

Opinion Mining and MCDT Based Ranking and University Recommendation System

by

Wasim Mahmud Surjo

19341019

Nazifa Khanom

17101334

Sadman Sakib Nadvi

16101076

Sayan Manoranjan Roy

13201063

Moury Moontaha

14101036

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
BRAC University
October 2020

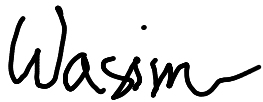
© 2020. BRAC University
All rights reserved.

Declaration

It is hereby declared that

1. The thesis submitted is our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:



Wasim Mahmud Surjo

19341019



Nazifa Khanom

17101334



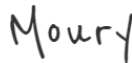
Sadman Sakib Nadvi

16101076



Sayan Manoranjan Roy

13201063



Moury Moontaha

14101036

Approval

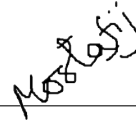
The thesis titled “Opinion Mining and MCDT based Ranking and University Recommendation System” submitted by

1. Wasim Mahmud Surjo (19341019)
2. Nazifa Khanom (17101334)
3. Sadman Sakib Nadvi (16101076)
4. Sayan Manoranjan Roy (13201063)
5. Moury Moontaha (14101036)

Of Summer 2020 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science.

Examining Committee:

Supervisor:
(Member)



Mostafijur Rahman Akhond

Lecturer, Department of Computer Science and Engineering
RAC University

Program Coordinator:
(Member)



Md. Golam Rabiul Alam

Associate Professor, Department of Computer Science and Engineering
BRAC University

Head of Department:
(Chair)

Prof. Mahbub Majumdar
Chairperson
Dept. of Computer Science & Engineering
Brac University

Mahbubul Alam Majumdar

Professor and Chairperson, Department of Computer Science and Engineering
BRAC University

Abstract

Opinion mining is an approach to natural language processing (NLP) that identifies the emotional tone behind a body of text. This is a popular way for organizations to determine and categorize opinions about a service or idea. It involves analyzing data set from different aspects to mine text for sentiment and subjective information which can be used to increase efficiency. Students are unaware of the best institution suited for them due to insufficient information. This paper explores recommendation of perfect universities which the students desire by using data of opinion mining through a Multi Criteria Decision Technique (MCDT) based ranking and university recommendation system to tackle university admission problems.

Keywords: opinion mining, data analysis, natural language processing, recommendation, topsis.

Dedication

This research is wholeheartedly dedicated to our beloved parents who have been a source of our inspiration, supported us and gave us strength at each turn of our life. We also want to remember our friends with whom we have lots of memories to cherish with.

Acknowledgement

Firstly, all praise to The Great Allah for whom our thesis has been completed without any interruptions.

Secondly, to our supervisor Mostafijur Rahman Akhond and coordinator Md. Golam Rabiul Alam for their kind support and advice in our work. They provided us with guidance, encouraged, advised and helped us whenever we needed help throughout our time as his students. We have been extremely lucky to have a supervisor and coordinator who cared deeply about our work, and who responded to our questions and queries so promptly.

And finally to our parents without their throughout support it may not be possible. With their kind support and prayer we are now on the verge of our graduation.

Table of Contents

Declaration	i
Approval	ii
Abstract	iii
Dedication	iv
Acknowledgment	v
Table of Contents	vi
List of Figures	viii
List of Tables	ix
List of Acronyms	ix
1 Introduction	1
1.1 Background	1
1.2 Motivation	2
1.3 Research Problem	2
1.4 Research Objectives:	3
1.5 Thesis Orientation	4
2 Literature Review	5
3 Data Collection And Analysis	8
3.1 Initial Data Processing	8
3.1.1 Initial Stage	8
3.1.2 Second Stage of Collecting Data	10
3.1.3 Data exploration and profiling	11
3.2 Data Operations	11
3.2.1 Data Analysis	11
3.2.2 Data Profiling	11
3.3 Transforming Data and Data Modeling	12
4 Methodology	13
4.1 Data Manipulation	13
4.1.1 Tokenization	13

4.1.2	POS Tagging	13
4.2	Opinion Mining	16
4.2.1	Opinion Mining	17
4.2.2	Opinion Wise Cardinality Determination	18
4.3	MCDT Ranking and University Recommendation System	19
4.3.1	Optimized Demand Based Dataset	19
4.3.2	Use TOPSIS Decision Making Algorithm For MCDT Ranking	19
4.4	Preliminary Results And Comparison	21
5	Conclusion	23
	Bibliography	24

List of Figures

3.1	Survey Form	9
3.2	Survey Form	10
3.3	Survey Form	10
3.4	Data Analysis	11
3.5	Data Profiling	12
4.1	Methodology	14
4.2	Grammar For Opinion Mining	17
4.3	Results: Top 7 issues/demands regarding universities	18
4.4	University Scores Per Demand	19
4.5	TOPSIS decision making algorithm	20
4.6	Obtained Results	21
4.7	Private University Rankings according to 4icu.org (Unirank BD) until 6th September,2020	22

List of Tables

4.1	NLTK POS TAGGING	15
4.2	NLTK POS TAGGING 2	16

List of Acronyms

NLTK - Natural Language Toolkit

TOPSIS - Technique of Order Preference Similarity to the Ideal Solution

NLP – Natural Language Processing

CSV – Comma Separated Values

POS – Parts Of Speech

MCDT - Multi Criteria Decision Technique

Chapter 1

Introduction

1.1 Background

As days are passing, opinion is becoming an indispensable part of our lives. Opinion is not necessarily based on fact or knowledge, rather it can also be a view or judgment formed by people. People give different types of opinion regarding different things. With the help of those opinions a person can make a decision. Opinion is very important because it helps to decide between right or wrong. It confronts a number of challenges to analyze the public views. It is seen that difficulties are seen for getting admission in a University according to the desired needs of people. Some have a huge amount of money and therefore can get admitted in the most popular private universities whereas, the people who have less money faces problems in getting admission according to their needs. Here, the opinions of all the people regarding admission and recommendation according to the criteria's of the students and their parents is very much necessary. People can say different types of things and facts about different types of universities. But it is hard to know which facts are true and which are false because choices and opinion varies from person to person[2] [3]. So, if a person likes a specific university then that does not mean that all the other people wants to get admitted or have the capability to get admission in that specific university. So, here comes the necessity in the opinion and the recommendation part.

Opinion mining is used for tracking the mood of the public about a particular product which is also called sentiment analysis. It involves building a system to collect and categorize opinions about a product [1]. It is the computational study of opinions, emotions and attitudes which is one of the key research field in Natural Language Processing (NLP). It helps to decide the mentality of a speaker or an author in regards to a certain assessment, which will be on the record. The mentality might be their judgment or assessment, of their current state of mind, which is the emotional state of the author while putting down the feedback of a certain thing. There is lack of study in opinions because of the lack of text opinionated before the World Wide Web. In the earlier days, when someone needed to take any decision then they would ask for opinions from friends, families and other nearby people. But with the growth of time and the user generated content on the web, the world changed. Now-a -days, for knowing any opinion about anything there are a lot of reviews on the Web.

Therefore, we thought about the difficulties of people in finding the desired private university and so we proposed opinion mining and university recommendation system using MCDT based algorithm, which is a new publicly available opinion mining for the University selection process. The opinion mining system will recommend a University selection based on the newly created resource that contributes to an automatic opinion. After the collection of all the valid inputs from the authors/users, the data is trained by using NLTK (Natural Language Toolkit) in python for the opinion mining section and for the recommendation section we used TOPSIS (Technique of Order Preference Similarity to the Ideal Solution) on finding an ideal and an anti-ideal solution and comparing the universities according to the criteria's of the student and his/her family for getting admission in their desired university.

The paper is organized into following sections: Literature Review, Proposed Approach, Dataset Analysis, Methodology, Result Analysis and Conclusion with future works.

1.2 Motivation

Our primary motivation was to research about opinion mining for getting admitted in the desired universities so that people can rely on the reviews that are found from the opinions of different people. We explored several research papers, journals and articles and finally decided to work on the opinion mining of universities. Before opinion mining, the people mainly relied on the ratings and comments of different people which were given on social media and also people would ask others about the best universities. But it was seen that some of the opinions were fake and so it was very difficult to get admitted in a suitable university and it was difficult to meet the exact requirements of the people. So, in that case, opinion mining solves all the problems in getting admission smoothly meeting all the requirements of people since then there is a proof of the opinions to be right or wrong.

After analyzing various research approaches, we found that only opinion mining is not enough. So, we decided for making a recommendation system so that the most suitable university is found. People can have different types of problems. Some may want a good campus where others may want a good quality of education. As a result, recommendation system helps to categorize the demands of different people and then it shows a result regarding the criteria's given by the users. A lot of attention should be paid in designing an opinion mining and recommendation system. Opinion mining should be done very carefully so that the analysis that comes out is efficient and also recommendation depends on that result crucially. In this paper, we tried to design our system by keeping all the significant constraints in mind.

1.3 Research Problem

Opinion mining from the reviews of the user aims to algorithmically identify product aspects and their corresponding opinions from a collection of reviews which are opinionated[3]. Since there is no availability of work done specifically for a recom-

mendation system using opinion mining, so we wanted to use both of them for this system. Also, one may ask about the benefits of opinion mining for creating a service recommendation. So, it has two main subtasks.

For the implementation part, we can just manually add data to the system for training but some issues must be considered in this approach. Also, the data from the websites are not always true and also for adding it manually, it requires a lot of details. Some may be spam and fake reviews[4]. For this reason, we must filter out unnecessary and inefficient data before recommendation. Services are not always in constant demand with time or advancement of technology rather it takes a drastic change in mind. Therefore, it becomes a necessity where opinion is important so that students can get admitted to their desired university with their desired features.

But it is seen that by only relying on the opinions of people, students can never find a solution because without any credibility they will not rely on it. An item is recommended based on the ratings of other users who have similar interests as the user under consideration[5]. Moreover, by facts alone, several students are unable to enquire about certain requirements or problems that they are concerned about. Many of them will have different needs. Some may want a beautiful campus; some may want a good education and also some people may want to get scholarship. So, in that case, for earning the trust of the students, we have to collect relevant and genuine opinions by taking a professional approach with opinion mining. It is seen that by only believing in the opinions, many students get admitted to the universities but later face a lot of problems. So, a solution should be made for these types of issues. Moreover, we claim that with the advancement of technology people will expect many features to be considered and if they do then ordinary recommendation systems are bound to be extinct since people will not need it anymore. Considering all these issues, we have decided to work on recommending the best university to the students according to their desired needs.

1.4 Research Objectives:

- Building a whole new university recommendation system that will use the data generated by opinion mining.
- Including student preference-based recommendation feature using advanced natural language processing toolkits.
- Perform opinion mining by ‘tokenization’ and ‘chunking’ for efficiency.
- Making the recommendation system available for institutions like “Private Universities of Bangladesh”
- Collecting data professionally for generating relevant and efficient recommendations.
- Avoiding biased and improper recommendations from data found on the web or by other means.

1.5 Thesis Orientation

The rest of the thesis is organized as the following order:

- Chapter 1 – Introduction where background information of our research, the problem about which we are researching and research objectives are discussed.
- Chapter 2- Literature Review where previous and related works of the same problem are discussed and reviewed.
- Chapter 3- Data Collection And Analysis where we discuss collection methodologies of our dataset in details and preliminary analysis.
- Chapter 4- Methodology where we discuss the used technologies and workflow of our work in details. Also includes experimental Result Analysis, where we visualize our data and analyze the result obtained from our model.
- Chapter 5- Conclusion and Future works where the conclusion consists of our work till now and future works includes the scope of the improvement.

Chapter 2

Literature Review

Machine learning is a buzz word in the world of science and technology. The term was coined in 1959 by a computer scientist Arthur Samuel in the IBM lab. There are various approaches to solve a problem by utilizing different mechanisms, like predicting future trends and possibilities. Machine learning approaches are used very often to solve different kinds of scenarios. Machine learning is derived from Artificial intelligence as pattern recognition, and NLP is an important aspect of AI [26]. With the help of machine learning, opinion mining can be done. User-generated content can be used for opinion mining as it aims to discover collective and subjective information. This is particularly helpful as it provides user-generated information rather than factual ones, which helps in human decision making and collective information retrieval [35] [36] [37]. Furthermore, machine learning techniques are used to train data sets. The machine learning algorithm is often used as a mathematical expression that represents data in the form of a problem to solve different types of models. Opinion mining is a computational technique that is a subdivision of data mining and computational linguistics, which is used to categorize opinions which we gather from various sources [27]. Work on opinion mining has significantly brought improvement in the effectiveness of work. For example, automatic movie review mining and movies' summarization using the opinion mining system is very resourceful [28]. Opinion mining can also be used to show product reviews by rank according to the intention of the searcher [29]. Market characters can be predicted by monitoring twitter activity to better predict the flow of stocks [30]. There are many challenges when dealing with opinion mining while working on different kinds of data sets [31]. Different kinds of data sets like blogs, movie reviews, chatting data etc. are used for opinion mining [32]. Emotions can be detected from poems or lyrics [33]. According to Liu B. different opinion mining can be done on document level, feature level and sentence level [34]. Different kinds of classifications are used for opinion mining, which can be classified into Supervised and Unsupervised learning. Supervised learning method is used when we want to predict or explain certain data sets. To understand Supervised learning, we need to further break it down into classification and regression. A SVM or Support Vector Machine Classifier takes the data and sorts it into two different categories. The data points are closer to the hyper plane and therefore influence the position and orientation of the hyperplane. [39] [38] [40] [44] [42] [43] [41] [45]. K-nearest neighbor classifier or KNN consider similar types of point that shares similar features in the nearby proximity of a point. The distance to all the points is calculated and, the class is decided based on K nearest neighbors.

[46] [47] [48]. Another learning algorithm that is used for text classification is Naive Bayes classifier. It is based on Bayes' probability theorem, and it accepts the feature variable that is independent of each other [49] [50]. Regression is another model that is used to distinguish different kinds of data using continuous values or values based on a set of preliminary data. Some of the different types of regression are simple linear regression, polynomial regression, support vector regression, decision tree regression, random forest regression. There are a few Unsupervised machine learning methods as well. The two main Unsupervised models are clustering and dimensionality reduction. While clustering works with similar kinds of data, dimensionality reduction works with reducing the number of random variables taken into machine learning algorithm to set principal variables. It removes irrelevant data in order to manage better and reduces complexity. These learning methods are used to find similarity detection, automatic labeling and object segmentation. Natural Language Processing or NLP is not a machine learning method, but it is rather a widely used technique to prepare the text for machine learning. Computers cannot fully understand human language. The machines can be trained to do certain tasks like auto-correction, correct misspelled words, and more. One of the most popular packages for text processing is NLTK or Natural Language Toolkit, created by researchers at Stanford. NLTK is used with a python program to work with human language data as it can easily provide easy to use edges like WordNet, stemming, and more. For numerical representation, one of the easiest ways is to compute the frequency of each word within each text documents [55]. It can easily be explained by matrix. The matrix representation of the text frequencies is commonly called Term Frequency Matrix or TFM. Besides, there are new methods: Neural Network and Deep Learning, Transfer Learning, and Reinforcement learning used in machine learning.

The recommendation system is a significant class of AI calculations that offer applicable recommendations to users. Our work is not quite the same as past investigations regarding opinion mining and recommendation systems. We believe almost no work has been done on our topic. The perspective recognizable proof assignment controlled the free-text surveys in an unsupervised way with no help from polarity audits. In our research work, the identification of opinionated viewpoints begins by recognizing opinion words in the students and guardians survey, and afterward, we look for the angles adjacent to the opinion word. As far as we could know, there has been no endeavor to utilize opinion mining and MCDT to extricate student's and guardian's views in this prospect. In this manner, we use TOPSIS along with opinion mining in our exploration. Multi-criteria Decision Technique (MCDT) is considered an exceptionally dependable technique for positioning multiple choices, depending on a few criteria [51]. In light of a few criteria, Multi-criteria Decision Technique (MCDT) is considered as an excellent approach for positioning multiple other options. In various regions and areas, MCDT strategies have been effectively applied [52][53]. With a little minor departure from existing strategies, different MCDT techniques have been created, and as a result, new parts of exploration is made[52]. In light of the briefest separation to the positive ideal arrangement and the farthest good ways from the negative arrangement, the positioning cycle of techniques is made. Subsequently, a positive-ideal arrangement contains all the best estimations of the criteria in the decision matrix, though negative arrangement contains all the most noticeably terrible qualities [54]. Among various MCDT strategies, TOPSIS

has been widely used in research work. Notwithstanding its effortlessness, it has the least position inversion contrasted with other MCDT strategies, and the means of its cycle are similar, whatever the quantity of criteria is given [53].

Chapter 3

Data Collection And Analysis

We have made two different forms with a set of Questionnaires and collected information about our selected sample population. For this survey, we have used a data collection approach to know the opinion of recent university students, fresh graduates and even those who are going to take admission to universities soon.

In this era of science and technology, crucial features of any university include useful and effective lectures, tutorials, hand notes, library facilities, lab work and practical knowledge. Technical facilities like tools and types of equipment in labs, clubs and practice rooms are also essential.

From an article by Rowley (2003)[13], we were able to learn four crucial cause of getting student feedback:

1. To supply auditable evidence that students have had the chance to pass, discuss their courses which such information is employed to cause improvements
2. To motivate students to execute the implementation of their learnings
3. To give the students a chance to discover their talents and also detect and work with their weaknesses
4. To make sure that they can contribute to the reputation of the university and encourage the development of the others

3.1 Initial Data Processing

3.1.1 Initial Stage

At the very beginning we needed a huge chunk of opinions from various resources so that we can understand the point of the students from different phases from different Universities. To collect the data, we proposed a survey form through which we could get the opinions and later we can work with it.

Survey Form:

As we can see here the forms are segmented into four different sections for four different categories of students. As an example, for an ongoing University student

Bangladesh University Survey: 2020

Are You Planning To Get Admitted To An University?
Or, Are You Already Admitted To An University?
Or, Are You Planning To Transfer Credit To Another University?
Or, Are You Already A Graduate?

Either Way, We Hope That You'll Benefit Us & Other Upcoming Students With Your Valuable Feedback.

Your Current Status:

☐ Looking For An University

☐ University Student

☐ University Transfer Student

☐ A Graduate

Figure 3.1: Survey Form

were asked these set of questions based on their experiences and opinions.

Please Mention On What Basis Did You Choose Your University? (e.g. subject, fees, achievements, facilities, etc.) *

Long answer text

Tell Us About The Problems You're Facing In Your University *

Long answer text

What Is The Best Thing About Your University? *

Long answer text

Figure 3.2: Survey Form

We also decided to include an option to add additional comments. These comments will not only give us the opportunity to develop our database but also contribute to discover new issues regarding the universities.

Final Words

Thank You So Much. We Really Appreciate Your Feedback.

Is There Any Other Opinion That You Wish To Share With Us

Long answer text

Figure 3.3: Survey Form

3.1.2 Second Stage of Collecting Data

After we got our hundreds of unprocessed and unaltered datas we needed more opinions from the students so we took some more approaches to get more valuable datas for research.

1. Face to Face Interviews

We approached respondents personally, either in the university, in labs, and by calling them one to one. Then We asked the respondents a series of questions and noted their responses. With these interviews we were getting more information than the other methods but the major drawback was we had limited

amount of time to do that and getting a good range of data within a very short time was not possible. So, we tried online.

2. Blog and Group Crawling We crawled different Facebook groups that talk about university issues. We also looked for different blogs as well. We got a good chunk data from these online platforms and most of these data are from anonymous sources.

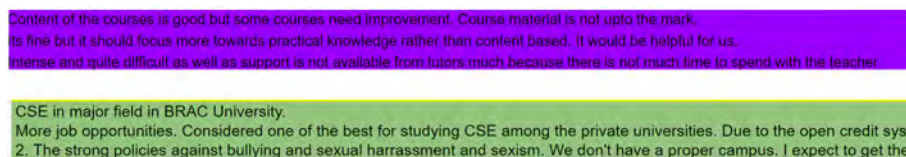
3.1.3 Data exploration and profiling

As we found a large number of unprocessed and unaltered data but these data were not in the format, we needed for our codes to work. So, we needed to format the Data which include looking for trends, outliers, exceptions, incorrect, inconsistent, missing, or skewed information. This part of the exploration was important so that we get the idea that our data isn't biased and have proper information from anonymous sources. We took 600 volunteers' opinions about the universities. We have asked them several unstructured questions to know their problems, challenges, their study satisfaction, facilities and their plan, thoughts, etc. Some of the data were invalid and didn't go with our objectives. So, we needed to validate the data.

3.2 Data Operations

3.2.1 Data Analysis

Data Analysis is an integral step in Machine Learning as the quality of data and the useful information that can be derived from it directly affects the ability of our model to learn; therefore, we must format our data before feeding it into our model. Initially the data we got was full of trends, outliers, exceptions, incorrect, inconsistent, missing, or skewed information. So, we needed to work with the data.



Content of the courses is good but some courses need improvement. Course material is not upto the mark.
Its fine but it should focus more towards practical knowledge rather than content based. It would be helpful for us.
Intense and quite difficult as well as support is not available from tutors much because there is not much time to spend with the teacher.

CSE in major field in BRAC University.
More job opportunities. Considered one of the best for studying CSE among the private universities. Due to the open credit sys
2. The strong policies against bullying and sexual harrassment and sexism. We don't have a proper campus. I expect to get the

Figure 3.4: Data Analysis

3.2.2 Data Profiling

In this cycle we surveyed existing information to improve the quality and bring structure through an organization. It helps in improving the quality and intelligibility to specific guidelines. Most machine learning models neglect to work when the datasets are imbalanced and not very much profiled.

The library should be the strength of a university. It is awesome that we have this perfectly. we can really learn
 Every semester the universities are increasing fees. This has gone beyond our affordability to pay.
 IT IS VERY MUCH USEFUL AND WE CAN PARTICIPATE IN OUR INTERESTS
 Education system should be more developed like world's other university. Too many students is for r
 My department is CSE. The authority doesn't care about how the people study. It only cares about g
 We are facing a lot of problems regarding the high amount of tuition fees so we need a scholarship y
 duration of issuing books from library should be changed from 1week to 2 week and that would be great.
 IT IS ALSO GOOD and interesting and so I like to study here an love to be here
 Descency is absent. People stare at us like we're something to eat. I feel like I'm in a
 School of Law. Always wanted to be a lawyer. Best private university for Law
 Facilities - 40/100 Good luck
 I hate my University. It is not good at all and so for this reason I do not want to stud
 Our exam pattern of University is very nice but paper checking is not at all go

Figure 3.5: Data Profiling

Data cleansing guarantees information is perfect, exhaustive, blunder free with exact data. We removed the opinion that we felt like was biased and obsolete. we also paraphrased some opinions, removed extra spacing, made grammatical corrections while working with all data.

Data transformation is changing information to make it homogeneous. Information like locations, names, and other field types are spoken to in various organizations and information change helps in normalizing and normalizing this.

Data anonymization is the way toward eliminating or encoding individual data from the datasets to ensure privacy. We removed personal information to maintain the security.

Data augmentation we need to diversify the data available for our training models. We rewrote new information without extracting any information.

3.3 Transforming Data and Data Modeling

After gathering all the data and information we tried to justify all the information and transformed data for our data model.

Chapter 4

Methodology

4.1 Data Manipulation

4.1.1 Tokenization

To extract opinions, our primary step will be to analyze each opinion and extract useful information from it. To do that, we will tokenize every opinion into tokens of words. Let us take the first opinion from our dataset.

Library should be the strength of a university. It is awesome that we have this perfectly. We can really learn practically. However, we lack open space and place to relax also.

We have used the NLTK library of python to tokenize each word from a sentence using its method 'word.tokenize' method. Other algorithms or methods can be used as well but we have to make sure that every single word is split into tokens and appended to certain indexes of an array or list. Because we will need these tokens for the upcoming procedures. Once every opinion is split into tokens of words, we can now analyze each token and fetch relevant information.

4.1.2 POS Tagging

To fetch relevant information from every opinion, we should also know which tokens are essential for us and which are not.

“My name is Rahim. I think a University should have proper facilities in a library.”

In the above opinion, the first sentence is not an opinion. We do not need such data for our algorithm. Moreover, “I think a” part of the second sentence is also irrelevant for opinion mining. But we cannot identify such tokens among thousands of opinions if we don't tokenize and tag each of them. Then with our algorithm expressions, we can filter out relevant opinions from all sentences. So, in this phase, we will tag all tokens from step 1 according to their grammar (Parts of Speech). We will use the NLTK method 'pos_tag'. This method classifies each token according

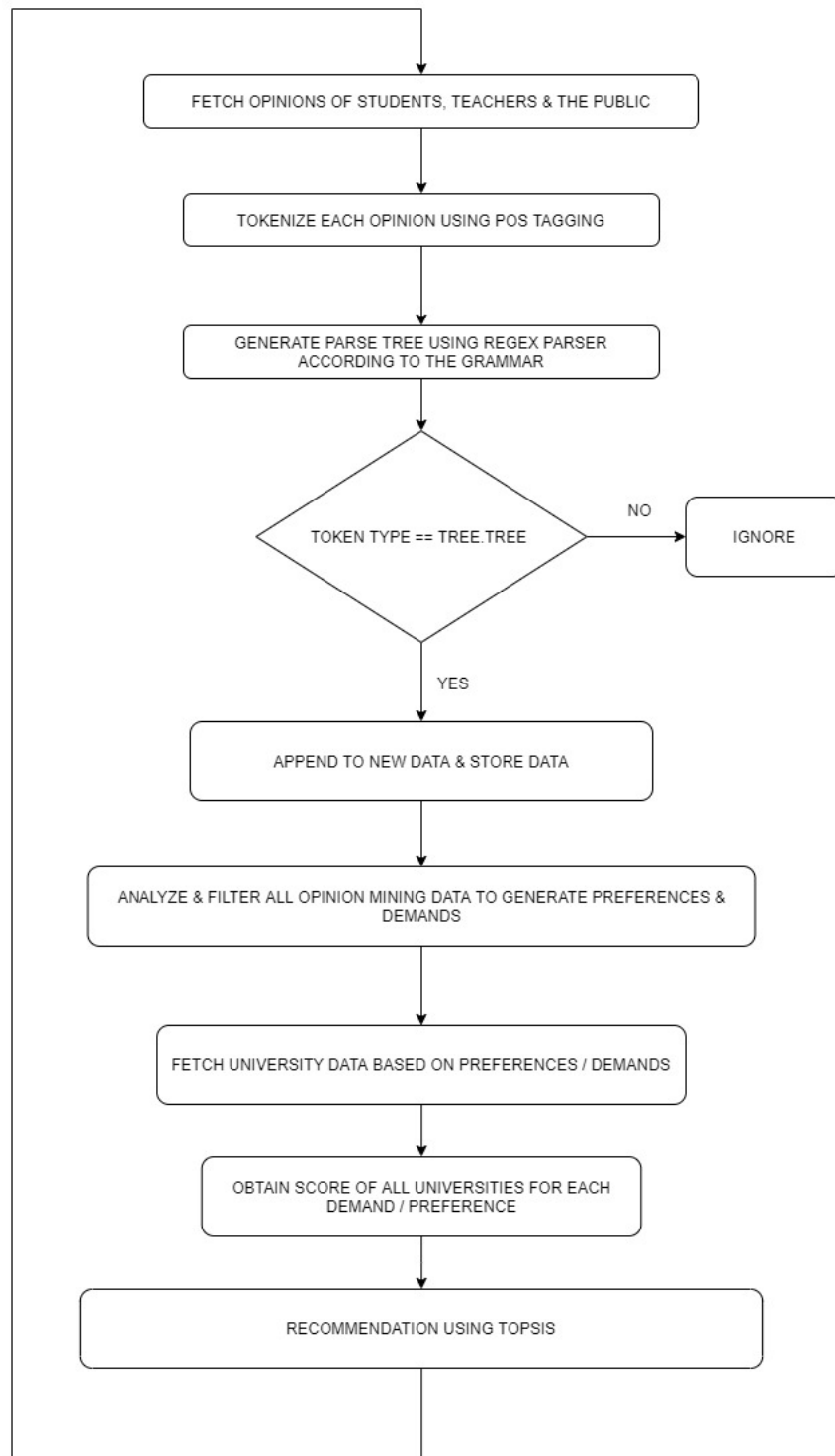


Figure 4.1: Methodology

to the grammar.

The tokens were generated from the original opinions in step 1 and then the tokens are tagged in step 2. The meaning of the classifications can be generated from the NLTK POS table.

NLTK POS TAGGING 1	
Attributes	Grammar
RB	ADVERB e.g. VERY, SILENTLY
RBR	ADVERB, COMPARATIVE e.g. 'BETTER'
RBS	ADVERB, SUPERLATIVE e.g. 'BEST'
RP	PARTICLE e.g. GIVE UP
TO	TO GO e.g. 'TO' THE STORE.
UH	INTERJECTION e.g. 'ERRRRRRRRM'
VB	VERB, BASE FORM e.g. TAKES
VBD	VERB, PAST TENSE e.g. TOOK
VBG	VERB, GERUND/PRESENT PARTICIPLE e.g. TAKING
VCN	VERB, PAST PARTICIPLE e.g. TAKEN
VBP	VERB, SING. PRESENT, NON-3D e.g. TAKES
VBZ	VERB, 3RD PERSON SING. PRESENT e.g. TAKES
WDT	WH-DETERMINER e.g. WHICH
WP	WH-PRONOUN e.g. WHO, WHAT
WP\$	POSSESSIVE WH-PRONOUN e.g. WHOSE
WRB	WH-ABVERB e.g. WHERE, WHEN
POS	POSSESSIVE ENDING e.g. PARENT'S
PRP	PERSONAL PRONOUN e.g. I, HE, SHE

Table 4.1: NLTK POS TAGGING

NLTK POS TAGGING 2	
Attributes	Grammar
CC	COORDINATING CONJUNCTION
CD	CARDINAL DIGIT
DT	DETERMINER
EX	EXISTENTIAL THERE e.g. 'THERE IS'
FW	FOREIGN WORD
IN	PREPOSITION/SUBORDINATING CONJUNCTION
JJ	ADJECTIVE e.g. 'BIG'
JJR	ADJECTIVE, COMPARATIVE e.g. 'BIGGER'
JJS	ADJECTIVE, SUPERLATIVE e.g. 'BIGGEST'
LS	LIST MARKER e.g. 1)
MD	MODAL e.g. COULD, WILL
NN	NOUN, SINGULAR e.g. 'DESK'
NNS	NOUN PLURAL e.g. 'DESKS'
NNP	PROPER NOUN, SINGULAR e.g. 'HARRISON'
NNPS	PROPER NOUN, PLURAL e.g. 'AMERICANS'
PDT	PREDETERMINER e.g. 'ALL THE KIDS'
PRP\$	POSSESSIVE PRONOUN e.g. MY, HIS, HERS

Table 4.2: NLTK POS TAGGING 2

Therefore, from the NLTK POS Table, We can Say that 'Library' is NNP (Singular Proper Noun), 'could' is Modal, etc. In this way, every single opinion is tokenized and tagged in this phase.

4.2 Opinion Mining

As we complete the data processing phase, we move on to the next phase of opinion mining. We will execute three more steps here.

1. Chinking
2. Parsing
3. Analysis

Chinking

To extract opinions, we have to filter out all unnecessary words from our opinions. There are many ways to do it but we will do chinking. 'Chinking' means filtering out certain words based on an expression or grammar. The grammar that we will use is shown below:

With this expression formula, we can fetch a particular segment or phrase from the sentences. We can implement it using regular expressions.

Expression: ‘<DT | PP | CD>? <JJ || JJR | JJS> * <NN |NNS |PRP |NNP |IN |PRP \ \$> + <VBD |VBZ |VBN |VBP |IN > * <JJ |RB>*<PRP |NN |NNS>*

Figure 4.2: Grammar For Opinion Mining

Here, < DT |PP |CD > ? means that we will filter out the sentence based on either a determiner, preposition or a cardinal number. Thereafter, < JJ ||JJR |JJS > * means 0 or more adjectives may follow the sentence afterward. Moreover, < NN |NNS |PRP |NNP |IN |PRP \ \$ > + means 1 or more nouns/prepositions occurrences will be the condition for the validity of the sentence. Furthermore, < VBD |VBZ |VBN |VBP |IN > allows one kind of verb and so on. This grammar will give us useful segments from the sentences.

Let us consider the opinion below:

“The Library should be the strength of a university. It is awesome that

we have this perfectly. we can really learn practically. However, we lack an open space and a relaxing place also.”

Opinion:

[‘The’, ‘library’], [‘the’, ‘strength’, ‘of’], [‘a’, ‘university’]
 [‘It’, ‘is’, ‘awesome’], [‘that’, ‘we’, ‘have’]
 [‘we’, ‘lack’, ‘open’, ‘space’], [‘a’, ‘relaxing’, ‘place’, ‘also’]

These are the results from one single opinion, we can get information like that a library is an important factor according to this person. The person is also expressing the opinion of the shortage of space in the university. In this way, we will get several opinions from all the sentences in the dataset.

4.2.1 Opinion Mining

We will save all the parsed data to a CSV file so that we can use them for analysis. In step 1, we received an opinion regarding the ‘library’. After we analyzing the whole data, we found out that ‘library’ is an important feature demanded by the students and we found some expected qualities like ‘silence’, ‘books’, ‘e-books’ and ‘reading room’ etc. Similarly, we found out concerns regarding ‘campus’. We analyzed the data based on the most repeated noun and adjective using regular expressions and counter. In this way, we can generate much more important information if we collect more data.

As we know, most universities promote their facilities, accommodation, fees, etc. But very few universities will provide concerns regarding the library. So, all the students who desire good library features will miss out universities having great library

facilities if he or she searches simply based on ranking or web search.

We can use regular expressions and the counter library to extract relevant features. However, there can be several other alternative strategies to extract information. The most important part is tokenizing, tagging the sentences properly and parsing the tokens efficiently for the opinion mining process.

4.2.2 Opinion Wise Cardinality Determination

From the above discussions, we know about one opinion extraction only. The same process is repeated for all the opinions in the dataset. This gave us a full dataset containing extracted opinions. The next step is to calculate which issue or topics are most often repeated. By doing so, we can figure out what are the main concerns regarding universities in the mindset of the mass.

Using **Textblob** in python, we extract all the noun and noun phrases from the whole dataset of extracted opinions. After a little filtering, we were able to generate a list of concerns that were repeated most in the whole dataset. For this demonstration, we will show 7 most repeated concerns that we obtained:

Opinion Issue	Frequency
Subject/Major	86
Proper Campus	72
Faculties	64
Education System	46
Tuition Fee	45
Features/Extra Curriculum	35
Labs & Facilities	32

Figure 4.3: Results: Top 7 issues/demands regarding universities

It is important to know that these values will change every time the dataset is updated. So, none of these results are constant. As shown in the above figure, Subject, Campus, Faculties, Education System, Tuition Fees, Extra Curriculum and Lab Facilities are respectively the most repeated concerns in our primary dataset containing 1000+ opinions. We will now use these demands on our recommendation system. It is extremely important to note that these demands may change every time the data set is updated. Therefore, one of the biggest advantages of our system is that it will always be up to date. We will now proceed to the development of a very intelligent recommendation system unlike normal university articles or ranking websites for both students and the teachers.

4.3 MCDT Ranking and University Recommendation System

4.3.1 Optimized Demand Based Dataset

Subject, Campus, Faculties, Education System, Tuition Fees, Extra Curriculum and Lab Facilities were the top 7 demands we obtained from our initial opinion mining dataset analysis. Now, We will collect ratings and opinions from core sources (e.g. official website, university officials and also the opinion of the students and faculties of the particular university).

Based on these data, all universities will get a rating or score for each demand separately. For this research we will take the top 14 universities from the list with the highest scores among all private universities. Up to 100 people will generate a datasheet containing scores for each demand for every university. Afterwards, a final optimized datasheet is generated based on the average of all the datasheets. It is shown below:

Result:

1	University	Subject (R)	Campus (R)	Faculties (R)	Edu. System (R)	Tuition Fee (R)	Extra Curriculum(R)	Lab Facilities (R)
2	BRAC University	8.5	6.5	8.5	8.4	6.6	9.5	7.8
3	North South University	8.7	9	8.5	8.8	6.8	8.5	7.5
4	East West University	8.4	9	7.5	8.2	7.5	8	7.5
5	AIUB	8.3	9.3	7.5	5.5	8.2	8.2	7.5
6	IUB	8.2	7.5	8.2	8.2	7.8	7.5	7.5
7	UIU	7.2	8	7.2	6.9	7.5	8.2	7.5
8	Daffodil University	7	7.5	7	6	7.5	6.9	7.5
9	University of Asia Pacific	6.9	6.5	6.9	5	7.5	6	5.5
10	IUBAT	6	6.8	6	5.5	6.9	6.9	5.5
11	Dhaka International University	6.6	7	6.6	4.5	8.5	4.5	5.5
12	Manarat International University	6	7.5	6	5	8	5	5.5
13	Southeast University	6.2	7	6.9	4.5	8	4.5	5.5
14	Green University	5	6.3	5	4	8.5	4	5.5
15	Prime University	4.3	4.2	5.8	4.5	8.5	5.5	4.5

Figure 4.4: University Scores Per Demand

4.3.2 Use TOPSIS Decision Making Algorithm For MCDT Ranking

Now that we have obtained a dataset ready to use for recommendation, any person seeking a university some particular demands can search or their own custom set of recommendations based on their required demands instead of a website that contains university ratings.

For example, according to 4icu.org (Unirank BD), BRAC University is ranked number 1 among the private universities in Bangladesh. But in the figure of university scores, we can see that BRAC University have a lower score then some universities in campus and tuition fees. Now suppose there is a student named Fariha who wants to get admitted to a university with a very good and permanent campus, education system and at the same time, tuition fees matter to her. According to her demands, AIUB and East West University are undoubtedly better choices then BRAC University. But if she only follows the ratings, then she'll get admitted to BRAC University and then find out that her demands aren't satisfied properly. For such scenarios, our recommendation system is undoubtedly a revolutionary solution!

But which Algorithm will we use for the recommendation? If the algorithm is not efficient then our recommendation cannot be worth the glory that it claims to have. Therefore, to make sure that the preference generation is reliable, we will use TOPSIS (Technique of Order Preference Similarity to the Ideal Solution).

This method is well recognized not only for finding the shortest distance to the most idea solution but also the longest distance to the most non-ideal solution. We will use TOPSIS method to determine the most idea solution among the 14 universities (Figure) for this demonstration.

For this demonstration we will use 3 choices of Fariha which are Campus (2), Edu. System (4 and Tuition Fee (5). The users can make as many choices as they want and the number of available demands will be increased further with time. After the choices are made, the values of the chosen columns will be sent to the TOPSIS method for the analysis to get 5 best choices among all 14 (The number of choices and Universities can be increased or decreased). The flowchart showing the algorithm of TOPSIS is provided below:

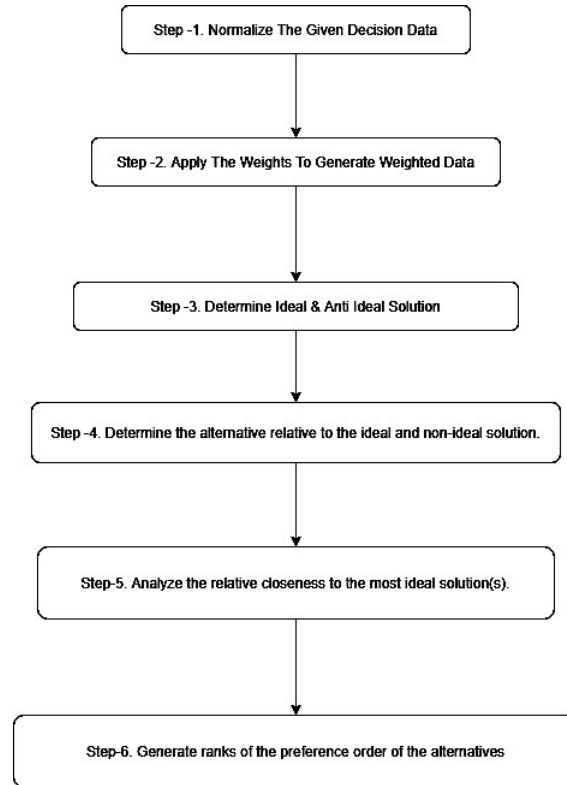


Figure 4.5: TOPSIS decision making algorithm

The step by step process of TOPSIS:

Step 1 - Normalize The Given Decision Data:

This step transforms various demand dimensions obtained from opinion mining into non-dimensional demands, which allows comparisons across criteria. For standardizing, each column of decision matrix, is divided by root of sum of square of the respective one.

Step 2 - Apply The Weights To Generate Weighted Data:

In this step, a weighted standardized decision matrix based on the demands by multiplying attributes weight to each of the ratings.

Step 3 – Determine Ideal Solution and Non-Ideal Solution:

The ideal solution is a set of maximum values for each criteria obtained after the calculations. Similarly, the set of minimum values for each criteria is the Non-Ideal solution.

Step 4 – Determine The Alternative Relative To The Ideal and Non-Ideal Solution:

The separation (S_i^*) is calculated from the difference between ideal and non-ideal solutions obtained from the previous step.

Step 5 - Analyze The Most Relative Closeness To The Most Ideal Solution:

This is simply the maximum value obtained from the differences of the relative separations to ideal and non-ideal solutions i.e. max closeness value.

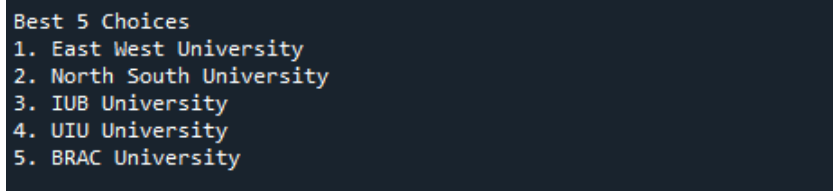
Step 6 - Generate Ranks of the Preference Orders of the Alternatives:

The entity with the maximum closeness value gets the maximum rank. So, if East West University gets the max closeness value compared to the others it will rank first place.

4.4 Preliminary Results And Comparison

We have successfully completed all steps of the entire project. Now we will look in to the output generated by the TOPSIS algorithm that we executed on Step -2 in the recommendation system.

Result:



```
Best 5 Choices
1. East West University
2. North South University
3. IUB University
4. UIU University
5. BRAC University
```

Figure 4.6: Obtained Results

These are the results obtained from our system generated by customized TOPSIS algorithm using the values of the preferences only. In TOPSIS, we can also include what the user doesn't prefer. For example, if Fariha desires a University without good faculties then she can choose faculties as an anti-ideal preference. But for this

demonstration, no anti ideal preference was provided but it is an extra feature our system. We will now compare these results with the rankings of these universities according to 4icu.org (Unirank BD).

sort by: rank a-z town oldest		
filter by: public private non-profit for-profit		
Rank ▼	University ▼	Town ▼
3	BRAC University	Dhaka ...
4	Daffodil International University	Dhaka
5	North South University	Dhaka
8	American International University-Bangladesh	Dhaka
10	Asian University for Women	Chittagong
11	Independent University, Bangladesh	Dhaka ...
13	United International University	Dhaka
14	Ahsanullah University of Science and Technology	Dhaka
17	East West University	Dhaka

Figure 4.7: Private University Rankings according to 4icu.org (Unirank BD) until 6th September,2020

According to 4icu.org (Unirank BD), East West is ranked 9th, North South is ranked 3rd, IUW is ranked 6th, UIU is ranked 7th and BRAC University is ranked 1st. As you can see, the lowest in the rankings happen to be the most ideal solution for Farina's demands rather than BRAC University although it is number 1 in the rankings! Moreover, BRAC University is the 5th in the result of our analysis as per Fariha's demands. In this way, anyone can get recommendation for universities as per their custom preferences and demands using our brand-new opinion mining and recommendation system.

The dataset of our system is generated from opinion of students, teachers and the public. With addition of more data, the system will constantly keep improving and stay up to date. If new categories of preferences are demanded in the future, people will certainly talk about it and thus eventually, it will be included in our dataset. When that happens, our recommendation dataset will also be updated and will include the new preferences. Similarly, if any preferences become old or unwanted, with time they will lose position in our datasets. In a nutshell, one of the biggest advantages of this system is that to stay up to date all we need to do is update our datasets from time to time.

As a result, we have already added features for the user to provide their own opinions regarding universities. These will be added to our dataset by our team members. Similarly, if any user has a preference that is not available on our system, our team will also make note of it and work on collecting data on that particular preference. We will constantly be working with our project to try our best to help students get the best possible university recommendations as per their demands.

Chapter 5

Conclusion

University life is a phase of life through which almost every student goes through. It is also extremely crucial for the road to academic or professional success. Such an impactful institute should be chosen with preference and according to self-demands as much as possible. Otherwise, not only is there a chance of backing out or slacking off but also a possibility of failure to make ambitions set for self-come true. Hence, a service must be available that guides the student properly. The system must be fast, accurate and light-weight to achieve higher efficiency. The system must be able to recommend a university based on the student's choices. Our system helps out every student and even teachers to reach the universities of their own field, preferences and demands which leads to accurate prediction and recommendation. Opinion Mining was done through Supervised Learning and TOPSIS was used to develop an efficient recommendation system of admission to university. This approach improves efficiency in classification of relevant universities and performs substantially better than the normal reviews or ratings which are not always suitable for specific individuals. This system will be developed and improved from time to time. Though positive reviews from students and teachers from demonstration of results were found but it does not yet meet the satisfaction level. Our goal is to ensure that people will actually be benefitted from our system and eventually we can become a part of their story on which they make their ambitions come true.

Bibliography

- [1] Dr.B.Radha and V.Meera 2017, "*International Journal of Innovative Research in Computer and Communication Engineering*", (An ISO 3297: 2007 Certified Organization)', Vol. 5, Issue 2, February 2017
- [2] Bing Liu 2008, "*Opinion Mining*", Association, A. L. (n.d.)
- [3] Amani Khalaf Samha 2016, "*ASPECT-BASED OPINION MINING FROM CUSTOMER REVIEWS*"
- [4] Haseena Rahman 2014, "*Opinion Mining and Sentiment Analysis - Challenges and Applications* ", Volume 3, Issue 5, May 2014
- [5] Satya Prakash Sahu, Anand Nautial and Mahendra Prasad 2014, "*Machine Learning Algorithms for Recommender System*", a comparative analysis Volume 6–Issue 2, 97-100, 2017, ISSN:-2319–8656
- [6] Altug Akay, Andrei Dragomir and Björn-Erik Erlandsson 2013, "*A Novel Data-Mining Approach Leveraging Social Media to Monitor Consumer Opinion of Sitagliptin*" , Volume: 19, Issue-1, Jan- 2015
- [7] Hsinchun Chen and David Zimbra 2010, "*AI and Opinion Mining*" , Volume: 25, Issue-3, May-June 2010
- [8] Emmanouil Perakakis, George Mastorakis and Ioannis Kopanakis 2019, "*Social Media Monitoring: An Innovative Intelligent Approach*"
- [9] Xiangyu Hu; Lemin Li, Tingmin Wu; Xiaoxiang Ai, Jie Gu and Sheng Wen 2018, "*Every word is valuable: Studied influence of negative words that spread during the election period in social media*"
- [10] Ha Huy Cuong Nguyen 2018, "*Opinion Mining: Using Machine Learning Techniques*"
- [11] Jeannette Paschen 2019, "*Investigating the emotional appeal of fake news using artificial intelligence and human contributions*"
- [12] Kate Kelley, Belinda Clark, Vivienne Brown and John Sitzia 2003, "*Good practice in the conduct and reporting of survey research*"
- [13] Aldridge, S. and Rowley, J. 1998, "*Measuring customer satisfaction in higher education*" , Quality Assurance in Education, Vol. 6 No. 4, pp. 197-204.

- [14] Angur, M.G. 1998, "*Service quality measurements in a developing economy: SERVQUAL versus SERVPERF*" , Journal of Customer Service in Marketing Management, Vol. 4 No. 3, pp. 47-60.
- [15] Asubonteng, P., McCleary, K.J. and Swan, J.E. 1996, "*SERVQUAL revisited: a critical review of service quality*" , The Journal of Services Marketing, Vol. 10 No. 6, pp. 62-81.
- [16] Banwet, D.K. and Datta, B 2003, "*A study of the effect of perceived lecture quality on post-lecture intentions*" , Work Study, Vol. 52 No. 5, pp. 234-43.
- [17] Crawford, F. 1991, Total Quality Management, Committee of Vice-Chancellors and Principals, occasional paper (London, December), cited in Hill, F.M. 1995, "*Managing service quality in higher education: the role of the student as primary consumer*" , Quality Assurance in Education, Vol. 3 No. 3, pp. 10-21.
- [18] Deming, W.E. 1982, "*Out of the Crisis*" , Massachusetts Institute of Technology, Cambridge, MA.
- [19] Dillon, W.R., Madden, T.J. and Firtle, N.H. 1993, "*Essentials of Marketing Research*" , Irwin, Boston, MA.
- [20] Martilla, J.A., James and J.C. 1977, "*Importance – performance analysis*" , Journal of Marketing, January, pp. 77-9.
- [21] Soujanya Poria, Alexander Gelbukh, Amir Hussain, Newton Howard, Dipankar Das and Sivaaji Bandyopadhyay 2013, "*Enhanced SenticNet with Affective Labels for Concept-Based Opinion Mining*" , Volume: 28, Issue-2, March-April 2013
- [22] Quan Fang; Changsheng Xu; Jitao Sang; M. Shamim Hossain; Ghulam Muhammad 2015, "*Word-of-Mouth Understanding: Entity-Centric Multimodal Aspect-Opinion Mining in Social Media*" , Volume: 17, Issue-12, Dec-2015
- [23] Xinyu Chen, Youngwood Cho and Suk-Young Jang 2015, "*Crime Prediction Using Twitter Sentiment and Weather*" , 24-24 April 2015
- [24] Rui Wang; Deyu Zhou, Mingmin Jiang, Jiasheng Si and Yang 2019, "*A Survey on Opinion Mining: From Stance to Product Aspect*" , Volume: 7, 21 March 2019
- [25] Zhen Hai, Kuiyu Chang, Jung-Jae Kim and Christopher C. Yang 2013, "*Identifying Features in Opinion Mining via Intrinsic and Extrinsic Domain Relevance*" , Volume: 26, Issue-3, March 2014
- [26] Witten and E. Frank, Data mining: Practical machine learning tools, San Francisco: Morgan Kaufmann, 2005
- [27] H. Chen and D. Zimbra, "AI and Opinion mining," IEEE Intelligent Systems, May/June 2010, pp.74, 80.
- [28] L. Zhuang, F. Jing and XY. Zhu, "Movie Review Mining and Summarization," Proceedings of the 15th ACM International Conference on Information and Knowledge Management, November 2006, pp.43-50.

- [29] Yune, H., H. Kim, J.Y. Chang, “An Efficient Search Method of Product Review using Opinion Mining Techniques,” *Journal of KIISE: Computing Practices and Letters*, Vol.16, No.2, 2010.2, pp.222-226.
- [30] J. Bollen, H. Mao and X. Zeng, “Twitter mood predicts the stock market,” *Journal of Computational Science*, Vol.2, 2011, pp.1–8
- [31] Pang B. and Lee L., “Opinion Mining and Sentiment Analysis”, *Foundations and Trends in Information Retrieval*, vol. 2, no. 1-2,2008, pp. 1–135
- [32] Kaur and Saini, ”On classifying sentiments and mining opinions,” *International Journal of Emerging Technologies in Computational and Applied Sciences (IJET-CAS)*, vol. 9, no. 3, pp. 210-214, 2014.
- [33] J. Saini, “Polarity Determination using Opinion Mining in Stocks and Shares-advertising Unsolicited Bulk e-mails”, *International Journal of Engineering Innovation Research (ISSN: 2277-5668)*, pp. 86-92., 2012.
- [34] B. Liu, *Sentiment Analysis and Opinion Mining*, Morgan Claypool Publishers, 2012.
- [35] B. Liu and L. Zhang, “A survey of opinion mining and sentiment analysis,” in *Mining Text Data*. New York, NY, USA: Springer, 2012, pp. 415–463.
- [36] K. L. Keller, “Conceptualizing, measuring, and managing customer-based brand equity,” *J. Marketing*, vol. 1, no. 1, pp. 1–22, 1993.
- [37] D. Carmel, N. Zwerdling, I. Guy, S. Ofek-Koifman, N.Har’El, I. Ronen, E. Uziel, S. Yogev, and S. Chernov, “Personalized social search based on the user’s social network,” in *Proc. CIKM*, 2009, pp. 1227–1236.
- [38] B. Pang, L. Lee and S. Vaithyanathan, ”Sentiment Classification Using Machine Learning Techniques,” *Conference on Empirical Methods In NLP*, pp. 79-86, 2002.
- [39] F. Can, P. Duygulu and M. Kalpakli, ”Automatic Categorization of Ottoman Literary Texts by Poet and Time Period,” *Computer and Information Sciences-II*, pp. 51-57, 2012.
- [40] S. Hamidi, M. Ghaemmaghami and F. Razzazi, ”Automatic meter classification in Persian poetries using SVM,” *Conference on IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, p. 563 – 567, 2009.
- [41] Jamal N., Mohd M. and Noah S.A., “Poetry Classification Using Support Vector Machines” *Journal of Computer Science*, vol. 8, no.9, 2012, pp.1441-1446.
- [42] N. Ghneim, O. Alsharif and D. Alshamaa, ”Emotion Classification in Arabic Poetry using Machine Learning,” *International Journal of Computer Applications*, vol. 5, 2013.
- [43] Poria S., Cambriab E., Wintersteinc G., Huang G.B., “Sentic Patterns: Dependency-Based Rules for Concept-Level Sentiment Analysis,” *Knowledge-Based Systems*, 2014, pp. 1–32.

- [44] Kumar V. and Minz S., "Poem classification using machine learning" Proceedings published in Advances Volume 236, 2014, pp. 675-682.
- [45] Z. He, W. Liang, L. Li and Y. Tian, "SVM-Based Classification Method for Poetry Style," International Conference on Machine Learning and Cybernetics, vol. 5, p. 2936 – 2940, 2007.
- [46] B. Li, J. Lu, J. Yao and Q. Zhu, "Automated Essay Scoring Using the KNN Algorithm," vol. 1, pp. 735-738, 2008.
- [47] Kumar V. and Minz S., "Poem classification using machine learning" Proceedings published in Advances Volume 236, 2014, pp. 675-682.
- [48] Lee H.Y. and Renganathan H. "Chinese Sentiment Analysis Using Maximum Entropy" Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP), IJCNLP 2011, pp.89–93.
- [49] A. Bagheri, M. Saraee and F. Jong, "Sentiment Classification in Persian: Introducing a mutual information-based method for feature selection," 21st Iranian Conference on Electrical Engineering (ICEE), pp. 1-6, 2013.
- [50] S. Noah and F. Ismail, "Automatic Classifications of Malay Proverbs Using Naive Bayesian Algorithm," Information Technology Journal, vol. 7, pp. 1016-1022, 2008.
- [51] Umm-E-Habiba and S. Asghar, "A survey on Multi-criteria decision-making approaches," 2009 International Conference on Emerging Technologies, ICET 2009, 2009, pp. 321–325.
- [52] M. Velasquez and P. T. Hester, "An Analysis of Multi-Criteria Decision Making Methods," Int. J. Oper. Res., vol. 10, no. 2, pp. 56–66, 2013.
- [53] H.-S. Shih, H.-J. Shyur, and E. S. Lee, "An extension of TOPSIS for group decision making," Math. Comput. Model., vol. 45, no. 7–8, pp. 801–813, Apr. 2007.
- [54] C. Bai, D. Dhavale, and J. Sarkis, "Integrating Fuzzy C-Means and TOPSIS for performance evaluation: An application and comparative analysis," Expert Syst. Appl., vol. 41, no. 9, pp. 4186–4196, Jul. 2014.
- [55] N. Hardeniya, J. Perkins, D. Chopra, N. Joshi, and. I. Mathur, Natural Language Processing: Python and NLTK, Birmingham: Packt Publishing Ltd., 2016.