Section A-Research paper

Sentiment Analysis of Employees Using Machine Learning Approach

Pratiksha L. Jagtap[1], Dr. Mukta Dhopeshwarkar[2], Nilofer Farhat[3], Pratik S. Jaiswal[4], Sarika K. Swami[5]

[1]M.Phil. Student, [2]Assistant Professor, [3]PhD. Student, [4]M.phil., Department of Computer Science & Information Technology, Dr.Babasaheb AmbedkarMarathwada University, Aurangabad.(Maharashtra), INDIA

Abstract: Twitter is a prominent social networking service where users freely communicate their thoughts, opinions, and emotions. These tweets are recorded and analysed in order to extract the emotions of people connected to an employee who works in several locations. A collection of commonly occurring words is created by recoding and pre-processing around 184 tweets. SPSS is used to investigate emotions and polarity. Navie-based classifiers were used to determine the correctness of the conclusion reached.

Keywords: Twitter, Sentiments, Emotions, Polarity, SPSS, ML.

Introduction: The terms "sentiment" and "emotion" do not mean the same thing. According to the definition, a "sentiment" is an opinion or point of view. The term "emotion" refers to "a strong feeling resulting from one's mood."

Sentiment analysis seeks to capture the overall feeling or impression that people have after viewing a piece of content. It does not concentrate on distinct articulate feelings.

It instead employs a simplified binary system of "positive" and "negative" replies. We are solely interested in whether the reader had a pleasant or negative experience with the information.

Emotions are important in human life because they serve as motivators. These feelings are complicated psychological states. Emotions, whether pleasant or bad, have a significant impact on us. According to Paul Eckman, there are six universal basic emotions. They are joy, sorrow, wrath, fear, surprise, and disgust. However, psychologists later added other feelings to this list, such as pride, enthusiasm, embarrassment, contempt, humiliation, and so on.

Machine Learning (ML):

Machine learning (ML) is a subfield of artificial intelligence (AI) concerned with the development of computational algorithms that enhance performance based on experience; in other words, an ML system learns from data. ML techniques contain three types of knowledge acquisition in their traditional definition: supervised learning, unsupervised learning, and reinforcement learning.

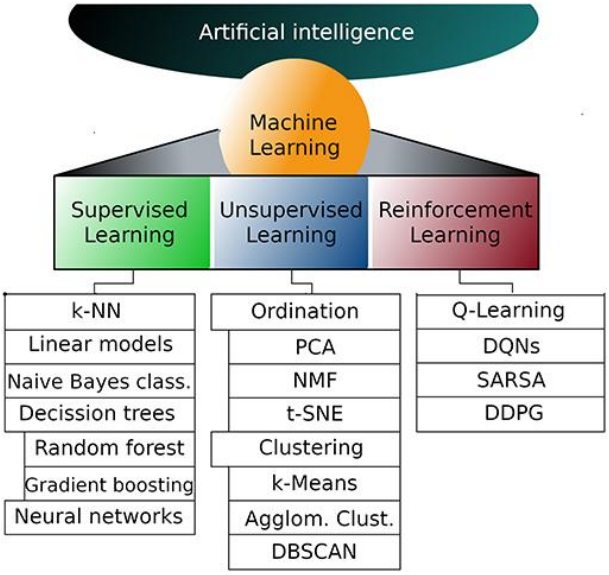


Figure 1: Overview of machine learning

In this research Naive Bayes classification algorithm is used.

Navie Bayes is a classification strategy based on Bayes' Theorem and the assumption of predictor independence. A Naive Bayes classifier, in simple terms, posits that the existence of one feature in a class is unrelated to the presence of any other feature.

$$\mathbf{P(A|B) = \frac{P((B|A).P(A)}{P(B)}}$$

Where,

P (A|B) is Posterior probability: Probability of hypothesis A on the observed event B.

P (B|A) is Likelihood probability: Probability of the evidence given that the probability of a hypothesis is true

Bayes Net: A Bayes Network classifier's base class. Data structures (network structure, conditional probability distributions, etc.) and facilities shared by Bayes Network learning algorithms such as K2 and B are provided.

b) Naive Bayes (classifier with estimator classes): Numeric estimator precision values are chosen based on training data analysis. As a result, the classifier isn't an Updateable Classifier (which in typical usage are initialized with zero training instances) Use the Nave Bayes Updateable classifier if you require the Updateable Classifier functionality. When the build Classifier is called with zero training instances, the Nave Bayes Updateable classifier will use a precision of 0.1 for numeric attributes.

c) Nave Bayes Multinomial Text: Only operates on String attributes. Other types of input attributes are permitted but are ignored during training and classification.

d) Nave Bayes Updatable: When the build Classifier is called with zero training instances, this classifier will use a precision of 0.1 for numeric attributes.

Summary of related work:

Following Table 1 shows the review of literature for this research.

Table 1: Literature Review

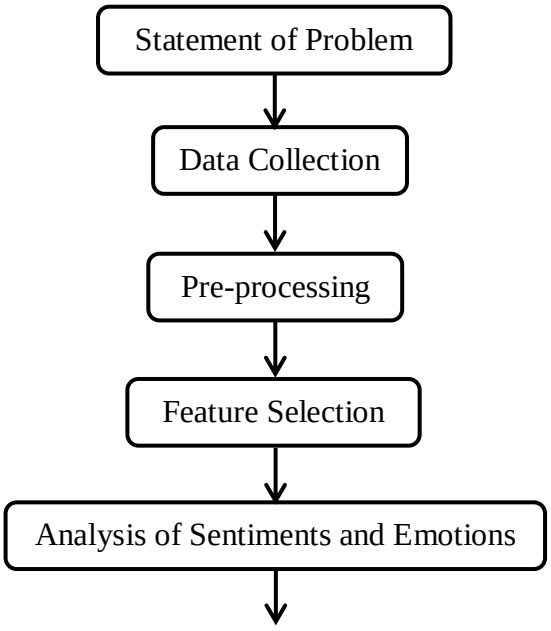
SNo	Author(s)	Inventions/Findings/Results
1	Boiy, E., & Moens, M.-F.	Classifying clients' emotions by means of approaches such as information processing, natural language therapy and machine learning systems, in English, French and Dutch helped them produce successful results.
2	Boiy, E., & Moens, M.	Integrated approach incorporating knowledge, information retrieval,recovery approaches, natural language processing and machine learning have shown strong results
3	Go, A., Bhayani, R., & Huang, L.	Describes the pre-processing steps to achieve high precision or accuracy.
4	Yang, C., Tang, X., Wong, Y. C., & Wei, C.P.	This sentimental analysis compares and describes the favourite goods that make customers happy and looks at unsupervised approaches to learning for sensational study.
5	Li, N., & Wu, D. D.	Empirical research relates the feeling of post text to the distribution of hotspots. SVM is used for hotspot prediction with exact data tests.

6	Kumar, A., & Sebastian, T. M.	Examines the significant study of emotions and discusses their fundamental terms, tasks and granularities, functional and future applications and challenges
7	Duric, A., & Song, F.	Learn the features of feedback automatically and identify them into positive, negative and neutral comments.
8	Vohra, S., & Teraiya	Exploring and explaining its policies in this field by contrasting the concept of emotional research in natural language processing.
9	Medhat, W., Hassan, A., & Korashy, H.	Gives the overview and review of the latest SA algorithms and software updates.
10	Axhiu, M., Veljanoska, F., Ciglovska, B., & Husejni, M.	Stimulus studies are joined together for tracking views and collecting qualitative information, for calculating and communicating in a palatable way.

Dataset: Twitter tweets were used to construct the dataset. A tweeter application was created to obtain Consumer Key, Consumer_Secret, Access_Token, and Access_Token_Secret in order to retrieve tweets. These keys are used to communicate between Python and the Tweeter application. Once the connection was established, a dataset of 184 tweets was generated by supplying a search phrase. This data collection has been pre-processed to remove duplicate tweets.

Methodology:

Following Figure 2 shows the research flow diagram of this research.



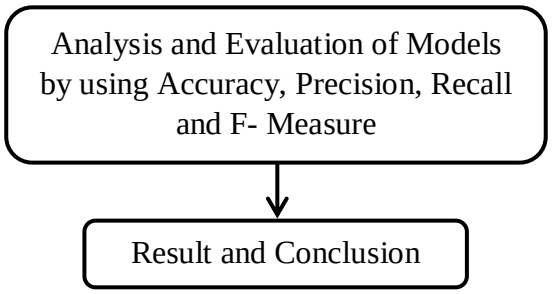


Figure 2: Research Workflow

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
Emotions * Sentiments	100	100.0%	0	0.0%	100	100.0%

Following Table 2 shows case Processing summary which shows that there is no any missing values.

Table 2: Case Processing Summary

Following Table 3 shows Sentiments frequency table with histogram.

Table 3: Case Processing Summary Emotions

Emotions	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Anxiety	4	100.0%	0	0.0%	4	100.0%
Satisfaction	72	100.0%	0	0.0%	72	100.0%
Depression	64	100.0%	0	0.0%	64	100.0%
Happy	21	100.0%	0	0.0%	21	100.0%
Confused	23	100.0%	0	0.0%	23	100.0%

Following Table. 4 shows the frequency table for sentiments with histogram figure 3

Table 4: Frequency table for sentiments

Valid Sentiments	Frequency	Percent	Valid Percent	Cumulative Percent
Negative	92	50.0	50.0	50.0
Positive	92	50.0	50.0	100.0
Total	**184**	**100.0**	**100.0**	

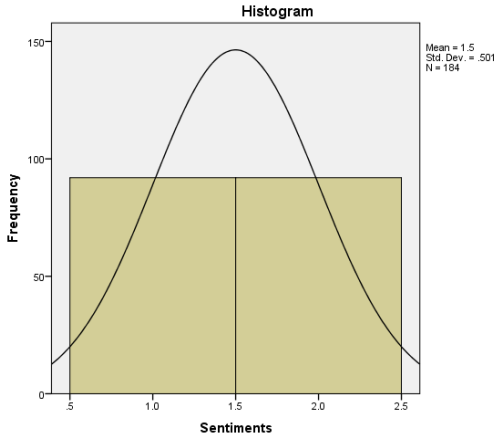


Figure 3: Histogram of Sentiments

Following Table 5 shows the frequency table of Emotions with histogram Figure 4.

Table 5: Frequency table for Emotions

Valid Emotions	Frequency	Percent	Valid Percent	Cumulative Percent
Anxiety	4	2.2	2.2	2.2
Satisfaction	72	39.1	39.1	41.3
Depression	64	34.8	34.8	76.1
Happy	21	11.4	11.4	87.5
Confused	23	12.5	12.5	100.0
Total	**184**	**100.0**	**100.0**	

Following that the data distributed.

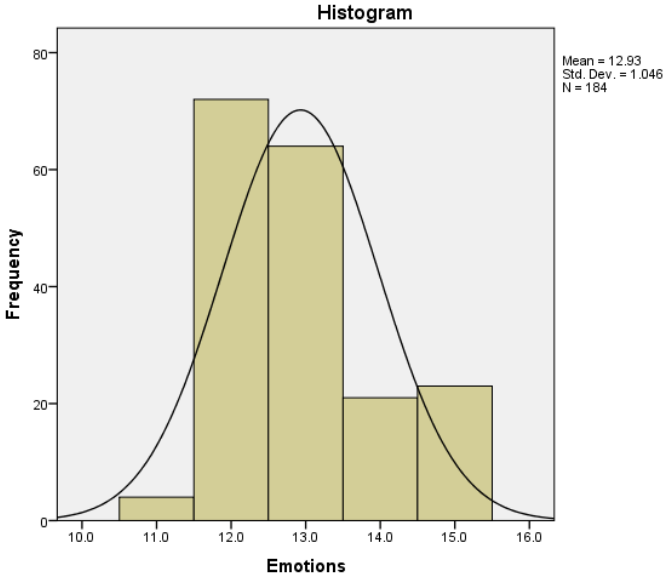


Table 6 shows are normally

Table 6: Tests of Normality

Figure 4: Histogram for Emotions

		Kolmogorov-Smirnov[b]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Sentiments	Emotions	.533	72	.000	.096	72	.000

Comparative Analysis:

Following Table 7 shows the comparative analysis of algorithms with attributes, test mode and time to taken build the model.

Attributes: 2 (Sentiments & Emotions)

Test mode: 10-fold cross-validation

Time taken to build model: 0.01 to 0 seconds

Table 7: Comparative Analysis of algorithms

Sr. No.	Algorithms	Correctly Classified	Incorrectly Classified
1	Naïve Bayes	73.36 %	26.63 %
2	Naïve Bayes Multinomial Text	39.13 %	60.86 %
3	Naïve Bayes Updatable	73.36 %	26.63 %
4	Bayes Net	73.36 %	26.63 %

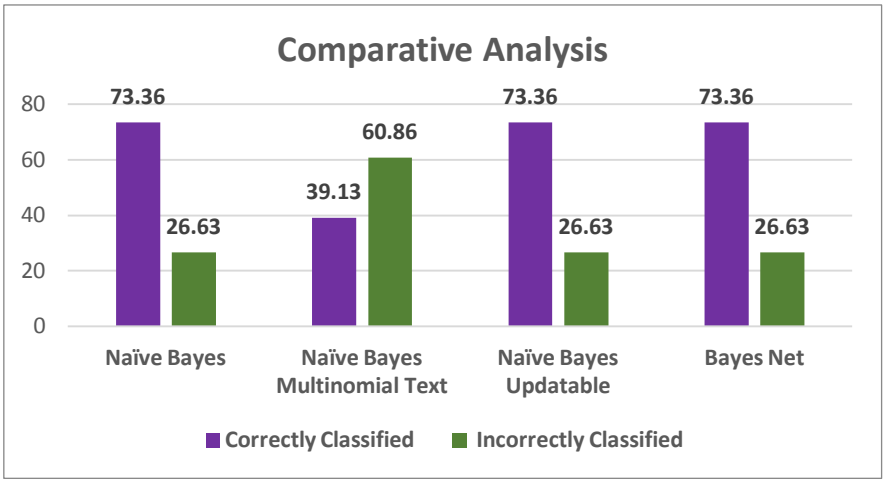


Figure 5: Graphical Representation of algorithms

Figure 5 shows that Graphical representation of comparative analysis of algorithm. Here the accuracy level of Naïve Bayes, Naïve Bayes Updatable and Bayes Net is 73.36% which is reliable for this dataset.

Result and conclusion:

Emotions are cantered on the person experiencing them, coming from a subjective experience and eliciting both a physical and behavioural reaction. Sentiment, on the other hand, can be defined as the result of interactions with other members of society. This study and literature review provides useful insights on the benefits and worth of evaluating and categorizing feelings. The main challenge is that all sentiment analysis algorithms must be precisely classified. This paper describes the machine learning approaches used in emotion analysis of employee at their work place. Firstly collected twits are analysed in the SPSS software and then give the machine learning approach to it is observed that Naive bayes algorithm supports this dataset and gives higher accuracy i.e. 73.36%. Emotions works like fundamental element of human behaviour so it played a very vital role at workplace where employee work in office.

References:

1. *https://www.researchgate.net/publication/343736541_A_Literature_Review_on_Application_of_Sentiment_Analysis_Using_Machine_Learning_Techniques*
2. *https://www.researchgate.net/publication/318368881_Sentiment_Analysis_Machine_Learning_Approach*

3. *https://www.usesignhouse.com/blog/sentiment-emotional#:~:text=Although%20synonyms%2C%20%E2%80%9CSentiment%E2%80%9D%20and,feeling%20deriving%20from%20one's%20mood%E2%80%9D.*
4. *https://www.differencebetween.com/difference-between-emotion-and-vs-sentiment/*
5. *Sarika K. Swami , Mukta G. Dhopeshwarkar , Pratik S. Jaiswal" Review on Emotional Intelligence" International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429 Volume 9 Issue III Mar 2021- Available at www.ijraset.com*
6. *Sarika K.Swami, MuktaG. Dhopeshwarkar,"Role of Emotional Intelligence Dimensions in Stress Detection", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-12, October 2020.*
7. *http://shodhganga.inflibnet.ac.in/bitstream/10603/46472/13/13_chapter3.pdf Joiceswarnalatha.R,"A study on the Emotional Intelligence Levels Of the urban Students and rural students-with special reference to SVIM", International Journal of Scientific and Research Publications, Volume 5, Issue 7, July 2015 ISSN 2250-3153.*
8. *Manasa Godati, Bhagyalakhmi. M and Hemalatha. S," EMOTIONAL INTELLIGENCE AND ACADEMIC STRESS AMONG ADOLESCENT BOYS ANDGIRLS", Eastern Academic Journal, ISSN: 2367–7384, Issue 3, pp.46-51, October, 2015.*
9. *Rahel Tajeddini," Emotional Intelligence and Self Esteem among Indian and Foreign Students – (A Comparative Study)" International Journal of Humanities and Social Science Invention ,ISSN (Online): 2319 – 7722, ISSN (Print): 2319 – 7714 www.ijhssi.org Volume 3 Issue 6 ǁ June. 2014ǁ PP.16-25.*
10. *Rohana Ngah, Wahyukaton, Zarina Salleh, Rahimah Sarmidy," Comparative study of Emotional Intelligence and Entrepreneurial Orientation between Malaysian and Indonesian University Students", Science Direct Procedia Economics and Finance 37 (2016) 100 – 107. [32] Saeid Farahbakhsh, "The role of emotional intelligence in increasing quality of work life in school principals", SciVerse Sciencedirect, Procedia - Social and Behavioral Sciences 46 (2012) 31 – 35.*
11. *Sanjay Dixit, Bhupendra Kumar Rohit, Veena Yesikar, Geeta Shivram," A Comparative Study To Find Out Emotional Quotient In Students Of MGM Medical College, Indore (MP), India", International Journal of Medical Science and Public Health | 2014 | Vol 3 | Issue 5, p.585-588.*

12. *Uzma Hanif Gondal and Tajammal Husain, " A Comparative Study of Intelligence Quotient and Emotional Intelligence: Effect on Employees' Performance" Asian Journal of Business Management 5(1): 153-162, 2013 ISSN: 2041-8744; E-ISSN: 2041-8752 © Maxwell Scientific Organization, 2013.*

ACKNOWLEDGEMENT :-

I am thankful to CS and IT Department of Dr. Babasaheb Ambedkar Marathwada University, Aurangabad (MS) for providing me a moral support.