

Original Research Paper

Sentence Classification Using Attention Model for E-Commerce Product Review

Nagendra N and Chandra J

Department of Computer Science, Christ Deemed to be University, Bengaluru, India

Article history

Received: 20-12-2023

Revised: 15-02-2024

Accepted: 20-02-2024

Corresponding Author:

Nagendra N

Department of Computer
Science, Christ Deemed to be
University, Bengaluru, India

Email: nagendra.n@res.christuniversity.in

Abstract: The importance of aspect extraction in text classification, particularly in the e-commerce sector. E-commerce platforms generate vast amounts of textual data, such as comments, product descriptions, and customer reviews, which contain valuable information about various aspects of products or services. Aspect extraction involves identifying and classifying individual traits or aspects mentioned in textual reviews to understand customer opinions, improve products, and enhance the customer experience. The role of product reviews in e-commerce is discussed, emphasizing their value in aiding customers' purchase decisions and guiding businesses in product stocking and marketing strategies. Reviews are essential for boosting sales potential, maintaining a good reputation, and promoting brand recognition. Customers extensively research product reviews from different sources before purchasing, making them vital user-generated content for e-commerce businesses. The current work provided an efficient and novel classification model for sentence classification using the ABNAM model. The automated text classification models available cannot categorize the data into sixteen distinct classes. The technologies applied for the mentioned work contain TF-IDF, N-gram, CNN, linear SVM, random forest, Naïve bays, and ABNAM with significant results. The best-performing ML method for the successful classification of a given sentence into one of the sixteen categories is achieved with the proposed model named the based Neural Attention Model (ABNAM), which has the highest accuracy at 97%. The research acclaimed ABNAM as a novel classification model with the highest-class categorizations.

Keywords: Text Classification, Natural Language Processing (NLP), TF-IDF, N-Gram, Neural Networks, Machine Learning, Deep Learning, Logistic Regression (LR), Random Forest (RF), Multinomial Naïve Bayes (MNB), Linear Support Vector Machine (LSVC)

Introduction

The field of artificial intelligence, known as Natural Language Processing (NLP), investigates how computers and human language interact. It comprises several formulas and procedures that allow computers to comprehend, decode, and produce human language. Text classification, on the other hand, is a specialist NLP task that involves classifying or assigning predetermined labels to textual input following its content. NLP approaches are used to analyze, assess, and extract valuable information from text data in the context of text classification. For precise classification, computers must comprehend the text's structure, semantics, and context.

Electronic commerce, or just "e-commerce," is the

term used to describe the purchasing and selling of goods and services via different computer networks. In addition to private platforms like extranets that use Internet technologies like TCP/IP, these networks cover the vast Internet domain. Furthermore, these transactions are made possible largely via Electronic Data Interchange (EDI) networks. To put it simply, e-commerce encompasses a wide range of online transactions carried out through various digital channels.

There is a clear disparity in the challenges associated with efficiently acquiring human language in human-computer communication. The perceived communication gap between computers and people may be inflated as a challenging task due to the non-availability of a single-store solution. The proposed work is to provide a

multiclass classification model to accommodate a maximum of sixteen classes. The novelty of the work is the dataset processing; the existing studies apply the already existing word embeddings available at open-source platforms, but the proposed work applied unstructured data for the word embedding. To date, none of the models provided a successful classification for 16 categories. Even though computers typically use binary sequences and code to communicate, developments in Machine Learning (ML) and Natural Language Processing (NLP) have significantly improved computers' comprehension of human language. Technological advancements have enabled computers to understand human intent, including subtleties like slang and colloquialisms and instances when context is missing. Natural language comprehension allows technology to function autonomously on many tasks, like breaking down intricate information into manageable portions and deciphering queries without human support. Because modern technology automates related activities, humans are relieved of the tedious work of contextualizing language. Before discussing specific examples, the significance of natural language processing and its significance to the e-commerce business needs to be explained.

Natural Language Processing (NLP) to text categorization analysis (Beausoleil, 2019) has attracted much interest because of the quick expansion of user-generated text data online. NLP research investigates how people perceive goods, services, activities, problems, etc., by looking at people's actions, text classification, assessments, and appraisals. It can be used in various ways to help people and organizations locate, summarize, and comprehend relevant information from texts that reflect multiple points of view to support decision-making in both the corporate and social sectors (Khurana *et al.*, 2023).

Text classification analysis has been the subject of numerous studies. The preliminary study's main goal was text classification at the sentence and document levels. The current study aimed to determine the class or category for a given statement or sentence among the available sixteen classes defined in further sections in detail. The work proposed an accurate classification of a sentence into one of the sixteen classes such as restaurant, ambiance, clothing, beauty products, food prices, fragrance, home products, beauty-hair, food, health and wellness, home improvements, cosmetics-make-up, personal care, price, cosmetics-skincare, and staff. Machine Learning methodologies are implemented for the classification with the product review dataset collected from open-source repositories. There are various points of view when evaluating the traits of an aspect being in the actual world. Create categories relevant to the texts rather

than just making a generic comment about the quality of the depiction. For such categorization, there is a need for a system that is suitable to accommodate more than one or two categories. E-commerce is a platform used by numerous people in various ways and roles that need to be defined with clarity to provide the services they ask for. Consumers are spending more time online than ever chatting, browsing, and shopping, which is suitable for e-commerce. The state of the economy is changing. Many physical stores are closing right now. Online retailers must interact with customers anywhere they are, not just on the payment page, as social media browsing has exceeded window shopping better ROAS and ROI from an active social media presence online (Gasparetto *et al.*, 2022).

Product reviews both assist customers in making more informed purchases and assist businesses in deciding which goods to stock. Reviews are essential for boosting sales potential, upholding a good reputation, and promoting brand recognition. Reviews enhance the marketing strategy of your e-commerce firm by giving important product information. Reviews are one of the most crucial tools in your toolbox for e-commerce because they are user-generated content that you can utilize for free with the possibility of making a substantial profit. Before making a purchase, customers research numerous sources and types of product reviews, such as those on social media platforms, review websites, product listing pages on e-commerce websites, and customer endorsements (Xu *et al.*, 2020).

Looking at people's actions, text classification, assessments, and appraisals. It can be used in various ways to help people focus. It explains that the information given addresses a mutual information feature selection technique called MIFS-C. The MIFS-style approach looks for a significant parameter known as the redundancy parameter. One of the MIFS algorithms, the MIFS-C algorithm, is promoted as an effective and time-saving technique that enables quick execution without sacrificing accuracy (Bakus and Kamel, 2006).

The author's research paper compared and provided distinctions between four methods of document transformation. To apply label weights to multi-labeled documents, the author advises utilizing Entropy-based Label Assignment (ELA) based on label entropy. This method was tested on benchmark text sets in the paper using an SVM classifier and multi-label evaluation. The investigation showed that the document transformation technique significantly affects the effectiveness of multi-class categorization. According to the report, ELA did better in the evaluation than the alternatives (Chen *et al.*, 2007).

A thorough analysis of the feature selection criteria to categorize the material. The regional strategy functions better, which considers local and global regulations when only a few or no keywords are used. As the number of

keywords rises, the global strategy outperforms the regional strategy. When the current feature selection outcomes are reviewed, particularly in datasets with few terms, high success rates are observed. The Adaptive Keyword Selection (AKS) policy is mentioned in the paragraph as another technique for choosing keywords (Tasci and Gungor, 2008).

The statistical approaches explain data sparsity and how it affects the precision of categorizing short texts. By creating a feature thesaurus called Solid Feature Thesaurus (SFT) based on Latent Dirichlet Allocation (LDA) and Information Gain (IG) models, the author offers a solution. The strategy seeks to solve the issue of data sparsity in brief words. On two small text datasets, the author experiments and evaluates their approach's performance against cutting-edge techniques like Support Vector Machine (SVM) and Naive Bayes multinomial. The outcomes show that the suggested solution performs better than the current approaches, improving classification accuracy for short texts (Wang *et al.*, 2012).

The paper compared six existing techniques to Correlation-based Feature Selection (CMFS) in an experimental investigation. The study finds that CMFS outperforms other approaches, such as Information Gain (IG), Orthogonal Centroid Feature Selection (OCFS), CHI statistic (CHI), Document Frequency (DF), and DIA association factor when used in conjunction with the Naive Bayes algorithm. However, when paired with the Support Vector Machine (SVM) strategy, IG, DF, DIA and OCFS perform admirably. The findings indicate that when paired with Naive Bayes, CMFS is highly effective, while SVM is a superior choice for the other techniques (Meng *et al.*, 2011).

The importance of feature selection in text categorization highlights the need to consider the semantic relationship between phrases to improve classification accuracy. Accuracy is reported to suffer because semantic links are typically disregarded in the high-dimensional space required for document representation. The study offers a two-stage feature selection strategy as a solution to increase the efficacy and accuracy of categorization. To add new semantic space between concepts, the study suggests using the latent semantic indexing technique (Liu and Yang, 2012).

The importance of using the proper feature selection techniques to improve text categorization efficiency. The suggested feature selection approaches seek to improve text categorization through improved ambiguity metrics. The approaches are then contrasted with four currently employed techniques: Information gain, odd ratios, ambiguity measure, and mutual information. Three distinct datasets Reuters 21578, web KB and 20-newsgroups are used in the comparison, which is done using the Naive Bayes and SVM algorithms (Zhao *et al.*, 2020).

The bidirectional attention strategy with location encoding for creating aspect-specific visualization prototypes. The impetus for this study is that numerous sentence components are depicted in current methodologies without considering the significance of sentiment dependencies. Aspect-level categorization accurately depicts sentiment polarities across several aspect terms in a sentence. The authors offer a bidirectional attention mechanism that forges links between the context words and the aspect to prototype aspect-specific descriptions. Additionally, emotional relationships between different sentence parts are discovered using Graph Convolutional Networks (GCN) over the attention mechanism. The paragraph provides a high-level summary of the suggested approach, emphasizing its focus on aspect-specific representations and sentiment correlations in the text (Rohidin *et al.*, 2022).

The experiment she conducted used a dataset of 20 newsgroups and 20 student essays. The CBFSA (fuzzy soft set association rules mining) and different soft set classifiers, such as the Hybrid Fuzzy Classifier (HFC), Soft Set Classifier (SCC), and Fuzzy Soft Set Classifier (FSSC), are compared in this study for their effectiveness. The results show that the CBFSA approach works better than the present classifiers in terms of producing accurate results. The fuzzy decision set of an FP-soft set is used to construct a classifier and assign unclassified articles to a class (Huan *et al.*, 2022).

It is outlined regarding the goal of enhancing text classification accuracy as to why sentiment data addition and multi-dimensional document representation should be researched. The author advocates combining document sentiment with vector sequences or matrices. Additional sentiment vectors could be layered on top of the word vectors as a fully connected layer to increase accuracy even more. The best traditional method, the Naive Bayes Bernoulli classifier, has a 71.3% accuracy rate when applied to the dataset of suicide notes. The classifier uses sentiment and semantic data and achieves an accuracy rate of 75%. The primary focus of the paragraph is the potential to increase text categorization precision by including sentiment data and multi-dimensional representations (Koksal *et al.*, 2022).

The e-commerce industry has experienced significant growth in the dynamic environment brought about by post-pandemic shifts in consumer behavior. Lengthening a customer's online experience has become increasingly valuable. A strong search engine is one of the most crucial instruments for enhancing the user experience since it may provide valuable insights into the motivations of online buyers. A good search engine should deliver results promptly and accurately understand user intent to generate effective recommendations. This study integrates pre-training libraries in Turkish with language models such as BERT, Electra, and RoBERTa, which have demonstrated performance in similar tasks. These

are utilized, especially for e-commerce product categories, to construct models for text classification and customer intent analysis. Give a comparative analysis of the outputs from these deep learning models, which were further optimized using product descriptions from a particular e-commerce company and comment libraries. Expand on the findings from the e-commerce end-user intent analysis model by looking at past user searches and product categories. In addition to giving Turkish e-commerce products precedence, the proposed approach makes room for developing intent analysis algorithms that anticipate the main product consumers would look for using a range of products (Luo *et al.*, 2022).

The feature fusion strategy blends statistical data with significant semantic variables to suit the need for thorough, brief text categorization. Deep learning models and a weighting technique are used to accomplish this integration. The suggested technique uses Bidirectional Encoder Representations from Transformers (BERT) to automatically build sentence-level word vectors. Afterward, Convolutional Neural Network (CNN), Bidirectional Gate Recurrent Unit (BiGRU), and Term Frequency-Inverse Document Frequency (TF-IDF) weighting are used to construct statistical features, local semantic features, and overall semantic features. For classification, the fusion characteristic that results is employed (Soni *et al.*, 2023).

This study introduces TextConvoNet, a novel architecture for binary and multiclass text classification tasks that are based on Convolutional Neural Networks (CNNs). Unlike existing CNN models that use one-dimensional convolving filters that can only extract intra-sentence N-gram features, text convo net goes above and beyond by extracting inter-sentence N-gram characteristics from the input text data. This new method describes the input matrix using a different approach and employs a two-dimensional multi-scale convolutional operation. Experimental evaluations of text convo net's text classification performance are conducted on five datasets, encompassing binary and multi-class classification scenarios. Evaluation metrics include things like gmean1, gmean2, specificity, recall, accuracy, precision, and Mathew's Correlation Coefficient (MCC). Moreover, a thorough comparison study is carried out, in which TextConvoNet is contrasted with attention-based, machine-learning, and deep-learning models. It is possible to conclude from the experimental findings that TextConvoNet outperforms other models that are used for text classification (Wang and Li, 2022).

In the field of natural language processing, text classification is a hot topic. The application of Graph Neural Networks (GNN) to text categorization problems has garnered increasing attention in recent times. On the other hand, existing graph-based methods frequently miss important information that is concealed by text grammar

and sequence patterns. Additionally, because the text graph is built using the complete corpus, including the test set, using these models to analyze fresh documents creates difficulties. In order to address these problems, provide a text categorization model that combines the benefits of Graph Attention Networks (GAT) with Long Short-Term Memory Networks (LSTM). Each document is given its own unique graph by model, which then uses LSTM to create word embeddings enhanced with contextual information, GAT to learn inductive word representations, and finally, a consolidation of all nodes in the graph into the document embedding. The approach outperforms current text categorization techniques, as shown by experimental findings on four datasets. Compared to previous graph-based methods, it demonstrates quicker convergence and uses less memory. Interestingly, the model significantly improves even with small amounts of training data, highlighting the importance of text syntax and sequence information for better classification outcomes (Huang *et al.*, 2023).

In the series of most recent work, some of the strongly related work is highlighted here. Sentiment analysis is based on aspects, where the Aspect phrase extraction and aspect sentiment categorization are performed as the two subtasks. Current techniques have issues with fine-grained information extraction and aspect term border identification. It is challenging for the sentiment aspect classification task to adapt to text and detect local context. To overcome these issues, the proposed adaptive semantic relative distance method uses dependent syntactic analysis to determine each text's local context and enhance sentiment analysis accuracy. The model's mix of local and global information features outperforms state-of-the-art techniques in terms of aspect sentiment categorization. It boosts accuracy and F1 scores on the semeval-2014 task 4 restaurant and laptop datasets (Mujawar and Bhaladhare, 2023).

Text classification using labeled datasets for training with ML methods is implemented in the study. Feature selection, or the process of finding relevant words or phrases, enhances efficiency and interpretability by focusing on important aspects. The study introduces the wRMR method, which is used to extract aspect phrases from product assessments by combining the Gini index, information gain, and ML classifiers. In aspect phrase extraction from customer testimonies, the proposed solution outperforms both traditional approaches and state-of-the-art techniques, not only for aspect term extraction but also for other NLP issues (Kaur and Sharma, 2023a).

The increasing prevalence of text content from various sources, such as social media, messaging apps, and e-commerce websites, is important for company executives to understand the general public sentiments about their products and services. The project focuses on developing a consumer review summarization model

using NLP techniques and Long Short-Term Memory (LSTM) in order to deliver concise insights into customer behavior. Pre-processing, feature extraction, and sentiment classification are all combined in the proposed hybrid technique. The LSTM deep learning classifier is used for sentiment classification, while the hybrid approach generates unique feature vectors and the NLP approach eliminates superfluous input. The model was evaluated using average precision, recall, and F1-score on three datasets (Kaur and Sharma, 2023b).

The study highlighted the significance of online reviews for a company to assess its products and services. Still, the work at hand is sorting through the large number of reviews. The proposed Customer Review Summarization (CRS) methodology, also known as Hybrid Analysis of Sentiments (HAS), aims to provide organizations with a simplified understanding of consumer choices and behavior. HAS includes pre-processing, feature extraction, and review categorization. Hybrid feature extraction combines aspect-and review-related features, while NLP-based pre-processing removes unwanted data. For review classification, supervised classifiers like support vector machine, Naïve Bayes, and random forest are employed. With an impressive F1-score of 92.2%, the experimental results demonstrate that HAS outperforms the state-of-the-art techniques in sentiment analysis (Imron *et al.*, 2023).

The paragraph discusses the challenges posed by the different review formats-text, images, videos, and star ratings-that may be found on Bukalapak, a prominent Indonesian online marketplace. The manual analysis of these evaluations for potential clients takes time. The study acknowledges that entities exhibit a range of sentiments and offers aspect-based sentiment analysis as a potential remedy. The chosen approach combines CNN and LSTM, employing LSTM for aspect extraction and CNN for sentiment extraction. Prior research has demonstrated that BERT outperforms glove and word2vec as word embedding techniques. The LSTM-CNN method outperforms CNN or LSTM alone, achieving an impressive accuracy of 93.91%. The study emphasizes the impact of dataset distribution on model performance and observes that a larger dataset increases accuracy. Moreover, classification accuracy improves by 2.04% without stemming (Jiang *et al.*, 2023).

The paragraph introduces a unique technique for aspect-based sentiment analysis and emphasizes the importance of looking at grammatical relationships in a text. The proposed model called the Weighted Graph Attention Network aspect-based sentiment analysis model (WGAT), addresses limitations in existing models by leveraging dependency weighting in graph attention networks. By using pretraining, the model creates low-dimensional word vectors from the input text and analyses syntactic dependencies to create a dependency syntax graph. This graph is shown as a dependency-weighted

adjacency matrix to illustrate the relative importance of the various dependencies. The WGAT model uses graph attention networks to extract data and a classification layer to predict sentiment polarity. Test results over a wide range of datasets demonstrate that the WGAT model outperforms baseline approaches in terms of accuracy and F1 values, confirming its effectiveness in fine-grained aspect-based sentiment analysis tasks (Jorvekar *et al.*, 2023a).

The study is focused on the aspect-based sentiment classification approaches and the importance of text analysis, such as social media comments or product evaluations that examine interactions related to specific objects, such as restaurants or products. The conventional technique includes two approaches: Aspect extraction and sentiment evaluation. This study presents a framework for aspect-based sentiment categorization and recommender systems. This framework efficiently identifies items and produces high-accuracy classification using machine learning techniques like the random forest, naive bayes, decision trees, support vector machines, artificial neural networks, and a hybrid machine learning methodology. Experiments using real-time datasets validate the framework's capability to assist tourists in finding the best spots, places to stay, and restaurants in a given area. Compared to conventional classifiers, the hybrid machine learning methods outperformed the conventional classifiers (Jorvekar *et al.*, 2023b).

The authors discussed emotion type classification, sentiment magnitude detection, subjectivity detection, and polarity identification. Using a real-time dataset of customer reviews, the focus is on machine learning methods for aspect-based sentiment classification. Aspects are detected and the sentiment is assessed through the use of feature extraction and selection techniques and a range of machine learning classification methods, such as Artificial Neural Network (ANN), Random Forest (RF), Naïve Bayes (NB), and Support Vector Machine (SVM). TF-IDF, Bigram, and NLP features are a few feature extraction methods. In a comprehensive experimental study, SVM with NLP features outperforms other machine learning classifiers, demonstrating its effectiveness in aspect-based sentiment classification on custom (Hong *et al.*, 2022).

Various text classification techniques to improve classification efficiency and accuracy are applied in multiple studies. Some of the most significant methods observed in the literature are-Mutual Information Feature Selection (MIFS-C), Entropy-based Label Assignment (ELA), Adaptive Keyword Selection (AKS), Solid Feature Thesaurus (SFT) creation and Correlation-based Feature Selection (CMFS). Methods involved in the mentioned techniques are parameter selection, document transformation, feature selection criteria analysis, and statistical approaches to address challenges like data sparsity and semantic relationships between phrases. Strategies like bidirectional attention mechanisms, Graph

Neural Networks (GNNs), and fusion of statistical and semantic features are also applied using deep learning models. To enhance the classification accuracy role of Sentiment data incorporation, multi-dimensional document representation and aspect-based sentiment analysis are also highlighted in some studies.

Materials and Methods

Data Set Description

The tests made use of ten freely accessible and open-source datasets. The passage notably alludes to the examination of the Amazon review dataset, which comprises a range of data, including brand images, pricing, categories, and product descriptions with reference to views and purchases. Data is considered from 10-500 K records. Text, helpfulness ratings, and descriptions of product metadata, such as brand, price, and image features, are also included in the collection. The line highlights the range of information present in the dataset of Amazon reviews (Amazon Product Review, 2021) and suggests that it was used in the trials.

The dataset's nine auxiliary parts offer a variety of analysis choices for the text's many facets. To maintain confidentiality, all references to the real company mentioned in the reviews have been replaced with the word "retailer." Also, more information about other datasets. The section highlights the availability of a specific women's clothing e-commerce dataset as well as the efforts taken to anonymize the data to protect the shop's identity.

The data set was collected from multiple sources, which are available from open-source websites, hotel reviews (Datafiniti, 2019), the Nykaa dataset (Nykaa Product Review, 2020) considered from the data world website, and Amazon-related data sets like beauty products, clothing, office products, tools, and home improvements are considered from cseweb.ucsd.edu. All the raw data sets were downloaded in different file formats and were converted from a standard file format that is Comma-Separated Values (CSV) using Python. It also considered women's clothing (Women's Clothing Product Review., 2018), Emotional Intelligence boosters (Emotional Intelligence, 2022), and gym exercise (Gym Exercise Review, 2022) datasets from Kaggle. From the dataset considered, unstructured data is available in the review text of the dataset.

Figure 1 explains the stages of the workflow of the aspect-based Based Neural Attention Model (ABNAM) approach. Obtaining the data from numerous sources and eliminating English stop words, special symbols, excessive spaces, and duplicate sentences constitute the first stage of preprocessing. The second stage contains NLP with TF-IDF and N-Gram methods for word embedding, followed by the ABNAM processing

consisting of cluster creation for sixteen clusters and categorization of the input tokens into one of the clusters using weighted keywords. Stage 2 also tried the word2vec word embedding algorithm for creating tokens for implementation with the ABNAM model. ABNAM, which transforms unclassified utterances into classified sentences (Mao *et al.*, 2023), is created. Some of the NLP techniques employed in this step include developing vocabularies, word embeddings, and vector representations of the preprocessed data. In the third stage, the ABNAM model is evaluated using several statistical methods, such as accuracy, F1-score, ROC, and AUC curve. It is contrasted with other models currently in use to ascertain its efficacy. The final step involves testing the revised text using the ABNAM model to gauge the model's effectiveness. The results of TF-IDF and N-gram (max 2 grams) are considered as the inputs for machine learning models (Zhao *et al.*, 2022). The current work has used linear SVM, random forest, Naïve Bays, and Logistic regression for the classification.

Data Preprocessing

Data preparation is a crucial step in Natural Language Processing (NLP) for various reasons. One of the primary goals of data preparation is to reduce noise from the text data. In text data, noise can take the form of HTML tags, punctuation, special characters, and unnecessary symbols. Preprocessing helps to improve data quality by eliminating or cleaning up noise.

Describes the many steps involved in the research's data preprocessing stage. In the first step, stop words like "a," "the," "this," and "where," among others, are eliminated. The words in the phrases are then all changed to lowercase, even those that started out in the capital. Standardization ensures that the text is consistent. Word lemmatization is also used to turn the stop words into their root words. For efficient analysis and visualization, emphasize the significance of data pretreatment stages such as lowercase conversion, stop word removal, and word lemmatization.

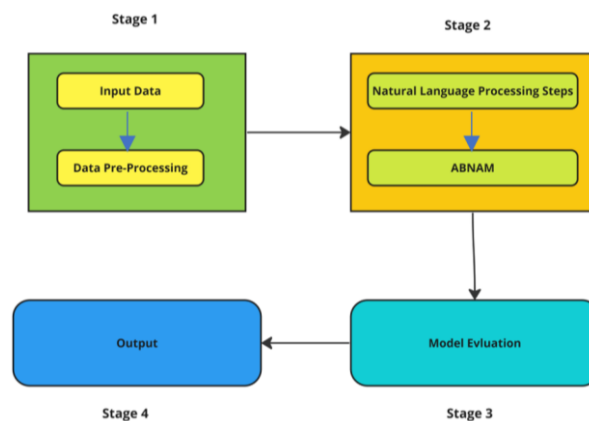


Fig. 1: Difference stages of ABNAM

Unusual terms are removed from the dataset in the second stage to prevent the creation of unnecessary vectors in word embeddings. The sentence offers information on locating and correcting spelling mistakes as well as considering keywords that are similar to the wrong term. In the third stage, word cases are normalized and sentences are punctuation-free. This normalization allows for word recognition regardless of the capitalization style used inside a sentence. The sentences are also free of emoticons, punctuation like "@," "#," and "\$," and other punctuation signs. Finally, words from many languages are transformed into a set of common characters and used in the specified sentences. The section describes key steps in data preprocessing, such as deleting rare terms, spelling corrections, punctuation removals, case normalization, removing special characters, etc.

The fourth step is the preparation of data. The goal of this step is to eliminate person names that have been changed to person keywords and to eliminate company names that have been changed to company keywords. Additionally, sentences with fewer than five keywords are disregarded. This step tries to secure sensitive personal and business data while removing sentences with insufficient keyword content. The paragraph stresses the value of data privacy and the need for a dataset with a minimal keyword threshold.

Machine Learning Techniques

Assessing how well a model performs while classifying text. Along with other high-performing machine learning models, the model is integrated with the Linear Support Vector Machine (LSVM), Multinomial Logistic Regression (MLR), multinomial Naive Bayes (NB), and Random Forest (RF). The goal is to evaluate the text categorization system's accuracy and contrast it with the performance of the ABNAM and ML systems. The sentence implies that, in terms of text categorization accuracy, the performance of the integrated models, LSVM, LR, NB, and RF, outperforms ABNAM and ML (Wang *et al.*, 2019). It also included TF-IDF, Word2Vec, and N-gram to improve the word vectorization of the corpus while implementing machine learning and ABNAM.

Linear support vector classifier: Since Linear Support Vector Machines (LSVMs) can be swiftly and easily applied to big datasets, their application in classification tasks has become more and more popular. Non-linearly separable problems, however, are challenging. Kernel-based SVMs are widely utilized for these kinds of scenarios; nevertheless, they have certain drawbacks regarding memory and computing efficiency when compared to their linear equivalents. More precisely, actual data indicates that their response is only provided as a function of the support vector set, which rises linearly in the size of the training set (Sperandei, 2014).

Multinomial logistic regression: Calculating odds ratios is the aim of logistic regression when multiple explanatory factors are involved. The process is quite similar to multiple linear regression, with the distinction that the response variable is binomial. The outcome displays how each variable affected the observed interest event's odds ratio. This approach is particularly useful in minimizing confusing effects because it simultaneously looks at the relationship between all variables (Rish, 2001). This article aims to simplify the logistic regression method through the use of informative examples:

$$Y = \frac{e^{(b_0 + b_1x)}}{1 + e^{(b_0 + b_1x)}} \quad (1)$$

Equation 1 describes that x stands for the input value. y represents the predicted output and b_0 represents the intercept or bias term. b_1 represents the coefficient for input (x). Where e is the base of the natural logarithm (about 2.718) a and b are the model's parameters and y is the probability of a 1 (the proportion of 1s, the mean of Y). When X is zero, the e value produces y and the b weight determines how rapidly the probability changes with a unit change in X .

Multinomial Naive Bayes: The effect of features is evaluated using naive Bayes performance. Use Monte Carlo simulations to perform a systematic analysis of the categorization accuracy for different classes of randomly produced issues. The results reveal the relationship between distribution entropy and classification error and they also indicate that low-entropy feature distributions yield the best performance for naive bayes type models. Surprisingly, it shows that naive bayes works effectively when some feature dependencies are almost functional. It performs best in two different scenarios: Totally independent qualities (as expected) and functionally dependent features (unexpected). Remarkably, class-conditional mutual information across features indicates the degree of feature dependencies, yet naive Bayes's accuracy does not correlate with it. A more accurate measure of Naive Bayes accuracy is the amount of class information lost due to the independence assumption (Kulkarni and Sinha, 2013).

Based on the prior probabilities $P(c)$, $P(x)$, and $P(x|c)$, Bayes' theorem provides a method for calculating the posterior probability, $P(c|x)$. Under the premise that the influence of a predictor's (x) value on a given class (c) is independent of other predictors' values, the Naive Bayes classifier functions. The term "class conditional independence" is frequently used to describe this premise:

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \quad (2)$$

$$P(c|X) = P(x1|c) \times P(x2|c) \times \dots \times P(xn|c) \times P(c) \quad (3)$$

Equations 2-3 describe that the posterior probability of class (target) given predictor (attribute) is represented by " $P(c|x)$." The prior probability of the class is " $P(c)$." The likelihood, or probability of the predictor given class, is expressed as " $P(x|c)$." The predictor's prior probability is denoted by " $P(x)$."

Random forest: Data mining is one area where random forest, an ensemble supervised machine learning technique, is used. Random Forest has the potential to be a top classifier in the future, with performance on par with ensemble methods like bagging and boosting. This study details the taxonomy built inside this framework for the random forest classifier. Furthermore, generate a comparative diagram comparing different random forest classifiers according to relevant parameters. Dynamically pruning a forest and finding the ideal subset of the forest can lead to improvements in performance. Moreover, semi-supervised learning and the domains of imbalanced data classification and stream data present chances for innovation. Building on these discoveries, the paper's conclusion makes numerous recommendations for future random forest classifier research (Chen *et al.*, 2022a; Jalal *et al.*, 2022; Tan *et al.*, 2022).

Term Frequency-Invert Document Frequency (TF-IDF): To calculate the relevance measure of a word in the series to the text. The frequency of a word is proportional to the total count of words that appear in the text. The IDF can be defined as the total count of documents in the dataset with distinct frequencies. Words in a document corpus are evaluated for possible query appropriateness using Term Frequency-Inverse Document Frequency (TF-IDF). As the name implies, TF-IDF calculates values for every word in a document by squaring the frequency of the word in that particular document against the proportion of documents that contain the word. Words with high TF-IDF values strongly correlate with the document they appear in:

$$tf(t, d) = \frac{n(t, d)}{\text{number of terms in the document}} \quad (4)$$

Equation 4 describes the most common Term Frequency (TF). The numerator, n , represents the number of times the phrase " t " appears in the document " d ." It measures how frequently a term, t , appears in a document, d . Every document and word would, therefore, have a unique TF value.

Word2Vec: Using distributed numerical representations of words' features, Word2Vec creates vectors to represent words. These characteristics include words that express the subtleties of context that certain words in the lexicon have. Because of the vectors they generate, word embeddings ultimately play a critical part

in forming the semantic linkages between a word and others with comparable meanings. It is a widely preferred NLP method that allows word representation in the form of word vectors. It helps to observe the semantic relationships among the words present in the vector by mapping the words to the multi-dimensional vectors. The Word2Vec works on the principle that similar meanings should be represented in a similar vector format.

The Architecture of ABNAM

Figure 2 the architecture consists of multiple processes, beginning with the data cleaning technique, which consists of two stages. The first stage removes special characters, names, sentences with less than 5 keywords, extra spaces, stop words, numbers, emojis, URLs, duplicate sentences, and company names. At the same layer, the next stage performs lemmatization, lowercase on words, spelling correction, rare word removal, tokenization, case normalization, word standardizing, performing stemming, removing Unicode, typo correction, and removing emoticons. The next step is to apply the TF-IDF, N-gram, and matrix formation for the preprocessed data. The final stage of the work takes the processed word embeddings as input, performs the data with a Convolutional Neural Network (CNN), followed by the attention layer, merging the layers, and finally, the output layer that defines the class for the input sentence (Chen *et al.*, 2022c; Ayetiran, 2022; Cruz *et al.*, 2014).

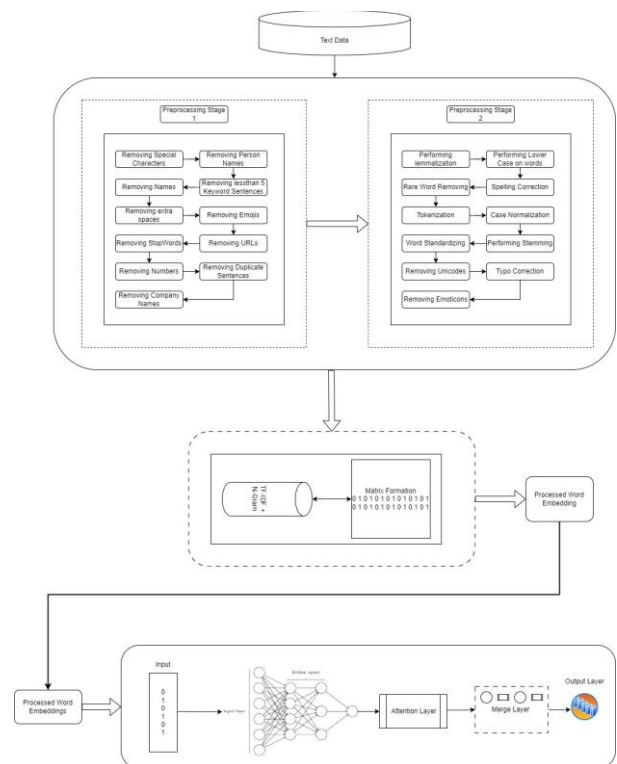


Fig. 2: Architecture of ABNAM model

The unprocessed data is gathered from various e-commerce websites, including those for restaurants, clothing companies, cosmetics and beauty products, home goods, and retail establishments like Nykaa and Amazon.

The stage before text data is processed. In this stage, various text cleaning techniques are used, including stemming and lemmatization to break down words into their simplest forms, the elimination of uncommon words, the removal of Unicode and emoticons, the conversion of words to lowercase, the correction of spelling errors, the use of case normalization and word standardization and the correction of keyword typos. Following preprocessing, word vectorization becomes the main focus, which turns the text into word embeddings. The Aspect Based Neural Attention (ABNAM) Model was developed using word embeddings. The significance of cleaning and converting the text data is emphasized in this paragraph in order to ensure that words are accurately represented in the ensuing modeling procedure.

The Aspect Based Neural Attention Model (ABNAM) uses word embeddings as input. A group of keywords within the model called "aspects" is created using the word embeddings. On the basis of their relatedness, these elements are then grouped into clusters. The weights of each keyword within each cluster serve as a representation of that keyword. With each cluster belonging to a different category, attention is centered on forming discrete clusters of aspects. With the use of this method, the keywords in the ABNAM model can be efficiently represented and organized, leading to improved text data interpretation and categorization (Brody and Elhadad, 2010; Ahmed *et al.*, 2022; Xia and Chen, 2022; Zhang *et al.*, 2022; Chen *et al.*, 2022b).

ABNAM Model

A set of aspect embeddings' intended purpose. According to this explanation, each vocabulary word can be connected to a feature vector by using the words that are closest to it in the embedding space. In a word embedding matrix, these feature word vectors stand in for the text rows. While fewer aspects have been defined than the total vocabulary, both words and aspects are contained in the same area. A neural attention mechanism is used to filter the vocabulary's aspect text words and these words have a strong relationship with the aspect embeddings. With the help of this method, text data may be effectively analyzed and understood because aspects can be represented in the embedding space.

The aspect-based neural attention layer. The layer lists the word indexes in a review sentence and seeks to produce a sentence embedding by applying weighted

word embeddings. After that, non-aspect words are given less weight and eliminated from the sentence embedding using an attention strategy. The sentence embedding is rebuilt by mixing aspect embeddings from a collection of aspect embeddings. The Aspect-Based Neural Attention Model (ABNAM) seeks to moderately modify the sentence embeddings of filtered sentences while retaining the majority of the data of aspect words. This strategy helps the model to concentrate on the important elements and enhance the representation of the input sentences (Hu and Liu, 2004; Ameer *et al.*, 2023; Wu, 2023; He *et al.*, 2018; Muennighoff *et al.*, 2022).

Here, T is an aspect embedding matrix, R_s is filtered reconstructions and P_t is a weight vector calculated using softmax; for the values of k , k is smaller than V and can be defined as no of aspects.

Figure 3 introduces this section's Aspect-Based Neural Attention Model (ABNAM) paradigm. The main goal is to obtain a collection of aspect embeddings, which can all be understood by looking at the words that are closest to the set (i.e., representative words) in the embedding space. To start this procedure, assign a feature vector " ew " in Rd to every word " w " in the vocabulary. Since word embeddings are specifically made to map frequently occurring words in a context to proximal places in the embedding space, using them for these feature vectors is essential. The vocabulary size " V " is represented by the feature vectors associated with the words, which correspond to the rows of a word embedding matrix " E " in " $RV \times d$." Aim to create embeddings for elements with embedding spaces similar to those of words.

An aspect embedding matrix " T " in " $RK \times d$ " is required to achieve this, where " K " is the number of declared aspects, a value much less than " V ," the vocabulary size. An attention mechanism that sifts through the aspect words aids in selecting these aspect embeddings, which are used to approximate the aspect words in the vocabulary. The Aspect-Based Neural Attention Model (ABNAM) uses two consecutive processes, in Fig. 3, for each input sample, which is effectively a list of word indexes within a review phrase. Using an attention mechanism, non-aspect words are filtered away by giving them lower weights. As a result, a sentence embedding called " zs " is produced using these weighted word embeddings. Next, try to rebuild this sentence embedding by expressing it as a linear combination of aspect embeddings from matrix " T ." Dimension reduction and reconstruction are involved in this process, where ABAE attempts to convert sentence embeddings of the filtered sentences (" zs ") into their reconstructions (" rs ") with the least amount of distortion. This strategy seeks to maintain a significant quantity of data on the aspect words present in the " K " embedded aspects.

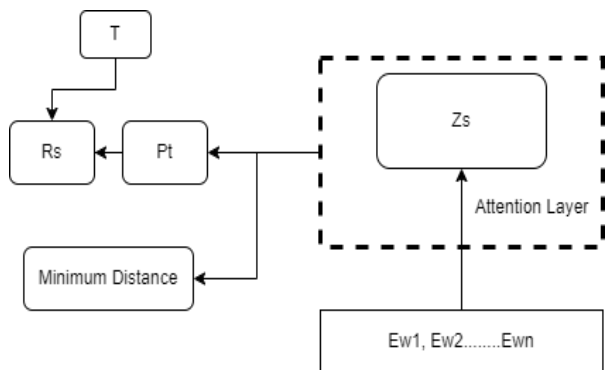


Fig. 3: ABNAM model

Results and Discussion

The result Table 1 highlights the excellent level of accuracy reached by the Aspect-Based Neural Attention Model (ABNAM) when used in conjunction with TF-IDF and N-Gram techniques. In 16 different categories, including price, staff, atmosphere, food, beauty products, Restaurant, makeup, food price, home improvements, clothing, personal care, health and wellness, skin, perfume, home goods, and hair, the model had an accuracy rate of 97%. The accuracy of ABNAM was compared to that of other well-known models, such as support vector machines, random forests, logistic regression, and Naive Bayes classifiers. The research results demonstrated that ABNAM with word2vec achieved 92% accuracy compared to Multinomial Naive Bayes classifiers. Using an ensemble technique by ABNAM with word2vec falls under the machine-learning realm. The ABNAM model demonstrated effective efficiency in sentence classification overall.

For the comparative analysis, we have performed the F1-score, Area Under the Curve (AUC), and Receiver Operating Characteristic curve (ROC) analysis for each of the 16 classes. The resulting graphs for AUC and ROC are presented in Fig. 4.

Figure 5 compares the performance of various machine learning models in text classification tasks. These models include ABNAM, LSVC, LR, RF, and NMB with TF-IDF and N-gram. The comparison also involves BERT and GTR-XL models. These models were evaluated based on the results from two papers: "Massive Text Embedding Benchmark" (MTEB): (Wahba *et al.*, 2022) and "A Comparison of SVM against Pre-trained Language Models (PLMs) for Text Classification tasks." (Datafiniti, 2018) The papers reported an accuracy of 86 and 85%, respectively.

Table 1: Model result comparison

Method	Accuracy in %
ABNAM + TF-IDF and N-gram	97
ABNAM + Word2Vec	92
L-support vector machine (TF-IDF + N-gram)	89
M- logistic regression (TF-IDF + N-gram)	90
Random forest (TF-IDF + N-gram)	81
Multinomial Naïve Bayes (TF-IDF + N-gram)	91

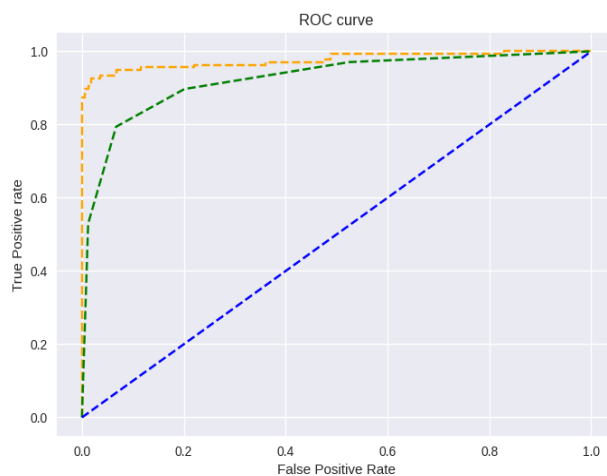


Fig. 4: Performance Evaluation of ABNAM model

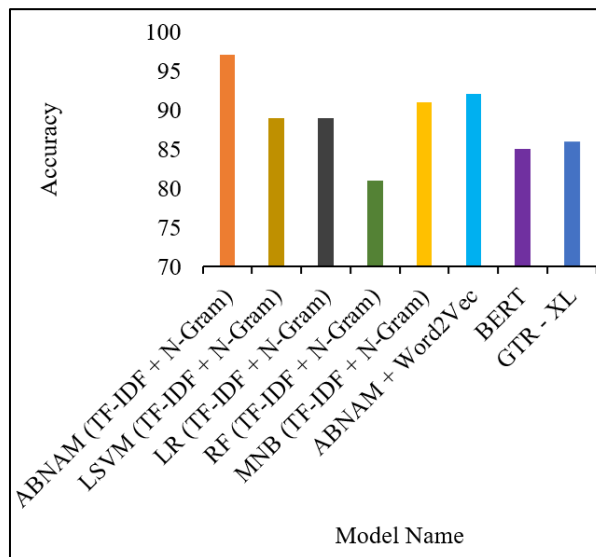


Fig. 5: Performance evaluation chart ABANM with other ML and hybrid models

Conclusion

Numerous text classification techniques have been explored for enhancing text classification accuracy using Mutual Information Feature Selection (MIFS-C) and Correlation-based Feature Selection (CMFS) approaches, which are significantly useful in addressing various challenges such as data sparsity and semantic

relationships. The implementation of sentiment data and multi-dimensional document representation plays an important role in NLP, aspect-based classification, and accuracy enhancement. The Aspect-Based Neural Attention Model (ABNAM) is a powerful aspect extraction model. ABNAM is superior to TF-IDF and N-gram models combined with ABNAM for capturing word co-occurrence patterns and overcoming data sparsity problems in review datasets. Compared to earlier techniques, experimental findings indicate that ABNAM effectively learns and collects more quality traits and review elements. To extract aspects, the authors advise employing unsupervised neural networks. In some situations, ABNAM might be used as an alternative to other approaches because of its straightforward and scalable architecture. Overall, ABNAM is an easy-to-use and robust method for aspect extraction in text analysis.

Acknowledgment

I thank almighty god and Christ University for giving me the opportunity to pursue my research work without any difficulties. Thank all my family members for the wonderful support.

Funding Information

The research work is self-funded and has not received any financial support for this manuscript's investigation, composition, and/or publication.

Author's Contributions

Nagendra N: Collected and curated the data. Developed the initial idea. Implemented the research. Authored the complete paper. Conducted thorough reviews. Created content for the paper.

Chandra J: Involved in the draft and implementation of the result analysis. Implemented the research. Conducted thorough reviews. Created content for the paper. Contributed additional ideas to the paper reviews.

Ethics

The research accurately represents the authors' findings, cites all references, acknowledges coauthors' contributions, and places the findings in a research context.

References

Ahmed, U., Jhaveri, R. H., Srivastava, G., & Lin, J. C. W. (2022). Explainable deep attention active learning for sentimental analytics of mental disorder. *Transactions on Asian and Low-Resource Language Information Processing*.
<https://doi.org/10.1145/3551890>

- Amazon Product Review. (2021). Large crawl of product reviews from Amazon.
https://cseweb.ucsd.edu/~jmcauley/datasets.html#amazon_reviews
- Ameer, I., Bölücü, N., Siddiqui, M. H. F., Can, B., Sidorov, G., & Gelbukh, A. (2023). Multi-label emotion classification in texts using transfer learning. *Expert Systems with Applications*, 213, 118534. <https://doi.org/10.1016/j.eswa.2022.118534>
- Ayetiran, E. F. (2022). Attention-based aspect sentiment classification using enhanced learning through CNN-BiLSTM networks. *Knowledge-Based Systems*, 252, 109409.
<https://doi.org/10.1016/j.knosys.2022.109409>
- Bakus, J., & Kamel, M. S. (2006). Higher order feature selection for text classification. *Knowledge and Information Systems*, 9, 468-491.
<https://doi.org/10.1007/s10115-005-0209-6>
- Beausoleil, K. (2019). The Nature of Decision-Making: Human Behavior vs. Machine Learning.
<https://doi.org/10.29173/aar49>
- Brody, S., & Elhadad, N. (2010, June). An unsupervised aspect-sentiment model for online reviews. In *Human language technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 804-812).
<https://aclanthology.org/N10-1122.pdf>
- Chen, H., Wu, L., Chen, J., Lu, W., & Ding, J. (2022a). A comparative study of automated legal text classification using random forests and deep learning. *Information Processing and Management*, 59(2), 102798.
<https://doi.org/10.1016/j.ipm.2021.102798>
- Chen, L. M., Xiu, B. X., & Ding, Z. Y. (2022b). Multiple weak supervision for short text classification. *Applied Intelligence*, 52(8), 9101-9116.
<https://doi.org/10.1007/s10489-021-02958-3>
- Chen, P. F., Chen, K. C., Liao, W. C., Lai, F., He, T. L., Lin, S. C., ... & Hung, F. M. (2022c). Automatic International Classification of Diseases coding system: Deep contextualized language model with rule-based approaches. *JMIR Medical Informatics*, 10(6), e37557.
<https://doi.org/10.2196/37557>
- Chen, W., Yan, J., Zhang, B., Chen, Z., & Yang, Q. (2007). Document transformation for multi-label feature selection in text categorization. In *Seventh IEEE International Conference on Data Mining (ICDM 2007)* (pp. 451-456). IEEE.
<https://doi.org/10.1109/ICDM.2007.18>
- Cruz, I., Gelbukh, A. F., & Sidorov, G. (2014). Implicit Aspect Indicator Extraction for Aspect-based Opinion Mining. *Int. J. Comput. Linguistics Appl.*, 5(2), 135-152.
<https://text2fa.ir/wp-content/uploads/Text2fa.ir-Implicit-Aspect-Indicator-Extraction-1.pdf>

- Datafiniti. (2018). Hotels review from Datafiniti's Business Database.
- Datafiniti. (2019). Hotels review from Datafiniti's Business Database.
- Emotional Intelligence. (2022). CC BY-SA 4.0, Emotional Intelligence Booster | Mistral-7B.
- Gasparetto, A., Marcuzzo, M., Zangari, A., & Albarelli, A. (2022). A survey on text classification algorithms: From text to predictions. *Information*, 13(2), 83. <https://doi.org/10.3390/info13020083>
- Gym Exercise Review. (2022). CC0: Public Domain, Gym Exercise Dataset, dataset created for analysing and evaluating workouts that one can do at the gym (or at home) to stay healthy. <https://www.kaggle.com/datasets/niharika41298/gym-exercise-data>
- He, R., Lee, W. S., Ng, H. T., & Dahlmeier, D. (2018). Exploiting document knowledge for aspect-level sentiment classification. *Arxiv Preprint Arxiv 1806.04346*. <https://doi.org/10.18653/v1/p18-2092>
- Hong, S., Kim, J., & Yang, E. (2022). Automated text classification of maintenance data of higher education buildings using text mining and machine learning techniques. *Journal of Architectural Engineering*, 28(1), 04021045. [https://doi.org/10.1061/\(asce\)ae.1943-5568.0000522](https://doi.org/10.1061/(asce)ae.1943-5568.0000522)
- Hu, M., & Liu, B. (2004, August). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 168-177). <https://doi.org/10.1145/1014052.1014073>
- Huan, J. L., Sekh, A. A., Quek, C., & Prasad, D. K. (2022). Emotionally charged text classification with deep learning and sentiment semantic. *Neural Computing and Applications*, 1-11. <https://doi.org/10.1007/s00521-021-06542-1>
- Huang, J., Cui, Y., & Wang, S. (2023). Adaptive Local Context and Syntactic Feature Modeling for Aspect-Based Sentiment Analysis. *Applied Sciences*, 13(1), 603. <https://doi.org/10.3390/app13010603>
- Imron, S., Setiawan, E. I., Santoso, J., & Purnomo, M. H. (2023). Aspect Based Sentiment Analysis Marketplace Product Reviews Using BERT, LSTM and CNN. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 7(3), 586-591. <https://doi.org/10.29207/resti.v7i3.4751>
- Jalal, N., Mehmood, A., Choi, G. S., & Ashraf, I. (2022). A novel improved random forest for text classification using feature ranking and optimal number of trees. *Journal of King Saud University-Computer and Information Sciences*, 34(6), 2733-2742. <https://doi.org/10.1016/j.jksuci.2022.03.012>
- Jiang, T., Wang, Z., Yang, M., & Li, C. (2023). Aspect-Based Sentiment Analysis with Dependency Relation Weighted Graph Attention. *Information*, 14(3), 185. <https://doi.org/10.3390/info14030185>
- Jorvekar, G. N., Arjariya, T., & Gangwar, M. (2023a). ABSC-HMLT: Aspect Based Sentiment Classification Using Hybrid Machine Learning Techniques. *International Journal of Intelligent Systems and Applications in Engineering*, 11(5s), 574-584. <https://ijisae.org/index.php/IJISAE/article/view/2818>
- Jorvekar, G. N., Arjariya, T., & Gangwar, M. (2023b). Hybrid Feature Selection Techniques for Aspect based Sentiment Classification using Supervised Machine Learning. *International Journal of Intelligent Systems and Applications in Engineering*, 11(2s), 211-217. <https://www.ijisae.org/index.php/IJISAE/article/view/2619>
- Kaur, G., & Sharma, A. (2023a). A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis. *Journal of Big Data*, 10(1), 5. <https://doi.org/10.1186/s40537-022-00680-6>
- Kaur, G., & Sharma, A. (2023b). Has: Hybrid analysis of sentiments for the perspective of customer review summarization. *Journal of Ambient Intelligence and Humanized Computing*, 14(9), 11971-11984. <https://doi.org/10.1007/s12652-022-03748-6>
- Khurana, D., Koli, A., Khatter, K., & Singh, S. (2023). Natural language processing: State of the art, current trends and challenges. *Multimedia Tools and Applications*, 82(3), 3713-3744. <https://doi.org/10.1007/s11042-022-13428-4>
- Koksal, D., Alacan, M. M., Olgun, E., & Sakar, C. O. (2022, May). Natural Language Processing-Based Product Category Classification Model for e-commerce. In *2022 30th Signal Processing and Communications Applications Conference (SIU)* (1-4). IEEE. <https://doi.org/10.1109/SIU55565.2022>
- Kulkarni, V. Y., & Sinha, P. K. (2013). Random forest classifiers: A survey and future research directions. *Int. J. Adv. Comput*, 36(1), 1144-1153. ISSN: 2051-0845.
- Liu, Z., & Yang, J. (2012, August). An improved ambiguity measure feature selection for text categorization. In *2012 4th International Conference on Intelligent Human-Machine Systems and Cybernetics (Vol. 1, pp. 220-223)*. IEEE. <https://doi.org/10.1109/IHMSC.2012.62>
- Luo, X., Yu, Z., Zhao, Z., Zhao, W., & Wang, J. H. (2022). Effective short text classification via the fusion of hybrid features for IoT social data. *Digital Communications and Networks*, 8(6), 942-954. <https://doi.org/10.1016/j.dcan.2022.09.015>
- Mao, Y., Liu, S., & Gong, D. (2023). A Text Mining and Ensemble Learning Based Approach for Credit Risk Prediction. *Tehnički Vjesnik*, 30(1), 138-147. <https://doi.org/10.17559/TV-20220623113041>

- Meng, J., Lin, H., & Yu, Y. (2011). A two-stage feature selection method for text categorization. *Computers and Mathematics with Applications*, 62(7), 2793-2800. <https://doi.org/10.1016/j.camwa.2011.07.045>
- Muennighoff, N., Tazi, N., Magne, L., & Reimers, N. (2022). MTEB: Massive text embedding benchmark. *Arxiv Preprint Arxiv:2210.07316*. <https://10.18653/v1/2023.eacl-main.148>
- Mujawar, S. S., & Bhaladhare, P. R. (2023). Effective feature selection methods for user sentiment analysis using machine learning. *International Journal on Recent and Innovation Trends in Computing and Communication*, 11, 37-45. <https://doi.org/10.17762/ijritcc.v11i3s.6153>
- Nykaa Product Review. (2020). E-commerce dataset from Nykaa database. <https://data.world/promptcloud/nykaa-product-review>
- Rish, I. (2001, August). An empirical study of the naive Bayes classifier. In *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence* (Vol. 3, No. 22, pp. 41-46). <https://doi.org/10.1039/b104835j>
- Rohidin, D., Samsudin, N. A., & Deris, M. M. (2022). Association rules of fuzzy soft set based classification for text classification problem. *Journal of King Saud University-Computer and Information Sciences*, 34(3), 801-812. <https://doi.org/10.1016/j.jksuci.2020.03.014>
- Soni, S., Chouhan, S. S., & Rathore, S. S. (2023). TextConvoNet: A convolutional neural network based architecture for text classification. *Applied Intelligence*, 53(11), 14249-14268. <https://doi.org/10.1007/s10489-022-04221-9>
- Sperandei, S. (2014). Understanding logistic regression analysis. *Biochemia Medica*, 24(1), 12-18. <https://doi.org/10.11613/BM.2014.003>
- Tan, T. P., Lim, C. K., & Rahman, W. R. E. A. (2022). Sliding Window and Parallel LSTM with Attention and CNN for Sentence Alignment on Low-Resource Languages. *Pertanika Journal of Science and Technology*, 30(1). <https://doi.org/10.47836/pjst.30.1.06>
- Tasci, S., & Gungor, T. (2008, October). An evaluation of existing and new feature selection metrics in text categorization. In *2008 23rd International Symposium on Computer and Information Sciences* (pp. 1-6). IEEE. <https://doi.org/10.1109/ISCIS.2008.4717900B>
- Wang, B. K., Huang, Y. F., Yang, W. X., & Li, X. (2012). Short text classification based on strong feature thesaurus. *Journal of Zhejiang University Science C*, 13(9), 649-659. <https://doi.org/10.1631/jzus.C1100373>
- Wahba, Y., Madhavji, N., & Steinbacher, J. (2022, September). A comparison of svm against pre-trained language models (plms) for text classification tasks. In *International Conference on Machine Learning, Optimization and Data Science* (pp. 304-313). Cham: Springer Nature Switzerland. https://10.1007/978-3-031-25891-6_23
- Wang, H., & Li, F. (2022). A text classification method based on LSTM and graph attention network. *Connection Science*, 34(1), 2466-2480. <https://doi.org/10.1080/09540091.2022.2128047>
- Wang, X., Yang, Z., Chen, X., & Liu, W. (2019). Distributed inference for linear support vector machine. *Journal of Machine Learning Research*, 20. <https://www.jmlr.org/papers/volume20/18-801/18-801.pdf>
- Women's Clothing Product Review. (2018). CC0: Public Domain. Women's e-commerce Clothing Reviews. <https://www.kaggle.com/datasets/nicapotato/women-s-e-commerce-clothing-reviews>
- Wu, M. (2023). Commonsense knowledge powered heterogeneous graph attention networks for semi-supervised short text classification. *Expert Systems with Applications*, 120800. <https://doi.org/10.1016/j.eswa.2023.120800>
- Xia, T., & Chen, X. (2022). Category-learning attention mechanism for short text filtering. *Neurocomputing*, 510, 15-23. <https://doi.org/10.1016/j.neucom.2022.08.076>
- Xu, F., Pan, Z., & Xia, R. (2020). E-commerce product review sentiment classification based on a naïve Bayes continuous learning framework. *Information Processing and Management*, 57(5), 102221. <https://doi.org/10.1016/j.ipm.2020.102221>
- Zhang, J., Cui, Z., Park, H. J., & Noh, G. (2022). BHGAttN: A Feature-Enhanced Hierarchical Graph Attention Network for Sentiment Analysis. *Entropy*, 24(11), 1691. <https://doi.org/10.3390/e24111691>
- Zhao, H., Xie, J., & Wang, H. (2022). Graph convolutional network based on multi-head pooling for short text classification. *IEEE Access*, 10, 11947-11956. <https://doi.org/10.1109/ACCESS.2022.3146303>
- Zhao, P., Hou, L., & Wu, O. (2020). Modeling sentiment dependencies with graph convolutional networks for aspect-level sentiment classification. *Knowledge-Based Systems*, 193, 105443. <https://doi.org/10.1016/j.knosys.2019.105443>