ISSN: 2229-6956 (ONLINE)
DOI: 10.21917/ijsc.2025.0536

COMPARATIVE ANALYSIS OF OXYMORONIC SENTENCES CLASSIFICATION USING MULTIPLE FEATURES WITH MACHINE LEARNING ALGORITHMS

K. Seethappan[1] and K. Premalatha[2]
[1]Department of Computer Science and Engineering, Anna University Regional Campus, Coimbatore, India
[2] Department of Computer Science and Engineering, Bannari Amman Institute of Technology, India

Abstract

The effectiveness of oxymorons in social media hinges on their context, target audience, and usage frequency. Using oxymorons can help clarify nuanced semantic variations or highlight inherent conflicts. The primary objective of this research endeavor is to develop a state-of-the-art oxymoron classifier. To achieve this, a comprehensive feature extraction process was undertaken, encompassing N-grams, part-of-speech tags, and structural features from a meticulously balanced dataset. These extracted features were then integrated with various feature weighting schemes and evaluated using a suite of machine learning algorithms, including Random Forest (RF), Naïve Bayes (NB), Support Vector Machine (SVM), and K-Nearest Neighbors (KNN). The proposed KNN algorithm, when used in conjunction with all features and the TF-IDF weighting scheme, demonstrated superior classification performance, achieving precision, recall, F-measure, accuracy, and kappa scores of 99.38%, 99.64%, 99.51%, 99.50%, and 0.99, respectively. These results demonstrate the superior performance of the KNN classifier in the context of oxymoron classification. Future advancements in this research will focus on predicting oxymoronic phrases within mixed-language environments.

Keywords:

Machine learning, Sentiment Analysis, Natural Language Processing, Figurative Language, Oxymoron

1. INTRODUCTION

An oxymoron, a figure of speech, juxtaposes contradictory terms to create a paradoxical effect. These intriguing combinations are prevalent in literature, advertising, and everyday conversation. Oxymorons are a common literary device, employed in plays, novels, and poems to underscore themes, enrich character development, or heighten dramatic impact. In advertising, they are often used to craft memorable and attention-grabbing slogans. For instance, the slogan "Act Naturally" for a health food store exemplifies this, cleverly combining opposing ideas to convey a unique message. The online Cambridge English Dictionary defines an oxymoron as "two words or phrases used together that have or seem to have opposite meanings." Oxymorons serve several significant purposes. In informal contexts, they inject wit and humor into language and conversation. Moreover, by challenging assumptions and employing ambiguous language, they encourage deeper audience engagement and critical thinking. Ultimately, oxymorons compel listeners to engage in more nuanced reasoning and contemplation, revealing deeper, more complex layers of thought.

In social media, oxymorons can be effective marketing tools by grabbing attention, increasing engagement, and enhancing brand recall. However, their use requires careful consideration. Unfamiliar audiences may misunderstand them; Overuse can lead to clichés, and some oxymorons may even be potentially offensive. Two oxymoronic sentences are listed below for better understanding.

- The affair with the boss was an open secret that nobody dared discuss in the office, but everyone knew about it.
- He felt like a living dead after the accident, moving through life aimlessly.

Pleonasm is a rhetorical device characterized by the use of superfluous words or phrases. While it can be employed for stylistic emphasis, it generally adds no meaningful information to a statement. In everyday communication, pleonasms are often considered redundant and are typically discouraged to improve clarity and conciseness. The online Cambridge English Dictionary defines pleonasm as "the use of more words than are needed to express a meaning, done either unintentionally or for emphasis: redundant." While pleonasm is often viewed as a stylistic flaw, it can also serve a valuable purpose. Redundant words can occasionally enhance communication by providing emphasis, clarifying meaning, or adding a pleasing auditory and rhythmic quality to a phrase. Two pleonastic sentences are listed below for better understanding.

- The small tiny ant crawled across the ground.
- The baby is an adorable cute little thing.

Oxymorons and pleonasms are pervasive in modern communication. While scholars from various fields, including linguistics, anthropology, and sociology, have acknowledged their significance in real-world interactions, there remains a significant gap in research. To the best of our knowledge, no studies have yet employed machine learning algorithms to automatically categorize sentences containing oxymorons or pleonasms within the context of social media data. The primary objectives of this research endeavor are multifaceted. First, the aim is to automate the classification of oxymoronic and pleonastic sentences within a balanced dataset by leveraging the power of machine learning algorithms. Secondly, this research will carefully construct a complete dataset of diverse oxymorons and pleonasms, along with their sentence contexts. Third, this research will explore the distinctive features that effectively capture the essence of oxymorons and pleonasms within textual data. Finally, this research aspires to develop an optimized model that demonstrates superior performance in identifying oxymoron and pleonasm within the dataset.

The remainder of this work is structured as follows: Part 2 provides an overview of relevant prior research, including studies on oxymorons, feature extraction techniques, feature weighting schemes, and figurative language detection within the context of sentiment analysis. Additionally, this section discusses various machine learning classification algorithms. Part 3 delves into the methodologies employed in this research, covering dataset collection, feature extraction, feature vector weighting schemes, and the specific machine learning classification algorithms utilized. Part 4 presents and analyzes the evaluation results obtained from this research. Finally, Part 5 offers concluding remarks and outlines potential avenues for future research.

2. RELATED WORKS

Karp [1] analyzed juxtapositions in George Orwell's Animal Farm, examining their lexical, syntactic, and semantic contexts. The study also explored the similarities and differences between Ukrainian and English oxymorons, paradoxes, and antitheses. Employing deductive and taxonomic methodologies, the research investigated the occurrence of these figures of speech in Animal Farm. Contrastive analysis was utilized to identify and compare recognized speech figures in English and Ukrainian. Furthermore, the study revealed several structural models within oxymorons, including attribute, verbal, noun, and adverb pairings, as well as free syntactic patterns. By extracting antonymous pairs from Italian corpora, Pietra & Masini [2] identified a set of oxymorons. Their analysis, based on nine syntactic components and the identified antonymous pairs, resulted in the categorization of 376 oxymorons. This research underscored the crucial connection between contextual oxymorons and the detection of humor, irony, and sarcasm.

Safarovna [3] focused on the origins of the oxymoron in Uzbek literature. This study offered concrete examples of oxymorons in literary works and encouraged students to explore the effective use of contrast in their writing. While oxymorons and antonyms share certain lexical-semantic features, their purpose, creation, and function can differ significantly depending on the specific type of speech. The use of the term "contrast" often provided insights into these varying aspects. Tsao et al. [4] investigated the structural elements of oxymorons and developed a novel method for creating them. To understand the contradictory image components of these creative oxymorons, they conducted principal component analyses on data collected through contradictory-image testing, assessing overall usage perceptions, operating methods, and product feedback. By identifying the properties associated with different parts of speech, they were able to develop models for converting these insights into design applications.

Bageshwar [5] investigated the use of literary devices such as oxymorons, puns, repetition, wordplay, and metaphors to achieve specific effects, including emphasis, persuasion, and emotional impact, in Instagram postings. This research demonstrated how writers can effectively utilize these strategies to craft compelling and memorable messages that resonate with their audience. The study categorized Instagram postings based on their presentational methods and highlighted the potential for powerful reader engagement and emotional response.

Fazlitdinovna [6] explored the multifaceted uses and meanings of oxymorons in speech, drawing upon the wisdom of proverbs. The study delved into the significant impact of oxymorons within higher education systems, examining their role in enhancing teaching and learning processes. To gain a thorough understanding of oxymorons, the research analyzed examples from Uzbek and English literature and poetry. Furthermore, the study highlighted the increasing prevalence of oxymorons in contemporary advertising. Sakaeva & Kornilova [7] conducted a comprehensive study of oxymorons within Shakespeare's sonnets, classifying them into various structural categories. These categories included Adjective + Noun, Adjective + Adjective, Verb + Article + Noun, Verb + Adverb, Noun + Article + Noun, Adjective + Noun + Adjective, and a free type category. This analysis encompassed both Russian and English translations of the sonnets. The study delved into the lexical and semantic properties of these oxymorons. Notably, the authors argued that an excessive concentration of stylistic devices, specifically oxymorons, within a sonnet could potentially disrupt its aesthetic balance and diminish its overall artistic quality.

Alwana [8] conducted a comprehensive analysis of oxymoron phrases, examining them from both semantic and pragmatic perspectives. The study explored how these figures of speech function as persuasive devices within the context of poetry. Focusing specifically on a select group of English poems by prominent poets such as Wilfred Owen and Alfred Lord Tennyson, the research aimed to understand the intricate syntactic and semantic relationships between sentence structures. The study emphasized the crucial role of linguistic context in shaping the reader's interpretation of the author's intended meaning, highlighting its significance over the physical setting of the poem.

Cho et al. [9] introduced a novel approach to evaluating word vector representations by leveraging the inherent paradoxes found in oxymorons. This method capitalizes on the semantic discrepancies between word pairs within an oxymoron. To begin, a set of offset vectors is constructed for pairs of synonyms and antonyms. These offset vectors effectively capture the directional relationship between the vector embeddings of the two associated words. Subsequently, a deterministic technique is employed to analyze the relationship between a given words pair based on the established set of offset vectors. A key advantage of this method lies in its efficiency, requiring minimal computational time to assess the quality of input word vectors while obviating the need for additional training.

Ruiz [10] conducted an in-depth investigation into the mechanisms of production and interpretation of paradoxes and oxymorons. This research challenged the notion that paradoxes and oxymorons are confined to simple propositions with a limited number of predicates. Instead, Ruiz demonstrated that they encompass a broader propositional dimension. The study delved into several key aspects of paradox and oxymoron within their respective contexts, focusing on the most common occurrences. According to Haan et al.[11], oxymorons can significantly benefit social science scholars and students. By providing access to a system that encourages participation and the discovery of relevant literature within their fields of interest, oxymorons facilitate scholarly work. Furthermore, they promote the development of writing skills and foster the exchange of ideas, effectively teaching expression through the written word.

Gibbs & Kearney [12] conducted a study to investigate how the internal conceptual organization of oxymora influences human perception of their poetic quality. They explored the correlation between comprehension speed and aesthetic appreciation of oxymora. Specifically, they examined whether rapid understanding of these figures of speech is associated with higher ratings of their lyrical value. Finally, the researchers delved into the phenomenon of emergent meaning in oxymora, investigating whether the combined meaning of the constituent words transcends the sum of their meanings. Shen [13] investigated the internal semantic structures of poetic and non-poetic oxymorons, categorizing them into three types: direct, indirect, and metaphorical. He argued that metaphorical oxymorons present the greatest processing complexity due to their demand for the simultaneous consideration of multiple meanings. They evaluated the processing complexity of these three semantic constructs, finding that direct structures exhibited the lowest processing complexity, followed by indirect structures, with metaphorical structures demonstrating the highest level of processing complexity.

Muhammed & Meftin [14] identified a diverse range of pleonastic devices, including emphatic reflexive pronouns, multiple affirmations and negations, double possession, multiple quality gradations, overlap semantic pleonasm, and prolixity semantic pleonasm. They emphasized that pleonasm is a rhetorical strategy employed to achieve specific effects, such as heightened emphasis and enhanced clarity. Moreover, they argued that the use of these pleonasms can imbue writing with poignancy and passion, making it more engaging and compelling for the reader.

Stamborg & Nugues [15] developed an algorithm, grounded in statistical methods, for the automatic identification of pleonastic pronouns. This algorithm aims to facilitate a deeper understanding of natural language by enabling the efficient extraction of pleonastic pronouns. To address the challenge of co-reference resolution, a two-step approach was employed. Haider et al. [16] investigated the impact of different adverb types, including adverbs of degree, comparative degree, generality, comparative generality, location, preposition, and time, on the sentiment classification of tweets (positive, negative, or neutral). Their analysis revealed that general comparative adverbs and generic verbs were the most significant polarity-bearing words for neutral opinions.

Aytan et al. [17] investigated the impact of euphemisms and dysphemisms as linguistic devices employed in English-language newspaper articles covering political topics. Their analysis encompassed 393 instances of euphemisms and dysphemisms collected from a diverse range of political media texts. The study identified five primary categories of euphemisms: modern living, COVID-19, social media, economics, and politics. Oktaviyani & Licantik [18] proposed a novel approach for detecting redundant sentence pairs within software requirements specification documents. Their methodology leverages semantic similarity measures based on WordNet to assess the effectiveness of redundancy identification. Sethi [19] investigated the strategic use of pleonastic English words and phrases by Indian professionals in workplace email communication to influence the recipient's perception of them. The study explored pleonasm from a multifaceted perspective, analyzing its syntactic, semantic, and morphological characteristics.

3. PROPOSED ARCHITECTURE OF CLASSIFICATION OF OXYMORON

The implementation of proposed oxymoron classification systems using machine learning, as depicted in Fig.1, follows a structured approach. This methodology encompasses data collection, label assignment, data cleaning, feature extraction, dataset partitioning into training and testing sets, applying machine learning algorithms for classification, and subsequent performance evaluation.

3.1 DATASET COLLECTION AND PREPROCESSING

High-quality data is essential for effective text classification, particularly when dealing with the nuanced nature of oxymoronic sentences. This study focuses solely on English sentences, aiming to classify them as either oxymoronic or non-oxymoronic. The dataset was compiled from various sources, including tweets, news articles, comments, and product reviews gathered from social media platforms such as Twitter, news websites, and e-commerce sites. Data collection employed a combination of manual retrieval and automated methods, such as keyword-based searches for tweets. Twitter, a microblogging platform, provides a rich source of data, with tweets limited to 280 characters. The study utilized the Twitter API and search query methods to extract relevant tweets from the platform. Tweets containing oxymorons (identified by hashtags such as #oxymoron and related terms) and pleonasms (identified by hashtags such as #pleonasm and related terms) were collected from Twitter using Python's Tweepy and Pandas libraries. These tweets contained sentences ranging from 20 to 50 words in length. Three experts manually verified the oxymoron and pleonasm labels after data collection. Subsequent cleaning steps included converting all text to lowercase (e.g., "Book" and "book" became "book") to avoid misinterpretations by the vector space model. Contractions and negations were addressed to ensure accurate polarity, and words with two or fewer letters were removed as they were deemed irrelevant for oxymoron classification.

Removing hashtags and punctuation streamlines the data by eliminating extraneous characters that could hinder analysis, thereby improving oxymoron classification. This preprocessing step simplifies the content, enhancing readability and creating a more neutral, uniformly structured text that is better suited for various analyses. Since hashtags are primarily used for categorization or emphasis on social media, their removal, along with punctuation, ensures a consistent textual representation. Additionally, removing numbers, mentions, stop words, URLs, and duplicates significantly reduces the size of the dataset, improving computational performance and memory efficiency. Numbers often do not contribute directly to textual meaning in NLP tasks. The elimination of stop words further enhances performance, reduces data dimensionality, and focuses the analysis on more meaningful terms, as these words add little value to the overall understanding of the text. Following preprocessing, the final dataset comprised 3838 instances, balanced with 1919 oxymorons and 1919 pleonasms. This dataset was used to train and evaluate several machine learning classifiers, including J48, Random Forest, Multinomial Naive Bayes (MNB), Poisson Naive Bayes, Bernoulli Naive Bayes, K-Nearest Neighbors, and SVM with radial, polynomial, and linear kernels. 10-fold cross-validation was employed for evaluation.

3.2 FEATURE EXTRACTION AND REDUCTION

Text classification offers various segmentation methods, including character, word, sentence, and document segmentation. This study employed word segmentation to extract N-gram lexical features from each document, enabling classification as either oxymoronic or pleonastic. POS tags, or part-of-speech tags, are a common feature representation in NLP. They indicate the grammatical function of each word in a text (e.g., noun, verb, adjective). POS taggers assign these tags to words, creating a sequence of tags that can be used as features for machine learning algorithms. POS tags can be represented as either binary or count features. Binary features indicate the presence or absence of a tag (1 if present, 0 otherwise), while count features represent the frequency of each tag within the text. Structural features, derived from POS tags, are widely used in NLP applications. This study identified specific POS tag patterns as features for oxymoron classification, including "adjective + noun," "adjective + verb," and others. These 12 structural features, along with 36 POS tag features and N-gram features were combined to create the oxymoron classification model using a machine learning algorithm.

In these experiments, we extracted unigram and bigram features, creating three feature sets: unigrams combined with bigrams, unigrams and bigrams combined with POS tags, and unigrams, bigrams, POS tags, and structural features combined for subsequent analysis. Table.1 shows the categories and lists of features used for oxymoron classification. The R package 'tm' was used to extract features, followed by feature reduction. Features occurring fewer than 25 times in the dataset were removed, as they were hypothesized to have minimal impact on the selection of optimal classification models. Nearly 38% of the features were removed from this feature set based on this assumption.

Table.1. Categories and list of features

Category	List of features
N-Gram Features	Unigram, Bigram
POS tags feature	36 Features (Penn Tree-Bank) (Noun, Verb, Adjective, Adverb, Conjunction, Determiner, Preposition, etc.
Structural feature	Adjective + Noun, Adjective + Verb, Adjective + Adjective, Verb + Adverb, Verb + Article + Noun, Verb + Adjective, Adverb + Adverb, Adverb + Verb, Adverb + Adjective, Noun + Noun, Noun + Article + Noun, Adjective + Noun + Adjective

3.3 FEATURE WEIGHTING SCHEME

Feature space pruning was performed by assigning weights to different term sets. Specifically, weights were assigned to unigrams, the cumulative sum of unigram and bigram weights, and the cumulative sum of unigram, bigram, and trigram weights.

3.3.1 TF-IDF:

Term Frequency-Inverse Document Frequency is a measure of a word's importance within a document in a collection of documents (a corpus). It calculates this importance by considering both how often the word appears in the specific document (term frequency) and how uncommon it is across all documents in the corpus (inverse document frequency).

3.3.2 Term Occurrence:

This method simply counts how many times each word is present in each document, creating a vector of word counts to represent the document.

3.3.3 Binary Occurrence:

Documents are represented by binary vectors, where each element corresponds to a word token. 1 refers to the presence of the word token in the document, and a value of 0 refers to its absence.

3.4 CLASSIFICATION USING MACHINE LEARNING

Supervised learning, which requires labeled training data, can be applied to oxymoron detection. This process involves three steps: 1) Labeling each document in the training set as either "oxymoronic" or "pleonastic." 2) Training a model (G) to learn the relationship between the document's features and its assigned label. 3) Using the trained model (G) to predict the label (oxymoronic or pleonastic) for new, unlabeled documents. Several classification algorithms were evaluated for oxymoron classification: J48, Random Forest, Naïve Bayes, KNN, and Support Vector Machine. Naïve Bayes, based on Bayes' theorem and conditional probabilities, assumes feature independence.

This simple, efficient technique calculates the posterior probability of a class assigning a label, using the likelihood of term distribution within the document. Word position within the document is not considered. Naïve Bayes requires minimal memory and training time, and its laplace smoothing hyperparameter controls the degree of smoothing (lower values mean less smoothing). KNN, another straightforward method, classifies based on the similarity of data points. The crucial parameter k (the number of neighbors) is typically determined using techniques like cross-validation or grid search. R, utilizing packages like naive Bayes, caret, RWeka, randomForest, e1071, snow, and rjava, was used to implement the machine learning algorithms for training, testing, and labelling the dataset. A key research objective was to determine the optimal method for feature category selection, evaluation and analysis.

Table.2. Performance analysis of oxymoron classification using different machine learning algorithms with combination of different features set and feature weight scheme

Algorithm	**Criteria**	**BO**			**TO**			**TF-IDF**		
		U+B	**U+B+P**	**U+B+P+S**	**U+B**	**U+B+P**	**U+B+P+S**	**U+B**	**U+B+P**	**U+B+P+S**
J-48	Precision	66.34	68.33	71.66	70.58	72.12	73.47	73.16	74.67	**76.00**
	Recall	63.05	65.66	69.31	68.26	70.35	72.17	72.43	72.95	**74.26**
	F-measure	64.65	66.97	70.46	69.40	71.22	72.82	72.79	73.80	**75.12**
	Accuracy	65.53	67.61	70.95	69.91	71.57	73.06	72.93	74.10	**75.40**
	Kappa	0.31	0.35	0.42	0.40	0.43	0.46	0.46	0.48	**0.51**
RF	Precision	74.48	75.35	76.19	74.60	75.21	76.94	75.29	76.28	**77.31**
	Recall	72.69	73.27	75.35	73.48	74.31	76.50	73.37	74.41	**77.23**
	F-measure	73.58	74.29	75.77	74.04	74.76	76.72	74.32	75.34	**77.27**
	Accuracy	73.89	74.65	75.90	74.23	74.91	76.78	74.65	75.64	**77.28**
	Kappa	0.48	0.49	0.52	0.48	0.50	0.54	0.49	0.51	**0.55**
MNB	Precision	88.26	90.26	93.32	91.99	93.30	94.69	94.67	95.71	**96.76**
	Recall	86.19	88.90	90.98	88.59	89.94	92.97	90.78	93.07	**94.89**
	F-measure	87.21	89.58	92.14	90.26	91.59	93.82	92.68	94.37	**95.82**
	Accuracy	87.36	89.66	92.24	90.44	91.74	93.88	92.83	94.45	**95.86**
	Kappa	0.75	0.79	0.84	0.81	0.83	0.88	0.86	0.89	**0.92**
Poisson NB	Precision	86.32	88.56	90.13	90.51	91.88	92.50	91.66	92.56	**93.48**
	Recall	80.25	83.90	87.55	83.48	86.09	88.69	84.73	88.17	**89.63**
	F-measure	83.18	86.17	88.82	86.85	88.89	90.56	88.06	90.31	**91.51**
	Accuracy	83.77	86.53	88.98	87.36	89.24	90.75	88.51	90.54	**91.69**
	Kappa	0.68	0.73	0.78	0.75	0.78	0.82	0.77	0.81	**0.83**
Bernoulli-NB	Precision	81.41	84.32	86.66	85.71	87.04	88.41	86.20	88.03	**89.01**
	Recall	79.42	83.48	84.63	80.67	84.00	87.44	84.00	85.88	**88.59**
	F-measure	80.40	83.90	85.63	83.11	85.49	87.92	85.09	86.94	**88.80**
	Accuracy	80.64	83.98	85.80	83.61	85.75	87.99	85.28	87.10	**88.82**
	Kappa	0.61	0.68	0.72	0.67	0.71	0.76	0.71	0.74	**0.78**
SVM Radial Basis	Precision	91.78	93.17	94.74	94.19	95.00	95.59	94.46	95.41	**96.10**
	Recall	89.63	90.98	93.80	89.63	93.07	94.84	90.67	93.07	**96.40**
	F-measure	90.69	92.06	94.27	91.86	94.02	95.21	92.53	94.22	**96.25**
	Accuracy	90.80	92.16	94.29	92.05	94.09	95.23	92.68	94.29	**96.25**
	Kappa	0.82	0.84	0.89	0.84	0.88	0.90	0.85	0.89	**0.92**
SVM Polynomial	Precision	88.10	89.73	90.43	89.37	90.69	91.54	91.40	92.17	**93.45**
	Recall	84.11	87.44	89.63	87.65	89.32	91.30	88.59	90.78	**92.97**
	F-measure	86.06	88.57	90.03	88.50	90.00	91.42	89.97	91.47	**93.21**
	Accuracy	86.37	88.72	90.07	88.61	90.07	91.43	90.13	91.53	**93.23**
	Kappa	0.73	0.77	0.80	0.77	0.80	0.83	0.80	0.83	**0.86**
SVM Linear	Precision	89.44	90.40	91.06	91.22	91.92	92.40	92.45	93.36	**93.78**
	Recall	79.42	84.42	87.55	84.42	87.75	88.69	85.57	87.86	**89.63**
	F-measure	84.13	87.31	89.27	87.69	89.79	90.51	88.88	90.52	**91.66**
	Accuracy	85.02	87.73	89.47	88.14	90.02	90.70	89.29	90.80	**91.84**
	Kappa	0.70	0.75	0.79	0.76	0.80	0.81	0.79	0.82	**0.84**

4. RESULT ANALYSIS

The R programming language, along with data mining packages containing various functions, was used for preprocessing tasks such as tokenization, lowercasing, digit removal, stop word removal, and the elimination of plural words. To evaluate performance, four standard natural language processing metrics were employed: precision, recall, F-score, and the Kappa statistic.

Precision measures the accuracy of identified oxymoronic sentences, while recall measures the proportion of relevant oxymoronic sentences that were identified. The F-score, representing the harmonic mean of precision and recall, balances their relative importance. The experimental results presented were obtained using a high-performance hardware and software platform, demonstrating the effectiveness of the proposed algorithms. A detailed analysis of these results further validates the overall system's accuracy and reliability.

The proposed machine learning model was evaluated on a dataset of 3,838 instances, equally split between oxymorons and pleonasms (1,919 each). To optimize oxymoron classification, 72 experiments were conducted, varying three factors: machine learning algorithm (including J48 and Random Forest), feature weighting scheme (binary occurrence, term occurrence, and TF-IDF), and feature combination (Uni+Bi, Uni+Bi+PoS, and Uni+Bi+PoS+Struct features). Table.2 presents the performance results for oxymoron classification using J48, Random Forest, and MNB algorithms.

For each algorithm, nine experiments were performed, with three experiments for each feature weighting scheme. These experiments tested three different feature combinations: Uni+Bi, Uni+Bi+ PoS, and Uni+Bi+PoS+Structural features. The J48 model using TF-IDF weighting across all features achieved the best results, with 76.00% precision, 74.26% recall, a 75.12% F-measure, 75.40% accuracy, and a kappa statistic of 0.51. While the Random Forest classifier using TF-IDF weighting across all features achieved the highest accuracy (77.28%), with a precision of 77.31%, recall of 77.23%, and an F-measure of 77.27%, its Kappa statistic of 0.55 was considered too low, preventing it from being selected as the optimal model.

Multinomial Naïve Bayes (MNB) performed best, achieving 96.76% precision, 94.89% recall, a 95.82% F-measure, 95.86% accuracy, and a kappa statistic of 0.92. Poisson Naïve Bayes achieved slightly lower results, with 93.48% precision, 89.63% recall, a 91.51% F-measure, 91.69% accuracy, and a kappa of 0.83. Bernoulli Naïve Bayes exhibited lower accuracy and kappa compared to both MNB and Poisson Naïve Bayes. Table.2 presents the performance of the Support Vector Machine algorithm for oxymoron classification using various kernel functions: polynomial, radial basis function (RBF), and linear. The RBF kernel yielded the best results, achieving 96.10% precision, 96.40% recall, a 96.25% F-measure, 96.25% accuracy, and a kappa statistic of 0.92.

Determining the optimal k value for K-Nearest Neighbors (KNN) typically involves evaluating performance across a range of k values. For this dataset, k values of 1, 3, 5, 7, and so on were tested. Accuracy remained below 97% for k values less than 40, peaking at k = 50. Table.3 shows the performance of KNN with all combined features and various weighting schemes. Using TF-IDF weighting, KNN achieved excellent results: 99.38% precision, 99.64% recall, 99.51% F-measure, 99.50% accuracy, and a kappa statistic of 0.99. The Fig.2 shows the accuracy of oxymoron classification using the KNN algorithm with different k values.

Table.3. Performance analysis of proposed oxymoron classification using KNN algorithms with combination of different features set and feature weight scheme

Algorithm	Criteria	BO		
		U+B	U+B+P	U+B+P+S
KNN	Precision	96.5	97.2	98.13
	Recall	89.11	93.9	95.57
	F-measure	92.66	95.52	96.83
	Accuracy	92.94	95.6	96.87
	Kappa	0.86	0.91	0.94
Algorithm	**Criteria**	**TO**		
		U+B	**U+B+P**	**U+B+P+S**
KNN	Precision	97.87	98.34	98.94
	Recall	90.98	92.86	97.13
	F-measure	94.3	95.52	98.03
	Accuracy	94.5	95.65	98.05
	Kappa	0.89	0.91	0.96
Algorithm	**Criteria**	**TF-IDF**		
		U+B	**U+B+P**	**U+B+P+S**
KNN	Precision	98.68	98.93	**99.38**
	Recall	93.8	96.2	**99.64**
	F-measure	96.18	97.54	**99.51**
	Accuracy	96.27	97.58	**99.5**
	Kappa	0.93	0.95	**0.99**

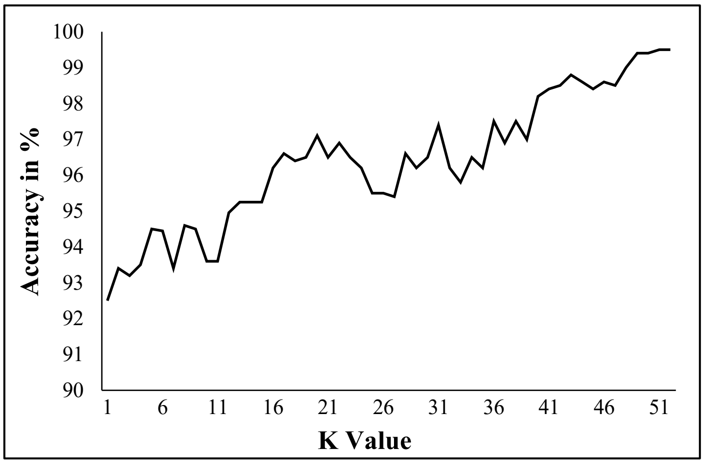
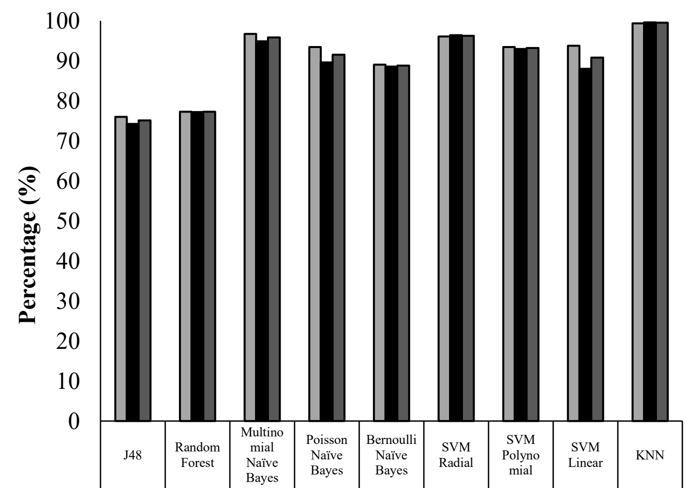


Fig.2. Accuracy of oxymoron classification using proposed KNN algorithm

The Fig.3 compares the overall performance of various machine learning models using the TF-IDF weighting scheme. The results demonstrate that the K-Nearest Neighbors (KNN) algorithm achieves superior performance compared to other algorithms tested, including J48, Random Forest, Multinomial

Naive Bayes (MNB), Poisson Naive Bayes, Bernoulli Naive Bayes, and Support Vector Machines (SVM) with three different kernel functions. KNN is the most effective machine learning classifier for oxymoron categorization, outperforming all other models tested.



	J48	Random Forest	Multinomial Naïve Bayes	Poisson Naïve Bayes	Bernoulli Naïve Bayes	SVM Radial	SVM Polynomial	SVM Linear	KNN
Precision	76	77.31	96.76	93.48	89.01	96.10	93.45	93.78	99.38
Recall	74.26	77.23	94.89	89.63	88.59	96.40	92.97	88.02	99.64
F-Score	75.12	77.27	95.82	91.51	88.80	96.25	93.21	90.81	99.51

Fig.3. Performance comparison of oxymoron classification using machine learning algorithms on all features with TF-IDF

5. CONCLUDING REMARKS AND FUTURE WORK

Oxymorons can capture attention and create surprises on social media, making posts and headlines stand out and potentially increasing user interaction. Their ironic nature also makes them more memorable and shareable. A new dataset of oxymorons and pleonasms has been created to aid research on the use of oxymorons in natural language. Experiments using this dataset demonstrate that the K-Nearest Neighbors (KNN) algorithm, in conjunction with the TF-IDF technique and a feature set combining unigrams, bigrams, parts of speech, and structural information, achieves the highest accuracy.

For oxymoron classification, the K-Nearest Neighbors (KNN) algorithm outperformed J48, Random Forest, Naïve Bayes, and Support Vector Machine. The best performance was achieved by combining n-grams, part-of-speech tags, and structural features. Given the small dataset size used in this study, future work could explore deep learning models for oxymoron classification with a larger, expanded dataset. Future scopes include exploring unsupervised methods to improve the Kappa statistics and reduce training time. Future research will focus on predicting oxymoronic phrases within mixed-language contexts.

REFERENCES

[1] M.A. Karp, "Figures of Contrast in Animal Farm by George Orwell: Lexical and Semantic Features of Oxymoron, Paradox and Antithesis", Baltija Publishing, pp. 232-249, 2021.

[2] M. La Pietra and F. Masini, "Oxymorons: A Preliminary Corpus Investigation", *Proceedings of Workshop on Figurative Language Processing*, pp. 176-185, 2020.

[3] S.N. Safarovna, "Oxymoron as a Method of Poetic Perception of the World", *Journal of Critical Reviews*, Vol. 7, No. 7, pp. 340-344, 2020.

[4] Y.C. Tsao, S.L. Chu, H.C. Liao and S.Y. Huang, "Application of Innovative Oxymorons in Product Design", *International Journal of Organizational Innovation*, Vol. 13, No. 1, pp. 340-355, 2020.

[5] B.R. Bageshwar, "A Study of Literary Devices used in the Posts on Instagram", *International Journal of English Literature and Social Sciences*, Vol. 6, No. 1, pp. 414-421, 2021.

[6] K.V. Fazlitdinovna, "Oxymoron and its Methodological-Semantic Features", *Bulletin of Science and Education*, Vol. 7, No. 2, pp. 114-119, 2020.

[7] L. Sakaeva and L. Kornilova, "Structural Analysis of the Oxymoron in the Sonnets of William Shakespeare", *Journal of History Culture and Art Research*, Vol. 6, No. 5, pp. 409-414, 2017.

[8] M.E. Alwana, "A Semantic and Pragmatic Study of the Oxymoron in Selected Poems", *International Journal of Innovation, Creativity and Change*, Vol. 13, No. 7, pp. 290-300, 2020.

[9] W.I. Cho, W.H. Kang, H.S. Lee and N.S. Kim, "Detecting Oxymoron in a Single Statement", Proceedings of International Conference of the Oriental Chapter of the International Coordinating Committee on Speech Databases and Speech I/O Systems and Assessment, pp. 1-5, 2017.

[10] J.H. Ruiz, "Paradox and Oxymoron Revisited", *Procedia-Social and Behavioral Sciences*, Vol. 173, pp. 199-206, 2015.

[11] C.B. De Haan, G. Chabre, F. Lapique, G. Regev and A. Wegmann, "Oxymoron, a Non-Distance Knowledge Sharing Tool for Social Science Students and Researchers", *Proceedings of International Conference on Supporting Group Work*, pp. 219-228, 1999.

[12] R.W. Gibbs and L.R. Kearney, "When Parting is Such Sweet Sorrow: The Comprehension and Appreciation of Oxymora", *Journal of Psycholinguistic Research*, Vol. 23, No. 1, pp. 75-89, 1994.

[13] Y. Shen, "On the Structure and Understanding of Poetic Oxymoron", *Poetics Today*, Vol. 8, No. 1, pp. 105-122, 1987.

[14] L.S.E. Muhammed and L.J.A. Meftin, "A Study of Pleonasm with Reference to Some Biblical Verses", *Journal of Al-Qadisiya University*, Vol. 19, No. 1, pp. 7-24, 2016.

[15] M. Stamborg and P. Nugues, "Statistical Identification of Pleonastic Pronouns", *Proceedings of International Conference on Swedish Language Technology*, pp. 67-68, 2012.

[16] S. Haider, M. Tanvir Afzal, M. Asif, H. Maurer, A. Ahmad and A. Abuarqoub, "Impact Analysis of Adverbs for Sentiment Classification on Twitter Product Reviews", *Concurrency and Computation: Practice and Experience*, Vol. 33, No. 4, pp. 1-15, 2021.

[17] A. Aytan, B. Aynur, P. Hilal, E. Aytac and A. Malahat, "Euphemisms and Cho Dysphemisms as Language Means Implementing Rhetorical Strategies in Political Discourse",

Journal of Language and Linguistic Studies, Vol. 17, No. 2, pp. 741-754, 2021.

[18] E.D. Oktaviyani, "Redundancy Detection of Sentence Pairs in the Software Requirements Specification Documents with Semantic Approach", *Proceedings of International Conference on Informatics, Engineering, Science and Technology*, Vol. 879, No. 1, pp. 1-8, 2020.

[19] D. Sethi, "Workplace Impression Management through Pleonastic English Words/Phrases in E-Mail Communication? Evidence from India", *International Journal of Indian Culture and Business Management*, Vol. 18, No. 4, pp. 391-404, 2019.