Proceedings of the Fourth International Conference on Inventive Systems and Control (ICISC 2020)
IEEE Xplore Part Number: CFP20J06-ART; ISBN: 978-1-7281-2813-9

Integrated Feedback Analysis And Moderation Platform Using Natural Language Processing

1st Ambarish Moharil
Department of Instrumentation and Control Engineering
Vishwakarma Institute Of Technology
Pune, India
ambarish.moharil17@vit.edu

2nd Shreya Singh
Department of Instrumentation and Control Engineering
Vishwakarma Institute Of Technology
Pune, India
shreya.singh17@vit.edu

3rd Yash Dravid
Department of Instrumentation and Control Engineering
Vishwakarma Institute Of Technology
Pune, India
yash.dravid17@vit.edu

4th Harshal Dharap
Department of Instrumentation and Control Engineering
Vishwakarma Institute Of Technology
Pune, India
harshal.dharap17@vit.edu

5th Prof. Vijaykumar Bhanuse
Department of Instrumentation and Control Engineering
Vishwakarma Institute Of Technology
Pune, India
vijaykumar.bhanuse@vit.edu

***Abstract*— Feedback and review moderation has become a common practice by service providers and universities. Reviews and feedbacks have the unexplored potential that can affect an organization's success as well as it's functioning. Hotels especially realize the potential of reviews, they help them to improvise their customer service and improvise as a service provider, but what they lack is an integrated service that aims towards personal care of each and every client that would address and resolve their grievances on an individual scale. This model not only analyses the sentiment of feedbacks and reviews from users but also helps the administrators track grievances with the minutest details. This model will make review analysis and moderation comfortable for service providers like hotels and educational institutions and they could know the sentiment and also the actual problem or compliment of the review in one word without reading it and having a properly maintained database for the same. This model aims towards improvising the client experience with minimum cost and utmost care. This idea stands out different from other existing platforms in such a way that it can be used anywhere, offices, hotels, educational institutions, etc. and will definitely reduce the time and cost of customer/student review and help address the grievances quickly and efficiently.**

Keywords*— *NLP, Support Vector Classification, Flask, Selenium, Spacy, machine learning.

I. INTRODUCTION

Most of the industries, service providers and educational institutions realize the importance of feedbacks. Feedbacks and reviews are a very integral part of these organizations and reflect and act as a mirror that reflects the quality of service and clearly distinguishes the room for improvement. With the rising tools and techniques for sentiment analysis, it is possible to review the sentiment and rate these reviews but the whole system very uses specific.[6] Most industries still lack a tool that is integrated and can be used by anyone irrespective of the industry. There are tools to review products, applications and even reviews that go online about hotels, this may include malpractices to increase the ratings and the main purpose of feedbacks is ignored. It is important to bring into light that feedbacks act as a mirror and should only be used to increase the functioning efficiency of organizations and to provide clients or users the best service and experience they deserve. Usually, university students face grievances about the campus infrastructure, food, etc. and the feedback system is slow and inefficient. A tool is required that not only quickly analyses this feedbacks but also directs administrators to address these issues. Industries and organizations are in dire need of a system that will direct the administrators about the sentiment as well as about the whole complement or the problem in a couple of words. Using efficient NLP techniques and mathematical algorithms implemented using machine learning it was possible for us to have achieved this goal of cost-effective review moderation and analyses and problem tackling.[3] For service industries like hotels, restaurants, customers providing constantly good reviews are assets and they don't want to lose them, a timely updated list of these customers can help these industries to improve and suggest their products and services on an individual scale. This model not only aims to solve the problems faced by customers or people but also aims to solve them efficiently and provide a solution so that they can be addressed on an individual scale. Traditional review systems only analyze the sentiment, This algorithm aims to offer a platform that not only rate and analyze the feedback but also provide the impact features of the review and form a categorized database of users. It also provides to receive feedbacks in a categorized manner and not in a generalized way that helps the organization to address the issue more efficiently. NLP tools like spacy, and vectorization techniques like TF-IDF, count-vectorization can help us achieve the goal. Organizations will be able to have feedback about each and every sector in an organized manner with the sentiment with the problem and compliment clearly highlighted with support of vast languages.

II. NATURAL LANGUAGE PROCESSING

[3] Natural Language Processing also termed as NLP is a branch of artificial intelligence that helps machines to understand and process human language. Humans communicate in different languages and language plays a

Authorized licensed use limited to: Carleton University. Downloaded on August 23,2020 at 14:21:46 UTC from IEEE Xplore. Restrictions apply.

key role in communication. Humans interpret the language and learn to understand what is being communicated. NLP is a tool that provides machines or computers with this ability. It has been around a decade since NLP is being used in numerous fields to solve problems and come up with effective and smart solutions, but the scope of this field is extremely vast and is quite a lot still unexplored. Modern NLP toolkits have been designed on python and libraries like NLTK, Spacy help innovators to come with solutions to solve modern-day problems. [2] Natural language processing uses mainly two techniques syntax and semantic analysis. When words are arranged in a proper order considering the grammatical sense of the language, it is called syntax. NLP makes use of syntaxes to interpret and understand the meaning of the language based on the grammatical rules. Parsing (analysis of a certain sentence), segmentation, lemmatization, morphological segmentation, stemming (breaking the word into its root form), word segmentation is some of the syntax techniques that natural language processing uses. Semantics is also a crucial part and plays a vital role, apart from the usage of a particular word it is very important to understand where and how the word was used. The placement of a particular word can totally change the meaning of a sentence. Semantic NLP techniques analyze the use and meaning behind the words. It is quite possible to apply algorithms that understand this structuring of a particular language and can help in deeper and accurate interpretation. Word sense disambiguation (deriving the meaning of the word based on the context), named entity recognition, natural language generations are some of the sematic techniques that NLP uses.

Extensive research about a lot of libraries and their integrations, feasibility, and simplicity was conducted. Explicitly developing algorithms in various NLP libraries available for python was one of them. [6] The most popularly used libraries today are NLTK, Genism, Intel NLP Architect and natural language toolkit, etc. But one library with extensive potential remains unexplored today, it is spacy. Spacy is a very powerful library built for python to perform NLP operations. Complex operations used or scripted in NLTK can be very easily performed in spacy.

Spacy has a different set of decisions than that of NLTK and are proved to be more accurate. The major difference that makes spacy stand out from other libraries is that Spacy tries to avoid asking the user to choose between multiple algorithms that deliver equivalent functionality.

The language support of spacy is extensive and it supports a variety of languages. While solving the issues and considering the linguistic diversity of India, spacy was selected which has extensive language support for Indic languages like Tamil, Hindi, Marathi, Kannada, Malayalam, etc. this made our NLP engine flexible and it became not only possible but also easier to analyze sentences from multiple languages.

III. Dataset

A huge database of reviews for the research was created. A lot of popular datasets are available online but to make a platform that can be integrated into multiple fields it was challenging to use a specific set of data.

The dataset that was used consisted of 4 datasets and a manually developed data scrapper to scrap data. When a platform has to be used in various fields it is very important for the NLP engine to train on different kinds of sentences and structure formations, also it is very important is to understand comments and reviews used for various cases.

The amazon, yelp and IMDb datasets of movie reviews along with a dataset of feedbacks were obtained from Stanford University. Considering the requirements of the model, product reviews and user reviews, reviews, and feedback from government websites and open-source platforms were fetched using a dynamically scripted scrapper using selenium and beautiful soup and were classified as positive and negative represented by 1 and 0 respectively.

The dataset consisted of 6942 rows and two features viz message (review) and target (sentiment).

IV Website

The user/customer will use the website to provide his/her reviews. The reviews, which are entered at the front end, are sent to a machine learning algorithm that is implemented at the back end. For demonstrating a use case, the website is implemented for college application, in which reviews(feedback) regarding the functioning of different equipment in a lab are entered on the website.
The implementation of the website is done using Flask, Bootstrap CSS and HTML. Flask Framework is used for back end implementation whereas Bootstrap CSS and HTML are used for implementing the front end for the website.

The passwords were encrypted using SHA (secure hash algorithm) 256 encryption method to maintain privacy so that the admin can't access or know the user passwords. The SHA-256 encryption is an encryption technique used in message authentication, digital signatures, etc.

The website contains a user-friendly navigation bar for the user to visit different web pages. For the use case application, a dropdown menu for selecting a Department is provided. After selecting the department, the user can then choose the desired lab/room. Then, the user can select the equipment for which review/feedback is to be provided.
Also, the impactful words which actually contribute to sentiment analysis will be collected from the reviews and stored in a TF-IDF matrix to generate a TF-IDF score. This score can be accessed by a 'Score Generator' button on the website, only by the admin.

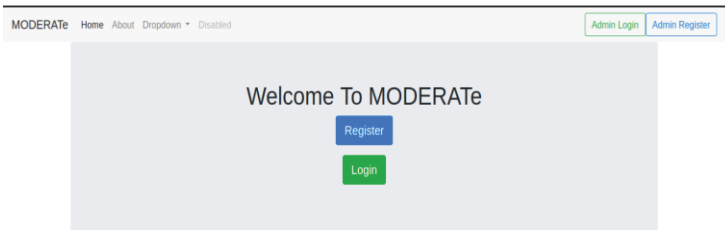

Fig. 1. Home Page

Authorized licensed use limited to: Carleton University. Downloaded on August 23,2020 at 14:21:46 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Fourth International Conference on Inventive Systems and Control (ICISC 2020)
IEEE Xplore Part Number: CFP20J06-ART; ISBN: 978-1-7281-2813-9

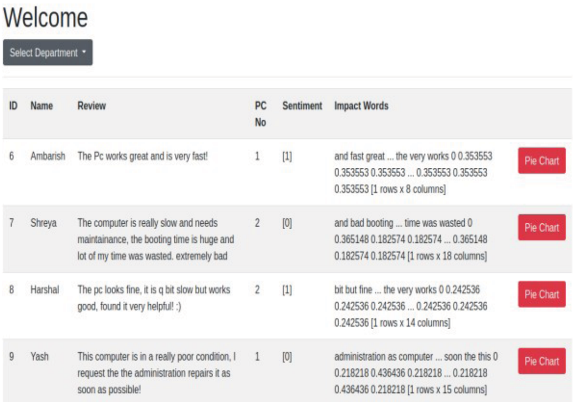

Fig. 2. Admin Dashboard

Fig. 3. Feedback Entry

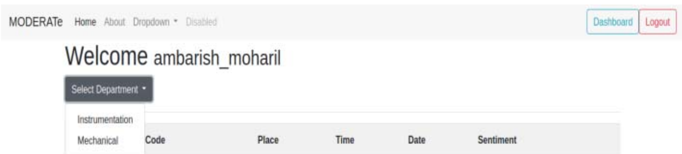

Fig. 4. User Dashboard

V. Architecture

The architecture consists of a pipeline of a text cleaner, word tokenizer, and finally the classifier

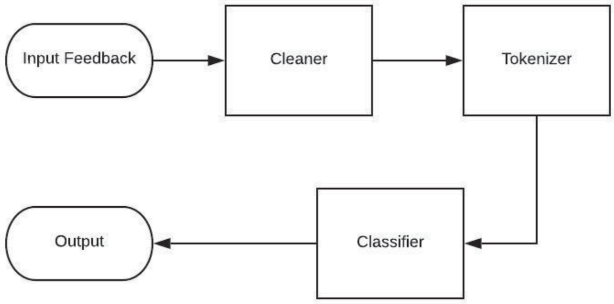

Fig. 5. The pipeline of the architecture

A. Cleaner

Pre-Processing of raw data is a very important and unavoidable step in designing an NLP engine and a machine learning classifier. [6] Raw data consists of numerous garbage elements which are of no use in the analysis of the data or can't be interpreted by the machine. The first step towards text cleaning is finding and removing all special characters and symbols. Symbols like ",", "!", "@ "etc. have no interpretive value and are unnecessary and will be of no use even if they are tokenized in the next step, so it becomes quite evident to remove such character from the raw data. Furthermore, words like "and" "or" "the" basically articles prepositions termed as "stopwords" are repetitive in nature and hold no individual meaning. So, all these words were removed by using suitable regular expressions. Spacy consists of a huge stopwords dataset and hence it becomes quite easier to remove these stopwords from the dataset to be tokenized. [9] Another crucial step that was carried in the cleaning of raw data after removing the special characters and stopwords was stemming and lemmatization of the data Stemming is a very important step and hence it was carried out in the cleaning part of the architecture itself. What happens in stemming is that any gerund or a verb in a continuous form to a root form, it removes the affixes and makes the word in a root structure which makes it easier to analyze during vectorization. It can be referred to a crude process that chops off the ends of the words, it can be considered to remove derivational affixes.

caresses → caress
ponies → poni
caress → caress
cats → cat

Fig. 6. Stemming

Lemmatization usually refers to the use of vocabulary and morphological analysis of the words, it aims to return the base or dictionary form of a word which is usually referred to as a lemma. If confronted with the token *saw*, stemming might return just *s*, whereas lemmatization would attempt to return either *see* or *saw* depending on whether the use of the token was as a verb or a noun.

Authorized licensed use limited to: Carleton University. Downloaded on August 23,2020 at 14:21:46 UTC from IEEE Xplore. Restrictions apply.

B. Tokenizer

[2] Tokenization is used to vectorize a given text corpus. Tokenization usually works by converting a word into a token more specifically a token vector i.e. a sequence of integers where each integer is the index of the token from a text dictionary. Tokenization also takes place by converting a string type into a vector where the coefficient of each token can be binary, based on the word count or based on TF-IDF.

Each and every word of the sentence after cleaning was broken up and was converted into a token which helped in the dependency parsing

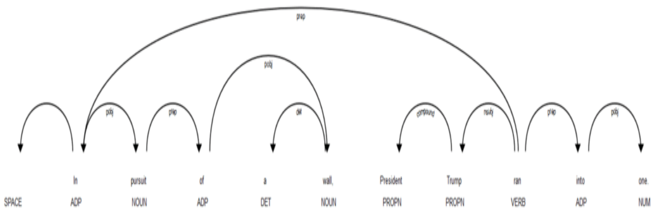

Fig. 7. Dependency parsing

The texts are tokenized into sentences that further are tokenized into individual words. These tokenized word vectors help in deciding the dependency in the stamen formation and also helped in computing the frequency of occurrence.

The major two tokenizers used in our case were count vectorizer and TF-IDF tokenizer.

[9]TF-IDF calculates the term frequency i.e. how many times the word appears in the given document but instead of calculating the raw frequency of occurrence of words, TF-IDF also calculates something called inverse frequency, like some words which are common like “are” that were not removed during pre-processing but are important may have high frequency of occurrence, so what TF-IDF did was divided it with document frequency i.e. the inverse frequency, so a particular word that occurred frequently in every document will have a lesser score than the word that occurs rarely in a document stressing the importance of that feature.

Now while analyzing the sentiments the admin could view the TF-IDF score matrix for words having a threshold greater than 0.5 implying that those were the impact features and would summarize the whole review or feedback into two to three words instead of reading it completely.

For a term t in document d in n documents, TF and IDF is computed as follows

tf-idf(t, d) = tf(t, d) * idf(t)

idf(t) = log [n / df(t)] + 1

Countvectorizer calculated nothing but the raw frequency of the occurrence of words in a frequency matrix. Both of these algorithms were tested and successfully implemented.

C. Classifier

[11] To classify the reviews as positive and negative an SVC Linear classifier was used. SVC linear classifier is used in multiclass classification and gave better results than Naïve Bayes probabilistic classification approach. Though the Naïve Bayes classifier is used in word count cases SVC was selected due to its ability to classify multimodal classes. As stated earlier the raw data had to be tokenized in order to process it later, due to this tokenization of 6942 reviews each of those 6942 reviews was represented into 14324 features, representing the TF-IDF score for different bigrams and unigrams. Naïve Bayes used a radial basis function as the kernel, basically a gaussian kernel or more commonly known as a bell curve while SVC has a linear kernel against a nonlinear kernel in case of the gaussian curve, in our case the linear kernel helped us better and produced more accuracy than that of the non-linear kernel. The SVC has a linear hyperplane that easily classifies data into multiple classes. As the classification in our case was into only 2 classes it made more sense in utilizing the linear hyperplane as it would obviously yield more accuracy instead of complicating and producing inaccurate results by using a gamma hyperplane and a nonlinear hyperplane.

The cleaner, tokenizer and the classifier were connected together in a compact pipeline.

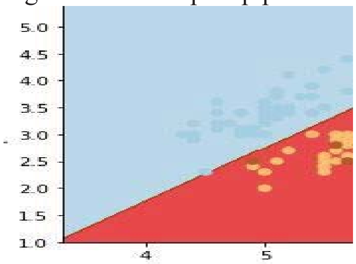

Fig. 8. Linear Hyperplane

VI. ENSEMBLE LEARNING

Ensemble learning is a process that is practiced quite often in machine learning. [12] Various models like decision trees and random forests work on the algorithm of ensemble learning. Ensemble learning can be described as combining the output of multiple models into a single output i.e. various models are cascaded in a parallel connection and the output of these models is converted into a singular output combining several predictive models into one.

In order to make the model more accurate considering the varied use case two models were created, rather two pipelines were formed and then combined the output in order to increase the accuracy. In certain cases, considering the length and vocabulary used in the review or feedback it may be possible that feedback is misclassified. So, in order to reduce this discrepancy and misclassification, two algorithms with different parameters were used.

Certain parameters in the vectorizer like min_df, max_df, etc. were tweaked and changed after numerous iterative cases and these two distinct values were selected for two algorithms, also in some cases count vectorizer was used in another pipeline where tf-IDF failed to produce accurate results.

Authorized licensed use limited to: Carleton University. Downloaded on August 23,2020 at 14:21:46 UTC from IEEE Xplore. Restrictions apply.

In the end the output of these two pipelines was combined and the weighted average of these two models was obtained as the final output.

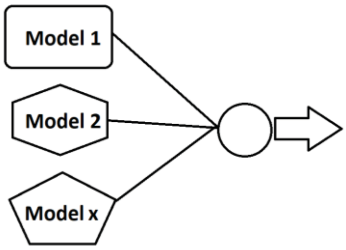

Fig.8. Representation of ensemble learning

VII. Results

A classification testing accuracy of 80.34% was attained The Following are the visuals for the analysis of the reviews.

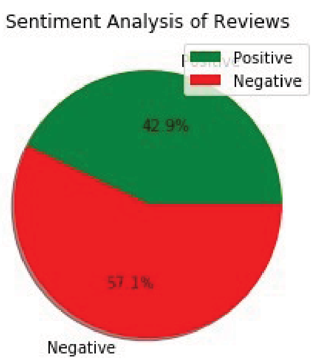

Fig. 9. Pie chart representing positive and negative reviews

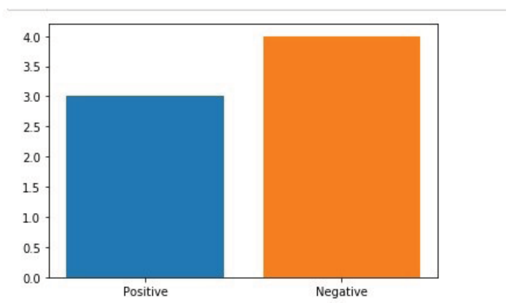

Fig. 10. Bar Graph of the reviews analysed

	amazing	and	beautiful	clean	dirty	experience	good \
0	0.389555	0.296267	0.296267	0.389555	0.00000	0.389555	0.000000
1	0.000000	0.354996	0.354996	0.000000	0.00000	0.000000	0.466778
2	0.000000	0.000000	0.000000	0.000000	0.52004	0.000000	0.000000

	it	loved	playground	room	the	very	was
0	0.296267	0.296267	0.00000	0.296267	0.230077	0.00000	0.230077
1	0.354996	0.354996	0.00000	0.354996	0.275686	0.00000	0.275686
2	0.000000	0.000000	0.52004	0.000000	0.307144	0.52004	0.307144

Fig. 11. Impact words for a particular review.

All these results can be obtained for the feedback of a particular area or a sector of an organization that will help in responsive and better administration.

VIII. Conclusion

A platform to analyze categorized feedbacks was successfully developed. It is possible to integrate and implement this platform almost everywhere. The feedbacks were successfully classified into the two categories and the impact score of each review was represented. All the reviews were categorically sorted in a constantly updating database with the author of the feedback. It was possible to access all the reviews posted by any author or user on the platform with their sentiment. A visual and graphical representation for every review for each sector of an enterprise or organization was provided. The admin could add respective sectors according to his requirements and the reviews or feedbacks for those sectors could be then analyzed accordingly with mentioned features A website for this platform was successfully implemented which acted as a powerful API for the machine learning model running at the backend or cloud. The model reduces the processing time of feedbacks as the admin can analyze the impact words and sentiment and extract the exact grievance or compliment. The model eradicates manual processing of feedbacks, it saves the resources and eliminates the need to analyze paper-written reviews of certain institutions saving months of time to review and act on the issues.

References

[1] Erik Kauffman, Jesus Pearl, David Gill, Antonio Fernandez, Ricardo Sellers, Higinio Mora, “A Framework for big data analysis in commercial social networks: A case study on sentiment analysis and fake review detection” August 2019.

[2] Sudanshu Kumar, Mahendra Yadav, Partha Pratim Roy “Fusion of EEG response and sentiment analysis of products review to predict customer satisfaction”,information fusion ,vol 2, pp 41-52

[3] Kusrini, Mochamad Mashuri, “Sentiment Analysis In Twitter Using Lexicon Based and Polarity Multiplication”

[4] Yasir Ali Solangi, Zulfiquar Solangi, Samreen Aarain,Amna Abro,Ghulam Ali Mallah , Assadullah Shah “Review on Natural Language Processing (NLP) and Its Toolkits for Opinion Mining and Sentiment Analysis,”

[5] R. Nicole, “Evaluating the performance of sentence level features and domain sensitive features of product reviews on supervised sentiment analysis tasks., 12 september 2019.

[6] Rachna Bandana “Sentiment Analysis of Movie Reviews Using Heterogeneous Features”, in press.

[7] Tiago de mello, Altigran da silva, Edlenao S. de Moura “An aspect-driven method for enriching product catalogs with user opinions”, 28 November 2018.

[8] Ana Valdiva, Emiliya Hrabova, Iti Chaturvedi, M. Victoria Luzon, Luigi Troiano, Erik Cambaria, Fransisco Herrera“Inconsistencies on TripAdvisor reviews: A unified index between users and Sentiment Analysis Methods”, 11 August 2019.

[9] RizkaPutriNawangsari[a] RetnoKusumaningrum[a] AdiWibowo “Word2Vec for Indonesian Sentiment Analysis towards Hotel Reviews: An Evaluation Study”, in press

Authorized licensed use limited to: Carleton University. Downloaded on August 23,2020 at 14:21:46 UTC from IEEE Xplore. Restrictions apply.

[10] Jian-WuBi[a]YangLiu[a]Zhi-PingFan[ab] "Representing sentiment analysis results of online reviews using interval type-2 fuzzy numbers and its application to product ranking", in press

[11] Characterizing Social Relations Via NLP-Based Sentiment Analysis Georg Groh, Jan Hauffa

[12] NLP based sentiment analysis on Twitter data using ensemble classifiers, Monisha Kanakaraj ; Ram Mohana Reddy Guddeti

Authorized licensed use limited to: Carleton University. Downloaded on August 23,2020 at 14:21:46 UTC from IEEE Xplore. Restrictions apply.