Journal of Theoretical and Applied Information Technology
31st March 2022. Vol.100. No 6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

RESILIENCE HIDDEN MARKOV MODEL BASED SUPPORT VECTOR MACHINE FOR CROSS - DOMAIN SENTIMENT CLASSIFICATION IN BIG DATA

RAMAJAYAM G[1] , Dr. R. VIDYA BANU[2]

[1]Research Scholar, Department of Compture Science, LRG Government Arts College for Women, Tirupur, India

[2]Asst. Professor, Department of Computer Science , LRG Government Arts College for Women, Tirupur, India

ramajayamgurusamy@gmail.com

ABSTRACT

Sentiment Analysis (SA) utilizes text contextual mining to identify and extract core subjective information from the source material. Businesses can use SA to monitor online comments or reviews about their brand, product, or service to understand how people feel about their organization. Currently, available classifiers are developed to classify sentiments only in a specific domain, and if it is applied in different domains, it will never give its better performance. General machine learning-based algorithms cannot give the best performance when applied in big data. This paper proposes a Resilience Hidden Markov Model based Support Vector Machine (RHMM-SVM) to classify sentiments in different big product review datasets. RHMM - SVM makes use of a forward - backward strategy to attain better classification accuracy in different big product review datasets. RHMM - SVM is compared and analyzed against existing classifiers with benchmark metrics, namely Precision, Matthew Correlation Coefficient, F1 - Score and Classification Accuracy. Results make a clear indication that RHMM - SVM has better performance than previous classifiers.

Keywords: *Classification, Product Review, HMM, Big Data*

1. INTRODUCTION

In recent years, the fast growth of social media has resulted in an enormous volume of daily responses worldwide, including reviews, opinions, and comments. Several companies and individuals use feedback from customers and users to help determine the quality and effectiveness of a product or service [1]. As a result, this data must be analyzed to identify and recognize important information and polarizing opinions. Companies and individuals must be able to determine if customer feedback is favourable or unfavourable automatically. Sentiment analysis is a method for determining and categorizing people opinions based on the degree of polarity with which they are held. Analysis of sentiment is an activity to uncover subjective content, such as the author's feelings or thoughts [2]. This means that sentiment analysis has an essential role in recognizing and categorizing the polarity of user - generated material. When a user or organization needs to know if a specific opinion is good or bad, this feature comes in handy. Users can locate information more rapidly with the aid of Sentiment Analysis. Sentiment analysis has gotten a lot of attention lately because of all of the new information available.

The sheer volume of evaluations has made it nearly impossible for consumers to keep up with and understand the public discourse. As a result, vast amounts of opinionated writing must be mined for their viewpoints and synthesized into a digestible form. We anticipate that a sentiment analysis model will be helpful to users in solving this issue [3]. As a result of this requirement, a sentiment analysis model must identify and deliver beneficial information about the representative opinion sentence. Taking the time to analyze customers thoughts before making a decision is helpful for the surveyed people and the businesses collecting the data. Sentiment analysis is used to determine the polarity of a piece of writing, such as whether a text expresses positive, negative, or neutral feelings. Text mining, web mining, and other data mining techniques rely on sentiment analysis to get the most out of their resources. With natural language processing, you may analyze and extract useful information from text by combining statistics, text analysis, machine learning and computational linguistics [4] –[7]. Sentiment analysis has been approached in a variety of ways throughout the last few decades. Support – Vector - Machines (SVM) and Nave -Bayes (NB) are two of the most popular approaches based on computational linguistics and machine learning, respectively (SVMs). However, although both had excellent outcomes, research has demonstrated that a machine learning technique outperforms a computational linguistic one [8].

The current world requires new technology/mechanisms to deal with the information overload, get exact information present in the reviews quickly and effectively, and extract the most relevant and vital information from the massive amounts of data readily available electronically. Nevertheless, using sentiment analysis as a tool to deal with the issues outlined above is possible.

1.1 Problem Statement

Sentiment analysis includes numerous difficulties and obstacles ranging from data collection to processing, which might be examined. Sentiment analysis topics reviewed for the study include: (i) Discovering overall sentiment from an unstructured dataset, (ii) Majority of reviews are unstructured, making it difficult to glean useful information from them, and (iii) When a classifier is trained on one domain, it fails to deliver adequate results when applied to another.

1.2 Motivation

Customer opinions regarding the products they buy might vary greatly while shopping for them online. Various social and networking services allow users to post and share their thoughts, ideas, and assessments. Increasingly intelligent technology will be developed as people get better at making use of limited resources. Thanks to the Internet, everyday tasks like paying bills, grocery shopping, and paying rent have been made much more accessible. The rise of social media and the availability of online shopping have piqued the general public's curiosity. To name a few, the field of sentiment analysis has made an essential contribution. However, using a categorization system created for one domain in another isn't always the case. It is impossible to generalize emotional expressions to all fields. It's mind - boggling how much labelling costs and how long it takes to finish in each sector.

1.3 Objective

The main objective of this research work is:

(i) To enhance the classification accuracy in sentiment analysis by identifying sentiment and topic concurrently from consumer product reviews supplied on the amazon website, using multiple machine learning classifiers

(ii) To analyze existing classifiers accuracy compared to that of a proposed classifier.

1.4 Organization of the Paper

The current section of the paper has given a brief introduction to sentiment analysis and discussed the problem statement, motivation and objective. Section 2 reviews the literature related to sentiment classification. Section 3 proposes the classifier to meet the challenges faced in classifying the sentiments. Section 4 discusses the performance metrics. Section 5 discusses the dataset used for evaluating the proposed classifier against the existing classifier. Section 6 discusses the results and Section 7 concludes the paper with future enhancement.

2. LITERATURE REVIEW

"Multi - Modal Aspect - based Sentiment Analysis (MMASA)" [9] is incorporated to analyze sentiments by grouping images and text. A multi - modal interaction - based model is used to study the relationship among aspects, text, and pictures through training and layer of interaction. A large - scale data set is constructed, and experiments are conducted. Results prove that better enhancement is provided for textual based analysis of sentiments. "Disambiguate Intonation for Sentiment Analysis" [10] network is developed for enabling reinforcement learning strategy. Features for encoding the phonetic information are learnt and fused with visual and textual - based elements to enhance performance. Chinese datasets are used for evaluation and representations at the Chinese level. "Attention - based Bidirectional CNN - RNN Deep Model" [11] is proposed for utilizing the bidirectional layers for extracting the future and past contexts with a flow of temporal information flow for both directions. The mechanism of attention is applied with the bidirectional layer of ABCDM. The dimensionality of features is reduced for fetching the invariant of local feature positions and the polarity of sentiments. Experiments were conducted, and results were brought by comparing with different DNNs.

"Dependency Relation Embedded Graph Convolutional Network" [12], a dependency - based augmented structure of a sentence, is explained. The recognition of sentiments along with aspect term extraction is carried out in the proposed study. A well - structured network is built, and the message - passing process is ensured for learning the model through a framework called Multi - task learning. Benchmarked datasets are used for illustrating the efficiency of the technique. "Knowledge enabled BERT" [13] is proposed in the current study to facilitate sentiment analysis. The embedding vectors are fetched, and sentiment domain knowledge is injected into a language representation model where the text in words is brought in vector space. Data training is performed by incorporating the aspect based sentimental data, and effectiveness is measured by performing the experimental analysis. "Machine Learning - based classifier" [14] is presented in the study to classify sentiments from platforms like Facebook and Twitter. Experiments were carried out for organizing the Sentiments based upon the conventional approach that is lexicon - based. Seven types of machine learning classifiers are applied and results are generated.

"Semantics Perception and Refinement Network" [15] is proposed in the present study for classifying the aspect based sentiments. Dual Gated Multichannel Convolution (DGMCC) is designed for retrieving the features of sentiments, and a Multichannel Convolution (MCC) is structured for fetching features of local semantics. For validating, a massive amount of experiments was carried out to prove its better performance. "Rule - based Methodology**"** [16] is developed in the attribute - level semantic analysis, fusing the deep learning technique in the present study. The sentiments are segregated into pre-defined categories, and expressions are analyzed in massive amounts fetched from e-commerce and social media platforms. Efficient detection of sentiments along with mapping of users to their attributes are performed. In the end, the sentiment attribute of individual expressions is examined to check the feasibility of the innovative technique. "Joint Term - Sentiment Generator (JTSG)" [17] is developed for identifying the polarity of sentiments that are expressed on terms extracted. A generative model generates all the pairs of aspect - term and training model encoders used for the encoding process. The decoder fetches the start and end positions, and the aspect - term is generated. For demonstrating the effectiveness of the model, an experimental study is performed and validated with its results.

"TF-IDF vectorization" [18] is performed in the field of data mining for fetching sentiments. A correlation test is carried out for survey scale results validation, and a comparison is performed. Different types of vectorization along with classification is done to enhance the prediction level of accuracy. "Topic identification and sentiment analysis" [19] is used to explore the massive amount of tweets, and the topics' effectiveness is ranked. The contents of the tweet are analyzed, and the assessment is performed. The trends of the sentiments are also studied, and the human behaviour of the geographical area are compared to observe the consensus, information dissemination and public reactions. "Natural Language Processing" [20] for sentimental analysis is performed to detect the sarcasm and its features for improving accuracy through positive words from a negative perspective. The lexicons and sentiments are built automatically and manually, and a balanced system with negative and positive tweets is also constructed. The classification with dialects of specificities are performed, and the accuracy is generated to improve its performance based on the Deep Learning technique. Optimization [4], [5], [7], [21]–[26],[27],[28] plays a significant role in different research work including sentiment analysis towards better results.

"BiLSTM model" [29] is built in the present study for expressing the analytical framework of sentimental analysis. The differences among sentiment orientation and words are detected with a technique called multi - polarity attention. To ensure discriminatory performance, the orthogonal restriction process for optimization and experimental results is generated to illustrate its performance. **"**Sentiment Padding" [30] is introduced in the proposed analytical study of sentiments. It is compared with the zero - padding technique to make the data input sample consistent for every review. CNN - LSTM model is integrated to analyze the model based on deep learning by fusing CNN with the LSTM/BiLSTM method. The performance of the proposed study is demonstrated with different challenging datasets. "Time - series technique" [31] called Massachusetts State of Emergency declaration, US State of Emergency declaration and Massachusetts public school

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

closure are incorporated in the study by exploring temporal structural patterns explored with the changes in polarity of sentiments. The Twitter data are validated for monitoring the sentiments and were compared with conventional techniques.

3. RESILENCE HIDDEN MARKOV MODEL BASED SUPPORT VECTOR MACHINE

3.1 Hidden Markov Models

Using two stochastic processes, a transition from one state to another and an output process, HMM is characterized as a probability - based model that uses two simultaneous stochastic processes. An initial set of probabilities and transformation probabilities are used to explain a Markov model's output. The second model, on the other hand, emits a single character from each alphabet. Since the variable states may only be inferred from the output symbols rather than directly from their values, state transitions are always hidden. An HMM may be broken down into states, probabilities, transition probabilities, and output probability.

The inputs are processed using an HMM, which is a kind of architecture. The mathematical formalism of HMM is a quintuple (E, R, π, D, V) with the following characteristics:

- $E = \{E_1, E_2, E_3, \ldots, E_T\}$ is the set of states, and T is the total number of states.
- The states of a Markov chain are represented as a triplet (E, π, D) in which they are not directly observable.
- The vocabulary is represented as a set R= $\{r_1, r_2, r_3, \ldots, r_C\}$.
- Individual states probabilities are given in the form of $\pi: E \rightarrow [0,1] = \{\pi_1, \pi_2, \pi_3, \ldots, \pi_T\}$ and it is the first stage of the probability distribution which is expressed in $Eq. (1)$

$$\sum_{e \in E} \pi(E) = \sum_{s=1}^{T} \pi_s = 1 \qquad (1)$$

- $D = (d_{sw})_{s \in E, w \in E}$ is the probability of transitioning from one transition state to another. $d_{sw} \in [0,1]$ is predicted for each transition state E_s and E_w which is represented in $Eq. (2)$.

$$\sum_{s \in E} d_{sw} = 1 \qquad (2)$$

- $V = (v_{sw})_{s \in R, w \in E}$ represents the output probability, where r_s represents the state existing in E_s.

When the system's state is unknown, HMM is a superior tool for process modelling. HMM's most important principle is to generate a random sequence of integers. As a general rule, while looking at the output events, HMM is regarded a generative model. HMM may have created the observed sequence $K = k_1, k_2, k_3, \ldots, k_F$, with $k_f \in R$.

3.2 Forward - Backward Strategy

HMM is used to calculate the post values of all variables in a hidden state $d_e(w) = X(g_e = F_w, M_1, .., M_s|\lambda)$ and it is considered as a forwarding variable that estimates F_w, or the reaching state, and it is observed at time t as an outcome of this approach known as Forward - Backward $(FD - BD)$. The backward process is referred to as a sequence of observed states $M_{f+1}, \ldots, M_e$, and the probability of seeing these states at t time is referred to as $M_1, \ldots, M_e$. $\beta_e(w) = X(M_{e+1}, \ldots M_e | g_e = F_w, \lambda)$. Smoothing is the current duty of interference. $FD - BD$ uses the dynamic programming approach to determine post values relating to distribution in two distinct phases. It starts with forwarding, then chooses features, and finally does categorization, all with distinct unique timings. Using $FD - BD$ to indicate any categorization approach that falls inside the relevant classes achieves its purpose. To get the probability state F_w at time f, the values of $d_e(w)$ and $\beta_s(w)$ The Expectation - Maximization $(Etp - Mxm)$ method, and the $FD - BD$ algorithms, are successively executed, and it must be determined by computation. The SVM method is used to classify the features as they are extracted, and the new feature vectors are created as a result of normalization. For the $FD - BD$ technique, the first step is to estimate the HMM parameters, which is done by computing $\lambda = (A, B)$.

There are several ways to estimate parameters, and one of them is to use an $Exp - Max$ method. The approximation function or value for maximum likelihood can be found iteratively using the $Exp - Max$ approach. A maximum - likelihood estimation can identify the best-fit model for the most data in a dataset. Some missing values (i.e., incompleteness concerns or latent variables not seen) in the dataset. In the $Exp - Max$ method, there are two steps: (i) the expectation step $(Ept - Step)$ and (ii) the maximization step $(Mxm - Step)$. A state's anticipated occupancy count is calculated in the $Exp - Step$, a state transition count based on prior probabilities is also calculated. For the feature extraction, probabilities are recalculated using γ and ξ, and in $Mxm - Stpp$.

3.3 SVM Classifier

An unsupervised learning algorithm, SVM is one of the best at classifying objects. Based on computational learning, SVM uses a risk - averse approach to solving problems. It is possible to use SVM's generalization to a wide range of datasets with many characteristics (i.e., features). Classification methods such as SVM are superior in several machine learning studies.

$\{(i_1, p_1), (i, p_2), \dots, (i_l, p_l)\}$ is considered a collection of samples where $p_l \in \{-1, +1\}$ and $i_l \in B^c$. Decision functions is defined as $sgn\big((N \bullet P) + v\big)$, where $(N \bullet P)$ where it represents product N and P. $Eq.(3)$ indicates decision - making function $j_{n,v}$ and it contains the protocols for performing sentiments classification.

$$p_w\big((N \bullet P) + v\big) \geq \frac{1}{2}, \frac{1}{2} \leq w \leq \frac{l}{2} \tag{3}$$

The separation that does not exist within the dataset can be achieved using the hyperplane in several ways. Slack variables $\xi_w \geq 0, w = 1, \dots, z$ must be included to conduct the probable violation of $Eq.(3)$. As a result, $Eq.(4)$ describes the problems with SVM mathematically.

$$\begin{aligned} &minimize \phi(N, \xi) = (N \bullet N) + V \sum_{w=1}^{z} \xi_w \\ &Subject\ to\ p_w\big((N \bullet P) + v\big) \geq 1 - \xi_w, 1 \leq w \leq z \end{aligned} \tag{4}$$

Constrained – quadratic - programming issues are the sort of problem indicated in $Eq.(4)$. Convex restricted quadratic programming issue is expressed mathematically as $Eq.(5)$

$$maximize \sum_{w=1}^{z} c_w - \frac{1}{2} \sum_{w=1}^{z} \sum_{a=1}^{z} u_w c_a p_w p_a (i_w \bullet i_a) \tag{5}$$

$$Subject\ to \sum_{w=1}^{z} p_w d_w = 0, 0 \leq d_w \leq V(w = 1, \dots, z)$$

where d_w represents the Lagrangian multipliers, V is a parameter used for assigning a penalty when the classification errors occur. While determining the $Eq.(5)$, it acts as an alternative to build a function for decision - making, as shown in $Eq.(6)$.

$$j(i) = ejt\left\{\sum_{w=1}^{z} d_w p_w (i_w \bullet i) + g\right\} \tag{6}$$

Where g is the bias term, d_w is the coefficient and it is the non-zero portion of the equation. The support vectors that characterize the function used in decision making are the elements that have equal pairings.

For mapping concerns, the non - linear issue may be generalized to a dimensional feature (L) by translating $(i_w \bullet i_a)$ to $\phi(i_w)\phi(i_a)$. This leads to the generalization of the non - linear issue, which is extremely feasible. $(i_w, i_a) = \phi(i_w)\phi(i_a)$ which is a definite kernel function, is used to define the mapping function in full. This equation, $Eq.(7)$, is a modified version of $Eq.(4)$, used to make the final decision.

$$j(i) = \frac{ejt\{\sum_{w=1}^{z} d_w p_w Z(i_w, i_a) + g\}}{i_a} \tag{7}$$

3.4 Hybridization

The prediction has grown in importance to manage and organize product reviews information because of advancements in online shopping. The following criteria are necessary for online shopping categorization and prediction. Assume that s_j signifies each text in the training set $E = \{(e_1, v_1), \dots, (e_c, u_c)\}$, and v_w denotes each text that is predicted to be in $V = \{v_1, \dots, v_t\}$. For the most part, the primary goal of prediction is to design a machine - learning algorithm that uses the training set E as input, which results in the creation of the machine learning classifier $g{:}E \rightarrow B$ and its ability to classify and perform prediction with greater accuracy. There are several classification methods for machine learning, but support vector machine and its related algorithms do better classification and deliver binary findings more quickly. $RHMM$ and SVM algorithms are used in this suggested study to enhance the classification of feelings found in online product reviews. It is used for feature extraction in the first stage and output creation in the second stage, utilizing discriminating information from the $RHMM$. The outputs of $RHMM$ are normalized in the second stage to yield new feature vectors. Unknown feature vectors are used as SVM input and sentiment analysis resulting in a binary value for classification. The $RHMM$ and SVM are combined in the suggested technique for sentiment data classification. Feature vectors are retrieved using $RHMM$s of the same dimension and embedded in the appropriate vector space. Size and performance complexities in SVM need a reduction in feature vector dimensions. The production of feature vectors is an essential part of the $RHMM$ process.

Consider I classes l= $\{l, \dots, l_Z\}$ with its suitable HMM set $\lambda' = \{\lambda_1, \dots, \lambda_Z\}$, where $\lambda^w = \{\lambda_1^w, \dots, \lambda_{C_w}^w\}$. The overall count of HMM will be $= \sum_{w=1}^{Z} C_w$. The probability of X$(M|\lambda)$ are computed by applying HMM's $\lambda^w = \{\lambda_1^w, \dots., \lambda_{C_w}^w\}$. Suppose $X_{max_{a_w}} = \max_{1 \le a \le C_w} X(M|\lambda_a^w)$, then the label l_w builds communication with the group having maximum probability. Therefore, it receives feature vector g= $\{R_1, \dots, R_w, \dots, R_T\}$ from $HMM\lambda_a^w$, where l_w and g are merged for a new feature vector formation, i.e., $\vec{g} = \{l_w, g\}$. Lastly attained feature vectors are performed with normalization by utilizing $\tilde{\vec{g}} = \frac{\vec{g}}{\|g\|^2}$, and it becomes mandatory for classifying with SVM.

Parameters λ are learned using HMM in the generation of set M= $\{M^1, M^2, \dots. M^Z\}$. To solve the issues present in the generation of set M= $\{M^1, M^2, \dots. M^Z\}$, Baum-Welch strategy [32], [33] is applied for estimating the parameter estimation, which is mathematically described in $Eq.$(8) and $Eq.$(9):

$$\overline{U_{wa}} = \frac{\sum_{z=1}^{Z} \sum_{E=1}^{E_{Z-1}} \varepsilon_e^z (w,a)}{\sum_{Z=1}^{Z} \sum_{E=1}^{E_{Z-1}} \gamma_e^z(w)} + M^Z \quad (8)$$

$$\overline{g_a}(w) = \frac{\sum_{Z=1}^{Z} \sum_{E=1, M_E^Z = R_t}^{E_Z} \gamma_e^z (w)}{\sum_{Z=1}^{Z} \sum_{E=1}^{E_Z} \gamma_e^z (a)} \quad (9)$$

where $\varepsilon_e^z(w,a)$ indicates the joint-event, $\gamma_e^z(w)$ indicates the state variable. With Z (i.e., observation sequence) $\varepsilon_e^z(w,a)$ and $\gamma_e^z(w)$ are merged.

Consider the output of HMM as a new feature vector pair. For each class w and a classifier, V_{wa} , the binary classification problem in SVM is specified in Class w which receives the positive labels, and Class a receives the negative labels. $Eq.$(10) is a mathematical description of the classification process, which involves the decision function.

$$j_{wa}(\tilde{\vec{g}}) = \sum_{c}^{\Lambda} p_c^{l_w l_a} d_c^{l_w a} \phi(\tilde{\vec{g}}^{l_w l_a}, \tilde{\vec{g}}) + g^{l_w l_a} \quad (10)$$
$$w \neq a, w = 1, \dots, T$$

where w-th and a-th class classes are shown as Λ. If the function of decision making forecasts and assigns it to class w, then the classifier V_{wa} provides a vote to class w, if not, then the vote is given to class a. It will be decided which class is assigned to each unknown sample once voting has finished.

Previous research works have focused only on data mining or machine learning algorithms, where RHMM-SVM focused on mathematical model and data mining algorithms for enhanced accuracy.

4. PERFORMANCE METRICS

Performance metrics are dependent on variables towards computing the results. The variables in data mining performance metrics are:

- True Positive (Tr_Ps): Positive sentiments accurately identified as positive.
- True Negative (Tr_Ng): Negative sentiments accurately identified as negative.
- False Positive (Fl_Ps): Positive sentiments inaccurately identified as negative.
- False Negative (Fl_Ng): Negative sentiments inaccurately identified as positive.

This research work considers the benchmark metrics to evaluate the proposed classifier's performance which are precision (PR), Matthew Correlation Coefficient (MCC), Classification Accuracy (CA) and F1 -Score (FS).

4.1 PR

It is the measure of relevant instances against the retrieved instances. Eq.(11) provides the mathematical expression of precision.

$$FNR = \frac{Tr\_Ps}{Tr\_Ps + 0.5(Fl\_Ps + Fl\_Ng)} \quad (11)$$

4.2 MCC

It is the measure of quality of classification. In other words, it is a correlation coefficient of target against the predictions. Eq.(12) provides the mathematical expression of MCC.

$$MCC = \frac{(Tr\_Ps \times Tr\_Ng) - (Fl\_Ps \times Fl\_Ng)}{\sqrt{(Tr\_Ps + Fl\_Ps)(Tr\_Ps + Fl\_Ng)(Tr\_Ng + Fl\_Ps)(Tr\_Ng + Fl\_Ng)}} \quad (12)$$

Journal of Theoretical and Applied Information Technology
31st March 2022. Vol.100. No 6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4.3 FS

It is the measure of harmonic mean between TPR and Precision. Eq.(13) provides the mathematical expression of F1 - Score.

$$FS = \frac{Tr\_Ps}{Tr\_Ps + 0.5(Fl\_Ps + Fl\_Ng)} \quad (13)$$

4.4 CA

It is the measure of exact classification against total number of instances. Eq.(14) provides the mathematical expression of CA.

$$CA = \frac{Tr\_Ps + Tr\_Ng}{Tr\_Ps + Tr\_Ng + Fl\_Ps + Fl\_Ng} \quad (14)$$

5. DATASETS

This research work makes use of four big product review dataset of Amazon which is a online shopping website. Among different product domains available in Amazon. This research work makes use of dataset from four domains, namely (i) Books, (ii) DVD, (iii) Electronics and (iv) Kitchen Appliances.
Table 1 provides the count of records available in big product review datasets.

Table 1. Record Count

Domain	Record Count
Books	146294
DVDs	142885
Electronics	88127
Kitchen Appliances	72839

6. RESULTS AND DISCUSSION

The results of the metric are calculated using Tr_Ps, Tr_Ng, Fl_Ps and Fl_Ng. Its better to know the results of Tr_Ps, Tr_Ng, Fl_Ps and Fl_Ng before going to the discussions of performance metrics results. Table 2 provides the results of Tr_Ps, Tr_Ng, Fl_Ps and Fl_Ng achieved by proposed classifier RHMM - SVM and existing classifiers EBC and MMASA.

Table 2. Result Values of Classifiers Against Big Product Review Dataset

Dataset	Variables	EBC	MMASA	RHMM-SVM
Books	**TP**	44131	48153	70933
	TN	43011	47374	69002
	FP	29490	25919	3009
	FN	29662	24848	3350
DVD	**TP**	40132	45704	68455
	TN	38019	44129	66481
	FP	32452	26243	4410
	FN	32282	26809	3539
Electronics	**TP**	23146	28134	42328
	TN	22366	27150	43452
	FP	21455	16472	1346
	FN	21160	16371	1001
Kitchen Appliances	**TP**	21375	24661	35789
	TN	20985	23903	35164
	FP	14139	12377	936
	FN	16340	11898	950

6.1 *PR* Analysis

Figure 1 analyzes the *PR* for the proposed classifier RHMM - SVM against the existing classifiers, namely EBC and MMASA. Figure 1 is plotted with product review datasets on the x - axis, and the percentage of obtained results for precision is plotted on the y - axis. From Figure 1, it is evident and easy to understand that the proposed classifier RHMM-SVM attains better precision than EBC and MMASA. RHMM-SVM perform classification with its different state of probability present in it, and it leads a way to achieve better results with different domain dataset, but EBC and MMASA perform classification only on a First – In – First - Out basis only. The result values of Figure 1 is provided in Table 3.

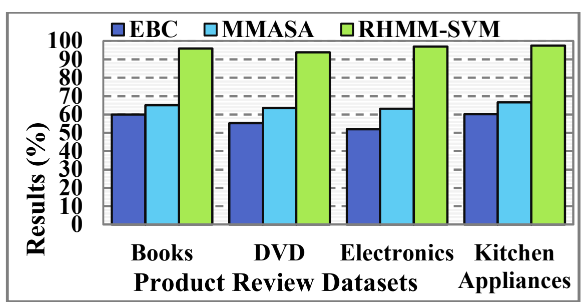

Fig 1. PR Vs Classifiers

Journal of Theoretical and Applied Information Technology
31st March 2022. Vol.100. No 6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 3. Result Values of PR Analysis

PRODUCT REVIEW DATASETS	EBC	MMASA	RHMM-SVM
Books	59.943	65.008	95.931
DVD	55.290	63.525	93.948
Electronics	51.896	63.072	96.918
Kitchen Appliances	60.188	66.583	97.451

6.2 MCC Analysis

Figure 2 analyzes the MCC for the proposed classifier RHMM - SVM against the existing classifiers, namely EBC and MMASA. Figure 2 is plotted with product review datasets on the x - axis, and the percentage of obtained results for precision is plotted on the y - axis. From Figure 2, it is conspicuous that RHMM - SVM has better MCC than EBC and MMASA. RHMM - SVM generates different random sequences before performing classification, and this leads a way to achieve better MCC, but the existing classifiers EBC and MMASA are entirely dependent on the features of the dataset, which leads to achieving low MCC. The result values of Figure 2 is provided in Table 4.

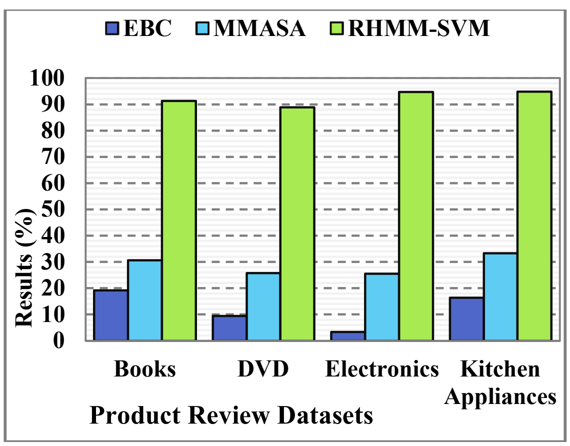

Fig 2. MCC Vs Classifiers

Table 4. Result Values of MCC Analysis

PRODUCT REVIEW DATASETS	EBC	MMASA	RHMM-SVM
Books	19.128	30.601	91.306
DVD	9.370	25.735	88.878
Electronics	3.281	25.455	94.676
Kitchen Appliances	16.415	33.345	94.821

6.3 FS Analysis

Figure 4 analyzes the FS for the proposed classifier RHMM - SVM against the existing classifiers, namely EBC and MMASA. Figure 4 is plotted with product review datasets on the x - axis, and the percentage of obtained results for precision is plotted on the y - axis. From Figure 4, it is obvious that RHMM-SVM has a superior FS than the existing classifier. In RHMM - SVM, Baum - Welch Strategy based parameter estimation leads the way in achieving better FS in all considered different domain datasets. In existing classifiers EBC and MMASA, parameters are not considered before performing classification. Hence, EBC and MMASA face low - level FS in all considered domain datasets. The result values of Figure 4 is provided in Table 6.

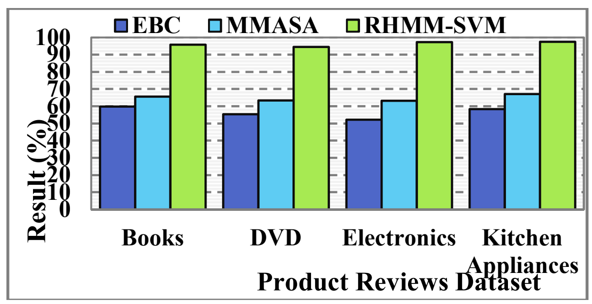

Fig 4. FS Vs Classifiers

Table 6. Result Values of FS Analysis

PRODUCT REVIEW DATASETS	EBC	MMASA	RHMM-SVM
Books	59.874	65.482	95.710
DVD	55.355	63.276	94.513
Electronics	52.068	63.144	97.302
Kitchen Appliances	58.379	67.016	97.433

6.4 CA Analysis

Figure 3 analyzes the CA for the proposed classifier HMM-SVM against the existing classifiers, namely EBC and MMASA. Figure 3 is plotted with product review datasets on the x-axis, and the percentage of obtained results for precision is plotted in the y-axis. RHMM-SVM is designed to perform classification in different domain big datasets where it uses a forward-backwards strategy to attain better results even if it is applied in different domain datasets. EBC and MMASA classifiers are designed to perform classification in a specific domain dataset only, and it leads a way for poor classification while applying in different domain datasets. The result values of Figure 3 is provided in Table 5.

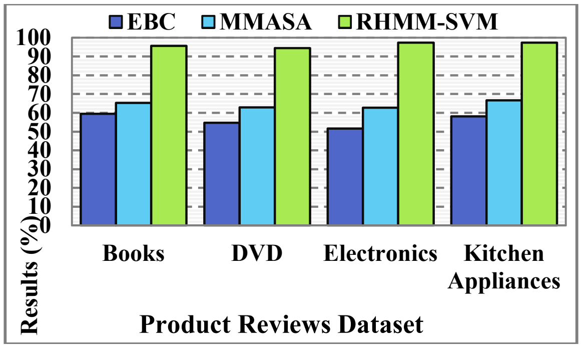

Fig 3. CA Vs Classifiers

Table 5. Result Values of CA Analysis

PRODUCT REVIEW DATASET	EBC	MMASA	RHMM-SVM
Books	59.566	65.298	95.653
DVD	54.695	62.871	94.437
Electronics	51.644	62.732	97.337
Kitchen Appliances	58.156	66.673	97.411

7. CONCLUSION

This paper has proposed Resilience Hidden Markov Model based Support Vector Machine (RHMM - SVM) to classify sentiments in a big product review dataset. Decision functions present in RHMM - SVM assist in classifying different domain product reviews that are unstructured. RHMM - SVM perform classification based on different state and their probability values which assist in collecting exact information from product reviews. The performance of RHMM - SVM is tested with four different big product review datasets using benchmark metrics. RHMM - SVM has achieved a classification accuracy of 95.653% with big product review books dataset, 94.437% with big product review DVD dataset, 97.337% with big product review electronics dataset, and 97.411% with big product review kitchen appliances dataset. Averagely RHMM - SVM has achieved a classification accuracy of 96.209%, where EBC and MMASA have achieved 56.015% and 64.394%, respectively. Future enhancement of this research work can be focused on audio and visual based sentiments present in product reviews, also optimization can be applied for increasing the classification accuracy even more.

REFERENCES:

[1] S. Bouktif, A. Fiaz, and M. Awad, “Augmented Textual Features - Based Stock Market Prediction,” *IEEE Access*, vol. 8, pp. 40269–40282, 2020, doi: 10.1109/ACCESS.2020.2976725.
[2] A. Vij and J. Pruthi, “An automated Psychometric Analyzer based on Sentiment Analysis and Emotion Recognition for healthcare,” *Procedia Comput. Sci.*, vol. 132, pp. 1184–1191, 2018, doi: https://doi.org/10.1016/j.procs.2018.05.033.
[3] H. Saif, Y. He, M. Fernandez, and H. Alani, “Contextual semantics for sentiment analysis of Twitter,” *Inf. Process. Manag.*, vol. 52, no. 1, pp. 5–19, 2016, doi: https://doi.org/10.1016/j.ipm.2015.01.005.
[4] J. Ramkumar and R. Vadivel, “Multi-Adaptive Routing Protocol for Internet of Things based Ad-hoc Networks,” *Wirel. Pers. Commun.*, pp. 1–23, Apr. 2021, doi: 10.1007/s11277-021-08495-z.
[5] J. Ramkumar and R. Vadivel, “Whale Optimization Routing Protocol for Minimizing Energy Consumption in Cognitive Radio Wireless Sensor Network,” *Int. J. Comput. Networks Appl.*, vol. 8, no. 4, doi: 10.22247/ijcna/2021/209711.
[6] J. Ramkumar and R. Vadivel, “Improved Wolf prey inspired protocol for routing in cognitive radio Ad Hoc networks,” *Int. J. Comput. Networks Appl.*, vol. 7, no. 5, pp. 126–136, 2020, doi: 10.22247/ijcna/2020/202977.
[7] J. Ramkumar and R. Vadivel, “Performance modeling of bio-inspired routing protocols in Cognitive Radio Ad Hoc Network to reduce end-to-end delay,” *Int. J. Intell. Eng. Syst.*, vol. 12, no. 1, pp. 221–231, 2019, doi: 10.22266/IJIES2019.0228.22.
[8] A. Reyes-Menendez, J. R. Saura, and F. Filipe, “Marketing challenges in the #MeToo era: gaining business insights using an exploratory sentiment analysis,” *Heliyon*, vol. 6, no. 3, p. e03626, 2020, doi: https://doi.org/10.1016/j.heliyon.2020.e03626.
[9] J. Zhou, J. Zhao, J. X. Huang, Q. V. Hu, and L.

He, "MASAD: A large-scale dataset for multimodal aspect-based sentiment analysis," *Neurocomputing*, vol. 455, pp. 47–58, 2021, doi: https://doi.org/10.1016/j.neucom.2021.05.040.

[10] H. Peng, Y. Ma, S. Poria, Y. Li, and E. Cambria, "Phonetic-enriched text representation for Chinese sentiment analysis with reinforcement learning," *Inf. Fusion*, vol. 70, pp. 88–99, 2021, doi: https://doi.org/10.1016/j.inffus.2021.01.005.

[11] M. E. Basiri, S. Nemati, M. Abdar, E. Cambria, and U. R. Acharya, "ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment analysis," *Futur. Gener. Comput. Syst.*, vol. 115, pp. 279–294, 2021, doi: https://doi.org/10.1016/j.future.2020.08.005.

[12] Y. Liang, F. Meng, J. Zhang, Y. Chen, J. Xu, and J. Zhou, "A dependency syntactic knowledge augmented interactive architecture for end-to-end aspect-based sentiment analysis," *Neurocomputing*, vol. 454, pp. 291–302, 2021, doi: https://doi.org/10.1016/j.neucom.2021.05.028.

[13] A. Zhao and Y. Yu, "Knowledge-enabled BERT for aspect-based sentiment analysis," *Knowledge-Based Syst.*, vol. 227, p. 107220, 2021, doi: https://doi.org/10.1016/j.knosys.2021.107220.

[14] K. Gulati, S. Saravana Kumar, R. Sarath Kumar Boddu, K. Sarvakar, D. Kumar Sharma, and M. Z. M. Nomani, "Comparative analysis of machine learning-based classification models using sentiment classification of tweets related to COVID-19 pandemic," *Mater. Today Proc.*, 2021, doi: https://doi.org/10.1016/j.matpr.2021.04.364.

[15] W. Song, Z. Wen, Z. Xiao, and S. C. Park, "Semantics perception and refinement network for aspect-based sentiment analysis," *Knowledge-Based Syst.*, vol. 214, p. 106755, 2021, doi: https://doi.org/10.1016/j.knosys.2021.106755.

[16] Y. Han and M. Moghaddam, "Analysis of sentiment expressions for user-centered design," *Expert Syst. Appl.*, vol. 171, p. 114604, 2021, doi: https://doi.org/10.1016/j.eswa.2021.114604.

[17] Z. Li, L. Li, A. Zhou, and H. Lu, "JTSG: A Joint Term-Sentiment Generator for Aspect-based Sentiment Analysis," *Neurocomputing*, 2021, doi: https://doi.org/10.1016/j.neucom.2021.06.045.

[18] O. Rakhmanov, "A Comparative Study on Vectorization and Classification Techniques in Sentiment Analysis to Classify Student-Lecturer Comments," *Procedia Comput. Sci.*, vol. 178, pp. 194–204, 2020, doi: https://doi.org/10.1016/j.procs.2020.11.021.

[19] K. Garcia and L. Berton, "Topic detection and sentiment analysis in Twitter content related to COVID-19 from Brazil and the USA," *Appl. Soft Comput.*, vol. 101, p. 107057, 2021, doi: https://doi.org/10.1016/j.asoc.2020.107057.

[20] I. Touahri and A. Mazroui, "Enhancement of a multi-dialectal sentiment analysis system by the detection of the implied sarcastic features," *Knowledge-Based Syst.*, vol. 227, p. 107232, 2021, doi: https://doi.org/10.1016/j.knosys.2021.107232.

[21] J. Ramkumar and R. Vadivel, "CSIP—cuckoo search inspired protocol for routing in cognitive radio ad hoc networks," in *Advances in Intelligent Systems and Computing*, 2017, vol. 556, pp. 145–153, doi: 10.1007/978-981-10-3874-7_14.

[22] R. Vadivel and J. Ramkumar, "QoS-Enabled Improved Cuckoo Search-Inspired Protocol (ICSIP) for IoT-Based Healthcare Applications," pp. 109–121, 2019, doi: 10.4018/978-1-7998-1090-2.ch006.

[23] J. Ramkumar and R. Vadivel, "Bee inspired secured protocol for routing in cognitive radio ad hoc networks," *INDIAN J. Sci. Technol.*, vol. 13, no. 30, pp. 3059–3069, 2020, doi: 10.17485/IJST/v13i30.1152.

[24] J. Ramkumar and R. Vadivel, "Improved frog leap inspired protocol (IFLIP) – for routing in cognitive radio ad hoc networks (CRAHN)," *World J. Eng.*, vol. 15, no. 2, pp. 306–311, 2018, doi: 10.1108/WJE-08-2017-0260.

[25] J. Ramkumar and R. Vadivel, "Meticulous elephant herding optimization based protocol for detecting intrusions in cognitive radio ad hoc networks," *Int. J. Emerg. Trends Eng. Res.*, vol. 8, no. 8, pp. 4549–4554, 2020, doi: 10.30534/ijeter/2020/82882020.

[26] M. Lingaraj, T. N. Sugumar, C. Stanly Felix, and J. Ramkumar, "Query Aware Routing Protocol for Mobility Enabled Wireless Sensor Network," *Int. J. Comput. Networks Appl.*, vol. 8, no. 3, p. 258, Jun. 2021, doi: 10.22247/IJCNA/2021/209192.

[27] J. Ramkumar and R. Vadivel, "Intelligent Fish Swarm Inspired Protocol (IFSIP) For Dynamic Ideal Routing in Cognitive Radio Ad-Hoc Networks," *Int. J. Comput. Digit. Syst.*, vol. 10, no. 1, pp. 1063–1074, 2020, Accessed: Dec. 02, 2020. [Online]. Available:

https://journal.uob.edu.bh/bitstream/handle/123456789/3961/paper 96.pdf?sequence=4&isAllowed=y.

[28] S. Boopalan and S. Jayasankari, "Dolphin Swarm Inspired Protocol (DSIP) for routing in underwater wireless sensor networks," *Int. J. Comput. Networks Appl.*, vol. 8, no. 1, pp. 44–53, Feb. 2021, doi: 10.22247/IJCNA/2021/207981.

[29] J. Wei, J. Liao, Z. Yang, S. Wang, and Q. Zhao, "BiLSTM with Multi-Polarity Orthogonal Attention for Implicit Sentiment Analysis," *Neurocomputing*, vol. 383, pp. 165–173, Mar. 2020, Accessed: May 06, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S092523121931656X.

[30] W. Li, L. Zhu, Y. Shi, K. Guo, and E. Cambria, "User reviews: Sentiment analysis using lexicon integrated two-channel CNN–LSTM family models," *Appl. Soft Comput.*, vol. 94, p. 106435, 2020, doi: https://doi.org/10.1016/j.asoc.2020.106435.

[31] D. T. Huerta, J. Hawkins, J. Brownstein, and Y. Hswen, "Exploring discussions of health and risk and public sentiment in MA during COVID-19 pandemic mandate implementation: A Twitter analysis," *SSM - Popul. Heal.*, p. 100851, 2021, doi: https://doi.org/10.1016/j.ssmph.2021.100851.

[32] T. Y. Lam and I. M. Meyer, "Efficient algorithms for training the parameters of hidden Markov models using stochastic expectation maximization (EM) training and Viterbi training," *Algorithms Mol. Biol.*, vol. 5, no. 1, pp. 1–16, Dec. 2010, doi: 10.1186/1748-7188-5-38/TABLES/2.

[33] A. Puengnim, N. Thomas, J. Y. Tourneret, and J. Vidal, "Classification of linear and non-linear modulations using the Baum–Welch algorithm and MCMC methods," *Signal Processing*, vol. 90, no. 12, pp. 3242–3255, Dec. 2010, doi: 10.1016/J.SIGPRO.2010.05.030.