Journal of Theoretical and Applied Information Technology
15th December 2024. Vol.102. No. 23

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

ENHANCED LDA MODEL FOR SENTIMENT ANALYSIS (ELDASA)

ALURI SUGUNA RATNA PRIYA[1], A . MARY SOWJANYA[2]

[1]Research Scholar, Department of CS&SE, Andhra University College of Engineering, Visakhapatnam Andhra Pradesh, India.

[2]Associate Professor, Department of CS&SE, Andhra University College of Engineering, Visakhapatnam Andhra Pradesh, India.

Email: [1] sugunaratnapriya.rs@andhrauniversity.edu.in, [2] dr.amsowjanya@andhrauniversity.edu.in

ABSTRACT

In the contemporary era, participative Internet communication or Social media platforms are widely used by people from all walks of life. Sharing opinions on various products and services has become common over social media. Unlike the conventional approach, opinions freely expressed over social media are goldmines to businesses. Analysing public sentiments has the potential to leverage business intelligence. Many researchers exploited Natural Language Processing (NLP) and Machine Learning (ML) to mine and ascertain opinions in online user-generated reviews. However, processing large text corpora is still challenging and prone to deteriorated performance. In this paper, we proposed a Generative Framework using an enhanced version of Latent Dirichlet Allocation Model considering sentiment polarities and latent aspects and we developed an algorithm named Enhanced Dirichlet Allocation Model for Sentiment Analysis (ELDASA) to realize the framework. This model is supported by a learning-based approach with ML toward the identification of sentiments and classifying them. Our empirical study using three social media datasets, consisting of reviews on hotels, music, and games, revealed that the proposed algorithm supports effective sentiment analysis.

Keywords – *Latent Dirichlet Allocation, Machine Learning, Sentiment Classification, Sentiment Analysis, Natural Language Processing.*

1. INTRODUCTION

Sentiment classification in textual documents, particularly online reviews, has attracted researchers due to the utility of discovering sentiments. With the emergence of cloud computing and social media, sentiment classification assumed unprecedented significance due to its ability to render required intelligence to businesses [1]. The rationale behind this importance is that online reviews can influence the decisions, of people from all walks of life. In this context, businesses can no longer ignore the opinions of customers or the general public on social platforms. Sentiment classification can be performed on large volumes of online reviews. It is made possible by the emergence of Artificial Intelligence (AI) which encapsulates NLP, machine learning, and deep learning.

The remainder of the paper covers the following details. Section 2 reviews the literature on the existing techniques used for sentiment analysis. Section 3 presents the methodology proposed along with the algorithm for sentiment analysis. Section 4 displays experimental results with three datasets. Section 5 wraps up our work and directs the scope for future research.

2. RELATED WORK

This section reviews the literature on existing techniques used for sentiment classification. Namcheol et al. [1] proposed a methodology based on unsupervised learning and NLP for classifying case studies about building information modelling (BIM). Machine learning is widely used for sentiment classification as explored in [3], [7], [11], [12], and [19]. In [3] a hybrid approach is developed based on SVM and RF to classify sentiments while different supervised ML approaches along with, text classification techniques are used in [7]. A

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

combination of ML and DP models is used in [11] for classifying drug review documents considering sentiment polarities. ML approaches are exploited in [12] based on Twitter tweets in the Covid-19 pandemic situations for sentiment classification. SVM along with the recursive feature optimization method is used in [19] for opinion mining.

Deep learning is another important approach used recently for sentiment classification as studied in [8], [10], [16]. LSTM is used in [8] for sentence representation to study the process of sentence-level sentiments in textual documents. Deep learning and weak supervision approaches along with the text classification method are used in [10] for classifying sentiments. ConvBiLSTM is the deep learning technique proposed in [16] for classifying sentiments in Twitter-generated documents. Single-layered bi-LSTM is used in [13] while CNN and LSTM combination with multi-task learning is explored in [14]. For processing text documents, it is important to have NLP and also feature engineering as investigated in [2], [22], [25], [29]. The impact of feature extraction on the accuracy of sentiment classification is investigated in [2]. A kind of feature selection known as the wrapper approach is followed in [22] for feature engineering to leverage sentiment classification. Multi-channel features along with self-attention coupled with bi-LSTM are explored in [25] for opinion mining. Feature engineering and term weighting approaches along with ML techniques are used in [29] for analysing sentiments textual corpora.

Sentiment classification methods used with the Arabic language are explored in [4] and [6]. Hybrid learning approaches are also found in the literature as studied in [5] where CNN and LSTM are combined to reap the benefits of both the deep learning techniques. Sentiment classification is found useful when it is incorporated into product reviews as investigated in [9] and [29]. Product reviews in the e-commerce domain are used for sentiment classification using Naïve Bayes techniques with a continuous learning phenomenon in [9] however the product reviews are analysed using feature selection and term weighting approaches in [29].

Sentiment classification of research publications of a clinical nature is done in [15] in terms of citation sentiment analysis. The Ensemble learning approach which is cost-effective reflecting a three-way combination is employed in [17] for sentiment classification. Tourist reviews are used in [18] for aspect-based sentiment classification while CNNs are enhanced to have sentence-level sentiment mining explored in [20]. Other important contributions found in the sentiment classification literature are the domain attention model [21], Gradient Boosting Machine [23], CNN with weakly supervised approach [24], context-aware approach [26], multi-task learning approach [27], [28] and ML based binary classification [30]. From the literature review, it is understood that many researchers exploited Natural Language Processing (NLP) and Machine Learning (ML) to mine and ascertain opinions in online user-generated reviews. However, processing large text corpora is still challenging and prone to deteriorated performance. [36] Proposed model identifies the polarity of sentiment from the topic-document and document-word.

3. MATERIALS AND METHODS

This section presents our proposed methodology, datasets used, enhanced LDA model, proposed algorithm, and performance evaluation procedure.

3.1 Dataset Collection

Popular websites such as Amazon and TripAdvisor are used to collect review datasets. The Hotel reviews dataset is collected from TripAdvisor while the video game reviews dataset and music CD reviews dataset are collected from Amazon product reviews. The Hotel reviews dataset has 1256 reviews, 585216 words, and 408 average number of words per review. The Music CD reviews dataset has 1521 reviews, 281050 words, and 167 average several words per review. The Video game reviews dataset has 2488 reviews, 443385 words, and 202 average number of words per review.

3.2 Methodology

The proposed Generative framework for sentiment Analysis is shown in Figure 1. It is based on an enhanced Latent Dirichlet Allocation (LDA) [31] with topic modelling. The enhanced LDA model, shown in Figure 2, is a statistical model or generative model that systematically deals with textual corpora. The framework takes one of the datasets described in Section 3.1 as input. The dataset is then subjected to NLP-based pre-processing before being assigned to an enhanced LDA model. Pre-processing includes data cleaning, exploratory data analysis, and NLP. NLP is used for resolving contractions, slang, and text cleaning by removing digits, stop words, duplicate letters, special characters, extra spaces, punctuation, white spaces, and HTML tags. Numbering and words with three or fewer are removed. The data is also subjected to stemming, lemmatization, and meaningless words.

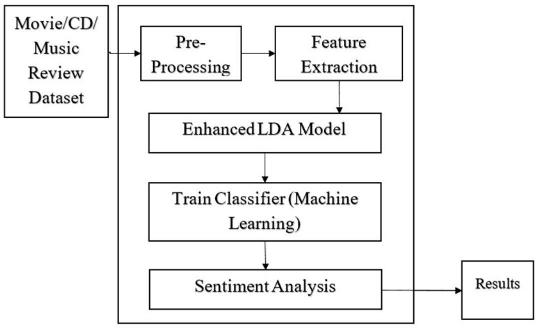

Figure 1: Generative Framework for Sentiment Analysis

Once pre-processing is completed, two variants of TF-IDF are used to extract features from pre-processed textual content. In the first variant, only stop-words are the argument considered while the second variant includes n-gram range and binary arguments. The N-gram range is set to (1,2) to generate one-word and two-word tokens while the binary argument is set to True reflecting the fact that total counts are shown in binary. Once features are extracted, the enhanced LDA model processes text corpora along ML models toward sentiment classification. The notations used in the enhanced LDA model are provided in Table 1. The boxes visualized in the model are plates that reflect replicates. The review documents are denoted by an outer plate while the inner place indicates words in each review document along with sentiment orientations and latent aspects.

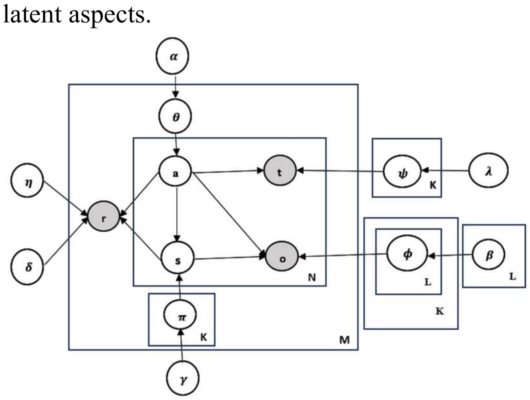

Figure 2: Illustrates the enhanced LDA model

The Traditional LDA aims at topic and word distributions whereas the proposed Enhanced LDA Model aims at both aspect and sentiment distributions. The modelling is done so that review documents are processed to find aspects and sentiments associated with aspects. Since user-generated reviews, do have opinions when compared with other documents, we approached the processing in terms of opinion pairs instead of using a bag of words. The enhanced LDA, as a probabilistic model, has provision to detect sentiments at the aspect level. It also exploits sentiment labels that come with online reviews in the form of ratings. Based on the constructed probabilistic framework, ML models are employed for sentiment classification. The latest aspects and associated sentiments in a given review are called hidden variables. Such variables or empirical frequencies are defined in Eq. 1.

$$\bar{z}_m = \frac{1}{C}\sum_{n=1}^{N}(a_{mn} \times (\omega^T \times s_{mn})) \quad (1)$$

Where ω is a normalization coefficient, d_m indicates a review, r_m is the rating response and C denotes a constant used for normalization. This definition is derived from $N(\eta^T \bar{z}_m \delta)$, a linear model, where variables associated with rating response are δ and η. Hidden aspects have covariates and regression coefficients like $\bar{z}_m$and η are linked to opinions. As discussed in [32], the non-linear model is a generalized form of linear model. As per that *r* is obtained from a distribution that is part of a family of distributions characterized by parameters like ρ and δ. This proposition is expressed in Eq. 2.

$$p(r|\rho,\delta) = h(r,\delta)\exp\{\frac{\rho r - A(\rho)}{\delta}\} \quad (2)$$

The natural parameter is denoted as ρ, which is associated with the distribution mean, while the other parameter, known as the dispersion parameter associated with the variance of the distribution. According to the work in [33], for a given review document, it is possible to substitute $\rho = \eta^T \bar{z}_m$in Eq. 2. This phenomenon gives rise to Eq. 3.

$$p(r_m|\bar{z}_m,\eta,\delta) = h(r_m,\delta)\exp\{\frac{(\eta^T\bar{z}_m)r_m - A(\eta^T\bar{z}_m)}{\delta}\}. \quad (3)$$

With the linear model's generalization, it has become flexible and has resulted in various models rating given review document in terms of exponential dispersion form. In the proposed framework, large-range probability distributions are supported. Each one of such distributions is linked to $h(r_m,\delta)$ and $A(\eta^T\bar{z}_m)$ as expressed in Eq. 4 and Eq. 5 respectively.

$$h(r_m,\delta) = \frac{1}{\sqrt{2\pi\delta}}\exp\{\frac{-r_m^2}{2\delta}\} \quad (4)$$

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

and

$$A(\eta^T \bar{z}_m) = \frac{(\eta^T \bar{z}_m)^2}{2}. \quad (5)$$

As far as normal distribution is concerned, σ^2 and μ are the parameters linked to δ and $\eta^T \bar{z}_m$. The novelty of the proposed LDA model lies in its intuitions. As the reviews associated with services and products, do have various aspects, each aspect has a different utility. Often simple regression on aspects and associated sentiments of reviews can help in finding meaningful information. Moreover, regression reflects relative contributions associated with aspects. The proposed framework uses a sentiment lexicon besides rating data available to ascertain latent and semantic sentiments in the reviews.

Table 1: Notations used in ELDASA

Notation	Meaning
$N_{k,l,v}$	Number of opinion words v linked to k and l
$N_{k,l}$	Number of opinion words linked to k and l
$N_{k,u}$	Number of aspect words linked to k
N_k	Number of aspect words lined to k
$N_{m,k,l}$	Number of words in the review linked to aspects k and l
$N_{m,k}$	Number of words in the review linked to aspect k
O_{mn}	In the given review d_m, it indicates an opinion word of n^th opinion pair
S_{mn}	Indicates sentiment assignment
$a^{\neg i}$	Indicates aspect assignment except a_i
a_{mn}	Indicates aspect assignment
r_m	For the given review d_m, it indicates the rating response.
$s^{\neg i}$	Indicates sentiment assignment except s_i
t_{mn}	In the given review d_m, it indicates an aspect term of n^th opinion pair
K	Indicates several semantic aspects
L	Indicates the number of semantic sentiments
M	Total number of documents present in the corpus
N	Indicates several opinion words in a given review document
U	Aspect words vocabulary
V	Opinion words vocabulary
α	Hyperparameter associated with θ
β	Hyperparameter associated with ϕ
γ	Hyperparameter associated with π
δ	Parameter pertaining to rating response
η	Parameter pertaining to rating response
θ	Dirichlet prior for aspects
λ	Hyperparameter associated with ψ
π	Dirichlet prior for sentiments
ψ	Dirichlet prior for aspect words
ϕ	Dirichlet prior for opinion words

The Model's inference plays an important role. The model inference is obtained, from each review and its latent variables. The hidden variables in terms of their posterior distribution are expressed in Eq. 6.

$$p(\boldsymbol{a}, \boldsymbol{s} | \boldsymbol{t}, \boldsymbol{o}, \boldsymbol{r}, \alpha, \gamma, \lambda, \beta, \eta, \delta) = \frac{p(\boldsymbol{a}, \boldsymbol{s}, \boldsymbol{t}, \boldsymbol{o}, \boldsymbol{r} | \alpha, \gamma, \lambda, \beta, \eta, \delta)}{p(\boldsymbol{t}, \boldsymbol{o}, \boldsymbol{r} | \alpha, \gamma, \lambda, \beta, \eta, \delta)}. \quad (6)$$

Obtaining inference accurately is intractable. Therefore, we employed the Gibbs sampling method found in [34]. Accordingly, each opinion pair is used to derive conditional distribution as in Eq. 7.

$$p(a_i = k, s_i = l | \boldsymbol{a}^{\neg i}, \boldsymbol{s}^{\neg i}, \boldsymbol{t}, \boldsymbol{o}, \boldsymbol{r}, \alpha, \gamma, \lambda, \beta, \eta, \delta), \quad (7)$$

Where $\boldsymbol{s}^{\neg i}$ and $\boldsymbol{a}^{\neg i}$ denote sentiment orientations and aspect assignments. The conditional distribution process is carried out for each opinion pair and aspect except s_i and a_i. The conditional distribution can be expanded as in Eq. 8.

$$p\left(a_i = k, s_i = l | \boldsymbol{a}^{\neg i}, \boldsymbol{s}^{\neg i}, \boldsymbol{t}, \boldsymbol{o}, \boldsymbol{r}, \alpha, \gamma, \lambda, \beta, \eta, \delta\right) = \frac{p(\boldsymbol{a}, \boldsymbol{s}, \boldsymbol{t}, \boldsymbol{o}, \boldsymbol{r} | \alpha, \gamma, \lambda, \beta, \eta, \delta)}{p(\boldsymbol{a}^{\neg i}, \boldsymbol{s}^{\neg i}, \boldsymbol{t}, \boldsymbol{o}, \boldsymbol{r} | \alpha, \gamma, \lambda, \beta, \eta, \delta)} \propto p(\boldsymbol{a}, \boldsymbol{s}, \boldsymbol{t}, \boldsymbol{o}, \boldsymbol{r} | \alpha, \gamma, \lambda, \beta, \eta, \delta). \quad (8)$$

The proposed probabilistic models related opinions, aspects, and aspect terms are reflected in Eq. 9.

$$p(\boldsymbol{a}, \boldsymbol{s}, \boldsymbol{t}, \boldsymbol{o}, \boldsymbol{r} | \alpha, \gamma, \lambda, \beta, \eta, \delta) = p(\boldsymbol{a} | \alpha) \cdot p(\boldsymbol{s} | \boldsymbol{a}, \gamma) \cdot p(\boldsymbol{t} | \boldsymbol{a}, \lambda) \cdot p(\boldsymbol{o} | \boldsymbol{a}, \boldsymbol{s}, \beta) \cdot p(\boldsymbol{r} | \boldsymbol{a}, \boldsymbol{s}, \eta, \delta). \quad (9)$$

The Eq. 10 is derived from Eq. 5 along with its first term and then integration of the same with θ.

$$p(\boldsymbol{a} | \alpha) = \prod_m \frac{\Gamma(\sum_k \alpha_k)}{\prod_k \Gamma(\alpha_k)} \frac{\prod_k \Gamma(N_{m,k} + \alpha_k)}{\Gamma(N + \sum_k \alpha_k)}. \quad (10)$$

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

The N indicates several words in the document while $N_{m,k}$ indicates the number of times the words are assigned to aspect k. Γ(x) indicates the Gamma function while the second term and its integration with π results in Eq. 11.

$$p(\boldsymbol{s}|\boldsymbol{a},\gamma) = \prod_m \prod_k \frac{\Gamma(L\gamma)}{\Gamma(\gamma)^L} \frac{\prod_l \Gamma(N_{m,k,l}+\gamma)}{\Gamma(N_{m,k}+L\gamma)}. \quad (11)$$

where $N_{m,k,l}$ denotes the number of times the words are linked to k and *l*. When the third term is considered for integration of ψ, we arrive at Eq. 12.

$$p(t|a,\lambda) = \prod_k \frac{\Gamma(|U|\lambda)}{\Gamma(\lambda)^{|U|}} \frac{\prod_u \Gamma(N_{k,u}+\lambda)}{\Gamma(N_k+|U|\lambda)}. \quad (12)$$

where u indicates aspect word, U denotes vocabulary and $N_{k,u}$ refers to a total number of times in which a given aspect is associated with k while the sum of $N_{k,u}$ is denoted by N_k. Regarding the fourth order, ϕ integration results in Eq. 13.

$$p(o|a,s,\beta) = \prod_k \prod_l \frac{\Gamma(\sum_v \beta_{l,v})}{\prod_v \Gamma(\beta_{l,v})} \frac{\prod_v \Gamma(N_{k,l,v}+\beta_{l,v})}{N_{k,l}+\sum_v \beta_{l,v}}. \quad (13)$$

The term in Eq. 9 can be expanded further leading to expression in Eq. 14.

$$p(r|a,s,\eta,\delta) = \prod_m \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(r_m-\eta^T \bar{z}_m)^2}{2\delta}\right). \quad (14)$$

Considering all the equations related to different orders and integrations, the terms not associated with opinion and aspect words are discarded. Thus, full conditional distribution opinion pairs and their index is expressed in Eq. 15.

$$p\left(a_i = k, s_i = l \middle| \boldsymbol{a}^{\neg i}, \boldsymbol{s}^{\neg i}, \boldsymbol{t}, \boldsymbol{o}, \boldsymbol{r}, \alpha, \gamma, \lambda, \beta, \eta, \delta\right) \propto \frac{N_{m,k}^{\neg i}+\alpha_k}{N^{\neg i}+\sum_{k`} \alpha_{k`}} \cdot \frac{N_{m,k,l}^{\neg i}+\gamma}{N_{m,k}^{\neg i}+L\gamma} \cdot \frac{N_{k,u}^{\neg i}+\lambda}{N_k^{\neg i}+|U|\lambda} \cdot \frac{N_{k,l,v}^{\neg i}+\beta_{l,v}}{N_{k,l}^{\neg i}+\sum_{v`} \beta_{l,v`}} \cdot \frac{1}{\sqrt{2\pi\delta}} \exp\left(-\frac{(r_m-\eta^T \bar{z}_m)^2}{2\delta}\right) \quad (15)$$

Here u is assigned k and the number of times it is done is denoted by $N_{k,u}^{\neg i}$. Related to parameter estimation that exploits Gibbs sampling, for each review, aspect distribution is computed as in Eq. 16.

$$\theta_{m,k} = \frac{N_{m,k}+\alpha_k}{N+\sum_{k`=1}^{K} \alpha_{k`}}. \quad (16)$$

Then, the aspect-specific sentiment distribution for each review document can be computed as in Eq. 17.

$$\pi_{m,k,l} = \frac{N_{m,k,l}+\gamma}{N_{m,k}+L\gamma}. \quad (17)$$

Concerning aspect word distribution, the distribution dynamics are expressed as in Eq. 18.

$$\psi_{k,u} = \frac{N_{k,u}+\lambda}{N_k+|U|\lambda}. \quad (18)$$

Having obtained the aspect word distribution, opinion word distribution is expressed as in Eq. 19.

$$\phi_{klv} = \frac{N_{k,l,v}+\beta_{l,v}}{N_{k,l}+\sum_{v`=1}^{|V|} \beta_{l,v`}}. \quad (19)$$

In the proposed approach different Dirichlet priors are used. Asymmetric Dirichlet prior is denoted by α and is computed, as explored in [35], with the help of the fixed point iteration approach. Word distribution of sentiment prior is used with the help of β for ϕ. Prior knowledge is obtained using a public lexical dictionary known as MPQA. In the proposed framework L value is set to 2 indicating two sentiment orientations such as positive and negative. The positive and negative orientations are set to $\beta_{lv} = 0.95$ and $\beta_{lv} = 0.05$ respectively. Symmetric priors like γ and λ are considered for given π and ψ and their values are set to 1/L and 0.01.

3.3 Algorithm Design

We proposed an algorithm named Enhanced LDA Model for Sentiment Analysis (ELDASA). In the proposed algorithm, a given document known as d_m has its rating r_m are two important variables for modelling. For each review document, ML-based enhanced LDA is employed.

Algorithm: Enhanced LDA Model for Sentiment Analysis (ELDASA)

Inputs

Sentiment orientations L

Aspects K

Output

Rating to documents for sentiment classification

1. Begin
2. For each k in K
3. Aspect word distribution $\psi_k \sim Dir(\lambda)$
4. For each *l* in L

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

Opinion word $\phi_{kl} \sim Dir(\beta_l)$
End For
End For
For each review d_m in r_m
Aspect distribution $\theta_m \sim Dir(\alpha)$
For each aspect k in r_m
Sentiment distribution $\pi_{mk} \sim Dir(\gamma)$
End For
End For
For each Opinion Pair $\langle t_{mn}, O_{mn} \rangle$ in N
Aspect assignment $a_{mn} \sim Mult(\theta_m)$
Assigning sentiments $s_{mn} \sim Mult(\pi_{ma_{mn}})$
Retrieve aspect terms $t_{mn} \sim Mult(\psi_{a_{mn}})$
Retrieval of opinion words $O_{mn} \sim Mult(\phi_{a_{mn}s_{mn}})$
End For
Rating computation $r_m \sim N(\eta^T \bar{z}_m, \delta)$
End

Algorithm 1: Enhanced LDA Model for Sentiment Analysis (ELDASA).

Algorithm 1 takes L and K as inputs and performs the required processing toward sentiment analysis. The algorithm has three iterative processes. After drawing aspect word distribution, the first iterative approach is meant for computing opinion word distribution. The second iterative process is meant for knowing aspect distribution and sentiment distribution. The third iterative process works on each opinion pair to draw aspect sentiment and its assignment besides drawing aspect terms and opinion words. Finally, the algorithm produces ratings and other required computations for effective sentiment analysis.

3.4 Performance Evaluation

Depending on the confusion matrix, the evaluation of the proposed algorithm is compared with the state-of-the-art. Table 2 shows different metrics used in the evaluation process.

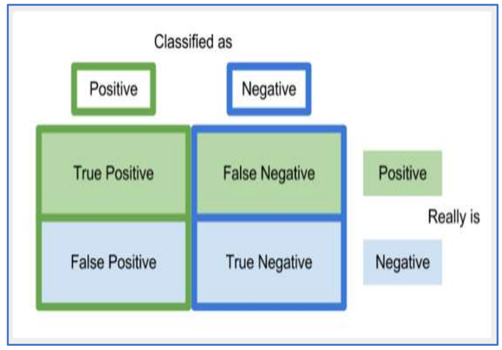

Figure 3: Confusion matrix

Depending on the confusion matrix presented in Figure 3 shows measures like true positive (TP), false positive (FP), false negative (FN), and true negative (TN). These are determined by comparing the result of the ML algorithm with the ground truth.

Table 2: Performance metrics used for evaluation

Metric	*Formula*	*Value Range*	*Best Value*
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$	[0; 1]	1
Precision (p)	$\frac{TP}{TP+FP}$	[0; 1]	1
Recall (r)	$\frac{TP}{TP+FN}$	[0; 1]	1
F1-Score	$2 * \frac{(p*r)}{(p+r)}$	[0; 1]	1

Precision refers to positive predictive value while recall refers to true positive rate. F1-score is the harmonic mean of both precision and recall which is used to have a measure without showing imbalance while the accuracy measure may show imbalance.

4. EXPERIMENTAL RESULTS

The proposed framework and underlying ML models with the proposed algorithm are evaluated with a prototype application. The Enhanced LDA model along with parameter tuning of ML models, the proposed algorithm is found to have significantly better performance with ML models. In the proposed framework, the L value is set to 2 indicating two sentiment orientations such as positive and negative. The positive and negative orientations are set to β_{lv} =

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

0.95 and $\beta_{lv} = 0.05$ respectively. Symmetric priors like γ and λ are considered for given π and ψ and their values are set to 1/L and 0.01. For experiments, ML models used along with the enhanced LDA model are Logistic Regression (LR), Naive Bayes (NB), AdaBoost, and Support Vector Machine (SVM). With every model, TF-IDF is used with two variants. The first variant uses the stop-words parameter while the second variant uses three parameters such as stop_words, binary (True), and ngram_range (1,2).

4.1 Results with Hotel Reviews Dataset

This section presents experimental results with the Hotel Reviews dataset. All the models used in the algorithm's pipeline are evaluated with different performance metrics.

Table 3: Shows experimental results using the Hotel Reviews dataset

Models	*Performance (%)*			
	Precision	*Recall*	*F1 Score*	*Accuracy*
LR1-TFIDF	94.7	99.1	96.9	94.6
LR2-TFIDF	90	99.9	94.7	90.7
Naive1-TFIDF	83	100	90.7	83.1
Naive2-TFIDF	83	99.7	90.6	82.9
AdaBoost1-TFIDF	93.3	96.5	94.9	91.3
AdaBoost2-TFIDF	93	97	95	91.6
SVM1-TFIDF	95.4	98.8	97.1	95
SVM2-TFIDF	92	99.8	95.7	92.6

As per Table 3, diverse models are used to perform sentiment analysis using the enhanced LDA approach and their respective performance statistics provided.

4.2 Results with Music CD Reviews Dataset

This section presents experimental results with the Music CD Reviews dataset. All the models used in the algorithm's pipeline are evaluated with different performance metrics.

Table 4: Shows experimental results using the Music CD Reviews dataset

Models	*Performance (%)*			
	Precision	*Recall*	*F1 Score*	*Accuracy*
LR1-TFIDF	93.74	97.6	95.67	93.37
LR2-TFIDF	91	98.4	94.7	90.7
Naive1-TFIDF	87	98	92.5	84.9
Naive2-TFIDF	85	96	90.5	82.8
AdaBoost1-TFIDF	94	89	91.5	87.9
AdaBoost2-TFIDF	89	91	90	86.6
SVM1-TFIDF	91	94	92.5	90.4
SVM2-TFIDF	87	91	89	85.9

As presented in Table 4, diverse models are used to perform sentiment analysis using the enhanced LDA approach, and their performance statistics are provided. These observations are made using the Movie CD reviews dataset.

4.3 Results with Games Reviews Dataset

This section presents experimental results with the Games Reviews dataset. All the models used in the algorithm's pipeline are evaluated with different performance metrics.

Table 5: Shows experimental results using the Games Reviews dataset

Models	*Performance (%)*			
	Precision	*Recall*	*F1 Score*	*Accuracy*
LR1-TFIDF	94.75	96.5	95.67	93.37
LR2-TFIDF	93	93.7	94.7	90.7
Naive1-TFIDF	91	89.7	92.5	88
Naive2-TFIDF	90	92.75	90.5	84
AdaBoost1-TFIDF	87	92.7	91.5	86
AdaBoost2-TFIDF	94	89	90	82
SVM1-TFIDF	82	96	92.5	91

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

SVM2-TFIDF	89	91	89	87

As presented in Table 5, several models are used to perform sentiment analysis using the enhanced LDA approach and their respective performance statistics provided. These observations are made using the Games reviews dataset.

4.4 Performance Comparison

This section compares the performance of the different models when applied to the three online review datasets. The performance of each model concerning every dataset is provided in terms of precision, recall, F1-Score, and accuracy.

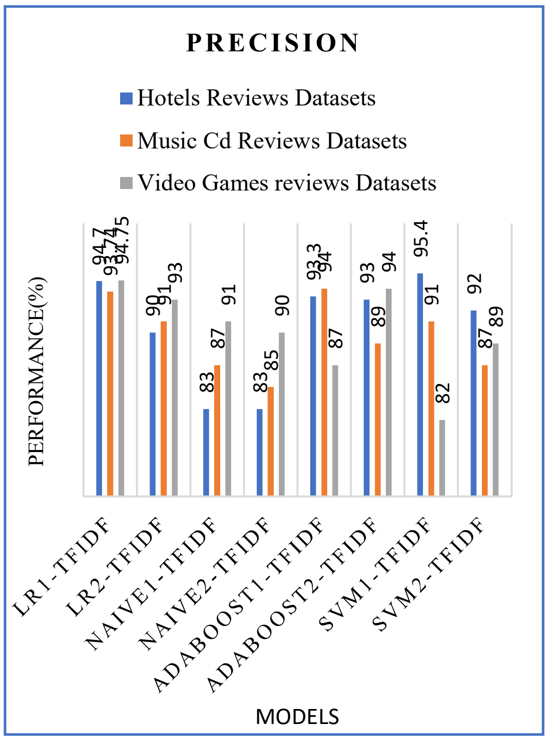

Figure 4: Performance comparison with all datasets in terms of precision

As presented in Figure 4, the results of four ML models used with enhanced LDA model are provided. The observations are made with experiments using all three datasets in terms of precision. Higher precision indicates better performance. The Highest precision of LR1 is 94.75% using the hotel reviews dataset. LR2 exhibited the highest precision 93% using the games dataset. NB1 and NB2 showed the highest performance 91% and 90% with the games dataset respectively. AdaBoost1 could achieve the highest precision 94% with the music dataset and AdaBoost2 94% with the games dataset. SVM1 showed the highest precision 95.4% with the hotel dataset while SVM2 showed 92% precision with the hotel dataset. From the results, it is observed that the SVM1 model showed the highest precision 95.40% using the hotel dataset in sentiment analysis in terms of precision.

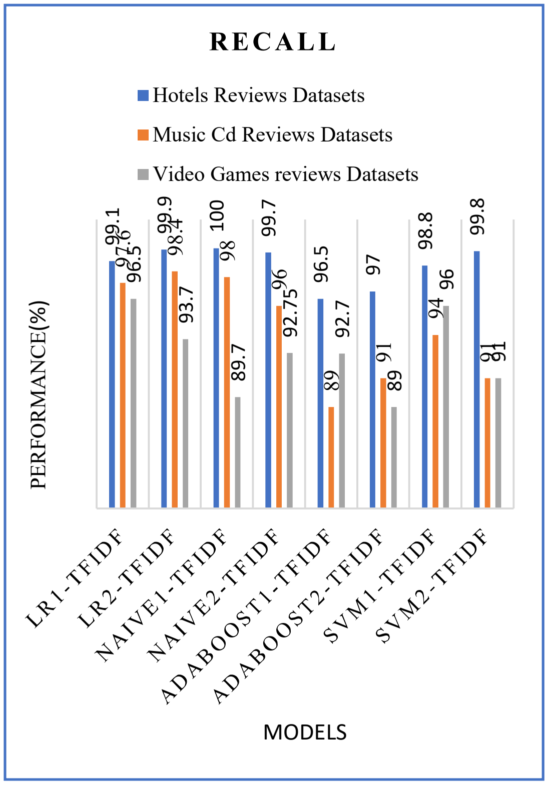

Figure 5: Performance comparison with all datasets in terms of recall

As presented in Figure 5, the results of four ML models used with enhanced LDA model are provided. The observations are made with the experiments, using all three datasets in recall. Higher recall indicates better performance. The Highest recall of LR1 is 99.10% using the hotel reviews dataset. LR2 exhibited the highest recall 99.9% using the hotel dataset. NB1 and NB2 showed the highest performance of 100% and 99.7% with the hotel dataset respectively. AdaBoost1 could achieve the highest recall of 96.5% with the hotel dataset and AdaBoost2 of 97% with the hotel dataset. SVM1 showed the highest recall 98.8.4% with the hotel dataset while SVM2 showed 99.8% recall with the hotel dataset. From the results, it is observed that the NB1 model showed the highest recall 100% using the hotel dataset in sentiment analysis in terms of recall.

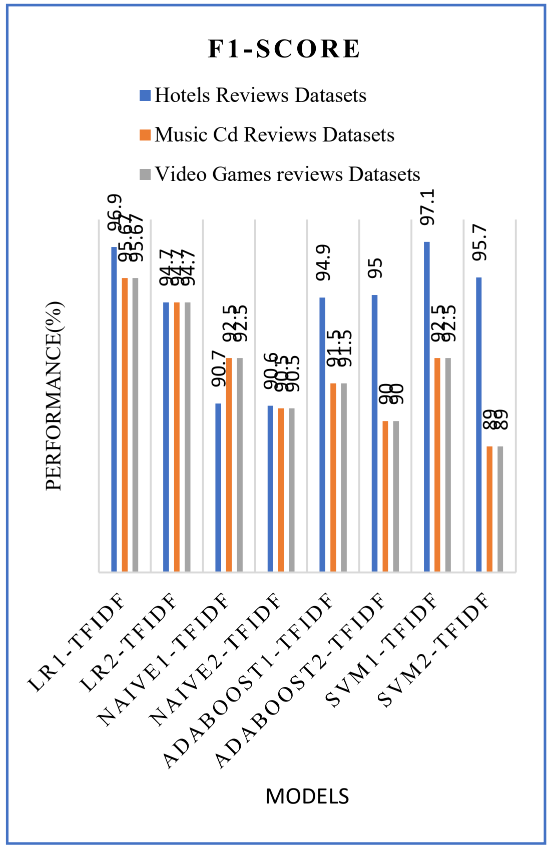

Figure 6: Performance comparison with all datasets in terms of F1-Score

As presented in Figure 6, the results of four ML models used with enhanced LDA model are provided. The observations are made experiments using all three datasets in terms of F1-Score. A Higher F1-Score indicates better performance. The Highest F1-Score of LR1 is 96.9% using the hotel reviews dataset. LR2 exhibited the highest F1-Score 94.7% with all datasets. NB1 and NB2 showed the highest performance 90.6% and 94.9% with the hotel dataset respectively. AdaBoost1 could achieve the highest F1-Score 94.9% with the hotel dataset and AdaBoost2 95% with the hotel dataset. SVM1 showed the highest F1-Score 97.1% with the hotel dataset while SVM2 showed 95.7% F1-Score with the hotel dataset. From the results, it is observed that the SVM1 model showed the highest F1-Score 97.1% using the hotel dataset in sentiment analysis in terms of F1-Score.

Journal of Theoretical and Applied Information Technology
15th December 2024. Vol.102. No. 23

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

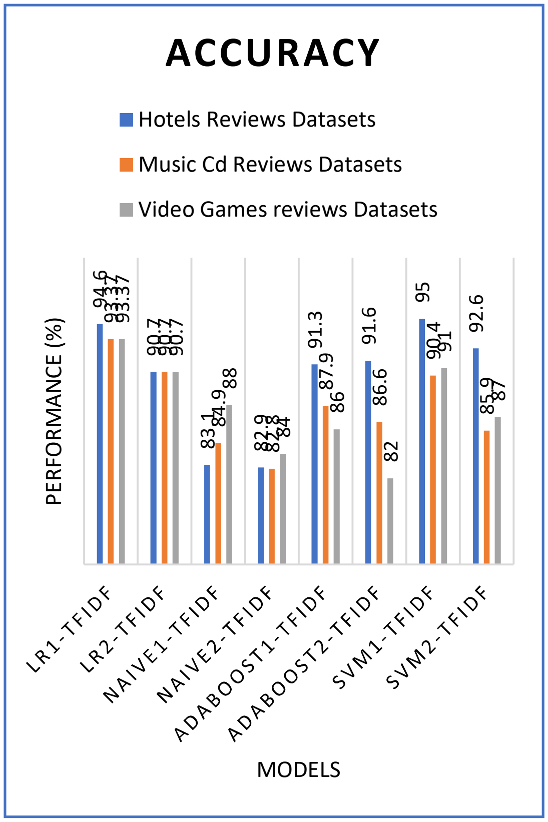

Figure 7: Performance comparison with all datasets in terms of accuracy

As presented in Figure 7, the results of four ML models used with enhanced LDA model are provided. The observations are made with experiments using all three datasets in accuracy. Higher in accuracy indicates better performance. The Highest accuracy of LR1 is 94.6% using the hotel reviews dataset. LR2 exhibited the highest accuracy 90.7% with all datasets. NB1 and NB2 showed the highest performance 88% and 84% with the games dataset respectively. AdaBoost1 could achieve the highest accuracy of 91.3% with the hotel dataset and AdaBoost2 91.6% with the hotel dataset. SVM1 showed the highest accuracy 95% with the hotel dataset while SVM2 showed 92.6% accuracy with the hotel dataset. From the results, it is observed that the SVM1 model showed the highest accuracy 95% using the hotel dataset in sentiment analysis in terms of accuracy.

4.5 Comparison with Prior Approaches

The sentiment classification results of the proposed model on the hotel review dataset are compared with those of the TDS model [36]. The ELDASA, along with SVM1, outperforms the TDS model across Precision, Recall, and F1-Score. Table 6 depicts the improvement shown by the proposed model compared to the TDS model.

Table 6: Experimental results compared to prior approach

Models	*Precision*	*Recall*	*F1- Score*
Proposed Model with SVM1	95.4	98.8	97.1
TDS Model	93.21	90.50	91.83

5. CONCLUSION AND FUTURE WORK

We proposed a Generative framework, the Enhanced Latent Dirichlet Allocation Model for Sentiment Analysis, considering sentiment polarities and latent aspects. The enhanced LDA is a generative model that systematically deals with textual corpora. The framework takes one of the datasets as input and then the dataset is then subjected to NLP-based pre-processing before giving it to enhanced LDA model. Besides our framework incorporates a strong pre-processing methodology based on NLP to improve the online review dataset before the generative process. Our empirical study using three social media datasets, consisting of reviews on hotels, music, and games, revealed that ELDASA supports effective sentiment analysis. From the results, it is observed that the SVM1 model showed the highest accuracy 95% using the hotel dataset in sentiment analysis in terms of accuracy. In the future, we intend to improve our framework with further research on feature engineering and other means of hyperparameter optimization.

REFERENCEs

[1] Jung, Namcheol and Lee, Ghang (2019). “Automated classification of building information modeling (BIM) case studies by BIM use based on natural language processing (NLP) and unsupervised learning”, Advanced Engineering Informatics, S1474034618304981–. http://doi:10.1016/j.aei.2019.04.007

[2] Ahuja, Ravinder; Chug, Aakarsha; Kohli, Shruti; Gupta, Shaurya and Ahuja, Pratyush, “The Impact of Features Extraction on the Sentiment Analysis”, Procedia Computer Science, 152, 341–348. http://doi:10.1016/j.procs.2019.05.008

[3] Al Amrani, Yassine; Lazaar, Mohamed and El Kadiri, Kamal Eddine, “Random Forest

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

and Support Vector Machine based Hybrid Approach to Sentiment Analysis", Procedia Computer Science, 127, 511–520. http://doi:10.1016/j.procs.2018.01.150

[4] Oueslati, Oumaima; Cambria, Erik; HajHmida, Moez Ben and Ounelli, Habib, "A review of sentiment analysis research in Arabic language", Future Generation Computer Systems, 112, 408–430. http://doi:10.1016/j.future.2020.05.034

[5] Rehman, Anwar Ur; Malik, Ahmad Kamran; Raza, Basit and Ali, Waqar , "A Hybrid CNN-LSTM Model for Improving Accuracy of Movie Reviews Sentiment Analysis", Multimedia Tools and Applications. http://doi:10.1007/s11042-019-07788-7

[6] Al-Ayyoub, Mahmoud; Khamaiseh, Abed Allah; Jararweh, Yaser and Al-Kabi, Mohammed N, "A comprehensive survey of arabic sentiment analysis", Information Processing & Management, S0306457316306689–. http://doi:10.1016/j.ipm.2018.07.006

[7] Kadhim, Ammar Ismael, "Survey on supervised machine learning techniques for automatic text classification", Artificial Intelligence Review. http://doi:10.1007/s10462-018-09677-1

[8] Huang, Weihang; Rao, Guozheng; Feng, Zhiyong and Cong, Qiong, "LSTM with sentence representations for Document-level Sentiment Classification", Neurocomputing, S092523121830479X–. http://doi:10.1016/j.neucom.2018.04.045

[9] Xu, Feng; Pan, Zhenchun and Xia, Rui, "E-commerce product review sentiment classification based on a naive Bayes continuous learning framework", Information Processing & Management, 102221–. http://doi:10.1016/j.ipm.2020.102221

[10] Wang, Yanshan; Sohn, Sunghwan; Liu, Sijia; Shen, Feichen; Wang, Liwei; Atkinson, Elizabeth J.; Amin, Shreyasee and Liu, Hongfang, "A clinical text classification paradigm using weak supervision and deep representation", BMC Medical Informatics and Decision Making, 19(1), 1–. http://doi:10.1186/s12911-018-0723-6

[11] Basiri, Mohammad Ehsan; Abdar, Moloud; Cifci, Mehmet Akif; Nemati, Shahla and Acharya, U. Rajendra, "A novel method for sentiment classification of drug reviews using fusion of deep and machine learning techniques", Knowledge-Based Systems, 198, 105949–. http://doi:10.1016/j.knosys.2020.105949

[12] Kamal Gulati; S. Saravana Kumar; Raja Sarath Kumar Boddu; Ketan Sarvakar; Dilip 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", Materials Today: Proceedings. http://doi:10.1016/j.matpr.2021.04.364

[13] Hameed, Zabit and Garcia-Zapirain, Begonya "Sentiment Classification Using a Single-Layered BiLSTM Model", IEEE Access, 8, 73992–74001. http://doi:10.1109/ACCESS.2020.2988550

[14] Jin, Ning; Wu, Jiaxian; Ma, Xiang; Yan, Ke and Mo, Yuchang, "Multi-Task Learning Model Based on Multi-Scale CNN and LSTM for Sentiment Classification", IEEE Access, 8, 77060–77072. http://doi:10.1109/ACCESS.2020.2989428

[15] Kilicoglu, Halil; Peng, Zeshan; Tafreshi, Shabnam; Tran, Tung; Rosemblat, Graciela and Schneider, Jodi, "Confirm or Refute?: A Comparative Study on Citation Sentiment Classification in Clinical Research Publications", Journal of Biomedical Informatics, 103123–. http://doi:10.1016/j.jbi.2019.103123

[16] Sakirin Tam; Rachid Ben Said and O. Ozgur Tanriover, "A ConvBiLSTM Deep Learning Model-Based Approach for Twitter Sentiment Classification," IEEE Access. http://doi:10.1109/access.2021.3064830

[17] Zhang, Yuebing; Miao, Duoqian; Wang, Jiaqi and Zhang, Zhifei, "A cost-sensitive three-way combination technique for ensemble learning in sentiment classification", International Journal of Approximate Reasoning, 105, 85–97. http://doi:10.1016/j.ijar.2018.10.019

[18] Afzaal, Muhammad; Usman, Muhammad and Fong, Alvis, " Tourism Mobile App with Aspect-Based Sentiment Classification Framework for Tourist Reviews", IEEE Transactions on Consumer Electronics, 1–1. http://doi:10.1109/TCE.2019.2908944

[19] Nur Syafiqah Mohd Nafis and Suryanti Awang, "An Enhanced Hybrid Feature Selection Technique Using Term Frequency-Inverse Document Frequency and Support Vector Machine-Recursive Feature Elimination for Sentiment Classification", IEEE Access. http://doi:10.1109/access.2021.3069001

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

[20] Zhang, Yuebing; Zhang, Zhifei; Miao, Duoqian and Wang, Jiaqi, "Three-way Enhanced Convolutional Neural Networks for Sentence-level Sentiment Classification", Information Sciences, S002002551830848X–. http://doi:10.1016/j.ins.2018.10.030

[21] Yuan, Zhigang; Wu, Sixing; Wu, Fangzhao; Liu, Junxin and Huang, Yongfeng, "Domain Attention Model for Multi-Domain Sentiment Classification", Knowledge-Based Systems, S0950705118302144–. http://doi:10.1016/j.knosys.2018.05.004

[22] Ansari, Gunjan; Ahmad, Tanvir and Doja, Mohammad Najmud, "Hybrid Filter–Wrapper Feature Selection Method for Sentiment Classification", Arabian Journal for Science and Engineering. http://doi:10.1007/s13369-019-04064-6

[23] Alzamzami, Fatimah; Hoda, Mohamad and Saddik, Abdulmotaleb El, "Light Gradient Boosting Machine for General Sentiment Classification on Short Texts: A Comparative Evaluation", IEEE Access, 8, 101840–101858. http://doi:10.1109/ACCESS.2020.2997330

[24] Lee, Gichag; Jeong, Jaey un; Seo, Seungwan; Kim, CzangYeob and Kang, Pilsung, "Sentiment classification with word localization based on weakly supervised learning with a convolutional neural network", Knowledge-Based Systems, S0950705118301710–. http://doi:10.1016/j.knosys.2018.04.006

[25] Li , Weijiang; Qi, Fang; Tang, Ming and Yu, Zhengtao, "Bidirectional LSTM with self-attention mechanism and multi-channel features for sentiment classification", Neurocomputing, S0925231220300254–. http://doi:10.1016/j.neucom.2020.01.006

[26] Masood, Muhammad Ali; Abbasi, Rabeeh Ayaz and Wee Keong, "Context-Aware Sliding Window for Sentiment Classification", IEEE Access, 8, 4870–4884. http://doi:10.1109/ACCESS.2019.2963586

[27] Akhtar, Md Shad; Garg, Tarun and Ekbal, Asif , "Multi-task Learning for Aspect Term Extraction and Aspect Sentiment Classification", Neurocomputing, S0925231220302897–. http://doi:10.1016/j.neucom.2020.02.093

[28] Zhang, Han; Sun, Shaoqi; Hu, Yongjin; Liu, Junxiu and Guo, Yuanbo, "Sentiment Classification for Chinese Text Based on Interactive Multitask Learning", IEEE Access, 1–1. http://doi:10.1109/access.2020.3007889

[29] Huiliang Zhao; Zhenghong Liu; Xuemei Yao and Qin Yang, "A machine learning-based sentiment analysis of online product reviews with a novel term weighting and feature selection approach", Information Processing & Management. http://doi:10.1016/j.ipm.2021.102656

[30] Ankita Sharma and Udayan Ghose. "Toward Machine Learning Based Binary Sentiment Classification of Movie Reviews for Resource Restraint Language (RRL)", IEEE. 11, pp.58546 - 58564. http://DOI:10.1109/ACCESS.2023.3283461

[31] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," J. Mach. Learn. Res., vol. 3, pp. 993–1022, March 2003.

[32] P. McCullagh and J. A. Nelder, "Generalized linear models," CRC Press, 1989, vol. 37

[33] D. M. Blei and J. D. McAuliffe, "Supervised topic models," Proceedings of the 21st Annual Conference on Neural Information Processing Systems - Volume 7, Vancouver, Canada, 2007, pp. 121–128.

[34] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 6, no. 6, pp. 721–741, 1984.

[35] T. Minka, "Estimating a Dirichlet distribution," in Technical report, MIT, 2000.

[36] AkhmedovFarkhod, Akmalbek Abdusalmoy, Fazliddin Makhmudov and Young ImCho, "LDA-Based Topic Modeling Sentiment Analysis Using Topic/Document/Sentence (TDS) Model", Appl. Sci. 2021, 11, 11091. https://doi.org/10.3390/app112311091