SERVICE BASED OPINION MINING APPLICATION FOR ANALYZING CUSTOMER FEEDBACK: CASE STUDY OF HUDUMA CENTRES

Prepared By: MOSES TILE

Reg No: P58/61728/2010

MSC COMPUTER SCIENCE

UNIVERSITY OF NAIROBI

2019

This research project report is submitted in partial fulfilment of the requirements for the degree of Master of Science Computer Science of the University of Nairobi. This report is the product of my own work except where indicated in the text.

DECLARATION

I declare that this project is my original work and that it has not been submitted elsewhere for examination, award of degree or publication. Where other people's work or my own work has been used, this has properly been acknowledged and referenced in accordance with University of Nairobi's requirements.

__________________________ __________________________

Moses Tile Date

P58/61728/2010

Declaration by Supervisor

This research project is submitted for examination with my approval as the research supervisor.

__________________________ __________________________

Dr. Wanjiku Ng'ang'a Date

School of Computing and Informatics

ACKNOWLEDGEMENT

I would like to thank the Almighty God for the strength and guidance He has given me throughout the whole research period. My sincere gratitude goes to my supervisor Dr. Wanjiku Ng'ang'a for the support and guidance she has shown me all through the project. Without her am sure I would not have attained my objectives. I would also like to sincerely thank my family for their moral and financial support without which I would not have been successful. Lastly I would like to thank all those who have assisted me in one way or the other in working on this project, including lecturers Dr. Mburu, Dr. Chepken and Prof. Miriti; my colleagues for their ideas and support all through and not forgetting all those who sacrificed their time. You have all made my project a success.

ABSTRACT

Huduma Centres are fast becoming the government front office where citizens are guaranteed access to a majority of citizen services under the same roof. It was awarded and recognised for offering the best customer service in the public sector in the year 2014. To maintain such high quality service, there is need to have a way of capturing an analyse customer feedback to enable management make decisions that are geared towards continuous improvement and maintenance of excellent services. Customer satisfaction is a critical factor that dictates whether an organization will be a success or not. Huduma centres need to implement strategies that can help them predict customer issues which should enable them identify and solve any slight change in customer feedback in real time. This study focuses on enabling Huduma Centres analyse customer feedback to understand the customer's opinion about their service being offered. The previous strategies used to capture feedback include the use of notes and the "One click feedback" strategy that requires a customer to rank or rate the service.

In this research, real time analysis of customer feedback is achieved through the utilisation of a hybrid model that incorporates both the supervised and the unsupervised machine learning techniques. This hybrid model relies on the Lexicons and the use of Naïve Bayes Machine learning that assumes that every feature and each word in a review being classified is independent of any other feature. The data is first extracted from Twitter, then subjected to a SOCAL algorithm that generates a semantic orientation of either positive, negative or neutral of a given opinion or feedback. Once the semantic orientation is done, the data is divided into two sets; a training set and a test set. The training set is used train the Naïve Bayes model to classify texts and it was composed positive and negative reviews. A gold standard dataset was then used to evaluate and measure the accuracy, precise, recall and F1 score of the hybrid model. This research also explored the performance of the lexicons and the Naïve Bayes Models separately to ascertain the performance of Hybrid model in comparison to the two models.

The results show that in all the evaluation tests done on the hybrid model scored higher. On the accuracy, the hybrid model scored **67.27%** which showed a higher degree of accuracy than the Lexicons and the Naïve Bayes models by 4% points. The same also applied to the precision and recall where the hybrid model scored **66.99%** for precision and **66.57%** for recall which was higher by 3.53 and 4.6 percentage points respectively. The F1 score on the other hand gave a score of **66.78%.** Based on the above results, it is concluded that the hybrid model is best fit to be used for sentiment analysis due to its higher accuracy, precision, recall and F1 scores than the Lexicons and the Naïve Bayes when implemented separately. However further improvements on the above scores can be explored by use of ensembles where several models are combined through the use of boosting or bagging methods to smooth out predictions and combine them into one hybrid model with a best fit.

Keywords: *Opinion Mining, Sentiment Analysis, Supervised Machine Learning, Text classification, Customer feedback*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 Background of the Study

Organizations operating within the public sector-health care organizations, local government, police, emergency services, government agencies have come to realize that customer service and quality are critical strategic issues in the late 1990s. According to Hsiao & Jie-Shin (2008), public sectors are established to serve people and thus service quality it provides is dictated by people satisfaction or contact experience, amid demands of various people, administrative institution service could satisfy people only by innovation and continuous improvement of service quality.

In Kenya, when the 2010 Constitution was enacted, it placed an obligation upon the Government to provide quality public services and information as a fundamental and constitutional right of all Kenyans (The standard, 2016). In regard to this, The Huduma Kenya Programme was launched in November 2013 as a part of the Government of Kenya public service reform programme. The aim of the programme is to use innovation to transform the quality of public service delivery in Kenya. The Huduma Centres are fast becoming the government front office where citizens are guaranteed they will access a majority of citizen services under the same roof, receive good quality. It was also awarded and recognised for offering the best customer service in the public sector in the year 2014.

According to Beevers (2006), the levels of customer service provided by an organisation are directly dependent on the strength of the organisation's internal leadership and the ability of the organisation's leaders to foster a culture of customer service excellence and gain commitment to that culture throughout the entire organisation. This clearly points out a way in which an organisation can use to build trust and loyalty in its citizens. For Huduma Centres to foster and maintain their reputation with the Kenyan citizens they need to implement strategies that will enable them to identify pending citizen issues or complains and hence be able to continually offer quality service. By this they will not only painting their image but also the government's.

1.2 Problem statement

Through the years, research has shown that the main challenge facing most service industries has been real time identification and resolving of customer issues. Although customer satisfaction is a strategic weapon that can be used by enterprises to ensure increased profits and market share, most have not recognised it. Many of these enterprises continually use outdated and unreliable measures of customer satisfaction like listening to their customers and use of other review techniques that only describes the customers' state of mind. These methods currently being used to measure customer satisfaction are slow and customer issues are identified long after the customer has left. Further due to increased use of social media and other content review sites, information has increased and analysis of such has become hectic.

Most of the feedback received is at document level whereby a scaling system is used to determine the level of satisfaction of its clients. Huduma centres have implemented a five star rating technique to capture customer's sentiments and provided a contact number and email through which citizens can leave their opinions. Citizens can either call the number or send an email to make their issues known to these centres. This method has many challenges as one would not be able to have the exact details of one's opinion. It gives a general view that may not really assist in decision making in terms of improvement of customer experience. On the other hand, most customer feedback on services rendered are posted on their social media platforms such as facebook, twitter, website etc. This kind of feedback is not useful and may not assist in decision making as it is unstructured and cannot be analysed. Having several services housed in a one stop shop therefore requires a more advanced and fine grained analysis model that would allow the Huduma Centres analyse opinions expressed on the different aspects and features of entities, thus the need to have an aspect based opinion mining system for analysing customer feedback.

1.3 Scope of the study

This research is geared towards addressing the issues identified in the problem statement. In achieving this, the research will focus mainly on Huduma Centres within Nairobi City County.

The research will identify, design and develop a model and classifier to effectively analyse customer feedback. A prototype will then be developed to analyse the efficacy of the model. It will have the following capabilities:

i) Prompt users for their feedback in relation to the services received.

ii) Identify and extract relevant aspects and features within the feedback. It will also identify key issues that identify the opinion words and phrases.

iii) Determine the orientation of the feedback.

iv) Generate graphs and charts that will aid in decision making

1.4 Research outcomes and their importance

Statistics have shown that Huduma Centres are the most frequented offices in Kenya. Thus the need to have a well organized and automated way of acquiring customer feedback. The proposed system will assist in improving customer feedback acquisition and analysis and therefore aid in decision making. This will in turn improve the efficiency and effectiveness in provision of service delivery to its customers.

Huduma Centre managers will also benefit because they will be able to receive and monitor real-time customer feedback remotely.

1.5 Research Questions

i. Which technologies and models are currently in place for analysing customer feedback?
ii. Which is the best model for analysing customer feedback at Huduma Centres.
iii. How can the above model be designed and developed.
iv. How can one analyse and evaluate the effectiveness of the above model.

1.6 Research objectives

i. To review and analyse literature on existing models used in collecting and analysing customer feedback.
ii. To design a hybrid opinion mining model for analysis of customer feedback in Huduma Centres.
iii. To develop a classifier based on the above approach.
iv. To evaluate the above model for purposes of analyzing its efficacy.

1.7 Justification of the Research

People's opinions are an important issue for decision making, not only for individuals but also for government institutions. Huduma Kenya is a program by the Government of Kenya that aims to transform Public Service Delivery by providing citizens' access to various Public Services and information from One Stop Shop citizen service centres called Huduma Centres and through integrated technology platforms. This therefore necessitates the need to have a way of receiving, processing and analysing feedback from customers who access government services from the centres. Since the centres provide various services, it is important to analyse feedback received keeping in mind the different aspects of services provided by the centres.

No matter how well an information service appears to be progressing, the establishment should always look for innovative ways to make improvements. This is indicated by participants of a survey carried out by (Abdalla, 2015). Rutto (2015), the model employed by Huduma centres is based on goodwill and customer satisfaction. Each day a new customer encounters a new issue and hence there is need for more innovation to enable smooth and well timed handling of customer issues.

The proposed research will assist Huduma centres acquire and analyse customer feedback in real time. The acquired feedback will not only enable the Huduma centres to resolve pending customer issues at first contact but it will also enable them to evaluate their performance. Through it the citizens can raise the issues they encountered while being served through their mobile phone; at their preferred time and place.

1.8 Assumptions and Limitations of the study

1.8.1 Different language and writing styles

Customer's opinion writing is not formal and does not follow grammatical rules. Comments by different people in terms of writing, language, abbreviations and language varies. This may be a limiting factor but was solved by Yang (2008) who in his research found that a classification of product features that heavily rely on linguistics and Natural Language Processing (NLP) can solve the challenge. Context mining approach using mining class association rules will also assist.

1.8.2 Opinions change with time

This may be attributed to changes in officers operating the centres. This can be handled by having a time aspect when processing to assist in managing customer opinion.

2. LITERATURE REVIEW

2.1 Introduction

This section highlights the various models and technologies that are being used to collect and analyse customer feedback. It explains the various studies that have been done and section 2.5 discusses some applications currently in place that are being used to analyse customer feedback. Most of the highlighted studies relate to customer satisfaction and marketing.

2.2 Monitoring and Analyzing Customer Feedback through Social Media Platforms

Social media platforms such as Facebook and Twitter offer a convenient way for businesses to gather customer feedback about their products and services. This is so because people use social media to share their experiences with various consumer products and services they have used before. The factor driving businesses to use social media as a way to gather customer feedback is that it enables businesses to gather much relevant data as compared to traditional methods (Afroze, 2012).

Afroze proposed a solution that would assist organisations in monitoring feedback and detecting potential customers issues raised through Twitter and Facebook. The identification process of customer complaints involved framing the problem in question as an event detection problem. This framework begins with monitoring the customer feedback flowing though the social media platforms, after which sentiment analysis is performed and finally events of interests are detected and analyzed (Bhatia, 2013).

According to Ochieng (2016), analyzing of customer feedback derived from social media platform comes with its limitations. To begin with is the problem of heterogeneous nature of news groups commonly experienced when using Facebook where unlike Twitter, Facebook posts are independent of one another and therefore comments and replies are rarely in reaction to a specific post, which makes data lack continuity. Another limitation is that of time series analysis, where reactions or comments that were supposed to be analyzed along with its post might be made after analysis has been carried out.

2.3 A Linguistics-Based Approach to Customer Feedback Analysis

Analyzing of customer feedback employing text mining has mainly focused on developing on accurate models that automatically predicts sentiments or insights within text (Gräbner, 2012). Ordenes (2014), proposed a framework to fill the gap in text mining which still lacks focus on customer service experience identified mainly through analyzing customer comments or reviews. This framework adopted a holistic approach for analyzing feedback by taking into account three key components of value co-creation process including: activities, resources and context.

This framework not only classifies text as positive or negative in terms of specific attribute but it also maps it to a chain of activities and resources that help in describing how value is co-created in a particular context (Gronroos, 2012). This kinds of frameworks are encouraged so as to curb problems brought about by businesses using quantitative methods which result in incomplete understanding of customer experience. The use of qualitative methods which allows customers to give qualitative feedback in a textual and unstructured format is what technological advances are driving at. It gives the customer the chance to take the lead role by defining the process, timing of feedback and context with which information is provided or targeting (Witell et al, 2011).

2.4 Current Techniques commonly used in Sentiment Analysis

Sentiments Analysis involves studying people's opinions, attitudes and emotions toward a particular entity. An entity can be an individual, an organisation, an event or a topic. Sentiment Analysis traces sentiments expressed in a text and then analyses it. The main aim of Sentiment Analysis is to find an opinion, extract the sentiments it expresses and then classify its polarity. Text polarity can either be positive, negative or neutral (Kiritchenko, 2014). Figure below illustrates the various sentiment analysis techniques used and their relationships.

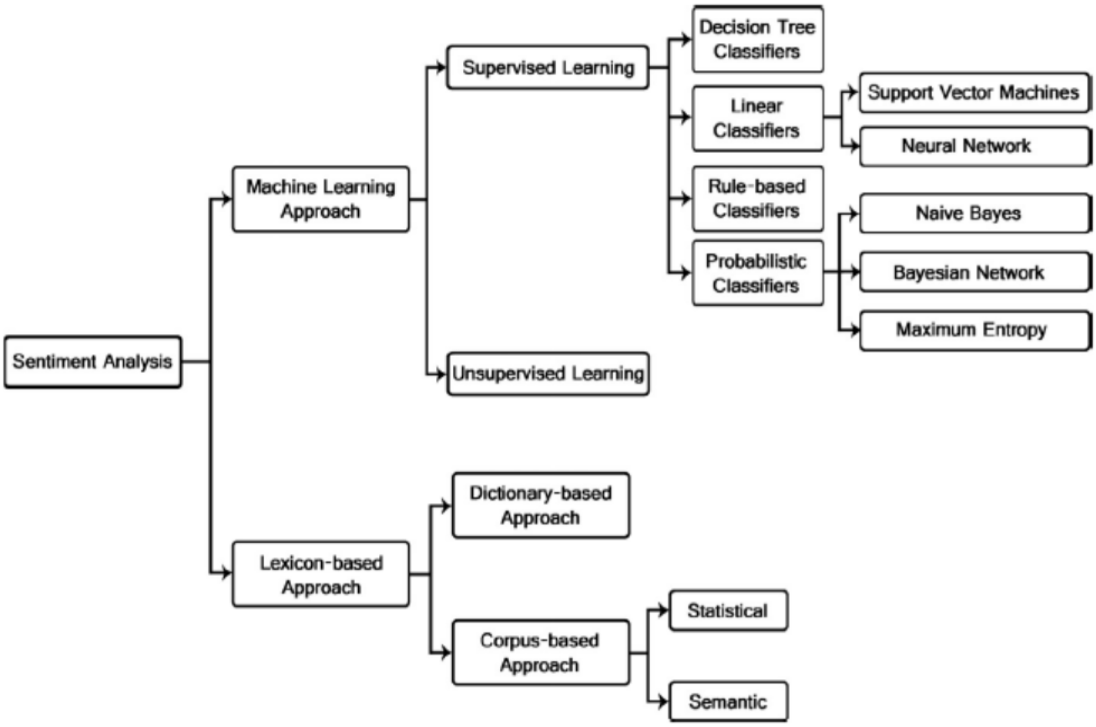


Figure 1: Sentiment Analysis Techniques by Walaa et al. (2014)

2.4.1 Machine Based Approach

2.4.1.1 Supervised Machine Learning

Support Vector Machine (SVM)

SVM uses a supervised leaning approach, which involves learning to classify data based on a set of training data. It requires both positive and negative training sets which is then used to seek the decision surface that distinguishes the positive from the negative data in the n-dimensional space also referred to as hyper plane. Its domain experts identifies the initial set of training data, used to build a model that can be applied to any other data not within the training set (Mertsalov, 2009). Vishal (2011), states that SVM-based approaches have the advantage of being able to handle large feature spaces with high classification accuracy. He highlights that SVM-based approaches are complex to implement as one of its limitations. The other limitation being that they do not scale well with the number of documents in text collections.

Decision Trees

Decision trees categorizes documents by constructing well-defined true or false-queries in the form of a tree structure. Internal nodes of a tree represents a document on test, branches represent the outcome of a test and leaf nodes holds a class label. (Vishal, 2011)The generated local dictionaries are able to differentiate meaning of each word for different categories. V ishal (2011), highlights the main limitation of using decision trees for text classification to be prone to overfitting when the tree is fully grown because some branches may be too specific to the training data provided. This happens because decision tree algorithms are based on heuristic algorithms where decisions are made locally at each node and hence cannot guarantee to return the globally optimal decision tree.

Naive Bayes

Naive Bayes Classifier is a probabilistic classifier used to classify text documents. It assumes that each feature word is independent of other feature words in a document. Depending on the level of preciseness required Naive Bayes classifiers can efficiently be trained and it requires a relatively small amount of training data in order to estimate parameters necessary for classification. Since independent variables are assumed, the only necessary items required to determine a class of text are variances of variables (Vishal, 2011). Following Naive Bayes approach, the probability of a review to belong to a particular class is calculated by calculating the prior probability which is the % of one class over the total data set multiplied by the likelihood of a review belonging to class given by number of words in a review belonging to a particular class divided by the total number of words belonging to that class in the data set. As illustrated by figure 2:

$$P(c \mid d) = \frac{P(d \mid c)P(c)}{P(d)}$$

Figure 2: Standard Naive Bayes Formulae (Abhishek, 2016)

According to Aurangzeb (2010), the key advantage of Naive Bayes is that it is robust enough to ignore serious deficiencies in its underlying Naive probability model. That is it correctly classifies text as long as the correct category is more probable than the others. The other advantages is that it is fast and easy to implement. Due to the pros highlighted, this research will utilise Naive Bayes Approach to classify text. Aurangzeb (2010), stated that Naive Bayes approach is limited by the fact that it is only manageable for low dimensions.

Survey by Burahudin et al. (2016) on the effectiveness of Supervised Machine Learning Technique indicated that the most commonly used ML mechanism is the Support Vector Machines and Naïve Bayes and that the out perform all the other techniques. On the other hand, supervised approach is useful in cases where there is no much data for analysis. Turrey (2008) introduced an unsupervised algorithm based on Pointwise Mutual Information and Information Retrieval (PMI-IR) to compute semantic orientation. Popescu (2005) used information extraction system called the OPINE which uses relaxation technique to determine orientation of words. The OPINE and PMI-IR have been known not to be good enough in extracting opinions, especially when using unstructured data.

2.4.1.2 Unsupervised Machine Learning

Unsupervised learning is a branch of machine learning that learns from test data that has not been labelled, classified or categorized. It identifies commonalities in the data, reacts based on the presence or absence of such commonalities in each new piece of data and classifies them as either positive or negative. Average sentiment orientations are used to predict the sentiment of each review document. Soundarya et al. (2015). Example of this is the Rule Based Approach.

Rule Based Method

Khan et al (2011) proposed a rule based semantic approach for text reviews. This was due to the fact that to classify a sentiment of a sentence, word dependency structures are used. They aggregate the sentence sentiments to predict the document level sentiment.

2.4.2 Lexicon Based Approach

This does not require prior training data sets. Liu (2012) explained this as a semantic oriented approach in which features are compared with semantic lexicons to determine the sentiment polarity of features present in a given document. Classification of a document is done by aggregating the semantic orientation of opinion words in a document. Documents with more positive sentiments are considered as positive while the ones with more negative sentiments negative document. Zhang et al (2009). It also uses a dictionary of sentiment words and phrases with their associated orientations to compute a sentiment score for each document.

The lexicon based approach is divided into two; the Corpus based and the dictionary based techniques.

2.4.2.1 Dictionary Based Technique

This is the use of a well-known corpora or thesaurus in determination of the semantic orientation of a sentiment. A set of few manually annotated words are passed through the thesaurus and through the use of synonyms and antonyms, their orientation is determined. According to Qui and He (2010), this approach has a major drawback as it does not work effectively with domain specific opinions.

2.4.2.2 Corpus Based Technique

This technique is divided into two i.e. Statistical and the semantic approach. Statistical approach, according to Fehrni and Klenner (2008), is where you find co-occurance patterns or seed opinion words from posterior polarities using co-occurences of adjectives in a corpus. Semantic approach on the other hand gives semantic values directly and relies on the different principles for computing similarities. It gives similar sentiment values to semantically close words Fehrni and Klenner (2008). Both techniques have been applied in very many applications used for sentiment analysis.

2.4.2.3 Semantic Orientation Calculator (SO-CAL) Algorithm

Maite et al (2011) in his research on Lexicon Based methods for sentiment analysis implemented the SOCAL algorithm using a dictionary of words annotated with polarity and strength. Previous versions of the SOCAL by Taboada et al (2004) used the simple aggregate and average method and relied on the adjective corpus to predict the overall sentiment orientation of a document. This was achieved by adding up and averaging the number of adjectives. But the recent version by Maite et al (2011) proposed the use of not only adjectives, nouns, verbs and adverbs in determining the SO value of a sentiment but also the use of intensifiers and negations in the determination of the semantic orientation value. Intensifiers are implemented with reference to the type of intensifier whether it’s an amplifier e.g. very or are downtoners e.g. slightly. Amplifiers increase the semantic intensity of the item while downtoners decrease it. These are then modelled into modifiers and are presented in percentages. Examples of intensifiers in SOCAL and their modifier percentages are as shown below:

S/No.	Name	Percentage
1.	Really	+14
2.	Somewhat	-30
3.	Most	+100

Table 1: Examples of intensifiers

Given a sentiment like '*Service was somewhat good'* and if good has a sentiment orientation value of +3, then the SOCAL computes the SO of the sentiment as:

+3 x (100-30)% = +3 x (70%) = 2.1

So after the inclusion of the downtoner (somewhat) the overall sentiment value of the sentiment becomes +2.1.

Negations are also implemented on SOCAL using the simple polarity shift technique also known as shift negations.

The main advantage of using this algorithm is that it incorporates the assigning of different weight to portions of text. According to Taboada (2004) this generally improves performance of the model. The other advantage is that it allows multiple cut-offs which makes it straightforward to understand.

Generally, corpus based approach alone is not effective as the dictionary based technique because it is very difficult to build a huge corpus to cover all the English words. However, this approach is very effective in finding domain specific words and their orientation.

2.5 Learning Algorithms

Different learning algorithms have been proposed and developed to implement this technique. Examples of these algorithms are the Transformational-based error driven algorithm, Transformation-based prepositional phrase attachment and the Pointwise Mutual Information Technique.

2.5.1 Transformational-based error driven algorithm

Ramshaw and Marcus (1994) defined this approach as where a set of simple learning rules is learned automatically to try and predict the semantic orientation of a sentiment. This has been applied widely to a number of national language problems including POS tagging and syntactic parsing. Figure below illustrates the learning process through the above approach.

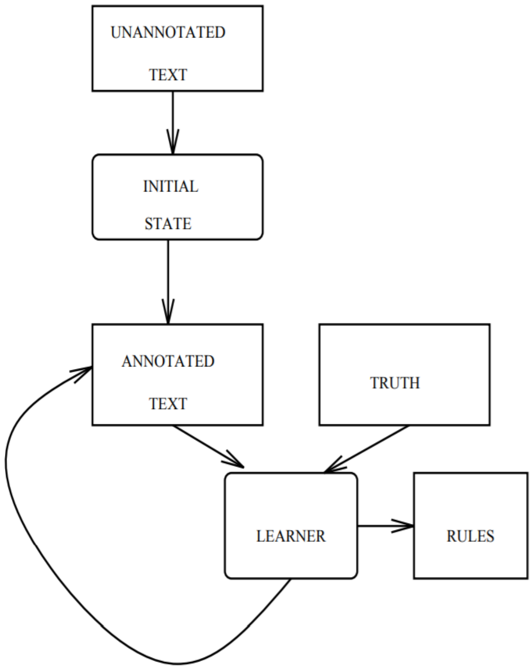


Figure 3: Transformational-based error driven algorithm

The process is that first, the unannotated text is put at the initial state annotator that ranges from complexity to quite trivial, it is then compared to the truth which is a manually annotated corpus and transformations are leaned that is applied to the output of the initial state annotator to make it similar to the truth. Once an ordered list of transformations is learned, the unannotated text is annotated by applying the initial state to it then applying the transformations.

2.5.2 Transformation-based Prepositional Phrase Attachment

This learns transformations from a corpus of 4 tuples of the form (v n1 p n2) where v is a verb, n1 is the head of its object noun phrase, p is the preposition and n2 is the head of the noun phrase governed by the preposition. The learning process involves first training set is processed to the start annotator, then each possible transformation is scored by applying it to the corpus and computing the difference (Increase or reduction) in error state. The process is iterated until no rule can be found that reduces the error rate.

2.5.3 Pointwise Mutual Information – Information Retrieval (PMI-IR) technique

The mutual information measure provides a formal way to model the mutual information between the features and the classes. It has been widely used in learning systems and is given as a measure which was derived from the information theory. Cover and Thomas (1991) defined the point-wise mutual information (PMI) Mi(w) between the word w and the class i is defined on the basis of the level of co-occurrence between the class i and word w.

The PMI-IR is defined as (Church and Hunks 1989):

$$\mathrm{PMI}(word_1, word_2) = \log_2\left(\frac{\mathrm{p}(word_1 \ \& \ word_2)}{\mathrm{p}(word_1)\ \mathrm{p}(word_2)}\right)$$

Where p($word_1$ & $word_2$) is the probability that $word_1$ and $word_2$ co-occur. The ratio between p($word_1$ & $word_2$) and p($word_1$) p($word_2$) is the measure of degree of statistical independence within the words.

The SO of a word *word* is calculated as:

$$\begin{aligned}\text{SO-PMI-IR}(word) = {} & \mathrm{PMI}(word, \{positive\ \ paradigms\}) \\ & - \mathrm{PMI}(word, \{negative\ \ paradigms\})\end{aligned}$$

2.6 Existing Solutions for Analyzing Customer Feedback

2.6.1 Tweview

Tweview is a system used to analyse sentiments from large volumes of data and generate a score that identifies the polarity of the sentiments. The score generated is referred to as a Tweview Score and is mostly used to analyze sentiments for products and services. Most marketing departments uses the service and monitors various social media sites including LinkedIn, facebook and twitter (Fearn, 2015).

The solution uses recurrent neural networks which is designed to predict the tree rather than a linear sequence in conventional recurrent neural networks. The model calculates the probability of a sentence by estimating the generation probability of its dependence tree. (Xingxing 2016).

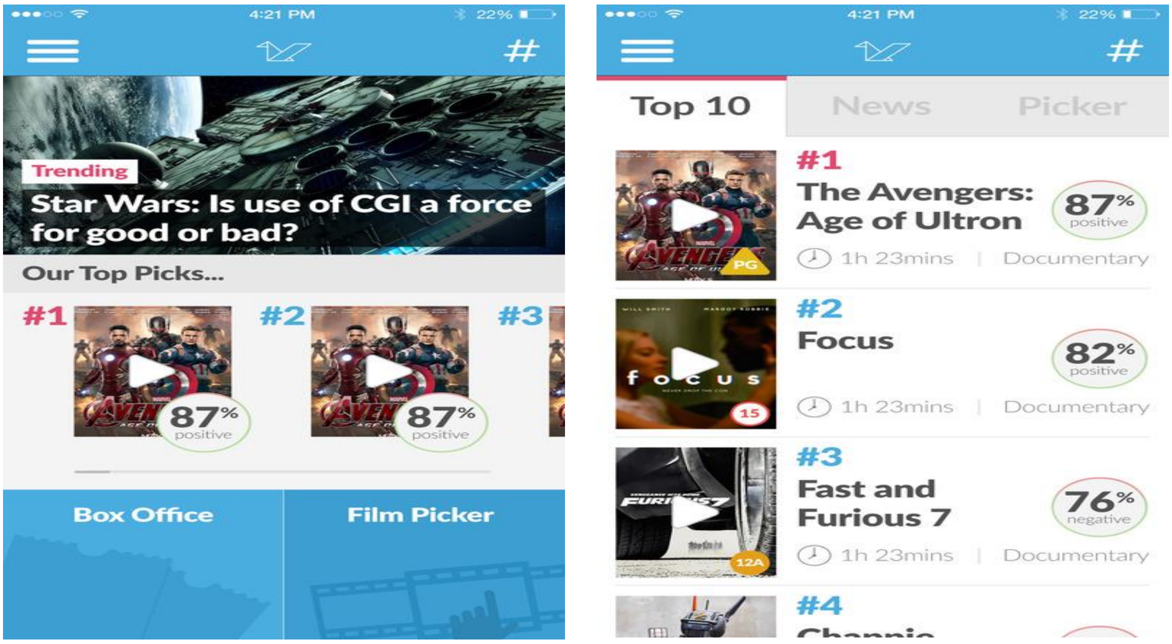


Figure 4: Screenshots of Tweview Application (Fearn, 2015).

2.6.2 Net Promoter Score Surveys

Net Promoter Score is a customer experience management program that measures customer experience and predicts business growth. It is used across businesses by calculating the Net Promoter Score using answers given to a question and gives a scale of between 1 and 10 to users on the likelihood of he/she to recommend the service to a friend or colleague.

The figure below shows the NPS calculator.

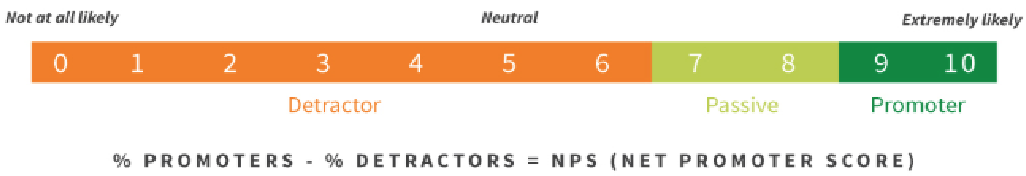


Figure 5: NPS Calculator

The responses are grouped into three:

- Promoters (score 9-10) These are dependable fans who will keep buying and refer others, leading to growth.
- Passives (score 7-8) are satisfied but unenthusiastic customers who are vulnerable to competitive offerings.
- Detractors (score 0-6) are unhappy customers who can damage your brand and impede growth through negative word-of-mouth.

Subtracting the percentage of Detractors from the percentage of Promoters yields the Net Promoter Score, which can range from a low of -100 (if every customer is a Detractor) to a high of 100 (if every customer is a Promoter). NPS learns the customer's general sentiment by measuring how many people would recommend the product in question (Kissmetrics, 2015).

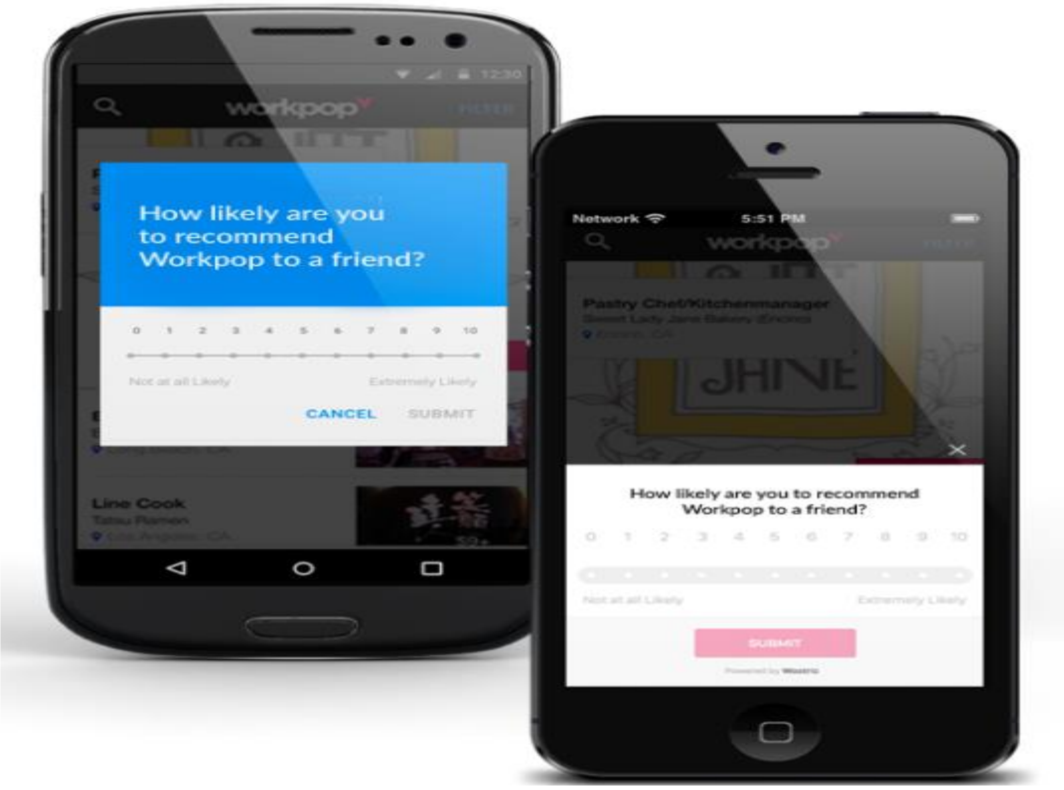


Figure 6: Mobile Application of Net Promoter Score Survey (Kissmetrics, 2015)

Once the customer has given his/ her score, they are further asked to give reasons for their score. This ensures personalized feedback and allows the organization act upon the reasons provided. This solution works well in measuring loyalty but can pose a challenge if the organization wants to set the real issues or the root cause of the issues be delving them. Since the measuring often happens after the customers' journey it can possibly mask underlying issues that are of concern and can be used to identify improvements necessary (Zaki, 2016).

2.6.3 The American Customer Satisfaction Index model

According to Mohammad (2008), the American Customer Satisfaction Index (ACSI) Model is a Framework used to examine the causal relationships among customer expectations, perceived quality, perceived values, customer satisfaction, customer complaint and loyalty. The system was developed at the University of Michigan's Ross school of Business and comprised of a multi-equation econometric model that uses customer interviews as input. It is a cause-effect model with indicies for drivers of satisfaction on the left side which could be customer expectations, perceived quality and perceived value, satisfaction (ACSI) in the center and outcomes of satisfaction on the right side which could be customer complaints and customer loyalty including customer retention and price tolerance.

According to Natural Resorces Conservation Service (2011), ACSI is a weighted average of three questions. The questions uses a 1 to 10 scale and converted to a 0 to 100 scale for reporting purposes. The three questions used measure the overall satisfaction, satisfaction compared to expectations and satisfaction compared to an ideal organization. This model assigns the weights to each question in a way that will ensure maximum ability of the index to predict changes in customer satisfaction. Figure below demonstrates the America Customer Satisfaction Index model.

Figure 7: American Customer Satisfaction Index Model

2.7 Gaps and Limitations on Existing Solutions

Existing solutions, measures customers' satisfaction by having them answer limited number of questions and whether they would recommend the product in question. This approach lacks the ability to identify and act upon driving factors behind customers' responses to the questions. Most of the solutions require structured texts and does not allow customers to express themselves without limitations.

2.8 Conceptual framework

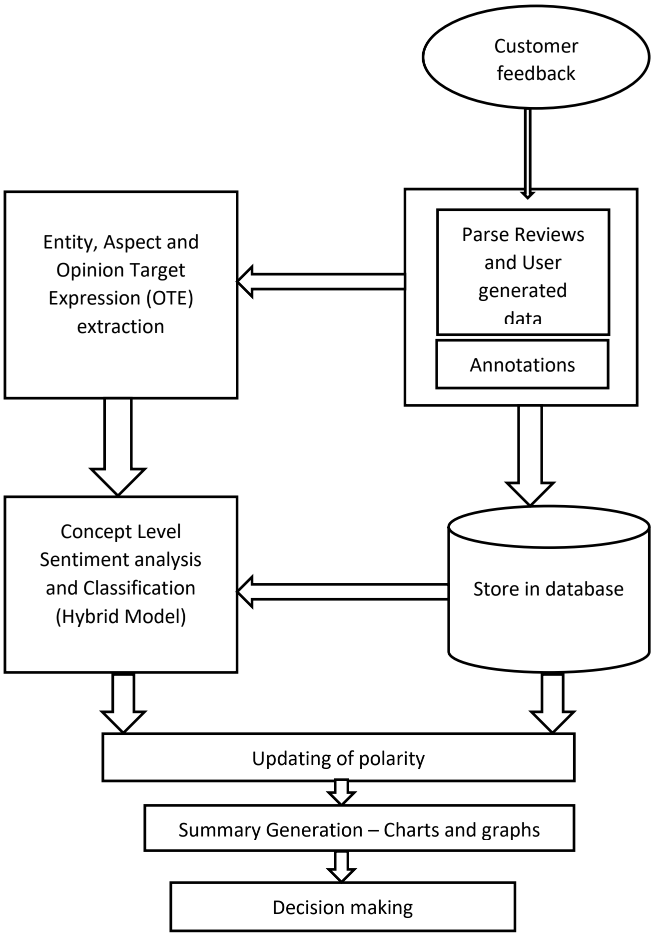


Figure 8: Conceptual framework

In the framework, the customer feedback from Huduma Centre will be inputted to the application. Here the tasks will be divided into two sub tasks:

a. In-domain
b. Out-domain.

For the in-domain, the application will identify all the opinion tuples with the following information.

i. Aspect Category – Rules will be used to identify the different Entities E and Attributes A from the tuple. Examples of Entities would be institutions represented at the Huduma Centres e.g. National Hospital Insurance Fund (NHIF), Directorate of Criminal Investigations (DCI) etc. while the attribute s would be Customer Service, Speed, Quality of Product, Privacy etc.
ii. Opinion Target Expression (OTE) – This is to extract the expression used to describe the attributes referring to an entity. If there exist no explicit mention of the expression, then it will take the value 'NULL'.
iii. Sentiment Analysis and Classification – This is where the entity and attribute pair will be assigned a polarity label. A hybrid Machine Learning algorithm will be applied for text classification. This will then generate sentiment orientation or the general polarity of the sentiment.

For out domain this is where there exists no training data in the database. This is the learning components and the proposed structure will allow annotation. Annotation is a process of adding information to already existing data and normally used when there is a big volume of unstructured data. Freie (2013). The system will apply the syntactic kind of annotations in and the Pointwise Mutual Information and Information Retrieval algorithm will be used to determine the orientation of the sentiments. The annotated data will then be stored in the database.

A web interface will be provided for management to monitor the opinions of the customers. This summary will be in form of charts and graphs for ease of interpretation and will assist in decision making.

3. METHODOLOGY AND IMPLEMENTATION

3.1 Introduction

This section highlights the methodology that was followed to achieve the overall objectives as outlined above. The previous chapter answers research question number one by reviewing the different models and techniques currently used to acquire and analyse customer feedback. The research methodology below describes how the data was collected, pre-processed and how the model was designed, developed and tested which will answer research questions two and three. It further describes how the model will be evaluated.

The methodology used for this research is the Design Science Research Methodology (DSRM). This methodology guided the development and evaluation of the hybrid model used in this research. This is because it is guided by several principles that make it preferable for the development of machine learning applications/ models. The principles include:

1. The goal of the methodology – which is to develop and evaluate a hybrid model to be used in the analysis of customer feedback
2. A process for designing models, evaluate designs and communicate the results to the relevant audiences.

Ken Peffers et al., (2008) described six research activities involved in DSRM. These activities are as shown in the figure 8 below.

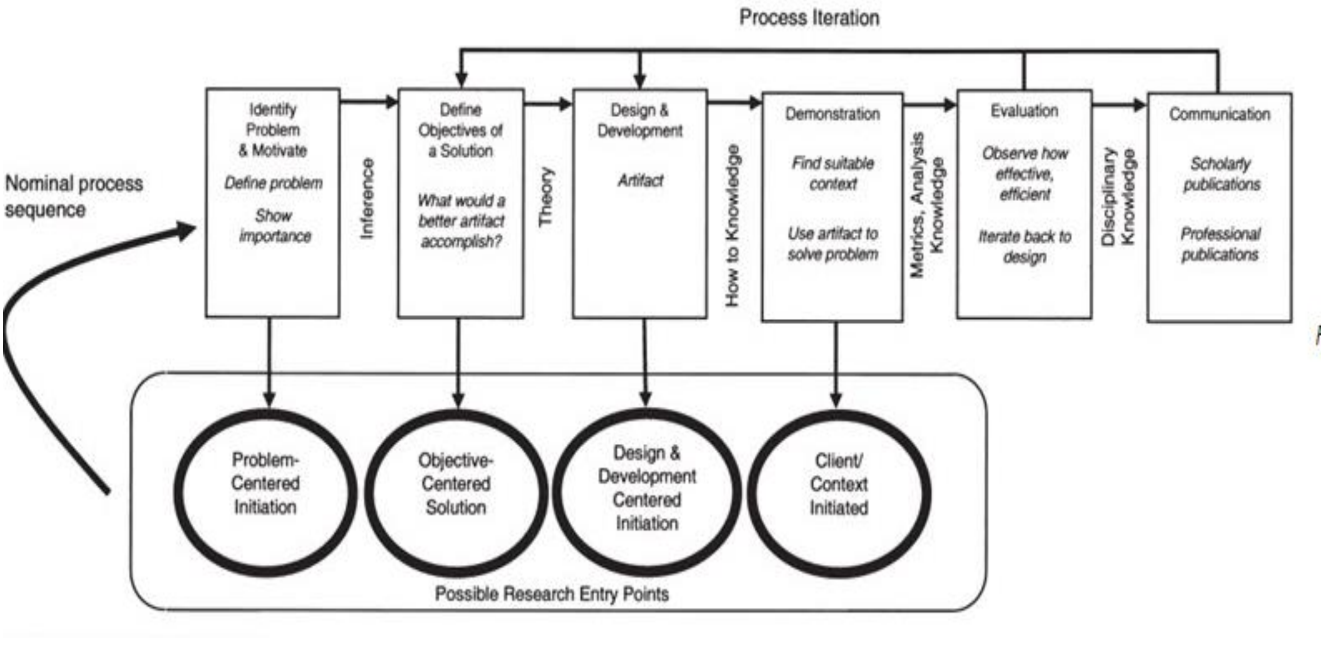


Figure 9: Design Science Research Methodology Model

DSRM describes the following research activities.

1. Problem identification and motivation – This is where the specific research problem is defined and the benefits of the model itself is explained. This may come from multiple sources including new developments in the industry or identification of a problem within a reference discipline.
2. Defining objectives of a feasible solution of the artefact. This activity is defined and derived from the problem definition. It is an objective centered solution and answers the question '*What would a better artefact accomplish".* The artefact may either produce quantitative or qualitative results..
3. Design and development of the model – This is the determination of the desired functionality and architecture of the model. The techniques of implementation varies depending on the artefact to be created. E.g. an algorithm may require construction of a formal proof to show the correctness while an expert system embodying assumptions of human cognition in an area of interest will require software using high level packages. The design and development of the hybrid model in this paper has been described in Chapter four below.
4. Demonstration of the artefact – This is done to exhibit the adoption of the artefact. It can be done through experimentation, case study or simulations. The various tests are done to the system to ensure that all components work as expected and achieves high performance rates. This would test the system unit, integration, usability and system functionality to identify errors and bugs in the system.
5. Evaluation of the artefact – Once the artefact has been developed, it is evaluated according to a criteria defined by the researcher. Deviations (either quantitative or qualitative) are noted and explained. It also includes an analytic sub-phase in which hypothesis is made about the behaviour of the artefact. In the evaluation of the model. Separate set of new labelled data will be inputted in the evaluation interface, and the results will be used to for calculating accuracy, precision and recall for the Classifier.
6. Communication/ Conclusion – This involves publishing study results to the relevant audiences e.g. other researchers or professionals. Communication is important in every research. Hevner, et al. (2004) thus importance of this activity. The results of the research effort are consolidated and reported to the stakeholders. Conclusion is the final of a specific research effort.

3.2 Problem identification and motivation

In this research, the problem identification has been outlined in chapter one above. Huduma centres have implemented a five star rating technique to capture customer's sentiments and provided a contact number and email through which citizens can leave their opinions. Citizens can either call the number or send an email to make their issues known to these centres. This method has many challenges as one would not be able to have the exact details of one's opinion. It gives a general view that may not really assist in decision making in terms of improvement of customer experience. Kagasi V. (2016) in his research concluded that methods currently being used to measure customer satisfaction are slow and customer issues are identified long after the customer has left and thus the need to have a real time service based opinion mining system for analysing customer feedback.

3.3 Defining objectives

Objectives of this research are based on the problem identified above. The objectives as listed above are:

i. To review and analyse literature on existing models used in collecting and analysing customer feedback.
ii. To design a hybrid opinion mining model for analysis of customer feedback in Huduma Centres.
iii. To develop a classifier based on the above approach.
iv. To evaluate the above model for purposes of analyzing its efficacy.

3.4 Design and development of the model

3.4.1.Requirements Specifications

Requirements specification is a comprehensive description of the intended purpose and environment for the application to be developed. It fully describes what the model will do and how it will be expected to perform. It can be categorized into functional and Non-functional requirements.

3.4.1.1 Functional Requirements

Functional requirements describe the activity process and the functions each and every component is expected to perform. Functional requirements for the application are as follows:

1. It should allow users to input feedback through an application into the model
2. It should be able to calculate a sentiment orientation value and map it to a given entity based on the hybrid model to be developed.
3. It should be able to communicate the analysis of the customer opinions in real time to users through a portal that will be developed.
4. The system should be able to learn new words that are not labelled.

3.4.1.2 Database Design

ER Diagram

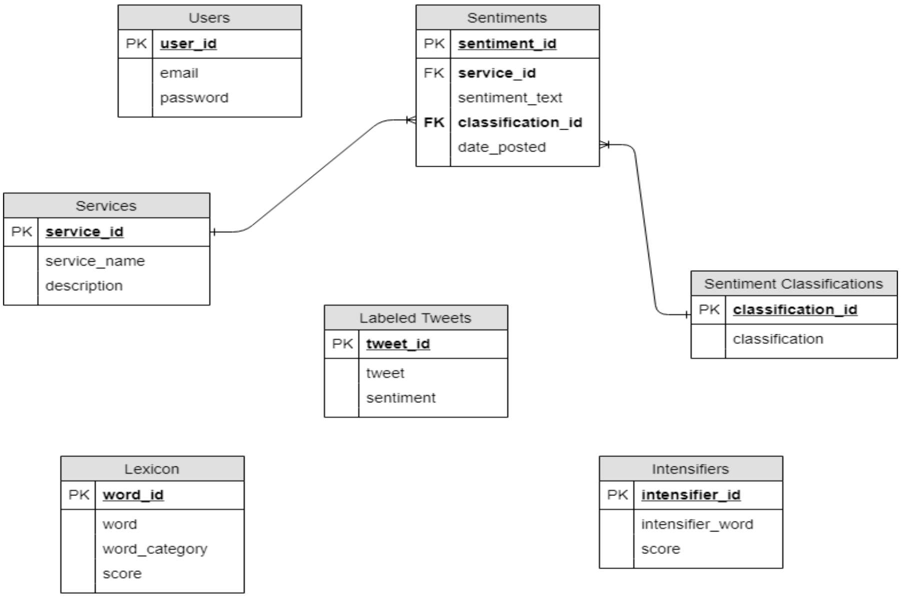


Figure 10: ER Diagram

3.4.1.3 Architectural Design

The goal of the study was to design and develop an Aspect Based Opinion Mining Application based on a hybrid model that can assist in the analysis of customer feedback. The figure below shows the architecture of the application.

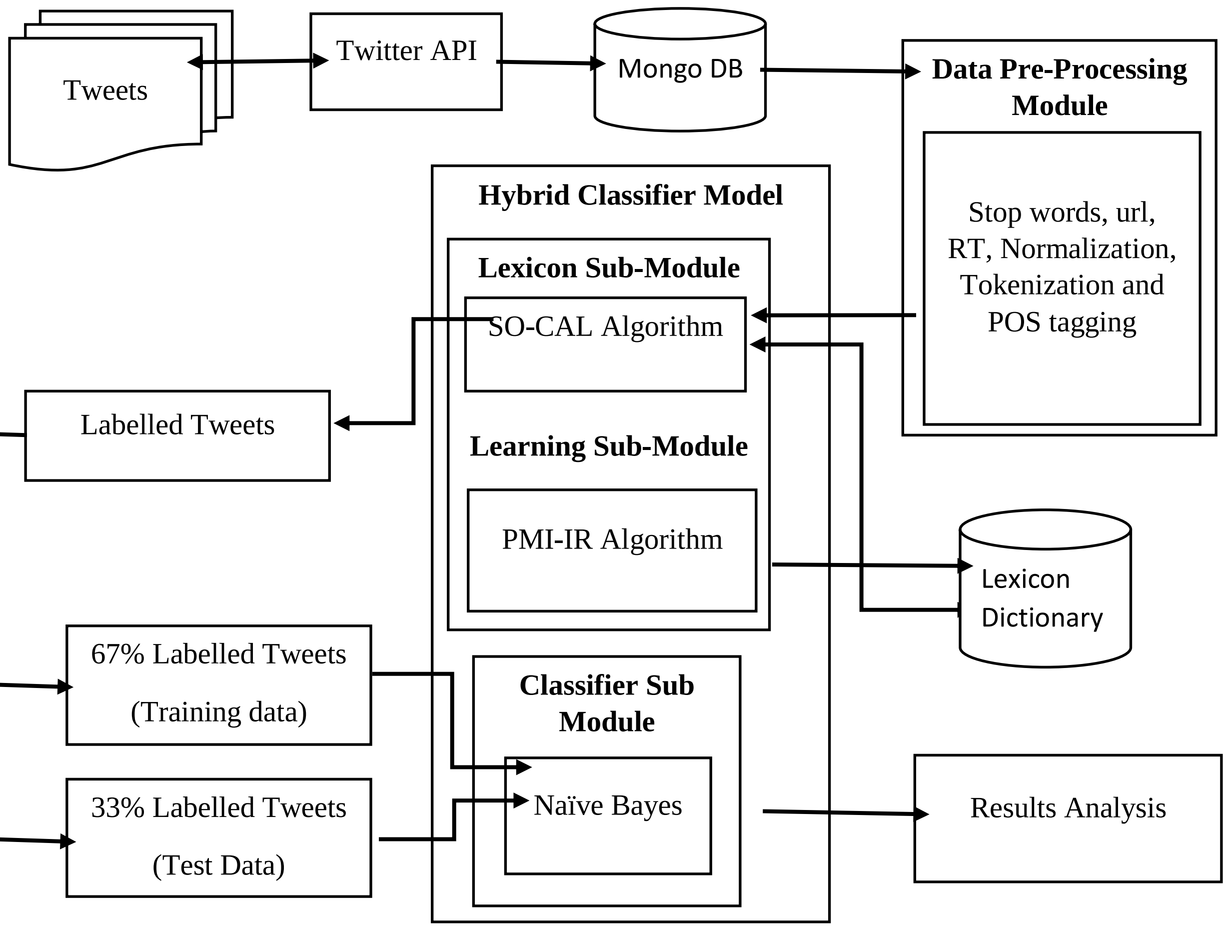


Figure 11: Architectural Design of the Aspect Based Opinion Mining Application

The hybrid model combines both the supervised and unsupervised machine learning techniques generally has five modules namely:

1. Data Pre-processing module
2. Aspect/ Feature extraction module
3. Hybrid classifier module
4. Learning module
5. Web interface module

3.4.2 Model Implementation

After the design as discussed in the above, this section explains the actual the implementation of the various modules, web application and the database. It describes the implementation procedures and platforms and later the various testing procedures performed on the application.

3.4.2.1 Data Acquisition

Data for the development of the solution was mined from social media site Twitter. Twitter is an online news and social networking service on which users interact. Twitter has a developer platform where developers can access the twitter API endpoints, requests and responses using a command line application that make authenticated request called twurl. Twitter APIs are readily available and provides two major types. One is the Twitter Streaming API which filters real time tweets and the Twitter Search API which provides the search function. The standard search API provides a search for 7 days while the premium versions provide searches for a longer period – upto year 2008.

Data in this case was retrieved using the Twitter Search API, in which it involved polling Twitter's data through a search or username. Through the Search API users request tweets that match some sort of "search" criteria and have different conventions. Examples of the twitter conventions are the Hashtags ("#") – used in organizing data, URLs- to track external sources, Emoticons (Smileys) and Colloquial expressions – expressions in an abnormal way e.g. "The service was awesommmeeeeeeeeeeee". In this case, the conventions used in the search was of Hashtags and usernames. For the users to access the services of the application, the device running the application should have active Internet connection. The page provides a text area where the search criteria is entered and a button for triggering the extraction of the data. Examples of the search criteria includes the Hashtags, usernames etc. and its aim is to assist in filtering the data that is needed. Once criteria has been entered, the search button queries the twitter endpoint with the request given.

The main tags were #HudumaKenya #Huduma #GOKService and the usernames used were the @HudumaKenya @Huduma. The search generated 700 reviews which were parsed to Pre-processing Module.

The figure below shows a screenshot of the module.

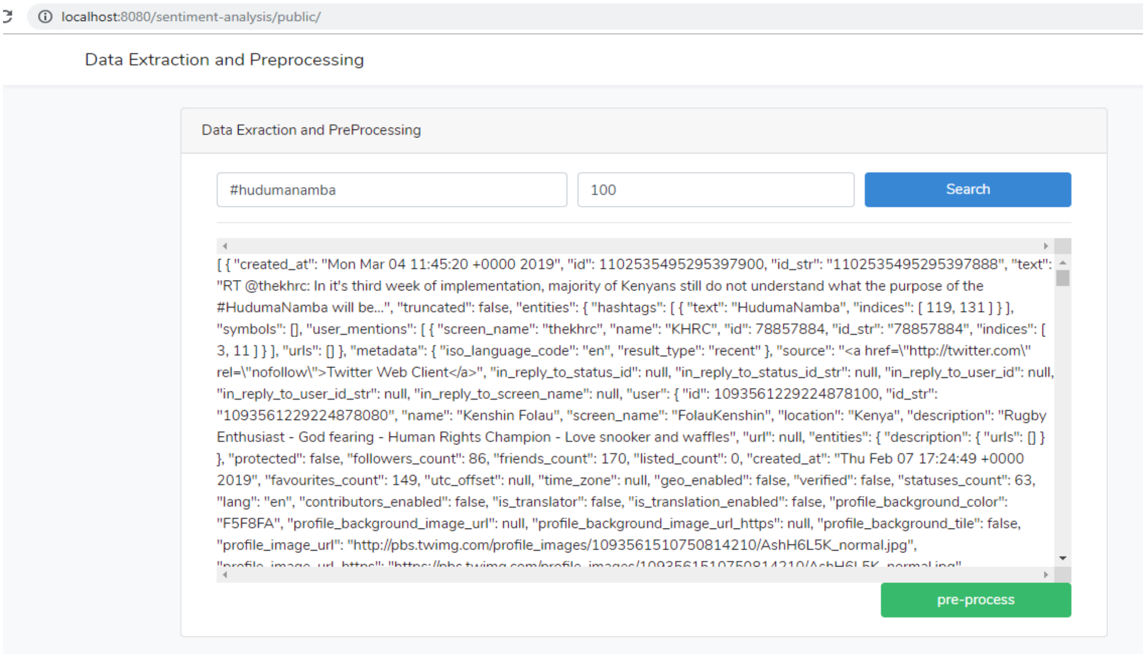


Figure 12 : Screen shot of Data extraction Module

3.4.2.2 Data Pre-Processing Module

The core functionality of the module is to process data from the twitter servers. Since it's a web-based application, the application runs on any software platform as long as it has a browser. This module was implemented using a PHP framework Laravel since it is accessible, yet powerful in providing powerful tools needed for large robust applications, superb inversion of control container, expressive migration system, and tightly integrated unit testing support that gives one the tools he/she needs to build any application.

Raw tweets normally contain audio, video, URL, emoticons, hashtags and contextual words that are not useful. Before processing the data extracted from social media, it was taken through the pre-processing stage where noisy data is removed from the data. Commonly used pre-processing techniques are removal of stop words and URLs, lemmatization, and stemming. In this case, the noisy data to be removed were data such as retweets which are repeated data and text starts with \RT. The URLs and stop words were also eliminated. After the cleaning, we performed segmentation which separates the collected tweets into individual sentences. The stage has two steps namely; Tokenization and Normalization.

1. Data Tokenization

This is the process of delimiting the reviews into single sensible units which maybe words, phrases, symbols, or other meaningful elements are called tokens. In this research, the reviews were delimited into segments that was meaningful to text classification.

It is triggered from the application and will generate the cleaned up tweets, segmenting them into sentences and reformatting them into a table format. The extracted processed data is then saved into a text file.

Figure below displays an example of segmented tweets.

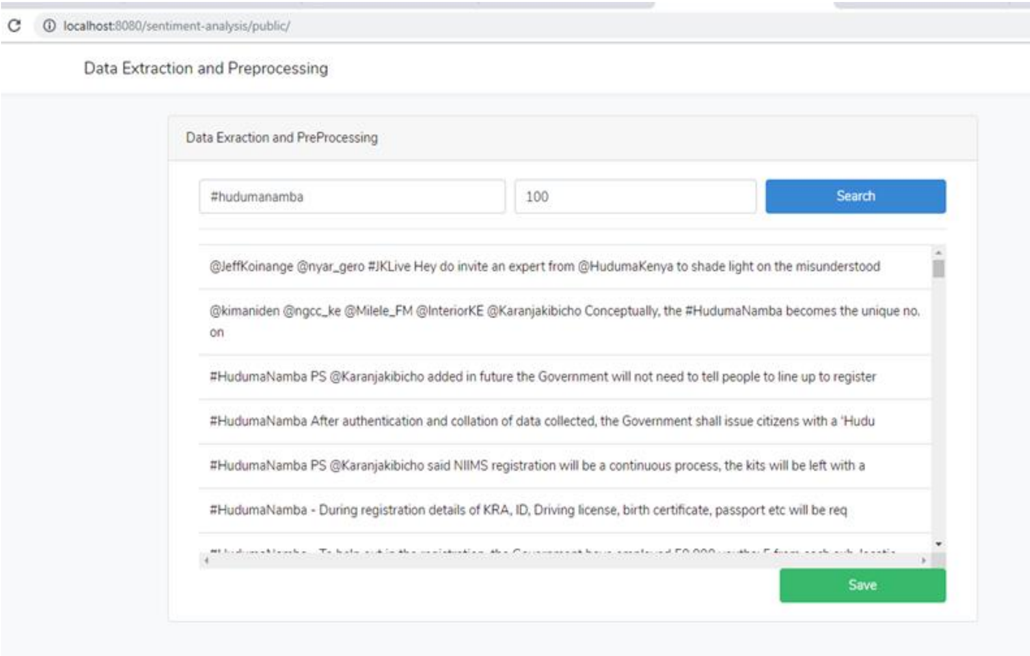


Figure 13: Segmented tweets

2. Data Normalization

This attempts to transform the text to make it consistent by Stemming and Lemmatization. It involves the removal of URLs, Retweets, emoticons, hashtags and stop words to make the sentiments consistent.

URL

Users at time share hyperlinks to other webpages in their tweets. ((www\.[\S]+)|(https?://[\S]+)) is a regular expression used to match the URLs. These hyperlinks are not important for classification as they lead to very sparse features thus the need to eliminate them from the tweets. This was achieved by replacing all the hyperlinks with the word URL.

Retweets

These are tweets that have already been tweeted by some other people and shared by others. They start with /RT. They don't add any value to text classification and therefore are eliminated at the data extraction using the expression –RT.

User mentions

Every twitter user has a handle associated with them. Users often mention other users in their tweets by @handle. In this research we replaced all user mentions with NULL.

Emoticons

Users always use different emoticons to express emotions. These can be mapped to an expression depending on whether they are denoting positive or negative emotions but for this research, the emoticons are eliminated due to its complexity.

Hashtags

These are unspaced phrases prefixed by the hash symbol (#) which is frequently used by users to mention a trending topic on twitter. In this research all the hashtags are replaced with the words with the hash symbol. For example, #NHIF is replaced by NHIF.

Stop words

These are words that do not change the semantic orientation of a sentiment. The words are searched and are marked and then not used during classification. In this research, the stop words remain in the sentiment but inactive during classification. The list of stop words used in this research are shown in appendix 2 below.

Some examples of tweets from the extracted dataset and their normalized versions are shown in table 2 below

S/No.	Raw	Normalized
1.	@noelkemunto Quality service and very accurate #Huduma #CitySquare.	Quality service and very accurate Huduma CitySquare
2.	/RT @njrogoge434 #Huduma couch was clean and comfortable. https://bit.ly/2I0AaJS	Huduma couch was clean and comfortable.URL
3.	@Mintejua Efficient and effective infrastuctures. Huduma	Efficient and effective infrastuctures. Huduma
4.	@Gerogertheman Polite and understanding staff from the enquiries depertment to the management	Polite and understanding staff from the enquiries depertment to the management URL Huduma

	https://bit.ly/2I0Aafds #Huduma	
5.	@mintish Affordable services and products. https://bit.ly/2IthehyewS This was just awesome. Cc @Jumabaya Huduma	Affordable services and products. URL This was just awesome. Cc Huduma
6.	@ronoben No queues and the staff are cheerful. What else can make you smile after a good service from a huduma you are loyal to.@fredmatiangi	No queues and the staff are cheerful. What else can make you smile after a good service from a huduma you are loyal to
7.	@jeremi234 Dishonest staff. #HELB @HelbKenya #Huduma	Dishonest staff. HELB Huduma
8.	@manueli Good service at #EACC #Huduma Server next time we need tea :)	Good service at EACC Huduma Server next time we need tea
9.	@Mikemmukera Quite good #Huduma #GPO at providing some good old fashioned spooks	Quite good Huduma GPO at providing some good old fashioned spooks
10.	@Rop Exceeds all expectations. #Hudumakenya	Exceeds all expectations. Hudumakenya
11.	@RisperMarley Good value good location good surroundings. #Huduma service	Good value good location good surroundings. Huduma service

Table 2: Sample of raw/ normalized data

3.4.2.3 Hybrid Classifier Module

This is the backbone of this new classifier. In this case, the solution uses a hybrid model that combines both the unsupervised and supervised machine learning approach. The unsupervised employs the use of lexicons and rule based techniques while the supervised approach will use the Naïve Bayes Algorithm approach. This was achieved through an ensambling technique called stacking. The two approaches are stalked together to improve accuracy of the model. The procedure is as shown in the table 3 below:

Procedure: Hybrid Model Input: Pre-processed tweets Output: Accuracy 1. Begin 2. Sentimentanalysis_lexicons() 3. Print Sentiment Orientation of tweets 4. Sentimentanalysis_Classifier()

5. Print SO of the sentences

Table 3: Hybrid model

This module consists of two sub modules. These are the Lexicon and Naïve Bayes Classifiers. The lexicon was implemented using the SOCAL algorithm as explained in the literature review. The extracted preprocessed data will be parsed through the lexicon sub module to determine the semantic orientation of the tweet. The output is then stored and will be used as training and test data for building the Naïve Bayes Classifier.

3.4.2.4. Lexicon Based Sub Module

The was implemented using the Semantic Orientation Calcularor (SO-CAL) which extracts first the semantic bearing words (including adjectives, verbs, nouns, and adverbs) and use them to calculate the semantic orientation, taking into account the valence shifters (i.e. intensifiers, downtoners and negations). The SO-CAL lexicons have separate dictionaries that define the semantic orientation of adjectives, adverbs, nouns and verbs. In addition, the corpus has a dictionary of intensifiers which is used to calculate the semantic orientation of context-dependent text.

The intensification was modelled using modifiers with each intensifying word having a percentage associated with it and is applied recursively.

Procedure: Sentimentanalysis_Lexicons Input: Pre-processed tweets Output: Labelled tweets 1. Begin 2. Segmented tweets 3. Apply unigrams Calculate the semantic orientation using SO-CAL 4. Apply N-grams Calculate semantic orientation of the N-grams 5. End

Table 4: Procedure for SO calculation using Lexicons

The output of the lexicon analysis is the labelled tweets. These are tweets with semantic orientation attached to each sentiment. The labelled tweets are then used to build a Naïve Bayes Classifier using two thirds of the data as training data while the remaining one third used as test data. The building of the classifier is a service and thus will be done automatically using the provided algorithm. Sample output of the lexicon module is as shown in the figure below.

ID.	Normalized	Semantic Orientation
1.	Quality service and very accurate Huduma CitySquare	Positive
2.	Huduma couch was clean and comfortable.URL	Positive
3.	Efficient and effective infrastuctures. Huduma	Positive
4.	Rude and un-understanding staff from the enquiries depertment to the management URL Huduma	Negative
5.	Affordable services and products. URL This was just awesome. Cc Huduma	Positive
6.	No queues and the staff are cheerful. What else can make you smile after a good service from a huduma you are loyal to	Positive
7.	Dishonest staff. HELB Huduma	Negative
8.	Good service at EACC Huduma Server next time we need tea	Positive
9.	Quite good Huduma GPO at providing some good old fashioned spooks	Positive
10.	Exceeds all expectations. Hudumakenya	Positive
11.	Good value good location good surroundings. Huduma service	Positive

After calculation of the SO of the statements, the output of the Lexicons are saved in a file and then parsed as input of the Naïve Bayes Machine learning algorithm.

3.4.2.5 Naïve Bayes Machine Learning Algorithm

This is where the labelled data set from the lexicons was used to build the machine learning based classifier.

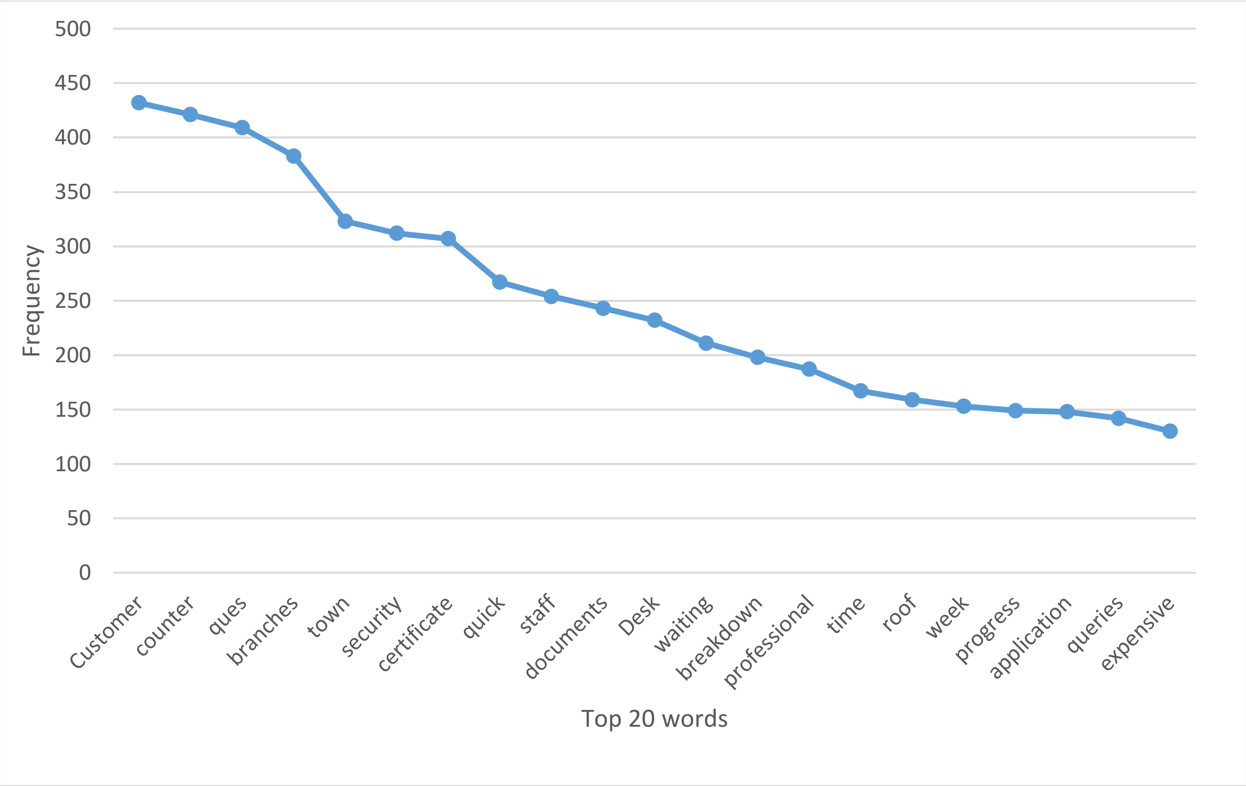
Procedure is as shown in the table below:

Procedure: Sentimentanalysis_Classification Input: Labelled Data set from the Sentimentanalysis_Lexicons Output: Accuracy

1. Begin
2. Extract features by use of a frequency distribution matrx
3. Train using training set with labelled dataset
4. Build Naïve Bayes Classifier
5. Analyze results
6. End

Table 5: Procedure for Machine based classifier

Unigrams were extracted from the training set obtained by tokenization. A total of **9387** unique words were extracted from the training dataset. The frequency of the unigrams were calculated through which a frequency distribution was created. For this research, we limited the number of unique words to form our vocabulary to **400** since most of the other words occurred less times and are noise thus was not useful in our text classification. Figure below shows the frequency distribution of the top 20 words.



The unigram was then assigned a unique index depending on the rank. The most common word is assigned index 1, second common word assigned 2 and so on. This is then used in the building of the classifier.

3.4.2.6 Learning Module

To expand the list of words in the corpus, the solution was developed to be intelligent enough to be able to learn new words and add to the dictionary. The new words are annotated before being stored in the corpus. This is implemented using the Semantic Orientation form Association (SO-A) technique. This uses the Pointwise Mutual Information – Information Retrieval (PMI-IR) technique to calculate the semantic orientation of a word. The rationale is that the technique has been emphatically tested using 80 synonyms test questions from the Test of English as a Foreign Language (TOEFEL) and obtained a very high accuracy values.

3.4.2.7 Web Application

The main function of the web application is to allow users interact with the system. Its other main function is assist in the evaluation of the model by having different web pages. It allows users to give feedback and generate reports through the system. Below are the components that make up the web application.

Login page

In order to access the web application, users have to login using correct user credentials made up of a username and a password.

Account Management

This component allows for distribution of user management roles. It enables the user manage their account, manage other users and add/delete when they need/ no longer need them into the system.

Sentiment Page

This is where the system is evaluated through. The page allows users to enter sentiments and generate score and add the details to the database. The results is then used to calculate the accuracy, precision and recall to determine the effectiveness of the system.

Aspects page

This is where all the sentiments provided are grouped by their respective institutions. From the page, one can tell the opinions of customers towards a given service or entity.

Evaluation page

This is where the actual feedback is inputted in order to get its semantic orientation and was implemented using a web page where it provides a text area where customers opinion are typed/ inputted and a trigger is set for analyzing the opinion using the hybrid model.

The model then outputs a score that determines the semantic orientation of the opinion/ feedback.

The figure 15 below is a screenshot of the module.

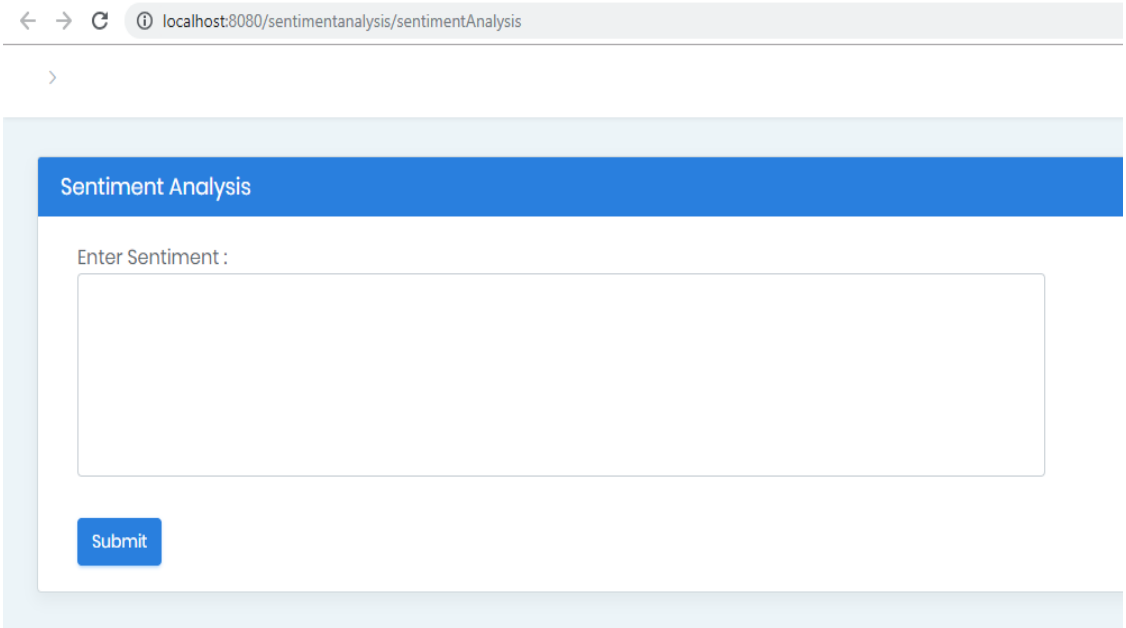


Figure 15: Evaluation page

Dashboard

The application summarizes all the customer feedback and based of the different aspects/ service, generates a dashboard page that displays the summary of all the customer feedback, indicating the total number of feedback per service, the total number of positive, neutral and negative feedback.

This is as shown in the figure 16 below.

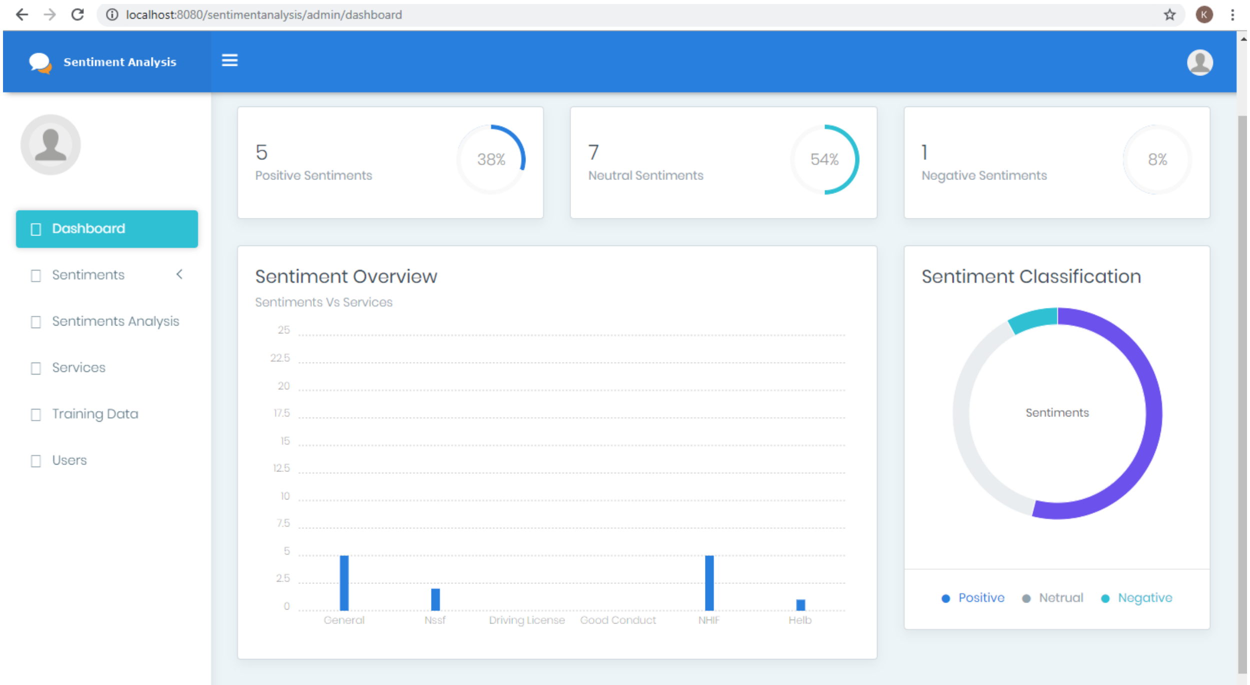


Figure 16: Dashboard

3.5 System Testing

This section details on tests carried out on both the model and web application. The tests were done to evaluate the functional and non-functional requirements of the application.

3.5.1 Login Functionality Testing

This test was done to ensure system user can successfully login and out of the application. Table 6 shows the result of the test.

Table 6: Login Functionality Test Results

Module	Logging in and out of the application
Test Parameters	Login with the correct user credentials
Expected Behaviour	Login success and access granted/success logout
Observed Behaviour	Login success and access granted/Success logout
Test Outcome	Pass

3.5.2 Data Retrieval Functionality Testing

This is where the retrieval of data from the source (twitter) to a Mongo database. The retrieval is done directly from the twitter servers by a web service that incorporates the Twitter API. This is triggered from a web page that was developed. The results of the tests is as shown in the table 7 below.

Module	Data Retrieval and storage
Test Parameters	Data through a search criteria can be retrieved and stored in the database
Expected Behaviour	List of tweets matching the search criteria retrieved and stored in the database.
Observed Behaviour	List of tweets with matching search criteria stored in the database
Test Outcome	Pass

Table 7: Data retrieval functionality tests

3.5.3 Data Pre-Processing Functionality Tests

This is done to ensure that the module functions as expected. The Pre-processing is done by a web service and is expected to remove all unwanted words i.e. stop words, URLs, RT, Videos and audios from the retrieved tweets. The table below shows the results of this test.

Module	Data Pre-processing
Test Parameters	Pre-processing of data
Expected Behaviour	Display a list of tweets that are segmented and have no unwanted words.
Observed Behaviour	List of segmented tweets with no unwanted data
Test Outcome	Pass

Table 8: Data Pre-Processing Functionality Tests

3.5.4 Hybrid Classifier Functionality Tests

This is divided into two sub modules namely;

i. Lexicon Based
ii. Classifier Based

3.5.4.1 Lexicon Based Functionality Tests

Tests were done to ensure that the sub module performed as expected. This was done by parsing different sets of processed tweets into the sub module and the output was noted. The test results are as shown in the table below.

Module	Lexicon Based Classifier
Test Parameters	Labelling of tweets using lexicon and SOCAL approach
Expected Behaviour	List of tweets that have sentiment orientation value attached
Observed Behaviour	Annotated list of segmented tweets
Test Outcome	Pass

Table 9: Lexicon Based Functionality Tests

3.5.4.2 NB Classifier Functionality Tests

This was carried out to ensure labelled tweets from the Lexicon sub module would be accurately parsed to the sub module for training and tests. The classifier was to be trained using 67% of the labelled data while the tests be done using the remaining 33%. 10 tweets were parsed through the sub module while performing the tests. The table below shows the results of the above tests.

Module	NB Classifier
Test Parameters	Correct allocation of labelled data
Expected Behaviour	Only two thirds of the labelled tweets be used as training data while a third as test data.
Observed Behaviour	7 out of the tem tweets were used as training data while 3 was used as test data.
Test Outcome	Pass

Table 10: NB Classifier Functionality Tests

3.5.5 Evaluation page Functionality Test

This was done to ensure that the evaluation module gives results when some tweets are parsed through it. This was done via a web interface that was developed that had a test area where a review is captures and submitted for processing. Tests on this is as shown in the table below.

Module	Evaluation Module

Test Parameters	1. Users can leave an opinion on the page 2. Calculation, display and storage of evaluation metrics
Expected Behaviour	1. Customer opinions be stored in the database 2. Evaluation metrics be displayed and stored in the database
Observed Behaviour	1. Successful storage of customer reviews and evaluation metrics in the database 2. Notification of the metrics was received.
Test Outcome	Pass

Table 11: Evaluation Functionality Tests

4. EVALUATION RESULTS, ANALYSIS AND DISCUSSIONS

4.1 Evaluation Results

Evaluation of performance was done using the method accuracy, precision, recall and the F1 score.

Accuracy is how close a value is to its true value. It is the fraction of number of predictions the model got right and calculated as shown below.

$$Accuracy\ =\ (TP + TN)/\ (TP + FP + FP + TN)$$

Precision is the exactness of a classifier. A higher precision means less false positives while a lower precision means more false positives. It is calculated as shown below.

$$precision = \frac{true\ positives}{true\ positives + false\ positives}$$

Recall (also known as sensitivity) is the fraction of relevant instances that have been retrieved over the total amount of relevant instances. It is calculated as shown below.

$$recall = \frac{true\ positives}{true\ positives\ + false\ negatives}$$

The F1 Score on the other hand is the harmonic mean of the precision and recall. It seeks a balance between precision and recall and its equation is given as follows:

$$F_1 = 2 * \frac{precision * recall}{precision + recall}$$

To evaluate the effectiveness of the model, a new set of **200** labelled reviews was used to evaluate the model, These labelled reviews were selected from a list of reviews or feedback given by Huduma clients. The selected reviews had to meet some conditions namely:

1. Had a service indicated in the review
2. Was written in English – This was due to the fact that 99% of the data extracted from twitter was in English language thus the model was developed using English language.
3. The feedback was relatively unambiguous – this was to minimize the risk of finding inappropriate context.

The annotation of the **200** reviews was manually done to generate a gold standard dataset that was to be used to validate/evaluate the model. The review was quite fast as the services of a Huduma Centre customer care representative was sought. This was because he had quite some knowledge of the domain. Figure below shows a sample of the gold standard dataset used in the evaluation.

S/No.	POSITIVE	NEGATIVE
1.	Good customer service at the NHIF counter	Long ques for ID services especially branches in town
2.	Friendly security guards at the Help Desk who direct you to the right table	Few personnel, so the NTSA services takes long
3.	One stop shop-All services under one roof hence time saver.	Reluctant staff who are not willing to help especially with HELB counter
4.	The service is quick especially with seeking certificate of good conduct from DCI	Poor customer service at CAJ desk
5.	Clean facility especially the branch in town	One has to go outside to do photocopies of documents. Poor services
6.	Short waiting time as there are many service points for NRB	Slow service at the NTSA most of the times there is slow network or computer breakdown
7.	EACC has professional staff unlike in other institutions	Very crowded especially for DL services at the branches in city center
8.	It's very convenient that they open on Saturdays especially for those people who don't get time during weekdays	Staff have attitude while serving – ID renewal
9.	I get emails on the progress of my application (certificate of good conduct) from NPS and am able to tell when it's out	It takes weeks before I can get my certificate of good conduct fro DCI
10.	Prompt response on queries on NHIF from Huduma center website	Some services e.g. CRB are damn expensive

Table 12: sample gold-standard dataset

The evaluation of the model was done by parsing a set of the labelled reviews to both the Naïve Bayes model and the hybrid model designed and developed above that combines both the lexicon analysis and the Naïve Bayes. The results of the evaluation were measured according to tables 14, 15, 16, 17 and 18 below.

Table 13: A confusion table

	True Positive Reviews	True Negative Reviews
Predict positive reviews	True Positive (TP)	False positive (FP)
Predict negative reviews	False Negative (FP)	True Negative (TN)

4.1.1 Accuracy

Positive and negative labelled reviews were manually inputted in the Lexicon, Naïve Bayes and the Hybrid models and the table 14 below shows the accuracy comparison on the models.

No.	Number of customer feedback reviews	Accuracy (Lexicon-SOCAL)	Accuracy (Naïve Bayes)	Accuracy (Hybrid Model)
1.	200	63.07	63.04	67.28

Table 14: Accuracy comparison

The results show that the Hybrid Model scored the highest with a percentage score of 67.28%, as compared with the Lexicon-SOCAL and the Naïve Bayes Model which scored 63.07% and 63.04% respectively.

4.1.2 Precision

200 reviews were inputted into the Lexicon- SOCAL, Naïve Bayes and the hybrid model separately to determine its precision. The table 15 below shows the results of the comparison of precision of both the positive and negative corpus.

Table 15: Precision comparison for both Positive and negative corpus on evaluation datasets.

No.	Number of customer feedback reviews	Precision (Lexicon-SOCAL)	Precision (Naïve Bayes)	Precision (Hybrid Model)
1.	200	63.26	63.46	66.99

The results show that out of the three models, only the hybrid scored a percentage score of more than 65%.It scored 66.99 % while the Lexicon-SOCAL scored 63.26% which was lower than the Naïve Bayes at 63.46%.

4.1.3 Recall

The same dataset of 200 reviews was inputted into the three models to determine and evaluate recall. The table 16 below shows the results of the comparison of both positive and negative corpus on evaluation datasets.

No.	Number of customer feedback reviews	Recall (Lexicon-SOCAL)	Recall (Naïve Bayes)	Recall (Hybrid Model)
1.	200	63.2	61.97	66.57

Table 16: Recall comparison for both positive and negative corpus

From the results above, the hybrid model scored 66.57% while the Lexicon-SOCAL scored 63.2%. The Naïve Bayes scored the lowest at 61.97%.

4.1.4 F1 score

Using the above metrics, the F1 score comparison of the three models is as shown below:

No.	F1 Score (Lexicon-SOCAL)	F1 Score (Naïve Bayes)	F1 Score (Hybrid Model)
1.	63.23	62.71	66.78

Table 17: F1 score comparison of the three models

Summary of the above results are as shown in the figure 17 below.

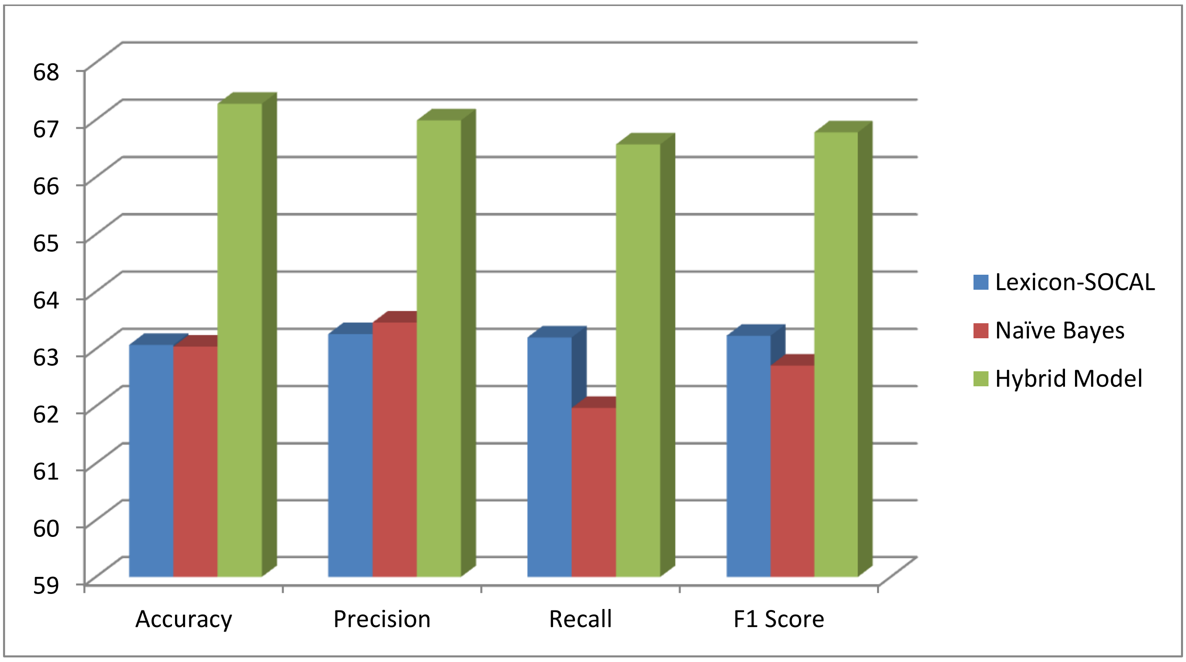


Figure 17: Summary of findings

4.2 Discussions

As shown in the section above, the evaluation carried out shows that the hybrid model outperforms the Naïve Bayes and the Lexicon-SOCAL models. It is best on all performance parameters evaluated. Lexicon-SOCAL on the other hand outperforms the Naïve Bayes in all the parameters but precision. Naïve Bayes has been known to always achieve the best possible accuracy especially when the independence assumption holds, and perhaps close to optimal when the attributes are only slightly dependent. In this case, it performed poorly on all the parameters.

The hybrid model, through the use of both the Lexicon-SOCAL and the Naïve Bayes has been able to perform better than the other models. This may be due to the fact that it combines the advantages of the lexicons and the Naïve Bayes, The use of lexicons in providing training data to the Naïve Bayes classifier improves the performance as shown by the results.

The above metrics of the hybrid model, though higher compared to the lexicons and Naïve Bayes models, was lower and this may have been due to the use of the limited datasets used in training the model. The tweets dataset was limited to a 7 day period and thus may have affected the accuracy, precision and recall scores. Premium twitter APIs that can extract data of up to 5 years would have assisted in improving the performance of the hybrid model. The main disadvantage is that it may lead to overfitting and thus negatively impact the system. Further, since the research did not employ the cross validation technique during training, this may have led to the decrease in the performance metrics.

Application of ensemble technique would have been another way to improving the performance of the hybrid model. Ensembles is where several classifier models are combined through the use of boosting or bagging methods to smooth out predictions and combine them into one hybrid model with a best fit.

5. CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

The aim of this research was to investigate the current technologies that are being used to capture and analyse customer feedback with the aim of designing and developing a hybrid model that would automatically classify customer feedback thus ensuring customer satisfaction. In the research objectives stated in Chapter 1, the first objective was to review and analyse literature on existing models used in collecting and analysing customer feedback. This was achieved through the review of the different models currently being used to prompt for and analyze customer feedback. As discussed in Chapter 2, the main models being used in the capture and analysis of customer feedback are Support Vector Machine and Decision Trees. The Huduma Centres though, have employed the "One click feedback" strategy which only requires a few seconds from a customer to leave a review. This strategy involves a customer clicking a button or a star to like, dislike or rank the content in question. Although this strategy is friendly and saves time for the customer it limits the Huduma Kenya Secretariat from acquiring sufficient, specific and more comprehensive feedback. This study hence concluded there was need for more comprehensive feedback because it enables the management of the various services within Huduma Centres to identify customer issues and also get suggestions, leading to product or service improvements tailored towards customer's preferences.

The second and third objectives were to design a hybrid opinion mining model for analysis of customer feedback in Huduma Centres. This objective was met by the design and development of the Service Based Opinion Mining Application (SBOMA) that integrates the lexicon analysis methods and the Naïve Bayes methods to classify tweets. The DSRM was used to develop the application.

The final objective was to evaluate the hybrid model to measure its effectiveness and efficiency. This was achieved through the use of a gold standard dataset for evaluation of the system. The dataset was inputted into the As shown in the chapter above, this was achieved through the use of a different set of labelled tweets which was used to determine the accuracy, precision and recall of the model.

5.2 Recommendations

Government entities are becoming more customer-centric than before. Most government entities have realised that they are and will be missing out if they do not work towards creating satisfied customers. The traditional ways of knowing the level of customer satisfaction is no longer tenable and the entities are also under pressure to make more data driven decisions. Technology has revolutionised the way government entities operate and has led to exponentially increased data at their fingertips. In order to maintain a high level of customer satisfaction, there is need to be more informed on what their customer expectations are.

Based on the currents trends, it is clear that entities/ organizations are embracing technology and it is the core of most operations. Due to these advancements, this study came up with a simpler way to analyse customer feedback in real-time. This model will ensure that Huduma Centres are aware of what customers are thinking about their services and products. It will promote first contact handling of customer issues, which is what customers appreciate than going round, in search for the management trying to air their issues out. It also lays a platform for management to attract more suggestions and ideas on how they can better their products and services. By doing so they will be able to satisfy customers and improve service delivery.

This application offers a wide variety of functionalities. To begin with it will enable the customer to leave a review of the service offered or general performance of the centre. It further enables the management get to know the specific service unlike other methods that give a general sentiment. The application enables centre managers to view customer reviews about their services after it has been analysed and classified by the sentiment analysis web service using the Hybrid model. This classification enables them tell whether the review was positive, negative or neutral, narrowing down the process of identifying issues pending. If for example the centre manager is after knowing what problems the customers are going through, it would be advisable to go through the negative reviews. Through it also the Centre managers are able to tell which service or staff is performing well than the other hence find a way to get them on track.

5.3 Recommendations for future studies

The use of ensemble techniques can be explored as it is always a better idea to apply it to improve the accuracy of the model. They are generally more complex than traditional methods but this research has provided a good base level from which the model can be improved and create ensembles. The use of other sources of data apart from tweets can also be explored to improve the accuracy levels of prediction. Another possible improvement is through the use of multi-lingual dataset that can be cross validated during training to be able to predict reviews given in different languages.

REFERENCES

Abdalla. (2015). Effect of Huduma centers (one stop shops) in service delivery – a case study of Mombasa Huduma Centre. International Journal of Academic Research. Retrieved 11-22-2018, from http://hrmars.com/ Effect_of_Huduma _centres in_Service_Delivery%E2%80%93 A_Case_Study_of_Mombasa_Huduma_Centre1. pdf

Abhishek, O. (2016). NLP for Text Categorisation/Classification. Retrieved 02 25, 2019, from https://www.slideshare.net/oswalabhishek/text-classificationcategorization-61462404

Afroze, B. I. (2012). Twitter Sentiment Analysis. National University of Sciences and Technology, PAKISTAN, 9. Retrieved 22-11, from https://arxiv.org/ftp/arxiv/papers/1509/1509.04219.pdf

Aurangzeb, K. (2010). A Review of Machine Learning Algorithms for Text-Documents Classification. JOURNAL OF ADVANCES IN INFORMATION TECHNOLOGY, 1(1), 9-10. Retrieved 11-22-2018, from http://www.jait.us/uploadfile/2014/1223/20141223050800532.pdf

Beevers, R. (2006). Customer Service Excellence in the Public Sector. Northern Housing Consortium, Retrieved 08 11, 2018, from

http://digitaltransform.org.uk/wp-content/uploads/2012/08/Public-Sector-Customer-Service-Excellence-report.pdf

Bhatia, S. (2013). Monitoring and Analyzing Customer Feedback Through Social Media Platforms for Identifying and Remedying Customer Problems. IEEE/ACM International Conference on Advances in Social Networks Anal, 1147-1148. Retrieved 02 24, 2017, from http://ieeexplore.ieee.org.ezproxy.library.strathmore.edu/xpls/icp.jsp?arnumber=6785848

Buharudin B., Khan K., Khan A. 2010. Sentence Based Sentiment Classificationfrom Online Customer reviews. Proceedings of the Conference on Frontiers of Information Technology. (F10), ACM, pp:1-6

Customer Champions. (2016). Net Promoter Score - a balanced view. Customer Champions. Retrieved 04 23, 2017, from http://www.customerchampions.co.uk/net-promoter-score-nps-a-balanced-view/

Donald, S. (2015). Understanding Data Flow Diagrams. Integrated Concepts, 3. Retrieved 04 23, 2017, from http://ratandon.mysite.syr.edu/cis453/notes/DFD_over_Flowcharts.pdf

Fahrni A, Klenner M. (2008) Old wine or warm beer: target-specific sentiment analysis of adjectives. In: Proceedings of the symposium on affective language in human and machine, AISB; p. 60–3.

Fearn, N. (2015). Social media analysis tool Tweview launches at Oi16. TechDragons, 1. Retrieved 05 01, 20017, from http://techdragons.wales/social-media-analysis-tool-tweview-launches-at-oi16/

Gera, P. (2015). Banking Customer 2020. Accenture Strategy, 5-8.

Gräbner, D. (2012). Classification of Customer Reviews based on Sentiment Analysis. 19th Conference on Information and Communication Technologies in Tourism, 460-470. Retrieved 02 23, 2019, from http://isbi.at/mzanker/ENTER2012_classification_preprint.pdf

Gronroos, C. (2012). Critical service logic: making sense of value. Academy of Marketing Science, 5. from www.researchgate.net/profile/ChristianGroenros/publication/256395545_Critical_Service_Logic_Making_Sense_of_Value_Creation_and_Co-Creation/links/004635232cc76f395a000000/Critical- Service-Logic-Making-Sense-of-Value-Creation-and-Co-Creation.pdf

Hassett, S., & Blunt, K. (2015). The American Customer Satisfaction Index and CSBG'S Perfomance Management Framework. NASCSP. Retrieved 09 23, 2018, from http://www.nascsp.org/data/files/events/annual%20conference/ocs_csbg_acsi%20presentation_nascsp%20conference_f.pdf

Hsiao, C.-T., & Jie-Shin , L. (2008). A Study of Service Quality in Public Sector. International Journal of Electronic Business Management, 29. Retrieved 08 07, 2018, from https://www.researchgate.net/publication/26542287_A_Study_of_Service_Quality_in_Public_Sector

Kagasi, V. (2016). Influence of Huduma Centre Innovation Strategies on Government Service Perfomance in Kenya. IJRDO- Journal of Business Management, 158. Retrieved 08 11, 2018, from http://www.ijrdo.org/International-Journal-of-Research-&-Development-Organisation-pdf/Business%20mgt/October-2016/Busniess%20Manag-October-2016-7.pdf

Ken Peffers, Tuure Tuunanen, Marcus A. Rothenberger, and Samir Chatterjee (2008) A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems, 24(3): 45–78.

Kiritchenko, S. (2014). Sentiment Analysis of Short Informal Texts. Journal of Artificial Intelligence Research, 1-2. Retrieved from https://www.jair.org/media/4272/live-4272-8102-jair.pdf

Kissmetrics. (2015, November 1). How to Get Customer Feedback from Mobile Users. Retrieved from Kissmetrics Blog: https://blog.kissmetrics.com/feedback-from-mobile-users/

Larman, C. (2007). An Introduction to Object-Oriented Analysis and Design and the Unified Process. Upper Saddle River. Retrieved from https://www.utdallas.edu/~chung/SP/applying-uml-and-patterns.pdf

Maite, Julian Brooke, Milan Tofiloski, Kimberly Voll and Manfred Stede (2011) Lexicon-Based Methods for Sentiment Analysis. Computational Linguistics 37 (2): 267-307.

Mertsalov, K. (2009). Document Classification with Support Vector Machines. Semantics Scholar. Retrieved, from https://pdfs.semanticscholar.org/10d7/0843a86918b19b941cc29d4dc8a9587d9245.pdf

Mohammad, S. A. (2008). An application of the American Customer Satisfaction Index (ACSI) in the Jordanian mobile phone sector. The TQM Journal, 4. Retrieved from http://www.emeraldinsight.com/doi/pdfplus/10.1108/17542731211270098

Natural Resorces Conservation Service. (2011). Customer Satisfaction Survey. USDA. Retrieved, from https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb1045886.docx

Ochieng, S. O. (2016). LIMITATIONS OF SENTIMENT ANALYSIS ON FACEBOOK DATA. International Journal of Social Sciences and Information Technology, 430 - 431. Retrieved from https://www.researchgate.net/publication/304024274_LIMITATIONS_OF_SENTIMENT_ANALYSIS_ON_FACEBOOK_DATA

Oracle. (2011). op 10 Reasons to Choose MySQL for Web-based Applications. A MySQL Strategy Whitepaper, 3. Retrieved from http://www.oracle.com/us/products/mysql/mysql-wp-top10-webbased-apps-461054.pdf

Ordenes, V. F. (2014). Analyzing Customer Experience Feedback Using Text Mining. Sage Journals, 279 - 282. Retrieved from http://journals.sagepub.com.ezproxy.library.strathmore.edu/doi/full/10.1177/1094670514524625

Popescu, A. Yates, and O. Etzioni. (2004). Class extraction from the World Wide Web. In AAAI-04 Workshop on Adaptive Text Extraction and Mining, pages 68–73.

Qiu Guang, He Xiaofei, Zhang Feng, Shi Yuan, Bu Jiajun, Chen Chun. DASA: dissatisfaction-oriented advertising based onsentiment analysis. Expert Syst Appl 2010;37:6182–91.

Ramshaw and M. Marcus. Exploring the statistical deviation of transformational rule sequences for part of Speech tagging. New Mexico State University, July 1994

Reliable Software Resources . (2011). Benefit of Agile Methodolog. Reliable Software Resources Inc, 3. Retrieved, from http://www.rsrit.com/Documents/AgileMethodology_ReliableSoftware.pdf

Rumpe, B. (2016). Sequence Diagrams. Springer International Publishing, 191. Retrieved, from http://www.readcube.com/articles/10.1007/978-3-319-33933-7_6

Rutto, D. (2015). Huduma Centres Revolutionizing Service Provision in Kenya: Lessons for the Information Profession. International Journal of Science and Research (IJSR) . Retrieved 11 22, 2018, from https://www.ijsr.net/archive/v6i6/ART20174146.pdf

Soundarya (2015). Opinion Mining and Feature Extraction – A survey. International Journal of Computer, Volume 14(2), 49.

Taboada, M. and J. Grieve (2004) Analyzing Appraisal Automatically American Association for Artificial Intelligence Spring Symposium on Exploring Attitude and Affect in Text. Stanford. March 2004. AAAI Technical Report SS-04-07. (pp.158-161). Download poster (pdf).

The standard. (2016). Huduma Centres and Power of Service. (p. 15). Nairobi: Ipsos Kenya. Retrieved from http://www.ipsos.co.ke/NEWBASE_EXPORTS/Vision%202030/161107_The%20Standard%20-%20Monday_15_c0daa.pdf

Vishal, G. (2011). Recent Trends in Text Classification Techniques. International Journal of Computer, Volume 35(6), 49. Retrieved from https://pdfs.semanticscholar.org/de05/4d3099d2668ae9f2dc5a28aacace866e883c.pdf

Wisniewski, M. (2010, 08 25). Measuring service quality in the public sector: The potential for SERVQUAL. Taylor and Francis group, 1-2. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/09544129650034710

Weichselbraun, A., Gindl, S., Scharl, A. (2013), "Extracting and grounding context-aware sentiment lexicons". IEEE Intelligent Systems 28(2), 39–46.

Zaki, M. (2016). The Fallacy of the Net Promoter Score: Customer Loyalty Predictive Model. Cambridge Service Alliance, 1. Retrieved from http://cambridgeservicealliance.eng.cam.ac.uk/resources/Downloads/Monthly%20Papers/2016OctoberPaper_FallacyoftheNetPromoterScore.pdf

APPENDICES

Appendix 1: List of stop words

a
able
about
above
abst
accordance
according
accordingly
across
act
actually
added
adj
affected
affecting
affects
after
afterwards
again
against
ah
all
almost
alone
along
already
also
although
always
am
among
amongst
an
and
announce
another
any
anybody
anyhow
anymore
anyone
anything
anyway
anyways
anywhere
apparently
approximate
ly
are
aren
arent
arise
around
as
aside
ask
asking
at
auth
available
away
awfully
b
back
be
became
because
become
becomes
becoming
been
before
beforehand
begin
beginning
beginnings
begins
behind
being
believe
below
beside
besides
between
beyond
biol
both
brief
briefly
but
by
c
ca
came
can
cannot
can't
cause
causes
certain
certainly
co
com
come
comes
contain
containing
contains
could
couldnt
d
date
did
didn't
different
do
does
doesn't
doing
done
don't
down
downwards
due
during
e
each
ed
edu
effect
eg
eight
eighty
either
else
elsewhere
end
ending
enough
especially
et
et-al
etc
even
ever
every
everybody
everyone
everything
everywhere
ex
except
f
far
few
ff
fifth
first
five
fix
followed
following
follows
for
former
formerly
forth
found
four
from
further
furthermore
g
gave
get
gets
getting
give
given
gives
giving
go
goes
gone
got
gotten
h
had
happens
hardly
has
hasn't
have
haven't
having
he
hed
hence
her
here
hereafter
hereby
herein
heres
hereupon
hers
herself
hes
hi
hid
him
himself
his
hither
home
how
howbeit
however
hundred
i
id
ie
if
i'll
im
immediate
immediately
importance
important
in
inc
indeed
index
information

instead
into
invention
inward
is
isn't
it
itd
it'll
its
itself
i've
j
just
k
keep
keeps
kept
kg
km
know
known
knows
l
largely
last
lately
later
latter
latterly
least
less
lest
let
lets
like
liked
likely
line
little
'll
look
looking
looks
ltd
m
made
mainly
make
makes
many
may
maybe
me
mean
means
meantime
meanwhile
merely
mg
might
million
miss
ml
more
moreover
most
mostly
mr
mrs
much
mug
must
my
myself
n
na
name
namely
nay
nd
near
nearly
necessarily
necessary
need
needs
neither
never
nevertheless
new
next
nine
ninety
no
nobody
non
none
nonetheless
noone
nor
normally
nos
not
noted
nothing
now
nowhere
o
obtain
obtained
obviously
of
off
often
oh
ok
okay
old
omitted
on
once
one
ones
only
onto
or
ord
other
others
otherwise
ought
our
ours
ourselves
out
outside
over
overall
owing
own
p
page
pages
part
particular
particularly
past
per
perhaps
placed
please
plus
poorly
possible
possibly
potentially
pp
predominant
ly
present
previously
primarily
probably
promptly
proud
provides
put
q
que
quickly
quite
qv
r
ran
rather
rd
re
readily
really
recent
recently
ref
refs
regarding
regardless
regards
related
relatively
research
respectively
resulted
resulting
results
right
run
s
said
same
saw
say
saying
says
sec
section
see
seeing
seem
seemed
seeming
seems
seen
self
selves
sent
seven
several
shall
she
shed
she'll
shes
should
shouldn't
show
showed
shown
showns
shows
significant
significantly
similar
similarly
since
six
slightly
so
some
somebody
somehow
someone
somethan
something
sometime
sometimes
somewhat
somewhere
soon
sorry
specifically

specified
specify
specifying
still
stop
strongly
sub
substantially
successfully
such
sufficiently
suggest
sup
sure
t
take
taken
taking
tell
tends
th
than
thank
thanks
thanx
that
that'll
thats
that've
the
their
theirs
them
themselves
then
thence
there
thereafter
thereby
thered
therefore
therein
there'll
thereof
therere
theres
thereto
thereupon
there've
these
they
theyd
they'll
theyre
they've
think
this
those
thou
though
thoughh
thousand
throug
through
throughout
thru
thus
til
tip
to
together
too
took
toward
towards
tried
tries
truly
try
trying
ts
twice
two
u
un
under
unfortunatel
y
unless
unlike
unlikely
until
unto
up
upon
ups
us
use
used
useful
usefully
usefulness
uses
using
usually
v
value
various
've
very
via
viz
vol
vols
vs
w
want
wants
was
wasnt
way
we
wed
welcome
we'll
went
were
werent
we've
what
whatever
what'll
whats
when
whence
whenever
where
whereafter
whereas
whereby
wherein
wheres
whereupon
wherever
whether
which
while
whim
whither
who
whod
whoever
whole
who'll
whom
whomever
whos
whose
why
widely
willing
wish
with
within
without
wont
words
world
would
wouldnt
www
x
y
yes
yet
you
youd
you'll
your
youre
yours
yourself
yourselves
you've
z
zero

Appendix 2: Sample extracted data

The gov should easen provision of services to the pple not treats over huduma number @ntvkenya #sidebar
@noelkemunto Quality service and very accurate #Huduma #CitySquare.
/RT @njrogoge434 #Huduma couch was clean and comfortable. https://bit.ly/2I0AaJS
@Mintejua Efficient and effective infrastuctures. Huduma
@Gerogertheman Polite and understanding staff from the enquiries depertment to the management https://bit.ly/2I0Aafds #Huduma
@mintish Affordable services and products. https://bit.ly/2IthehyewS This was just awesome. Cc @Jumabaya Huduma
@ronoben No queues and the staff are cheerful. What else can make you smile after a good service from a huduma you are loyal to.@fredmatiangi
@jeremi234 Dishonest staff. #HELB @HelbKenya #Huduma
@manueli Good service at #EACC #Huduma Server next time we need tea :)
@Mikemmukera Quite good #Huduma #GPO at providing some good old fashioned spooks
@Rop Exceeds all expectations. #Hudumakenya
@RisperMarley Good value good location good surroundings. #Huduma service
Soon 99% of services will require you have this crazy huduma number
Kenyans need jobs, food security, affordable housing, quality healthcare and education services. The despotic regim
@Emymun So we have no choice but to register with blind hope that we will get first class services from government with the huduma number.
@KTNKenya @StandardKenya Can someone kindly give a list of services one get with huduma number? Why is it that if
@KTNKenya @SophiaWanuna @SakajaJohnson @lindahoguttu @KTNNewsKE Ask @SakajaJohnson ,is a national register what aff
@kayteshierow @MediaMK @Mwirigi @kayteshierow hallo, we are in partnership with Huduma to bring our services close
@dailynation @NationBreaking @ntvkenya @KTNKenya @KTNKenya @citizentvkenya @RadioCitizenFM We would like to know wh
@EtalePhilip Huduma namba registration is the worst excuse this administration has come up with so far to manipulat
@M_otweyo @EtalePhilip @RailaOdinga Uhuru has stated categorically without Huduma No. one will be able to access Government services.
@KTNKenya Having used and enjoyed the services at Huduma Centre,It would have been more beneficial if they were lau
@MigunaMiguna @RailaOdinga @JubileePartyK Huduma center, devolution and ecitizen. Billions collected from e service
@MigunaMiguna @RailaOdinga @JubileePartyK The main reason to why the government hasn't been able to serve citizens
@FredMatiangi Provided your boys don't harass we who are opposed to this huduma namba thing, I hope you are very aw
High Court in Kenya has suspended implementation of significant aspects of the National Integrated Identification M
Does improving the deliverance of Services require #HudumaNamba really? Definitely there must be an inconspiciuo
If someone was wise enough, why wouldn't we incorporate the #HudumaNamba in our national ID? It would have been bet
#HudumaNamba then #CensusKE, the two would ideally have been done the same time but ideal is not a GOK language. Th
How does a number improve services? Mental transplant is what those huduma center guys need to offer better services... #HudumaNamba
Is not funder mental right for the government to force citizens to register Huduma No(nims) so that the may get the

Am not registering for Huduma number, serikali iwache kutuweka ukoloni, colonization is over. Am not getting gov't
@KenyanProcess @ItsMutai @ahmednasirlaw Alafu these benefits hazimake sense, we could still access the said service
HIGH COURT RULING Registration Exercise to continue on condition that: DNA or GPS information suspended Registra
No one will be able to access Government services bila Huduma Number.
@DavidNdii Every human being got moral right to oppose any sort of injustice.. the president should come out clearl
@DavidNdii Erm, @KanzeDena didn't edit his speech to reflect the limits the court set on implementation of Huduma N
How does huduma number help me get better services. If u cant pay nurses how will huduma number give me better heal

Appendix 3: Sample code

1. Naïve Bayes Classification

```
public function analyzeSentiment($sentiment){
        foreach ($this->arrTypes as $type)
        {
                $this->arrBayesDistribution[$type] = $this->arrSentenceType[$type]
/ $this->cntSentence;
        }
        $sentimentScores = array('positive', 'negative');
        $words = self::splitSentence($sentiment);
    $words = $words[0];
        foreach ($this->arrTypes as $type)
        {
                $sentimentScores[$type] = 1;
                foreach($words as $word)
                {
                        if (!isset($this->arrSentiments[$word][$type]))
                        {
                                $tracker = 0;
                        }
                        else
                        {
                                $tracker = $this->arrSentiments[$word][$type];
                        }
                        $sentimentScores[$type] *= ($tracker + 1) / ($this-
>arrWordType[$type] + $this->cntWord);
                }
                $sentimentScores[$type] *= $this->arrBayesDistribution[$type];
        }
        arsort($sentimentScores);
        if (key($sentimentScores) == 'positive')
        {
                $bayesDifference = $sentimentScores['positive'] /
$sentimentScores['negative'];
        }
        else
        {
                $bayesDifference = $sentimentScores['negative'] /
$sentimentScores['positive'];
        }
        $positivity = $sentimentScores['positive'] / ($sentimentScores['positive'] +
$sentimentScores['negative']);
        $negativity = $sentimentScores['negative'] / ($sentimentScores['positive'] +
$sentimentScores['negative']);
        if (in_array(round($bayesDifference, 1), $this->arrBayesDifference))
        {
                $sentiment = 'Neutral';
        }
        else
        {
                $sentiment = key($sentimentScores);
        }
        $this->sentresult = array('sentiment'=>$sentiment,
'accuracy'=>array('positivity'=>$positivity, 'negativity'=>$negativity));
```

```
return $this->sentresult;
        }
```

2. Lexicon –SOCAL

```
public function analysisSentimentLexicon($sentiment){
            $adj_dictionary = $this->getDictionary('adj');
            $adv_dictionary = $this->getDictionary('adv');
            $int_dictionary = $this->getDictionary('int');
            $noun_dictionary = $this->getDictionary('noun');
            $verb_dictionary = $this->getDictionary('verb');
            //dd($verb_dictionary);
            $analyzer = new Analyzer();
            $result = $analyzer->getSentiment($sentiment);
            return $result;
        }
```

3. Removing Stop words

```
public function removeStopwords($str = "")
        {
         $stopwords = array();
         $file = asset('resources/assets/stop_words.txt');
            $fp = @fopen($file, 'r');
            // Add each line to an array
            if ($fp) {
               $stopwords = explode("\r\n", fread($fp, 10000));
               //dd($stopwords);
             }
          $words = preg_split('/[^-\w\']+/', $str, -1, PREG_SPLIT_NO_EMPTY);
          if (count($words) > 1) {
            //  $words = array_filter($words, function ($v) use (&$stopwords) {
            //   return !isset($stopwords[strtolower($v)]);
            // }
            // );
             foreach( $words as $key => $word ) {
                     if(in_array(strtolower($word), $stopwords) ) {
                         unset($words[$key]);
                     }
                 }
            //dd($words);
          }
          if (empty($words)) {
            return $str;
          }
          return implode(" ", $words); }
```