Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 6 (2025)

Sentiment Analysis of Twitter Data on Procrastination Using Decision Tree Navie Base and Deep Learning

[1]Huda Begum, [2]Dr. Fouiza Jabeen, [3]Dr. Salma Noor

Article Details

Keywords: Sentiment Analysis, Decision Tree, Naive Base, Machine Learning, Deep Learning

Huda Begum
Computer Science & Shaheed Benazir Bhutto Women University Peshawar, Pakistan.
hudakhattak95@gmail.com
Dr. Fouiza Jabeen
Computer Science & Shaheed Benazir Bhutto Women University Peshawar, Pakistan.
fouiza.jabben@sbbwu.edu.pk
Dr. Salma Noor
Computer Science & Shaheed Benazir Bhutto Women University Peshawar, Pakistan.
dr.salmanoor@sbbwu.edu.pk

ABSTRACT

Today's social networking sites are the exploration, so large amount of data is generated. Millions of people are sharing their veiws daily on microblogs such as Facebook , Twitter and Instagram. In this paper, explore Sentiment analysis to identify the polarity (Positive, Negative and Neutral) along with emotion for procrastination using machine learning and deep learning algorithms. By analysing related research papers and articles to identify related factors and attributes and suggest 'how procrastination is cause of stress or not". The approach used to solve the above problem includes preprocessing and analysis of tweets using data mining tools such as RapidMiner with plug-in Aylien and Indico API in Python language using twitter data. The algorithm used at the backend of emotion API is LSTM (Long Short Term Memory). The proposed models were evaluated on two different tasks. First, a decision tree classifier was used to predict polarity on datasets of 300 and 500 samples, achieving accuracies of 72.22% and 68.67%, respectively. In contrast, a Naive Bayes classifier performed significantly worse on the same polarity task, yielding accuracies of 33.33% on the 300-sample dataset and 26.67% on the 500-sample dataset. Secondly, the models were assessed on their ability to distinguish between subjective and objective parts of tweets. For this task, the decision tree classifier achieved a precision of 86.67% when trained on 500 samples.

INTRODUCTION

Opinion mining (sometimes known as sentiment analysis or emotion AI) refers to the use of natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract, and study affective states and subjective information. Sentiment analysis mainly focuses on opinions which show positive and negative sentiments. Sentiment Analysis is the study of people views, judgment, attitudes, and emotion towards entities such as services, organisation, events, topics and their attributes[1]. Some SA jargon: Semantic orientation, polarity. For this task, machine learning and deep learning classifiers are used to determine the Polarity of people's actions with their feeling of how many people react positively or negatively.Sensitivity analysis is significant branch of information to determine the opinions of people. If an organisation conduct survey to recognize the feeling of people towards its product or any kind of services.

Sentiment analysis is also used to predict and analysis the opinion of people about different products, movie revives, product revives etc. Sentiment analysis then plays an important role in any subject (policy makers, stakeholders, companies, etc.) to perform different kinds of activities such as predict financial performance, understand consumer perception provide early warnings, define election result, etc.

The tools used for Sentiment Polarity to identify people's feelings are sentiWordNet, Emoticons, SentiStrengh, Happiness Index, PANAS-t, Sentiment140, EWGA and FRN, LIWC, AFINN, NRC. Sentiment analysis is used in different areas such as politics, marketing, etc. It is a new field in NLP. The main edge is to resolve the Polarity, Subject part of the sentences, Adjectives, Verbs and Nouns etc.

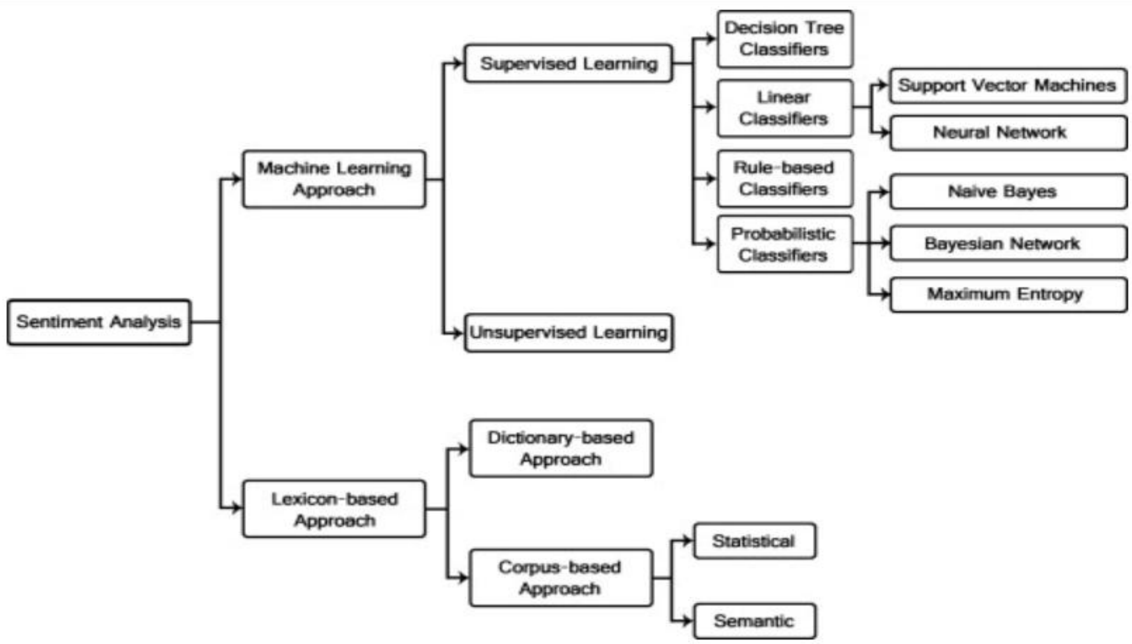

FIG.1: TECHNIQUES FOR THE SENTIMENT CLASSIFICATION APPROACH

In Sentiment Analysis Machine learning plays a vital role to identify the Sentiment Polarity which is based on test and train data sets. Machine learning has the ability to determine the performance of people's views. Machine learning consists of many of the algorithms and classifiers which trained the data using models to give the accuracy of the data with different percentages. However, it also has disadvantages due to the low capability of the methods on new data to perform the best accuracy. Machine learning has two approaches i.e. Supervised and Unsupervised methods. Machine learning has the ability to resolve the problem in changed areas of part with big data or corpus in image processing, face recognition, motion detection and natural language processing.

Machine learning plays important role because it helps us understand the people's behavior and their views to predict the views and pattern recognition of human behaviour which can't be predict by human itself. It has many applications for machine learning, and the most important portion is data mining.In SA Machine learning plays a vital role to identify the Sentiment Polarity which is based on test and train data sets. Machine learning has the ability to determine the performance of people's views. Machine learning consists of many of the algorithms and classifiers that train the data using models to give the accuracy of the data with different percentages. But it's also disadvantages due to the low capability of the methods on new data to perform the best accuracy. The machine learning has two approaches, i.e. Supervised and

UNSUPERVISED METHODS

Deep learning is a model which consists of lots of pre-processing layers to represent the data on multiple layers. It works in speech recognition, visual object recognition, object detection, and many other areas such as drug discovery and genomics. Deep learning presents a difficult structure to present large data sets by using the back-propagation algorithm to suggest how a machine change the internal parameters that are used in the representation of each layer from the representation in the previous layer. Deep convolution nets are used in images, video, speech, and audio whereas currently they are used in data such as text and speech.

Deep learning belongs to machine learning methods based on learning the representation of data. Learning can be supervised, semi-supervised, or unsupervised. In this literature review, we discuss about deep learning problems, techniques, and algorithms, used datasets and their size, and also the best accuracy.

Emoticons are the representation of face expression such as smile and sad by using the keyboard characters. We can detect feeling by identifying emoticons. We use Python language for identification of emoticons feeling.

In this paper, build a models for classifying 'tweets' into positive, nagative and neutral sentiment. We used two classifiers of machine learning such as Decision Tree and Navie Base for polarity and deep learning techniques for the subjective and objective part. This research uncovers the emotional triggers and common distractions about the behaviour of people by examining the polarity and contextual patterns of sentiment we need stratgies for reducing procrastination .

This study aims to develop a robust sentiment analysis model using machine learning and deep learning techniques such as Decision Tree, Navie Base and LSTM and evaluate their performance on real-world datasets. The objective of this analysis is to cover the area of sentiment analysis which deals with polarity of people's tweets using machine learning techniques. One of the challenge is to recognise polarity of text objectively and defining the text of opinion subjectively.

The rest of the paper follows as: In Section 2, discuss the literature review of sentiment analysis. In Section 3, discuss methodology such as data collection, pre-processing, tools and techniques. In Section 4, present experiments and tools used for implementation. In Section 5, we discuss the results and compare the results of each classifiers. In Section 6, we discuss the future direction of research.

LITERATURE REVIEW

In many literature reviews, the main focus of Sentiment is to resolve the problem domain in a definite area where most people behave such as state their feeling on specific tasks which are related to that area of the Sentiments. In our research we study different research papers that describe sentiment analysis the way people share knowledge, thoughts, feelings and opinions with the people using internet resources such as Twitter, Instagram, Facebook. In the following paper, we present an overview about SA. Below we present the important features of Sentiment Analysis which is based on an important topic. It should be noted that this paper is meant to be an introduction to the topic rather than a comprehensive and full overview. This study is about to discuss machine learning and deep learning problems, techniques and algorithms, used datasets and their size, and also the best accuracy.

This part provides an overview of the sentiment tasks, and the next will discuss some of the equipment. Pang, B. and Lee, L. [3] determine the Sentiment Polarity; present a concept of machine learning method. When performing on text data only take the subjective part of the data. For this purpose, used standard machine classifiers such as the support vector machine and Naive Byes to find minimum cut graph. The data taken are 5000 that are movie reviews and for objective part the size is 5000 sentences. The accuracy for the naive base is 86.4%, 85.2% and for SVM is 86.15%, 85.45%. Wilson, T., et al. [4] propose a method to find out the level of phrases. The main focus is to determine neutral and polar expressions. The datasets consist of 15,991 subjective portions from 425 documents. The classifiers used is BoosTexter AdaBoost.HM machine learning. The classifiers are evaluated in 10 features; 10-feature classifier achieves a Correctness of 65.7%, which is higher than the baseline.

TABLE 1: PROBLEMS, DATASETS, CLASSIFIERS, AND RESULTS FOR MACHINE LEARNING

Reference	Problem	Dataset/size	Algorithms	Result
[3]	Sentiment polarity	5000 Movie Review	Standard machine learning Nave Byes, Support vector machine (NB and SVM)	NB:86.4% SVM:86.15%
[4]	Phrase level Sentiment analysis	15,991 subjective	BoosTexter AdaBoost.HM machine	65.7% accuracy

		expressions	learning	
[5]	SENTIWORDNET 3.0	1,105 WORDNET synsets	Semi-supervised learning and random-walk process	19.48% relative
[6]	Semantic orientation of Adjectives	15,431 adjective	Log-linear regression model	90% accuracy
[7]	Semantic Classification of Product Reviews	600 sentences	baseline unigram method	80.6% accuracy
[8]	Sentiment expression (Claims) polarity	19748 adjectives and verbs	Word sentiment and sentence sentiment classifier	67% accuracy

Different researchers use different tools and methods in their work; the more accurate performance of the classifier in this table is log-linear regression model and the problem solved by this classifier is the semantic orientation of Adjectives.

Deep learning belongs to machine learning methods based on learning the representation of data. Learning can be supervised, semi-supervised, or unsupervised. In this literature review, I discuss deep learning problems, techniques and algorithms, used datasets and their size, and also the best accuracy.

Socher et al. [9] present a novel scheme over a Sentiment Treebank. For this purpose, they introduce sentiment detection supervised training. It is made up of Sentiment labels of 215,154 phrases in the trees of 11,855 sentences. For this, they introduce the Recursive Neural Tensor Network (RNTN). It works on positive / Negative classification from 80% up to 85.4%. The result of the sentiment labels is 80.7%, with an improvement of 9.7%.

A presentation and the Stanford Sentiment Treebank dataset are available[1] sentiment. Kim, Y. [10] uses convolutional neural network (CNN) classifiers. It is composed of one layer of convolutional from an unsupervised neural language model. The data set was used by Word2Vec that consists of 100 billion words from Google News[2]. Results are compared with different

[1](http://nlp.stanford.edu/)
[2](https://code.google.com/p/word2vec/)

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 6 (2025)

models with different datasets. The relative performance was 89.6%. Glorot et al. [11] present a problem in studying the domain of sentiment classifiers. They used the SVM classifier and the method denoted by SDAsh. They collect 340,000 reviews consisting of 22 types of products. The reduced data set contains 4 different domains: There are 1000 positive and 1000 negative values for each range. The best results were achieved by the SVM.

Tang, Y. [12] presents a problem to replace softmax layer using an LSVM. They used Softmax, Support Vector Machines, and Multiclass SVMs models. The data that they collect are 28,709 48x48 images of 7 different types of face expressions. The results of Softmax and DLSVM L2 are 69.4%.Mroueh, Y., Marcheret, E. and Goel, V. [13] show methods in deep multimodal learning for combining speech and visual modalities for Audio-Visual Automatic Speech Recognition (AV-ASR). For dataset they select 262 speakers which are in The "*IBM AV-ASR Large Vocabulary Studio Dataset*" contains 40 hours of audio-visual recordings. They use a back-propagation algorithm. The researcher show each audio+video frame is labelled with one of 1328 targets classes measure error rate at the level, this is refer to phone error rate (PER).there achieved result are DNNa (Audio Alone) 41:25% PER, DNNv (Visual Alone) 69:36% PER, Bimodal (DNN Fusion) has 35.77% PER, Bimodal (SoftMax Fusion) 35.83%.

Ngiam, J. et al [14] present a method of deep learning to learn the features. They use data set as AVLetters and CUAVE for supervised classification which are CUAVE takes 36 speakers, AVLetters 10 speaker, AVLetters2 5 speaker, Stanford dataset. 23 volunteers spoke and TIMIT also used for unsupervised audio feature pre-training.

Socher, R. et al. [15] used data of RGB-D dataset, which composed on 207,920 RGB-D IMAGES. The problem is to present a model consisting of convolutional and recursive neural networks for learning features and to classify RGB-D images. The classifier used is CNN, RNN, Linear SVM, Kernel SVM, Random Forest, CKM, and SP+HMP. The precision of these algorithms is 88%.

TABLE 2: PROBLEMS, DATASETS, CLASSIFIERS, AND RESULTS FOR DEEP LEARNING

References	Problem	Dataset/size	Algorithms /Classifier	Result
[9]	Semantic Compositionality	215,154phrases, Stanford Sentiment	Recursive Neural Tensor Network	80.7% accuracy

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 6 (2025)

		Treebank dataset	(RNTN)	
[10]	Sentence-level Classification	Word2Vec vectors,100 billion words from Google News	Convolutional neural network (CNN)	89.6% accuracy
[12]	To replace softmax layer with a linear support vector machine	28,709 48x48 images of 7 different types of face expressions.	Softmax, Support Vector Machines, Multiclass SVMs	Softmax: 69.4%, DLSVM L2: 71.2%.
[13]	Deep multimodal learning for the job of phonetic classification from audio and visual modalities	262 speakers	back-propagation	DNNa (Audio Alone) 41:25% PER, DNNv (Visual Alone) 69:36% PER, Bimodal (DNN Fusion) 35.77% PER, Bimodal (SoftMax Fusion) 35.83%
[14]	Deep networks to learn features over multiple modalities	CUAVE: 36 speakers, AVLetters 10 speaker, AVLetters 2 5 speaker, Stanford dataset. 23 volunteers spoke and TIMIT	canonical correlation analysis (CCA)	For audio, it decreased by 7.7% on CUAVE and 14.3% on AVLetters, audio and video10; performance decreased by 2.1% on CUAVE and 5.0% on AVLetters.

[15]	Combination of CNN and RNN for learning quality and to classify RGB-D images	207,920 images	RGB-D	CNN, RNN, Linear SVM, Kernel SVM, Random Forest, CKM, SP+HMP	88% accuracy

METHODOLOGY

PROPOSED SYSTEM

In our research, we solve a problem about sentiment polarity of tweets that is related to our topic i.e. Sentiment Analysis of tweets about Procrastination using machine learning and deep learning techniques. We used the Rapid miner software with three models and the indico emotion detection API.

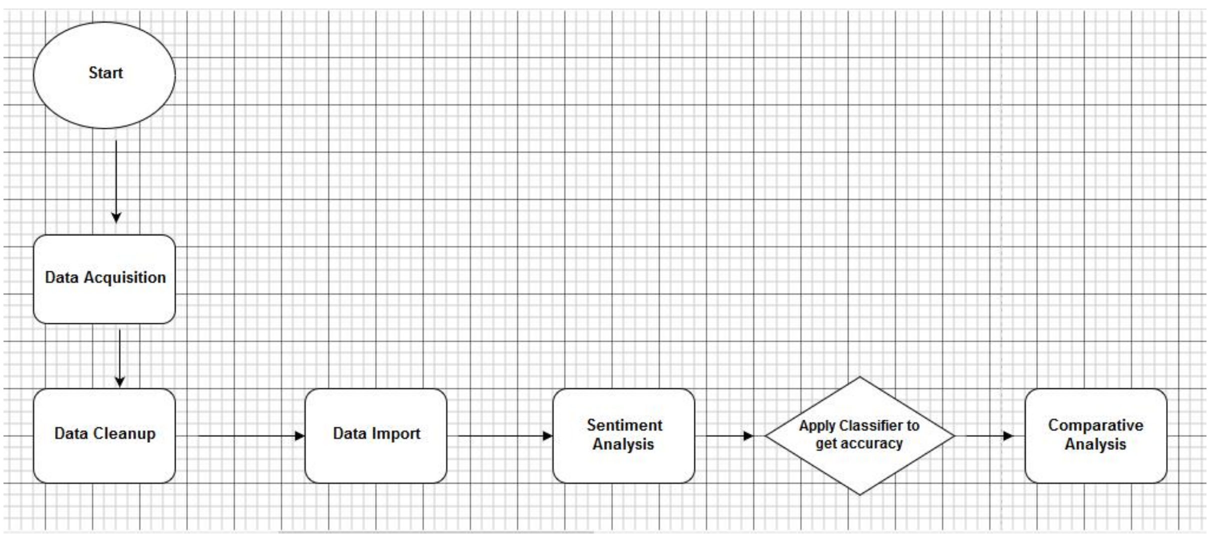

FIG. 2 A PROPOSED SYSTEM FOR SENTIMENT ANALYSIS USING MACHINE LEARNING AND DEEP LEARNING TECHNIQUES

DATA ACQUISITION

We have choose the data set from the AIDR (Artificial Intelligence for Disaster Response) platform through Indico API (Application Programme Interface), using twitter ID. AIDR is a free open source platform through which tweets can be automatically collects. We collect the data set about procrastination without retweet on twitter using AIDR. The name of the dataset is *data mining*; by select the dataset procrastination means SA is especially helpful in social media monitoring as it allows us to expand an indication of public attitude on the following assured topics. Also, applications of Sentiment Analysis are release and overprotective.

DATA CLEANUP

Here, we cleanup and remove unnecessary data for training datasets:

Removing URLs, eliminating stop words: "The stop words like a, an, this that are not useful in doing the sentiment analysis are removed in this phase.Unstructured to structure: Twitter comments are often unstructured i.e. 'aswm' is written 'awesome', 'happyyyyyy' to actually 'happy' Format conversion is done by unorganized dynamic data records of unstructured to structure and vowels adding.Emoticons: 'These are the most important ways to get ideas. The symbolic representation to change words at this stage.

DATA IMPORT

After cleaning the data are imported into Rapid Miner for processing. First, we drag retrieve operator to process panel window and load the data to process on it. When we connect the retrieve operator to the result (res) node, it shows the data in table form. The size of the datasets is 9082 with 14 attributes. Here we have some missing values in the table.Replace these missing values with null values. Now we have filter 394 out of 9082 examples.

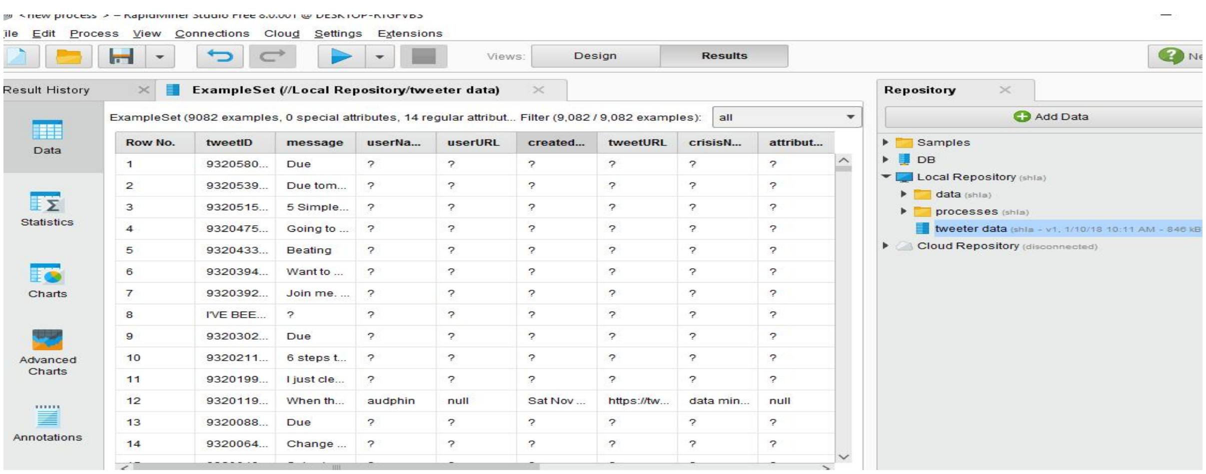

FIG.3: VIEW OF THE RESULTS RETRIEVE DATA FROM LOCATION

SENTIMENT DETECTION

To determine the text mining and sentiment analysis, we use the following operator and create the process diagram to find the polarity and subjectivity of text on people's behaviour using procrastination dataset.

ExampleSet (499 examples, 4 special attributes, 15 regular attributes)

Row No.	polarity_con...	subjectivity_...	polarity	subjectivity	tweetID	message
1	0.648	1.000	neutral	objective	9320580561...	Due #tomorr..
2	0.717	0.997	neutral	objective	9320539481...	Due tomorro..
3	0.820	1.000	neutral	objective	9320515638...	5 Simple Way
4	0.527	1.000	neutral	subjective	9320475623...	Going to go w
5	0.668	0.996	neutral	subjective	9320433074...	Beating #Pro.
6	0.952	1	positive	subjective	9320405183...	Wow, the #pr.
7	0.912	1.000	neutral	subjective	9320394906...	Want to Beat .
8	0.773	1	neutral	subjective	9320392493...	Join me. @m.
9	0.849	1	neutral	subjective	9320378755...	Just updated
10	0.808	1	neutral	subjective	9320372338...	"Okay. I need
11	0.868	1	neutral	subjective	9320358434...	Tim discusse
12	0.648	1.000	neutral	objective	9320302627...	Due #tomorr..
13	0.549	1	negative	subjective	9320251764...	Watching Dr. .
14	0.747	0.929	neutral	objective	9320211072...	6 steps to ov..
15	0.796	1	negative	subjective	9320202076...	"People often.

FIG.4: SHOW THE SENTIMENT OF EACH TWEET

IMPLEMENTATION

ANALYSES SENTIMENT OPERATOR

To extract Sentiment from a bit of passage such as a tweet, a review, or an article can provide us with helpful insight about the author's emotions and viewpoint: whether the attitude is *'Positive, Negative, or Neutral'* and whether the text is Subjective (mean it reflect the author's opinion) or objective (mean it's state a fact).

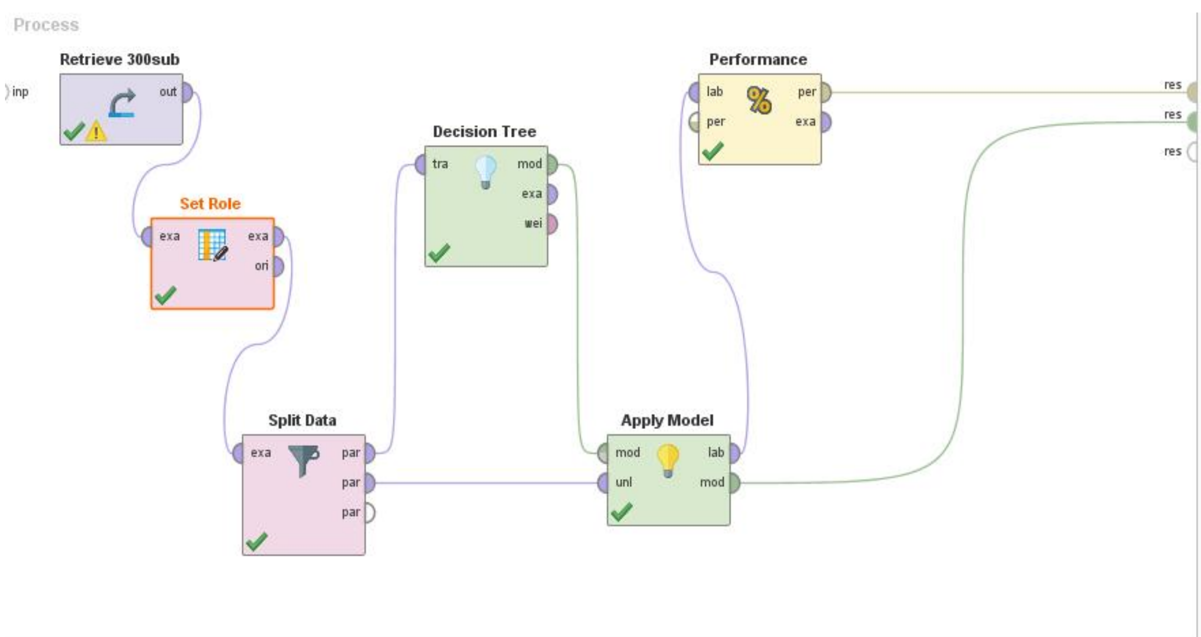

FIG.5: DATA PROCESSING TO GENERATE THE RESULT

DESCRIPTION OF THE PROCESS

1. We build a process so we first bring the datasets through retrieve operator in the process panel, and we are going to use all the variables other than the target variable as independent variables so we don't have to eliminate any variable or exclude any variable then we use the set role operator to tell the rapid miner which is our target variable.
2. Here the target variable in this example is 'polarity' and that is our label variable we indicate rapid miner that text is dependent variable, next we partition the data sets to evaluate the model to partition the data sets, we used split data operator to Split the data into two parts one use for building the model and the other to keep it for testing the model.
3. We can create the two partitions. One partition for 70% and the other partition with 30% of the data. Now we need the decision tree operator. The decision tree is used for regression, as well as classification is used. Here we are going to use this for classification problem. Because our target variables is polarity (i.e. positive, negative, and neutral).
4. Decision tree has many parameters. Rapid miner offers the operators to optimise the values or parameters in such a way that the accuracy of the model is maximize. The top partition has 70% data used to build the decision tree, the other 30% of the data come from the bottom port. We are going to use the data for testing and we have to apply the model, the operator is called the apply model operator.
5. The apply model requires two input one is the model and other is unlabelled data, so the apply model takes the unlabelled data and applies the model on it to find the missing labels for unlabelled data or to find the prediction for unlabelled data.
6. The 30% data come from the split data. This is not unlabelled because it comes from dataset where all the data are labelled, but apply model is going to pretended the data is unlabelled that the outcome is not known that is going to apply the model to create a label. The model comes from the decision tree operator and here we put the output of the decision tree model to the incoming port of the apply model.
7. In this case the labelled data is 30% of the label data now we have the last thing to do here is the performance operator which can be used to evaluate the model.
8. So we can use the classification performance model. It has one mandatory input and that input is a label data that come from the output of the apply model.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 6 (2025)

EXPERIMENT 1: (classification of sentiment polarity , subjective and objective portion)

How to classify sentiment polarity of peoples tweets through the machine learning method? How can we detect the subjective and objective portion of the data" through the deep learning method?

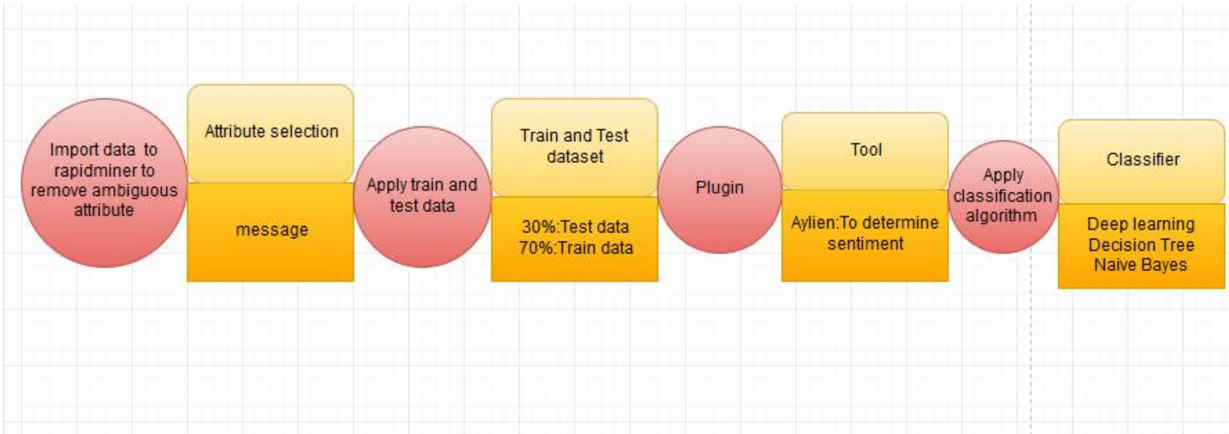

FIG. 6: CLASSIFICATION OF SENTIMENT OF MESSAGES OR SENTIMENT POLARITY

EXPERIMENT 2: (classification of sentiment through emotion)

How can we detect feelings through 'Emoticons' using deep learning techniques?

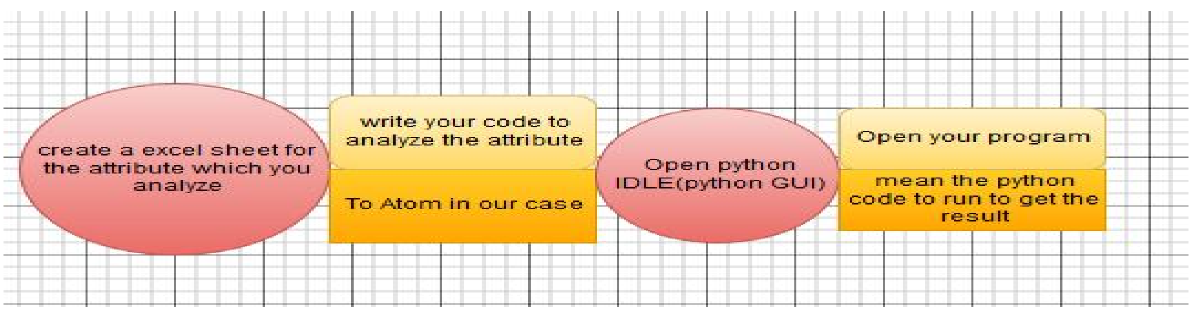

FIG.7: CLASSIFICATION OF SENTIMENT THROUGH EMOTION

We take the text mean message attribute from our datasets of size 300 and 500 tweets, and save it into an Excel sheet. Then we write the code required to analyse our data. After running the code, the function categories the emotion into five types, in our case, we have joy, fear, sadness, anger, and surprise. We can show the results of this coding after performing the codes.

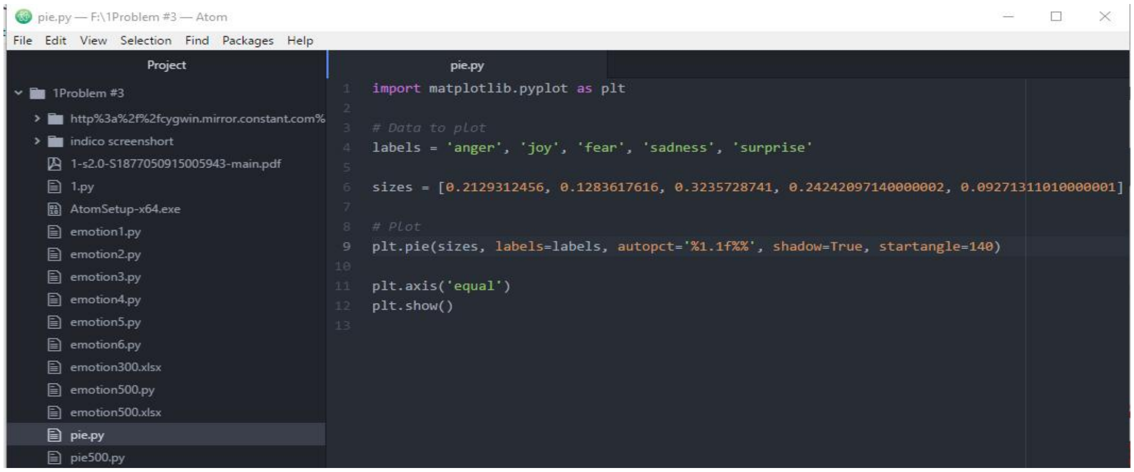

FIG. 8: CODE TO VISUALIZE THE RESULT IN TERM OF PERCENTAGE

The results from these experiments are discussed in Section 5 Table 4.

RESULTS AND DISCUSSION

To analyse the problem we perform different classifiers on the data and compare their results to find out which classifiers perform better and give the accurate result.

RESULTS COMPARSION

TABLE 3: TABLE OF COMBINED RESULTS OF ALL THE CLASSIFIERS ON 300 AND 500 SAMPLE OF DATA

Classifiers	Dataset	Results	Results
	Data samples	Problem1	Problem2
Decision tree	300	72.22%	68.97%
	500	68.67%	86.67%
Navie base	300	33.33%	34.44%
	500	26.67%	26.00%
Deep learning	300	60.00%	48.89%
	500	42.67%	87.33%

The above table shows the combined result of all the classifiers by using 300 and 500 sample of data. The deep leaning classifier gives the best result on problem 2 on 500 samples of data i.e. the accuracy is 87.33%.

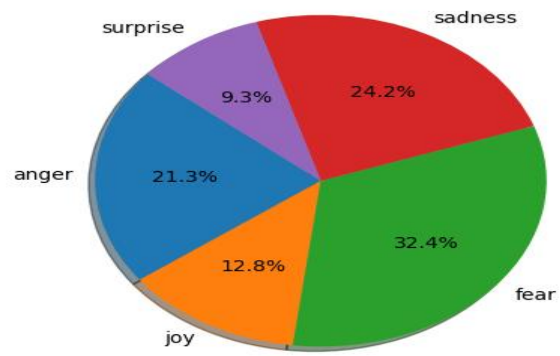

FIG. 9: PIE CHART OF EMOTION

TABLE 4: TABLE OF EMOTIONS FROM 300 TWEETS.

Emotion	Results
Anger	21.3%
Fear	32.4%
Joy	12.8%
Sadness	24.2%
Surprise	9.3%

CONCLUSION

The aim of this paper is to enhance sentiment analysis on publicly available tweets, which should not represent the full spectrum of procrastination. The contribution includes sentiment polarity and to classify the subjective and objective portions of the tweets by using machine learning algorithms such as Decision Tree and Navie Base. This study also develops a deep learning model like LSTM for emoji classification to understand the feeling of people in the form of text and emoticons.

LIMITATION AND FUTURE WORK

As our datasets consists of 11983 tweets, the software we used is Rapid Miner with Aylien data analysis API and Python language with indico emotion detection API. We analyse only a sample of 300 and 500 tweets because the software does not support or cannot give correct result for more than 500 tweets. So, in future work we will use such machine learning and deep learning software that supports more data to analyse and may classify complex emotions and user metadata for a richer analysis.

REFERENCES

[1] "Mejova, Y., 2009. Sentiment analysis: An overview. Comprehensive exam paper, available on http://www. cs. uiowa. edu/~ ymejova/publications/CompsYelenaMejova. pdf [2010-02-03].".

[2] LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. Nature, 521(7553), pp.436-444.

[3] Pang, B. and Lee, L., 2004, July. A sentimental education: Sentiment Analysis using subjectivity summarization based on minimum cuts. In *Proceedings of the 42nd annual meeting on Association for Computational Linguistics* (p. 271).

[4] Wilson, T., Wiebe, J. and Hoffmann, P., 2005, October. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the conference on human language technology and empirical methods in natural language processing* (pp. 347-354). Association for Computational Linguistics.

[5] Baccianella, S., Esuli, A. and Sebastiani, F., 2010, May. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. In *LREC* (Vol. 10, pp. 2200-2204).

[6] Hatzivassiloglou, V. and McKeown, K.R., 1997, July. Predicting the semantic orientation of adjectives. In *Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics* (pp. 174-181). Association for Computational Linguistics.

[7] Dave, K., Lawrence, S. and Pennock, D.M., 2003, May.Mining the peanut gallery: Opinion extraction and semantic classification of product reviews.In*Proceedings of the 12th international conference on World Wide Web* (pp. 519-528).ACM.

[8] Kim, S.M. and Hovy, E., 2004, August. Determining the sentiment of opinions. In *Proceedings of the 20th international conference on Computational Linguistics* (p. 1367).Association for Computational Linguistics.

[9] Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C.D., Ng, A. and Potts, C., 2013. Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 conference on empirical methods in natural language processing* (pp. 1631-1642).

[10] Kim, Y., 2014. Convolutional neural networks for sentence classification.*arXiv preprint arXiv:1408.5882.*

[11] Glorot, X., Bordes, A. and Bengio, Y., 2011. Domain adaptation for large-scale sentiment classification: A deep learning approach. In *Proceedings of the 28th international conference on machine learning (ICML-11)* (pp. 513-520).

[12] Tang, Y., 2013. Deep learning using linear support vector machines. *arXiv preprint arXiv:1306.0239*.

[13] Mroueh, Y., Marcheret, E. and Goel, V., 2015, April. Deep multimodal learning for audio-visual speech recognition. In *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on* (pp. 2130-2134). IEEE.

[14] Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H. and Ng, A.Y., 2011. Multimodal deep learning. In *Proceedings of the 28th international conference on machine learning (ICML-11)* (pp. 689-696).

[15] Socher, R., Huval, B., Bath, B., Manning, C.D. and Ng, A.Y., 2012. Convolutional-recursive deep learning for 3d object classification. In *Advances in Neural Information Processing Systems* (pp. 656-664).