Sentiment Analysis of Tweets using SVM and Maximum Entropy

[1]Sachin Chawla, [2] Kavita Tewatia

[1]Research Scholar, [2]Assistant Professor

[1] Department of Computer Science and Engineering, Delhi Institute of Technology Management and Research, Faridabad, Haryana, India,

[2] Department of Computer Science and Engineering, Delhi Institute of Technology Management and Research, Faridabad, Haryana, India.

Abstract : Sentiment analysis is a computational learning of masses opinion. Sentiment analysis will classify the text in a sentence or document to find out the opinions expressed in the sentence or document, which can be positive or negative. This scheme discusses the sentiment of someone on Twitter towards public figures. The data used is in the form of tweet data with the keywords "#NEP" or "National Education Policy". Tweets obtained are then processed by doing text preprocessing and then classified using the support vector machine algorithm with maximum entropy. The performance of this algorithm uses and obtained precision of 73.33%, accuracy of 77.47% ,recall of 87.76% and f-measure of 79.66% respectively.

***IndexTerms* - Sentiment Analysis, Opinion Mining, Data Mining, Machine Learning, Support Vector Machine, Maximum Entropy.**

I. INTRODUCTION

According to data released by the statista.com website in year 2019, the number of tweets in India is 34.4 million people. This number occupies the third position in the world, while data released by the a world of Tweets statista.com site places India as the third largest country in the world in writing tweets, which is 21.39% based on the recording of the total number of tweets worldwide. Since 2012 17 Million Twitter user profiles created before November 2019 [1]. The number of Twitter users in India is a promising market. So it is not surprising that various manufacturers ranging from small to large classes are competing to manage this huge economic potential so that their products sell well in the market at least as a reference. Sentiment analysis which is part of opinion mining [2-4]; Sentiment analysis is done to see opinions on a problem or it can also be used to identify trends in the market [5]. The magnitude of the influence and benefits of sentiment analysis causes research or application of sentiment analysis to grow rapidly, even in America approximately various categories of companies that focus on sentiment analysis services [6, 7]. Currently Twitter is a good indicator for influence in research [8-10]. These profit factors encourage the need to conduct sentiment analysis research on English language based tweets. This sentiment analysis research was conducted to find out public sentiment about something by using an approach in machine learning known as Support Vector Machine and Maximum Entropy Part of Speech Tagging which are specifically for English language text tweets with unigram features. The selection of the Support Vector Machine classification method is because it has the ability to generalize in classifying a pattern, excluding the data used in the learning phase of the method [12]. The Maximum Entropy (ME) model approach was chosen in Part of Speech because it has been shown to have a very efficient way to easily integrate a very large feature set in a model and has been used successfully in tasks such as Natural Language Processing (NLP) as part of tagging speech [13] or information extraction [14].

II. LITERATURE REVIEW

Social media is online content created using highly accessible and scalable publishing technology. The most important aspect of technology is the shift in how to recognize other people, share and read news, and seek information. There are hundreds of social media operating around the world today, one of them is Twitter. Twitter is a microblog that is quite popular in India by allowing every user to send and read messages called tweets in the form of 140 characters of text to be displayed on each user's page [33]. Twitter is one of the largest webs for people to express their feelings, share thoughts, opinions, provide comments and report events they experience in real-time [34]. The increasing popularity of Twitter and other social media are widely used by companies to improve their product services, and provide after sales services that allow their customers to review the quality of their products so that many companies rely on HootSuite tools to analyze data and provide customer service. However, this tool has shortcomings in detecting positive and negative emotional sentence forms in each user [35]. Several airlines are trying to gain a competitive advantage by continuously improving their services. There are many airline companies that focus on the quality of service and the experience of their service users as well as knowing the satisfaction of customers who use their products. By using social media twitter as an important source for tracking sentiment analysis [36]. Emotional life is indeed an area that can be handled with higher or lower skills, and requires its own expertise [37]. Emotion or feeling is a psychic atmosphere or inner mood that a person lives at one time.

Detection of emotional sentences in Twitter's online media content can be done by analyzing the sentiments of users conveyed through tweets on the Twitter social networking site. Sentiment analysis is the process of understanding, extracting, and processing textual data automatically to obtain information. so that this sentiment analysis can be used to obtain a person's emotional information contained in the messages of Twitter social network users on the topics discussed by users [38]. There has been much discussion about sentiment analysis on emotions such as adding features to the set of sentiment analysis trials using the Naive Bayes classification [39]. Evaluating the usefulness of existing lexical resources as well as features that capture information about informal and creative language used in microblogging by utilizing hashtags in twitter data to build training data [40]. Adding features in sentiment analysis such as pragmatic features related to hateful expressions and using them as additional features for hate detection with other features in detecting hate speech in short text messages on Twitter can make it easier to detect expressions of hatred [41]. Based on this research, we found a pragmatic feature that can combine counting, punctuation, capital letters, hashtags, the @ symbol and emoticons to emphasize one's emotions. Emoticons are used by users as an easy way to express emotions in a concise way. Thus, emoticons are an easy way to distinguish polarity (contradiction to words) from . For example, an exclamation point is used to express a strong emphasis which is usually a message of polarity [15-17]. The study discusses sentiment analysis on Facebook with the usual sentiment and to detect significant emotional changes in extracting information about the polarity of user sentiment (positive or negative). The results obtained through the approach using the Support Vector Machine method get the best accuracy. With sentiment analysis using Twitter on political opinion using the Support Vector Machine and Maximum Entropy methods, it is stated that the Support Vector Machine method has better results [18, 19]. In the sentiment analysis experiment, it was found that the Maximum Entropy method has a better level of effectiveness than the Support Vector Machine method and other methods [20-22]. Analysis of opinion sentiments regarding products and services in e-commerce using the Support Vector Machine and Maximum Entropy methods as classification methods shows the best results on the Maximum Entropy method [23,24].

Twitter is one of the most popular social media used by users today [2]. This has led to several studies being conducted using Twitter as the medium used. Previous research related to Twitter Sentiment Analysis was conducted by [25]. This study utilizes Twitter Sentiment Analysis for the classification of movie reviews using the Machine Learning algorithm. This study compares the use of two algorithms in Twitter Sentiment Analysis, namely the Naive Bayes Classifier and the Support Vector Machine. The dataset used is 21,000 tweets in English and the training process uses 1,800 tweets which are divided into 600 positive tweets, 600 negative tweets, and 600 neutral tweets. The dataset for testing is 150 tweets divided into 50 positive tweets, 50 negative tweets, and 50 neutral tweets. The conclusion from this research is that the use of machine learning algorithms is easier and more efficient than symbolic algorithms. The accuracy of the Support Vector Machine algorithm produces 75 percent and the Naive Bayes Classifier produces 65 percent. Other research related to Twitter Sentiment Analysis, namely research conducted by [26]. This research uses Support Vector Machine and Maximum Entropy algorithm. This study focuses on the training dataset stage that uses three models, namely Opinion Lexicon, Labeled Sentimental Tweets, and a combination of the two models. The dataset used in the training process is 30,000 positive tweets and 30,000 negative tweets. The dataset for the testing process is 1,000 tweets.

The word list used is 2,000 positive words and 4,000 negative words that will be used for the Opinion Lexicon as the basis for classification. The results obtained have an accuracy rate of 69 percent on unprocessed tweets, 73 percent on processed tweets, and 74, 2 percent on combination models. In 2015, [27] conducted a Twitter Sentiment Analysis research with a focus on the effect of the number of datasets on the training process for the Support Vector Machine algorithm and Abdelwahab's Naive Bayes Classifier [27]. The dataset used for the training process was 4,269 tweets and 782 tweets for the testing process. Training is carried out 10 times in order in the first training process using 10 percent of the training dataset, the second process using 20 percent of the training dataset, until the 10th training uses 100 percent of the training dataset. The results of this study indicate that the number of datasets will affect the level of accuracy of the Support Vector Machine algorithm and the Naive Bayes Classifier. The Support Vector Machine algorithm has a stable accuracy rate of between 73 and 76 percent in the entire training process. The Naive Bayes Classifier algorithm has an accuracy rate of 67 percent at the first time the training process, then the accuracy rate becomes stable between 73 and 75 percent at the next training process.

- ***Text Mining***

Text mining is the process of extracting information from unstructured source data. Data which has not been structured will be processed using specific techniques and methods of generating useful information for users. Text mining is a technique used for handle classification, clustering, information extraction, and information retrieval Feldman [28].

- ***Sentiment Analysis***

Sentiment analysis is the process of understanding, extract, and process textual data automatically to get sentiment information contained in an opinion sentence. Sentiment analysis is used to see opinion or opinion tendency towards a problem or object by which someone is headed to positive or negative opinions [29].

- ***Support Vector Machine (SVM)***

First time Support Vector Machine introduced by Vapnik in 1992 as a harmonious set of concepts excellent in the field of pattern recognition Feldman (2007) [28]. SVM is a machine algorithm learning that works on the principles of Structural Risk Minimization (SRM) with the aim of finding the best hyperplane that separates the two class in the input space.

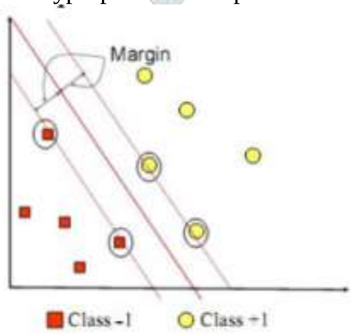

Figure 1: The Best Hyperplane Separating Both Class -1 and +1

a. *Kernel trick and non linear SVM*

To solve non-linear problems, SVM is modified by including the Kernel function. In the non-linear SVM, first the x data is mapped by the function ***Φ (x)*** to a higher dimensional vector space. In this new vector space, a hyperplane that separates the two classes can be constructed.

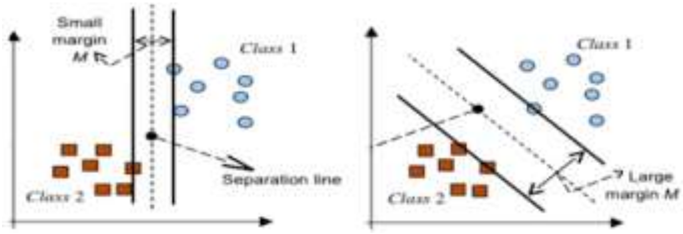

Figure 2: Mapping Dimensional Input Space Two with Mapping To High Dimensions

The learning process at SVM in find the points support vector, only depends on the dot product of the data already transformed on new space higher dimensions, namely:

$$\boldsymbol{\Phi}(\boldsymbol{x_i}).\boldsymbol{\Phi}(\boldsymbol{x_j}) \ \ (\mathbf{1})$$

Because generally this transformation $\boldsymbol{\Phi}$ is unknown, and it is very difficult to understand easily, the calculation of the dot product according to Mercer's theory can be replaced with a kernel function that defines it implicitly transformation $\boldsymbol{\Phi}$. This is known as the Kernel Trick, which is defined:

$$\mathbf{K}(\boldsymbol{x_i},\boldsymbol{x_j}) = \boldsymbol{\Phi}(\boldsymbol{x_i}).\boldsymbol{\Phi}(\boldsymbol{x_j}) \ \ (\mathbf{2})$$

Various types of kernel functions are known, as is summarized in the table.

Kernel Type	Definition
Polynomial	$\boldsymbol{K}(\boldsymbol{x_i},\boldsymbol{x_j}) = (\boldsymbol{x_j},\boldsymbol{x_j}+\mathbf{1})^{\boldsymbol{P}}$ $(\mathbf{3})$
Gaussian RBF	$\boldsymbol{K}(\boldsymbol{x_i},\boldsymbol{x_j}) = \boldsymbol{exp}(-\frac{\lVert x_i-x_j \rVert^2}{2\sigma^2}$ $(\mathbf{4})$
Sigmoid	$\boldsymbol{K}(\boldsymbol{x_i},\boldsymbol{x_j}) = \boldsymbol{tanh}(\boldsymbol{a}\boldsymbol{x_i},\boldsymbol{x_j}+\boldsymbol{\beta})$ $(\mathbf{5})$
Linear	$\boldsymbol{K}(\boldsymbol{x_i},\boldsymbol{x_j}) = \boldsymbol{x_i^t}\boldsymbol{x_j}$ $(\mathbf{6})$

Table 1: Commonly used Kernels

$$f\,(\boldsymbol{\Phi}(x))=w.\ \boldsymbol{\Phi}(x)+b\ (7)$$

$$= \sum_{\boldsymbol{i=1,x_i \in SV}}^{\boldsymbol{n}} \boldsymbol{\alpha_i y_i \Phi(x).\Phi(x)} + \boldsymbol{b}\ (\mathbf{8})$$

$$= \sum_{\boldsymbol{i=1,x_i \in SV}}^{\boldsymbol{n}} \boldsymbol{\alpha_i y_i}\ \boldsymbol{K}(\boldsymbol{x},\boldsymbol{x_i}) + \boldsymbol{b}\ (\mathbf{9})$$

b. Gaussian Kernel

The Gaussian Kernel is a promising kernel choice. This kernel non-linearly maps the sample into a higher dimensional space, so that unlike linear kernels it can handle cases where the relationship between class labels and their attributes is not linear. The second reason is that in the Gaussian Kernel, the hyperparameter complexity is less compared to other non-linear kernels such as the polynomial kernel with the equation:

$$\boldsymbol{K}(\boldsymbol{x_i},\boldsymbol{x_j}) = \boldsymbol{exp}\frac{\mathbf{-1}}{\mathbf{2\sigma^2}}\ (\boldsymbol{x_i}-\boldsymbol{x_j})^{\mathbf{2}}\ \ (\mathbf{10})$$

- ***Feature Selection***

Feature Selection is one of the relevant selection processes in terms of target learning problems. The purpose of feature selection is to remove redundant and irrelevant features [30]. Information Gain (IG) is a method for feature selection that is widely used by researchers to determine the limits of the importance of an attribute [31]. The value of IG is obtained from the entropy value before separation minus the entropy value after separation. This value is used to determine which attributes will be removed or used. Attributes that meet the weighting criteria will later be used for the classification process. In selecting features with Information Gain, it is carried out in 3 stages, namely:

1. Calculate the IG value for each attribute.
2. Determine the threshold (limit). This matter used to define which attribute its weight is smaller than threshold will thrown away.
3. Improve the dataset with subtraction attribute. Selection of Information Gain features $IG\ (t)$ formulated in equation (11).

$$\boldsymbol{G}(\boldsymbol{t}) = -\sum_{\boldsymbol{i=1}}^{|\boldsymbol{C}|} \boldsymbol{P}(\boldsymbol{C_i})\ \boldsymbol{logP}(\boldsymbol{C_i}) + \boldsymbol{P}(\boldsymbol{t})\sum_{\boldsymbol{i=1}}^{|\boldsymbol{C}|} \boldsymbol{P}(\boldsymbol{C_i}\,|\boldsymbol{t})\boldsymbol{logP}(\boldsymbol{C_i}|\boldsymbol{t}) + \boldsymbol{P}(\bar{\boldsymbol{t}})\sum_{\boldsymbol{i=1}}^{|\boldsymbol{C}|} \boldsymbol{P}(\boldsymbol{C_i}\,|\bar{\boldsymbol{t}})\boldsymbol{logP}(\boldsymbol{C_i}|\bar{\boldsymbol{t}})\ \ (\mathbf{11})$$

Where $\boldsymbol{C_i}$ is the data class, $\boldsymbol{P}(\boldsymbol{C_i})$represents the probability of the data class, $(\boldsymbol{t})$ and $\boldsymbol{P}\ (\bar{\boldsymbol{t}})$ is the opportunity term t that appears or does not appear in the document. In machine learning, information acquisition can used to help determine rating features [30].

- ***Classification Algorithm with Maximum Entropy***

The following is a text classification algorithm using the method Maximum Entropy [32]:

1. Identify specific words in the document (sentence).
2. Form a matrix containing the value of the appearance of specific words with the following index

$$\boldsymbol{f_j}(\boldsymbol{a},\boldsymbol{b}) = \begin{cases} \mathbf{1};\boldsymbol{if\ f_j\ appears\ in\ document\ b\ of\ the\ class} \\ \mathbf{0};\boldsymbol{if\ f_j\ does\ not\ appears\ in\ document\ b\ of\ the\ class} \end{cases}$$

3. Creating a Maximum Entropy model with training data that is calculating the value of α_j for each class with the GIS (Generalized Iterative Scaling) procedure $\alpha_j^{(0)} = 1$.

$$\boldsymbol{\alpha_j^{(n+1)} = \alpha_j^{(n)} \left[\frac{E_{\tilde{p}} f_j}{E^{(n)} f_j}\right]^{\frac{1}{2}}} \quad (\boldsymbol{eq.12})$$

Where

$$\boldsymbol{E_{\tilde{p}} f_j = \Sigma_{x\epsilon\varepsilon} p(\mathrm{x})\, f_j(x)}$$

$$\boldsymbol{E^{(n)} f_j = \Sigma_{x\epsilon\varepsilon} p(\mathrm{x})\, f_j(x)}$$

$$\boldsymbol{p^{(n)}(x) = \pi \prod_{j=1}^{k} (\alpha_j^{(n)})^{fj^{(x)}}}$$

$$\boldsymbol{\forall\, x\, \epsilon\, \sum_{j=1}^{k} f_1(x) = C}$$

4. Look for joint probability $p\ (a, b)$ for testing data $\boldsymbol{a = \{positive\ , negative\}}$

$$\mathbf{p} * (\mathbf{a}, \mathbf{b}) = \boldsymbol{\pi} \prod_{j=1}^{k} \boldsymbol{\alpha_j^{f_j^{(a,b)}}} \quad (\boldsymbol{eq.13})$$

5. Determination of the topic of the data testing document by looking at the value of $a *$ the greatest in a class

$$\boldsymbol{a * = \mathbf{argmax}\, p\ (a, b)} \qquad \text{(eq. 14)}$$

$a\ \epsilon\ (positive, negative)$

III. PROPOSED METHODOLOGY

The design of this system starts from performs text processing then results term weighting is used for IG feature selection after that do the calculations with the SVM with Maximum entropy method whose output is the result identification of tweets shown in Figure 3.

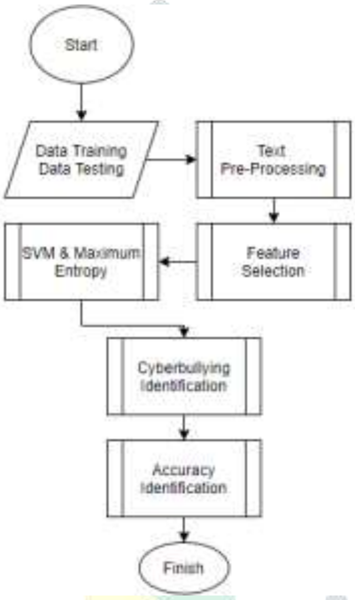

Figure 3 System Flow Diagram.

The system starts by entering data training and data testing, then doing text preprocessing process. Text preprocessing starting from the tokenization process, filtering, stemming and then term weighting. In terms of weighting count value term frequency, document frequency, inverse document frequency, weight term frequency and weight term document. Score weight term document is used for input from the SVM process. Before doing calculations using the SVM method were carried out feature selection first by counting term presence, then calculate the entropy value, after that sort the values in ascending order. Threshold is used to use the feature according to the specified threshold value. After the feature is selected, then takes a value weight term document of the feature for carried out using the classification process SVM. After the SVM classification process obtained the identification results

IV. RESULTS AND DISCUSSION

In this test using Accuracy, Precision, Recall and F-measure; This is because in this study only identify tweets that contains positive and negative words based on divergence, so retrieving the results identification is needed to find out whether the information requested by the user is in accordance with the information provided by the system. The amount of data used is as many as 300 tweets, of which 150 are positive tweets and 150 negative tweets. Data validated by scheme only for English Language. On testing the comparison scenario used are 240 training data and 60 data testing.

- ***Testing SVM Sequential Training Parameters along with Maximum Entropy***

There are 6 parameters tested in sequential training SVM with 10 different experimental values, namely lamda variable, gamma constant, epsilon, maximum iteration and complexity (C). The SVM sequential training parameter values used in the test are λ = 0.5, ϒ = 0.001, ε = 0.0001, C = 1, and iterMax = 100.

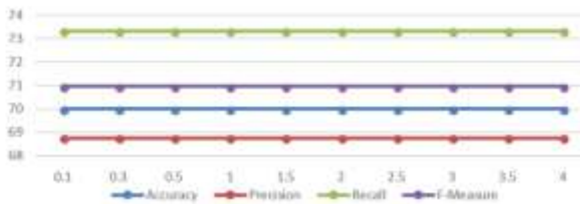

Figure 4: Graph of Lambda Test Results

In the graph in Figure 4, it is found that the results obtained are constant at all tested lambda values, namely accuracy 70%, precision 68.75%, recall 73.33% and f-measure 70.96%. This happens because the lambda value is only used to calculate the hessian matrix. And the hessian matrix is used when calculating the Ei value for sequential training in SVM. So that the results obtained do not really affect the bias value.

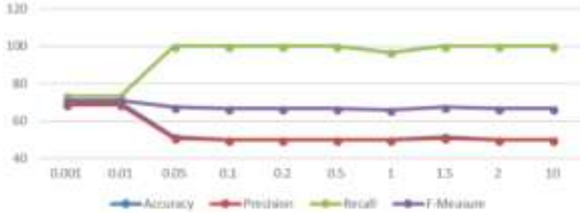

Figure 5: Graph of Gamma Constant Testing Results

In the graph in Figure 5 it is found that best results are obtained at gamma values 0.001 and 0.01, then the value decreases when the gamma value is getting higher. This is due because the gamma value is used for calculate delta alpha value, where delta value alpha is the value that determines whether convergent results or not. The gamma value taken is 0.001 with 70% accuracy, precision 68.75%, recall 73.33% and f-measure 70.96%.

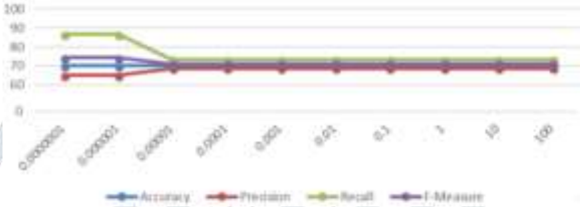

Figure 6: Graph of Epsilon Test Results

In the graph in Figure 6 it is found that best results at an epsilon value less than equal to 0.000001 and 0.0000001. This matter due to the value of the epsilon used as the maximum limit for convergent results; When the epsilon value gets higher, it yields will converge faster. Value selection epsilon 0.000001 with 68.33% accuracy, precision 64.10%, recall 83.33% and f-measure 72.46%.

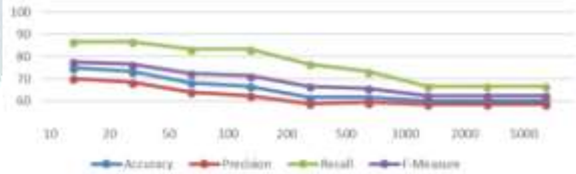

Figure 7: Graph of Maximum Iteration Test Results

In Figure 7, the best results obtained are in iterations of 20 with 76% accuracy value, 71.27% precision, 87.66% recall and 78.61% f-measure. Then in more than 1200 iterations the results are the same, this is because the calculation has entered a convergent state in iteration 1604.

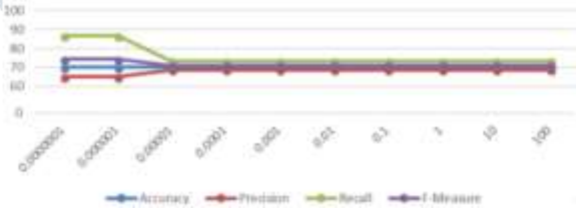

Figure 8: Graph of C Value Testing Results

The graph in Figure 8 shows that the results best on testing the C value is equal to 1 with 75% accuracy value, precision 70.27%, recall 86.66% and f-measure 77.61%. This matter because the higher the value of C then the polynomial kernel value will be higher also. When the polynomial kernel value is getting the higher the value of the Hessian matrix high anyway, so that causes on calculation of the value of the gamma constant the results are obtained smaller later can cause the iteration to become more and more converging quickly.

- **Testing Threshold Feature Selection Maximum Entropy**

After testing the sequential training for SVM with maximum entropy, the results obtained for the next test are the threshold with iterMax = 25, $\lambda = 1.5$, $\gamma = 0.001$, $\varepsilon = 0.000001$, and C = 1.

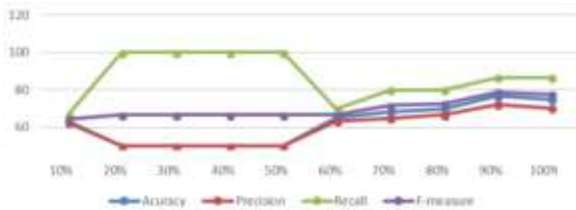

Figure 9: Graph of Maximum Entropy Threshold Test Results

In the graph Figure 9 the best results obtained at a threshold value of 90%. When all features are used the results obtained accuracy is 77.47%, precision 73.33%, recall 87.76% and f-measure 79.66%. When term selection using feature selection the information gain process produces a value which is high in terms which are unique to the term it appears once in one class. Score information gain will affect term used in the identification process, where is the term with the maximum entropy value high ones to use.

V. CONCLUSIONS AND SUGGESTIONS

Based on the results of the research, testing and analysis that has been carried out, it can be concluded that the results of the sentiments identification test using the SVM and Maximum Entropy method get the best results based on the iterMax test, parameter λ (lambda), γ (gamma constant), ε (epsilon) and C (complexity), the SVM sequential training affects changes in the weight value weight (alpha it-i) and the value of b (bias). The best results obtained from all testing of sequential training SVM parameters are iterMax = 25, $\lambda = 1.5$, $\gamma = 0.001$, $\varepsilon = 0.000001$, and C = 1.The accuracy results obtained are 77.47% accuracy, 73.33% precision, recall 87.76% and f-measure 79.66%. This is because the selection of the maximum entropy has a high values which represents a particular class, and has a low value if the feature appears in all classes. So the results of identifying sentiment analysis of tweets with maximum entropy get higher accuracy than using all existing features. While the suggestions that can be given are based on the research that has been done, namely it is hoped that it will increase the amount of data used, because it will affect the identification process to obtain optimal results. And in further research, it is necessary to consider the dimensionality of the training data. Because if you have high dimensionality training data, it is possible to cause an over-fitting state.

REFERENCES

[1] https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/

[2] Bhatia, Surbhi & Chaudhary, Poonam & Dey, Nilanjan. (2020). Opinion Mining in Information Retrieval.

[3] Singh, Sujitha & B, Dr & S, Santhoshini & Mary, Jasmine. (2020). Twitter Sentimental Analysis. Journal of Advanced Research in Dynamical and Control Systems. 12. 868-873. 10.5373/JARDCS/V12SP5/20201828.

[4] Manjappa, Rajani & Kumar, Aditya. (2020). Twitter Sentiment Analysis. SSRN Electronic Journal. 10.2139/ssrn.3633673.

[5] https://monkeylearn.com/sentiment-analysis/

[6] Ravi, Kumar. (2015). A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. Knowledge-Based Systems. 89. 14-46. 10.1016/j.knosys.2015.06.015.

[7] Walaa Medhat, Ahmed Hassan, Hoda Korashy, Sentiment analysis algorithms and applications: A survey, Ain Shams Engineering Journal, Volume 5, Issue 4, 2014, Pages 1093-1113, ISSN 2090-4479, https://doi.org/10.1016/j.asej.2014.04.011. (http://www.sciencedirect.com/science/article/pii/S2090447914000550)

[8] Vishal A. Kharde, S.S. Sonawane, Sentiment Analysis of Twitter Data: A Survey of Techniques, International Journal of Computer Applications (0975 – 8887), Volume 139 – No.11, April 2016

[9] Vishal A. Kharde, S.S. Sonawane, Sentiment Analysis of Twitter Data: A Survey of Techniques, International Journal of Computer Applications (0975 – 8887), Volume 139 – No.11, April 2016

[10] Reyes-Menendez, Ana & Saura, José & Alvarez-Alonso, Cesar. (2018). Understanding #WorldEnvironmentDay User Opinions in Twitter: A Topic-Based Sentiment Analysis Approach. International Journal of Environmental Research and Public Health. 15. 10.3390/ijerph15112537.

[11] Ficamos, Pierre & Liu, Yan & Chen, Weiyi. (2017). A Naive Bayes and Maximum Entropy approach to sentiment analysis: Capturing domain-specific data in Weibo. 336-339. 10.1109/BIGCOMP.2017.7881689.

[12] Afroze Ibrahim Baqapuri, Twitter Sentiment Analysis, Thesis, https://arxiv.org/ftp/arxiv/papers/1509/1509.04219.pdf

[13] Ratnaparkhi, Adwait. (2002). A Maximum Entropy Model for Part-Of-Speech Tagging. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP).

[14] Sridhar VK, Bangalore S, Narayanan SS. Exploiting Acoustic and Syntactic Features for Automatic Prosody Labeling in a Maximum Entropy Framework. IEEE Trans Audio Speech Lang Process. 2008;16(4):797-811. doi:10.1109/TASL.2008.917071

[15] Amritkumar Tupsoundarya, Prof. Padma S. Dandannavar, Sentiment Expression via Emoticons on Social Media: Twitter, International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887 Volume 6 Issue VI, June 2018- Available at www.ijraset.com

[16] Asli, Celikyilmaz & Hakkani-Tur, Dilek & Feng, Junlan. (2011). Probabilistic model-based sentiment analysis of twitter messages. IEEE. 79 - 84. 10.1109/SLT.2010.5700826.

[17] Oluwapelumi Aboluwarin, Panagiotis Andriotis, Atsuhiro Takasu, Theo Tryfonas. Optimizing Short Message Text Sentiment Analysis for Mobile Device Forensics. 12th IFIP International Conference on Digital Forensics (DF), Jan 2016, New Delhi, India. pp.69-87, ff10.1007/978-3-319-46279-0_4ff. ffhal-01758674f

[18] Hartono, Hartono & Sitompul, Opim & Tulus, Tulus & Nababan, Erna. (2018). Biased support vector machine and weighted-SMOTE in handling class imbalance problem. International Journal of Advances in Intelligent Informatics. 4. 21. 10.26555/ijain.v4i1.146.

[19] Alsaeedi, Abdullah & Khan, Mohammad. (2019). A Study on Sentiment Analysis Techniques of Twitter Data. International Journal of Advanced Computer Science and Applications. 10. 361-374. 10.14569/IJACSA.2019.0100248.

[20] Korovkinas, Konstantinas & Danėnas, Paulius & Garšva, Gintautas. (2019). SVM and k-Means Hybrid Method for Textual Data Sentiment Analysis. Baltic Journal of Modern Computing. 7. 10.22364/bjmc.2019.7.1.04.

[21] M., Bhumika & B., Vimalkumar. (2016). Sentiment Analysis using Support Vector Machine based on Feature Selection and Semantic Analysis. International Journal of Computer Applications. 146. 26-30. 10.5120/ijca2016910921.

[22] Gupta, B., M. Negi, K. Vishwakarma, G. Rawat and P. Badhani (2017). Study of Twitter Sentiment Analysis using Machine Learning Algorithms on Python. International Journal of Computer Applications. 165, 0975 – 8887.

[23] Yulinda Rizky Pratiwi,Edy Widodo, Comparison of Maximum Entropy and Support Vector Machine Methods for Sentiment Analysis of Pertalite Product Through Twitter Social Network, International Journal of Advances in Electronics and Computer Science, ISSN: 2393-2835 Volume-4, Issue-12, Dec.-2017

[24] Monali Bordoloi and S.K. Biswas, Sentiment Analysis of Product using Machine Learning Technique: A Comparison among NB, SVM and MaxEnt, International Journal of Pure and Applied Mathematics Volume 118 No. 19 2018, 71-83 ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version) url: http://www.ijpam.eu

[25] Akshay Amolik, et al., (2016), "Twitter Sentiment Analysis of Movie Reviews using Machine Learning Techniques", International Journal of Engineering and Technology" (IJET). vol 7. no 6, p-ISSN.2319 – 8613.

[26] Sahu, Sanket & Rout, Suraj & Mohanty, Debasmit. (2015). Twitter Sentiment Analysis -- A More Enhanced Way of Classification and Scoring. 67-72. 10.1109/iNIS.2015.40.

[27] Abdelwahab, Omar & Bahgat, Mohamed & Lowrance, Christopher & Elmaghraby, Adel. (2015). Effect of training set size on SVM and Naïve Bayes for Twitter sentiment analysis. 46-51. 10.1109/ISSPIT.2015.7394379.

[28] Feldman, Ronen & Sanger, James. (2006). The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data. 10.1017/CBO9780511546914.

[29] Bing Liu. Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers, May 2012.

[30] Bangsheng Sui, INFORMATION GAIN FEATURE SELECTION BASED ON FEATURE INTERACTIONS, A Thesis Presented to the Faculty of the Department of Computer Science University of Houston.

[31] Deng, Houtao & Runger, George. (2012). Feature Selection via Regularized Trees. International Joint Conference on Neural Networks (IJCNN). 10.1109/IJCNN.2012.6252640.

[32] https://lost-contact.mit.edu/afs/cs.pitt.edu/projects/nltk/docs/tutorial/classifying/nochunks.html

[33] Jansen, Jim & Zhang, Mimi & Sobel, Kate & Chowdury, Abdur. (2009). Twitter Power: Tweets as Electronic Word of Mouth. JASIST. 60. 2169-2188. 10.1002/asi.21149.

[34] John N. Fotisa, The Use of Social Media and Its Impactson Consumer Behaviour: The Context of Holiday Traveljohn N. Fotisa thesis, Bournemouth University, May 2015, Bournemouth University

[35] Icha Oyza, Agwu M. Edwin (2015) Effectiveness of Social Media Networks as a Strategic Tool for Organizational Marketing Management. J Internet Bank Commer S2:006.

[36] Hakh, Heba & Aljarah, Ibrahim & Al-Shboul, Bashar. (2017). Online Social Media-based Sentiment Analysis for US Airline companies.

[37] Elaine Houston, The Importance of Emotional Intelligence, https://positivepsychology.com/importance-of-emotional-intelligence/

[38] Jayasanka, Sachira & Madhushani, Thilina & Marcus, E. & Aberathne, I. & Premaratne, Saminda. (2013). Sentiment Analysis for Social Media.

[39] Lucía Gómez-Zaragozá, Javier Marín-Morales, Elena Parra, Jaime Guixeres, Mariano Alcañiz, Speech Emotion Recognition from Social Media Voice Messages Recorded in the Wild, HCI International 2020 - Posters, 10.1007/978-3-030-50726-8_43, (330-336), (2020)

[40] Shilpi Goyal, Nirupama Tiwari, EMOTION RECOGNITION: A LITERATURE SURVEY, International Journal For Technological Research In EngineeringVolume 4, Issue 9, May-2017 ISSN (Online): 2347 - 4718 www.ijtre.com Copyright 2017.All rights reserved. 1502

[41] Watanabe, Hajime & Bouazizi, Mondher & Ohtsuki, Tomoaki. (2018). Hate Speech on Twitter A Pragmatic Approach to Collect Hateful and Offensive Expressions and Perform Hate Speech Detection. IEEE Access. PP. 1-1. 10.1109/ACCESS.2018.2806394.