The Summarization Page

To view summary of the feedback, it is required for the user to be logged in as a lecturer or school admin. The green coloured bar represents the positive
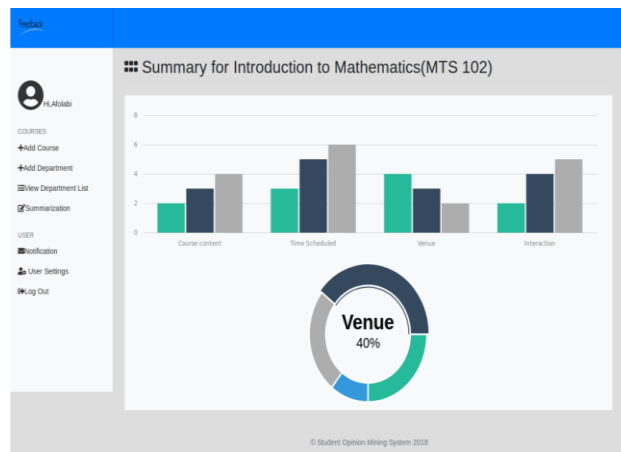
Fig.10. Summary view

5.5 The University Adminstrator Component

Functionalities of the university interface includes adding department, adding courses, View students list, view lecturer list

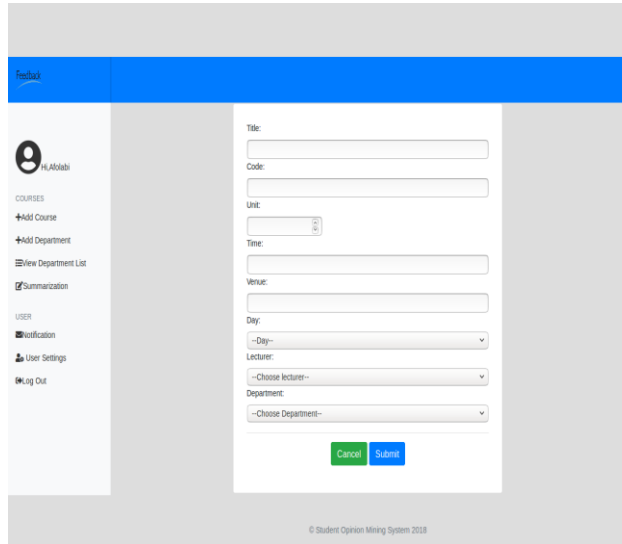

Fig.11. Ddd department form

Fig.12. Ddd course form

5.6 The Lecturer Component

Functionalities of the lecturer interface includes scheduling of feedback,, view courses assigned to lecturers.
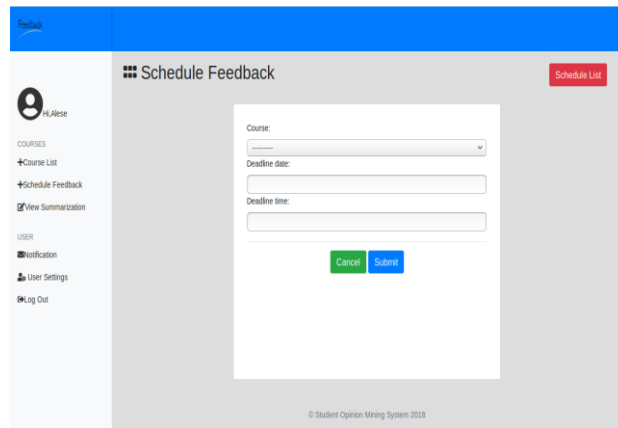

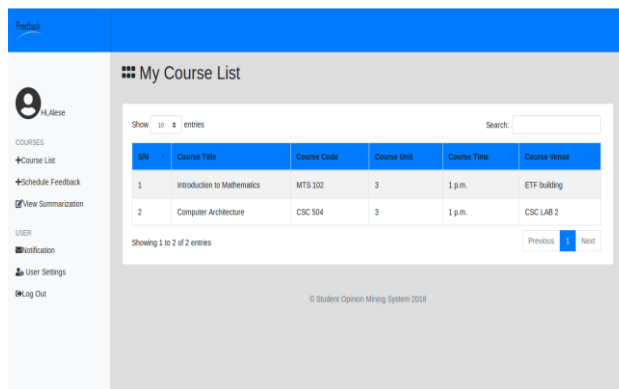Fig. 13. Schedule feedback form

Fig.14. List of courses assign to a lecturer for a semester

5.7 The Student Component

Functionalities of the student component interface include Taking feedback, viewing scheduled feedback.Add courses they are offering

Fig.15. List of schedule feedback

Fig.16. Feedback form

5.8 Sentiment Analysis Engine

The feedbacks are grouped into three classes: positive negative and neutral: The nature of all the feedback used is shown in the graph below
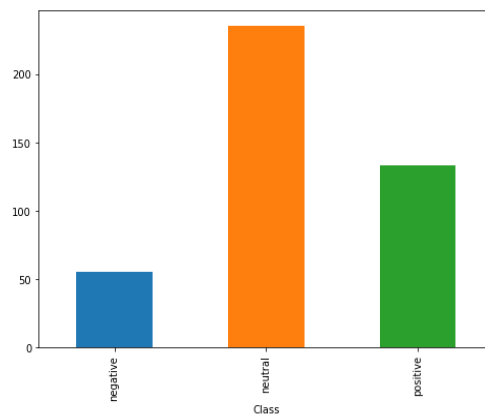

Fig.17. Chart showing the nature of the feedback gotten from students

After the preprocessing of the data, $\chi 2$(chi−square) is used for the extraction of important feature for training the classifier. The result of the feature extraction is shown in the graph below
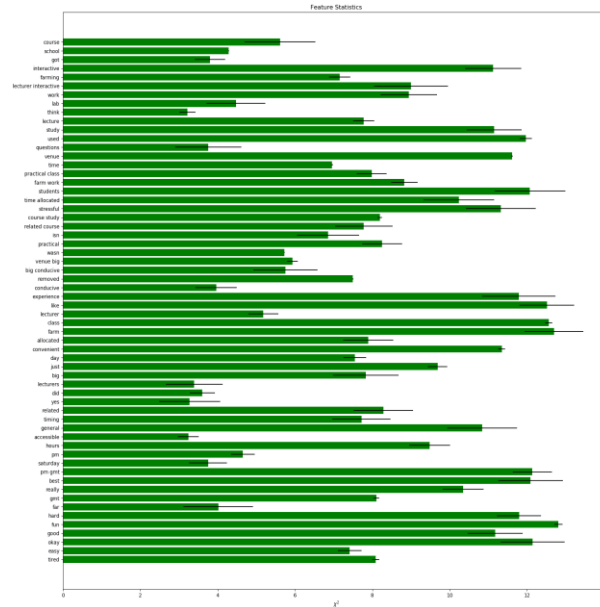
Fig. 18. Feature statistic

6. Conclusion

This project implements a web-based opinion mining system using Machine Learning approach. The application was tested using students of The Federal University of Technology, Akure and it shows an improvement over the existing systems. The design and implementation of the system was successful. Students, Lecturer and School administrators can greatly benefit from the using the platform. Also instead of lecturer waiting for examination period before they can get feedback from their students this platform comes in handy for them Instead of just getting feedback from few students the platform gets feedback from all students. And then gives them the result of the analysis.

Recommendation

After the successful implementation of the system using the student of the Federal University of Technology, as the case study, it is recommend that the application should be used by all Higher Institutions in Nigeria and all over the World, so as to foster a better student/lecturer interaction and also improve the quality of education.

References

[1]   Altrabsheh, Nabeela, Gaber, M. and Cocea, Mihaela 2013. SA-E: Sentiment Analysis for Education. In: 5th KES International Conference on Intelligent Decision Technologies, 26-28 June 2013, Sesimbra, Portugal
[2]   Bo Pang and Lillian Lee 2008. Opinion mining and sentiment analysis. Foundations and trends in information retrieval, 2(1-2):1–135, 2008.
[3]   Bo Pang and Lillian Lee 2004. A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts. Association for Computing Linguistics. Article No. 271.
[4]   Bo   Pang   and   Lillian   Lee   2005.   Movie   Review   Data.   Sentence   polarity   dataset v1.0.https://www.cs.cornell.edu/people/pabo/movie-review-data/ (05.05.2017)
[5]   Altrabsheh Nabeela 2016. Sentiment analysis on students' real-time feedback. PhD Thesis,Department of Computing,Faculty of Technology,University of Portsmouth.
[6]   Tang H, Tan S, and Cheng X 2009. A survey on sentiment detection of reviews. Expert Systems with Applications: An International Journal September 2009, 36(7):10760– 10773.
[7]   Nurulhuda Zainuddin,Ali Selamat 2014. Sentiment Analysis Using Support Vector Machine. IEEE International Conference on Computer, Communication, and Control Technology (I4CT 2014), September 2 - 4, 2014 - Langkawi, Kedah, Malaysia.
[8]   Nabeela Altrabsheh, Mihaela Cocea, and Sanaz Fallahkhair 2014. Learning sentiment from students' feedback for real-time interventions in classrooms. Third international Conference, ICAIS 2014 Bournemouth, UK,September 8-10 2014 Proceedings.[Adaptive and intelligent systems].
[9]   Rohini S. Rahate,Emmanuel M 2013.Feature Selection for Sentiment Analysis by using SVM. International Journal of Computer Applications (0975 – 8887).Volume 84 –No 5, December 2013.
[10]   Huan Liu and Lei Yu. 2005 .Toward Integrating Feature Selection Algorithms for Classification and Clustering. IEEE Transactions on Knowledge and Data Engineering. Volume 17 – Issue 4 2005 Page 491-502.

[11] Go, A., Bhayani, R., & Huang, L. 2009. Twitter sentiment classification using distant supervision. CS224N Project Report, Stanford, 1(12).

[12] Poulos, Ann, and Mary Jane Mahony (2008)."Effectiveness of feedback: the students perspective." Assessment and Evaluation in Higher Education 33.2 (2008): 143-154.

[13] J. Singh, G. Singh, and R. Singh (2016). A review of sentiment analysis techniques for opinionated web text, CSI Trans. ICT, 2016.

[14] H. Song and T. Yao (2010). "Active learning based corpus annotation," in Proc. IPSSIGHAN Joint Conference on Chinese Language Processing, Beijing, China, August 2010, pp. 28-29.

[15] Martin Ringsquandl and Dušan Petković (2013).Analyzing Political Sentiment on Twitter Analyzing Microtext: Papers from the 2013 AAAI Spring Symposium.

[16] Ayesha Rashid,Sana Asif,Naveed Anwer Butt,Imran Ashraf (2013).Feature Level Opinion Mining of Educational Student Feedback Data using Sequential Pattern Mining and Association Rule Mining . International Journal of Computer Applications (0975 –8887) Volume 81 – No.10, November 2013

[17] Francis F. Balahadia and Benilda Eleonor V. Comendador (2016).Adoption of Opinion Mining in the Faculty Performance Evaluation System by the Students Using Naïve Bayes. Algorithm International Journal of Computer Theory and Engineering, Vol. 8, No. 3, June 2016.

[18] Thorsten Joachims. (1999). Making large-scale SVM learning practical. In Bernhard Schölkopf and Alexander Smola, editors, Advances in Kernel Methods – Support Vector Learning, pages 44–56. MIT Press

[19] C. Cortes and V. Vapnik.(1995). Support-vector networks.Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.

[20] Grigori Sidorov, Francisco Velasquez, Efstathios Stamatatos, Alexander Gelbukh1, and Liliana Chanona-Hernández.(2012).Syntactic N-grams as Machine Learning Features for Natural Language Processing.Mexican International Conference on Artificial Intelligence.San Luis Potosi, Mexico.October 27-November 4,2012.

[21] Phayung Meesad, Pudsadee Boonrawd and Vatinee Nuipian.A Chi-Square-Test for Word Importance Differentiation in Text Classification.2011 International Conference on Information and Electronics Engineering IPCSIT vol.6 (2011) © (2011) IACSIT Press, Singapore.

[22] Web Forefront (2018). Set up content:Understand urls,templates and apps. Django Work flow for urls,templates and apps. https://www.webforefront.com/django/setupdjangocontent.html

[23] Zhixiong Hong (2009).Django Web Application Framework .https://www.slideshare.net/mobile/fishwarter/the-djang webapplicationframework-2-1221388

[24] Scikit-learn.org(2018).Choosing the right estimator. https://scikit-learn.org/stable/tutorial/machine_learning_map/index.html

Authors' Profiles

Olaniyi A. Ayeni (Ph.D): A lecturer in the School of computing, Federal University of Technology, Akure with interest in Information Security, cyber security and digital forensics. He is a member of Computer Professional Registration Council of Nigeria (CPN) and member of International Association of Engineers (IAENG).

Dr. Aderonke Thompson is an Associate Professor and the acting HOD of Cyber Security department, School of Computing. Federal University of Technology, Akure.

Alaba Mogaji (Ph.D) holds an M/Tech and a P.hD degree from the department of Computer Science, School of Computing, Federal University of Technology, Akure. Ondo-State. Nigeria

Akinkuotu Mercy (B.Tech. Computer Science) is a graduate of Computer Science from Federal University of Technology, Akure.