DETECTING SARCASM BASED ON QUANTUM VECTORIZATION

1 Niyozmatova Nilufar,
2 Turgunova Nafisakhon,
2 Turgunov Bakhodirjon,
2 Mamatov Abduvali,
2 Xoitkulov Abdumalik
1 Department of Digital technologies and artificial intelligence, PhD
"Tashkent Institute of Irrigation and Agricultural Mechanization Engineers"
National Research University
e-mail: n_niyozmatova@mail.ru
2 Department of Digital Technologies and Artificial Intelligence, Assistant Lecturer
"Tashkent Institute of Irrigation and Agricultural Mechanization Engineers" National Research University
e-mail: nafisa_turgunova@tiiame.uz

Abstract
This article proposes a neural network approach for sarcasm detection based on the formation of quantum bag of words features from text. In the proposed method, words are first converted into quantum states using the Grover algorithm. Then, quantum bag of words features are generated and used for classification. The obtained results were compared with those of classical approaches. According to experimental results, the quantum computing-based approach achieved 80% accuracy, while the classical computing-based approach showed 78% accuracy. Additionally, the quantum bag of words method produced results 2% faster than the classical approach. This demonstrates that the quantum approach, through the properties of superposition and entanglement, achieves higher performance. The limitations include the restricted vocabulary size, insufficient datasets in many cases, and the lack of access to real quantum technologies. Nevertheless, the findings show that computers capable of simulating quantum behavior indicate the promising potential of quantum computing in natural language processing.

Keywords: Sarcasm, quantum computing, quantum bag of words, LSTM, Grover algorithm, natural language processing, neural network.

Introduction

The introduction of informatization into all spheres has led to the rapid development of computational linguistics. It deals with the tasks of automatic processing of textual information. One of the important tasks in data processing is the development of classification methods. This is due to the need to process large volumes of electronic messages received by various organizations and institutions through internet resources (N. S. Mamatov et al., 2024; Дли et al., 2019).

Text data classification is the process of grouping it according to its content. Methods and algorithms for classifying textual information are currently used in many areas, including processing electronic requests, spam filtering, sentiment analysis, opinion, topic, and language detection, as well as information retrieval. At present, text classification widely employs methods such as Support Vector Machine (SVM), Decision Tree (DT), Naive Bayes (NB), K-Nearest Neighbors (KNN), Neural Network (NN), and Long Short-Term Memory (LSTM) (Assistant Professor, Rajarambapu Institute of Technology, Rajaramnagar, MS, India & Adamuthe, 2020; D. Muhamediyeva et al., 2025).

At present, solving problems related to classification is one of the most relevant tasks. In particular, tasks such as detecting sarcasm, irony, mockery, and performing sentiment analysis are among the key issues in classification. In this scientific work, the problem of sarcasm detection is addressed, and an effective solution is proposed based on a quantum feature formation algorithm combined with an LSTM neural network.

In spoken language, sarcasm can be detected through speech tempo and duration, variability, pitch level, and acoustic characteristics. However, in written form, sarcastic statements lack intonation or gestures, making them difficult to distinguish. To detect sarcasm in textual format, an effective Natural Language Processing (NLP) approach is required to classify ironic features within sentences (Bavkar et al., 2022; Маматов et al., 2025; Угли et al., 2023)[5-7].

In recent years, deep learning methods, particularly the LSTM neural network, have demonstrated high efficiency in text classification. However, challenges remain in feature extraction and preserving semantic richness during classification. Nevertheless, many studies based on LSTM and Bag of Words (BoW) have been conducted to date. The most significant of them are analyzed below.

Admuthe (2020) proposed a classification approach based on Word2Vec vectorization and LSTM. The author achieved approximately 91% accuracy on well-known datasets such as IMDB, Yelp, AG News, and others. The study provided an in-depth analysis of how optimizing vectorization parameters, the number of LSTM layers, and training algorithms affect model performance (Assistant Professor, Rajarambapu Institute of Technology, Rajaramnagar, MS, India & Adamuthe, 2020). Jamil et al. (2021) proposed a sarcasm detection approach using a CNN–LSTM architecture. In this approach, contextual features are first extracted using CNN, and then temporal dependencies are learned through LSTM. The multi-domain data-based study achieved ≈91.6% accuracy, outperforming classical models. This work demonstrated that combining CNN-based feature extraction with LSTM-based temporal analysis is an effective solution (Bavkar et al., 2022; Jamil et al., 2021; Маматов et al., 2025; Угли et al., 2023).

Gumelar et al. (2024) developed a toxic speech detection model based on multimodal data. In the study, text and speech signals were analyzed jointly, and an LSTM-based model achieved an F1-score of approximately 92.7%. This demonstrates that LSTM models are effective not only for text processing but also when working with various types of data. In addition, the study emphasized the importance of transcription quality and voice parameters (such as timbre, intensity, and speech rate) in classification tasks ('An Improved Toxic Speech Detection on Multimodal Scam Confrontation Data Using LSTM-Based Deep Learning', 2024; D. T. Muhamediyeva et al., 2025; Ниёзматова & Турғунова, 2025).

Zhou et al. (2015) proposed the C-LSTM architecture, which sequentially combines CNN and LSTM, aiming to achieve high performance in text classification by integrating deep convolutional features with temporal dependencies. This approach enables a deeper understanding of local relationships between words and contextual semantics (N. Mamatov & Nuritdinov, 2025; Zhou et al., 2015).

Johnson and Zhang (2016) proposed training LSTM in supervised and semi-supervised modes based on region embedding. This method significantly improves classification accuracy, especially when only a small amount of labeled data is available. Through the use of region embeddings, the model deeply captures contextual text blocks (O. J. Babomuradov et al., 2019; Johnson & Zhang, 2016).

The fastText model, developed by Joulin et al. (2016), uses Bag of Words and n-gram structures as a simple yet effective approach for text classification. This model provides high speed and reliability, competing with classical machine learning models and some deep learning architectures. The main strength of the fastText model is its lightweight and computationally efficient architecture. This approach is considered one of the effective strategies for use in combination with quantized BoW models and LSTM (O. O. Babomuradov et al., 2019; Joulin et al., 2016).

The analysis conducted by Asudani et al. (2023) examines the impact of various word embedding models (Word2Vec, GloVe, ELMo, BERT, etc.) on deep learning models, particularly LSTM. They demonstrated that contextual embeddings ELMo and BERT achieve the highest classification results. In particular, contextual embeddings combined with a transfer learning approach significantly improve performance (Asudani et al., 2023; Niyozmatova et al., 2021).

The analysis of the literature above showed that the effectiveness of LSTM models is highly dependent on the quality of vectorization. Models built on word embeddings can capture deep semantic structures, yet a relatively simple quantum BoW approach, when properly adapted to the data, can achieve high accuracy. This confirms that the combination of quantum BoW and LSTM used in this study is effective.

Methodology

In this study, the proposed approach is used to detect sarcasm by combining NLP methods, quantum BoW features, and neural network models. This approach consists of stages of data preprocessing, vectorization based on quantum BoW features, model architecture, training, and prediction of new data.

Stage 1. Data preprocessing. At this stage, the text database is loaded, punctuation and stop words are removed, text is converted to lowercase, extra spaces are cleaned, and similar operations are performed. For this purpose, sarcasm headline databases *V1* and *V2* were used. The set of headlines is denoted as h_i, and for binary classification, groups are created as $y_i \in \{0,1\}$, where $y_i = 1$ represents sarcasm and otherwise indicates non-sarcasm.

Stage 2. Tokenization and classical BoW. Using tokenization, the text data is split into tokens. Based on this, a dictionary of the most frequently occurring words is created:

$$Tokenizator(X_{train}, V) \rightarrow L = \{l_1, l_2, ..., l_V\},$$

here L – dictionary(25 words), l_i - the i-th word in dictionary.

The classical BoW vector is formulated as follows:

$$v_i = classical\_BoW(T_i, L) = [f_1, f_2, ..., f_V],$$

here T_i - i -th text, $f_j = \sum \mathrm{I}(word = l_j)$ - frequency of l_j th word(I is the indicator function).

Stage 3. Quantum search. Here, the search is carried out using the Grover algorithm, that is, the presence of words in the text is determined. It is performed in the following steps:

Step 1. Constructing the quantum circuit. Initially, the state is brought into a quantum superposition state:

$$\left|\psi_0\right\rangle = H^{\otimes n}\left|0\right\rangle^{\otimes n} = \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} \left|x\right\rangle,$$

here $n = |V|$ - qubit quantum circuit, H - Hadamard gate.

Step 2. Applying the oracle operator. The amplitudes of the target states of the oracle U_f, based on the target vector v_t, are modified as follows:

$$U_f\left|x\right\rangle = (-1)^{f(x)}\left|x\right\rangle, \; f(x) = \begin{cases} 1, \; if \; x = v_t \\ 0, \; otherwise \end{cases}.$$

A multi-controlled $X(MCX)$ gate is used to mark the target state.

Step 3. Applying the diffusion operator. To restore the superposition and amplify the amplitudes of the target states, the diffusion operator U_d reflects the amplitudes around the average.

$$U_d = H^{\otimes n}(2\left|0\right\rangle\left\langle 0\right| - I)H^{\otimes n}.$$

Step 4. Performing iterations. The number of Grover iterations is taken as optimal $k \approx \sqrt{2^n}$. In each iteration, U_f and U_d are applied as follows:

$$\left|\psi_k\right\rangle = (U_d U_f)^k \left|\psi_0\right\rangle.$$

Step 4. Measurement is performed. $n$qubits are measured, resulting in bitstrings and their occurrence frequencies $counts = \{s_i : c_i\}$, where $s_i \in \{0,1\}^n$ is the bitstring and c_i is the number of measurements.

Converting the results into a quantum BoW vector. For each bitstring s_i, if it matches $s_i[j] = 1$, then c_i is added to the word at index j. Normalization is performed as follows:

$$q_t[j] = round\left(\frac{\sum_{s_i : s_i[j]} c_i}{\sum c_i} \cdot v_t[j]\right),$$

here q_t is a quantum BoW vector.

The formed vectors are prepared for the LSTM neural network, and all vectors are brought to the same size. Padding is applied for this purpose, with the maximum length set to $L = 25$:

$$P = pad(S, L) = S \oplus 0^{L-m},$$

here, S represents the sequence of quantum BoW vectors, and $\oplus$ denotes padding with zero vectors. The training and validation sets are formed as follows:

$$P_{train} = \left[pad(Q_{BoW}(x_i, V), L) \middle| x_i \in X_{train} \right],$$

$$P_{test} = \left[pad(Q_{BoW}(x_i, V), L) \middle| x_i \in X_{test} \right].$$

This $(N, L, |V|)$ represents the input tensors in the form.

Step 4. Training the neural network. The classification model consists of a simple sequential neural network, which employs bidirectional LSTM layers to capture dependencies in sequences prepared using the quantum method, as well as dense layers for binary classification. $f_\theta(\cdot)$ model is next:

$$f_\theta(P) = \sigma(Dense_1(Dropout_{0.3}(\mathrm{Re}LU(Dense_{32}(BiLSTM_{64}(BiLSTM_{128}(P))))))),$$

here $BiLSTM_k$ k represents a bidirectional LSTM layer with neurons, the second layer has a dropout rate of 0.2 and a fully connected layer with $return\_sequences = True$, $Dense_k$ k neurons, ReLU rectified linear activation, where $\max(0, x)$, σ sigmoid activation $\sigma(x) = \frac{1}{1+e^{-x}}$.

The loss function is implemented as the binary cross-entropy

$$L(y, \hat{y}) = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i) \right].$$

In it, accuracy is observed as follows:

$$Accuracy = \frac{1}{N}\sum_{i=1}^{N}\left[y_i \log\left(\hat{y}_i\right) + \left(1-y_i\right)\log\left(1-\hat{y}_i\right) \right]$$

The model was trained up to the 40th epoch, early stopping was applied, and the model that achieved the best result was saved.

Training and predicting new data. For a new headline h_{new}, the following calculations are performed:

$$h_{new,c} = C(h_{new}),$$

$$P_{new} = pad(Q_{BoW}(h_{new,c}, V), L),$$

$$\hat{y}_{new} = f_{\theta*}(P_{new}),$$

here $\hat{y}_{new} > 0.5$ if so, then the title is sarcastic.

In this study, the *V1* and *V2* datasets, consisting of headlines from the Kaggle platform, were used. The *V1* dataset contains 26,709 headlines, of which 14,985 are sarcastic and 11,724 are non-sarcastic, while the *V2* dataset contains 28,619 headlines, of which 15,085 are sarcastic and 13,534 are non-sarcastic. These datasets were combined and split into training and validation sets in an 80/20 ratio, and classification was performed using the proposed algorithm.

Results

In this study, sarcasm detection was performed using an LSTM neural network based on quantum Bag of Words features. The main evaluation metrics were accuracy, precision, and F1 scores, while the classification time for both quantum and classical approaches was also taken into account (Table 1).

Table 1 Results of classification algorithms

t/r	Approach type	Training time (sec)	Classification time (sec)	Accuracy (%)	Precision (%)	F1 (%)
1	Quant BoW anddd LSTM	540	3	80	89	81
2	Classic BoW va LSTM	631	5	78	83	78

The quantum BoW algorithm with a BiLSTM architecture and the LSTM approach achieved 80% accuracy on the validation set. This demonstrates the capability of quantum computing to effectively model semantic relationships. The classical BoW with LSTM approach achieved 78% accuracy. Both approaches were trained for up to 40 epochs using a loss function and the Adam optimizer, with early stopping (patience=3) and best model saving strategies applied. These approaches were applied to the same LSTM architecture with both quantum and classical BoW algorithms and compared. The classical model achieved 78% accuracy, which is 2% lower than the quantum approach. The classical method used simple word-frequency-based vectorization, whereas leveraging quantum computing properties provided a relatively higher result. Overall, the LSTM approach based on quantum BoW feature construction achieved higher accuracy compared to the classical approach, demonstrating the potential of applying quantum computing in natural language processing.

Discussion

Quantum and LSTM approaches exist in NLP, but fully quantum BoW and LSTM methods for sarcasm detection have not yet been developed. One of the recent innovations is hybrid quantum-classical models. For example, QFNN in 2023 achieved 92% accuracy for sarcasm tasks. However, it used quantum unstructured or kernel-based methods rather than BoW. Chen et al. (2020) proposed a quantum LSTM model that represents classical LSTM in a quantum state, processing sequences through superposition using qubits. In that work, the proposed approach was tested on NLP tasks, including text classification and language generation. As a result, it showed 10–20% faster convergence compared to classical LSTM, though on real NISQ devices, accuracy slightly decreased due to noise. Bausch (2020) proposed the Recurrent Quantum Neural Networks approach, where quantum RNNs and LSTM were hybridized to model sequences in NLP. The approach was tested on text classification and language models, demonstrating that quantum superposition better captures contextual relationships. On small datasets, accuracy was 5–15% higher than classical methods, but on large datasets, the lack of real quantum computing devices created difficulties for solving these tasks on classical computers.

Conclusion

This scientific study investigated the effectiveness of the quantum bag of words model and the LSTM approach for sarcasm detection, and also compared it with the classical bag of words method. The quantum BoW approach showed 2% higher accuracy compared to the classical BoW. Several features of quantum computing, including superposition, entanglement, and interference phenomena, provide advantages in data analysis and classification. This study demonstrated the promising potential of quantum computing approaches in natural language processing. In the future, the approaches presented in this study should be explored on various datasets and real quantum devices. In particular, expansion to multilingual datasets, increasing the vocabulary size, and testing on real quantum devices such as IBM Quantum or Google Sycamore are necessary. Furthermore, integrating the model with modern architectures, such as Transformers, could improve accuracy and broaden its applicability. The results of this study may serve as an important step illustrating the future of quantum technologies in language analysis.

References

1. An Improved Toxic Speech Detection on Multimodal Scam Confrontation Data Using LSTM-Based Deep Learning. (2024). International Journal of Intelligent Engineering and Systems, 17(6), 880–904. https://doi.org/10.22266/ijies2024.1231.67
2. Assistant Professor, Rajarambapu Institute of Technology, Rajaramnagar, MS, India, & Adamuthe, A. C. (2020). Improved Text Classification using Long Short-Term Memory and Word Embedding Technique. International Journal of Hybrid Information Technology, 13(1), 19–32. https://doi.org/10.21742/IJHIT.2020.13.1.03
3. Asudani, D. S., Nagwani, N. K., & Singh, P. (2023). Impact of word embedding models on text analytics in deep learning environment: A review. Artificial Intelligence Review, 56(9), 10345–10425. https://doi.org/10.1007/s10462-023-10419-1
4. Babomuradov, O. J., Mamatov, N. S., Boboyev, L. B., & Otaxonova, B. I. (2019). Classification of texts using decision trees algorithms. Descendants of Muhammad Al-Khwarizmi, 4(10). https://scholar.google.com/scholar?cluster=5659375246760009871&hl=en&oi=scholarr
5. Babomuradov, O. O., Otaxonova, B., Mamatov, N. S., & Boboev, L. B. (2019). "Text documents classification in Uzbek language," International Journal of Recent Technology and Engineering, vol. 8, no. 2S11, pp. 3787–3789, 2019. International Journal of Recent Technology and Engineering, 8(2S11), 3787–3789.
6. Bavkar, D. M., Kashyap, R., & Khairnar, V. (2022). Multimodal Sarcasm Detection via Hybrid Classifier with Optimistic Logic. Journal of Telecommunications and Information Technology, 3(2022), 97–114. https://doi.org/10.26636/jtit.2022.161622
7. Jamil, R., Ashraf, I., Rustam, F., Saad, E., Mehmood, A., & Choi, G. S. (2021). Detecting sarcasm in multi-domain datasets using convolutional neural networks and long short term memory network model. PeerJ Computer Science, 7, e645. https://doi.org/10.7717/peerj-

cs.645

8. Johnson, R., & Zhang, T. (2016). Supervised and Semi-Supervised Text Categorization using LSTM for Region Embeddings (No. arXiv:1602.02373). arXiv. https://doi.org/10.48550/arXiv.1602.02373
9. Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2016). Bag of Tricks for Efficient Text Classification (No. arXiv:1607.01759). arXiv. https://doi.org/10.48550/arXiv.1607.01759
10. Mamatov, N., & Nuritdinov, N. (2025). MATNNI TAHLIL QILISH BO‘YICHA ILG‘OR TAJRIBALAR. COMPUTER LINGUISTICS: PROBLEMS, SOLUTIONS, PROSPECTS, 1(1). https://compling.navoiy-uni.uz/index.php/conferences/article/view/1172
11. Mamatov, N. S., Turg’unova, N. M., Turg’unov, B. X., & Samijonov, B. N. (2024). “Algorithm for analysis of texts in Uzbek language,” in Artificial Intelligence and Information Technologies, CRC Press, 2024, pp. 511–515. In Artificial Intelligence and Information Technologies. CRC Press.
12. Muhamediyeva, D., Niyozmatova, N., Turgunova, N., Ungalov, S., & Almuradova, N. (2025). Classification of Emoji in Text Documents of Users in Social Networks Using Machine Learning. 2025 6th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI), 1491–1496. https://doi.org/10.1109/ICMCSI64620.2025.10883250
13. Muhamediyeva, D. T., Niyozmatova, N. A., Ungalov, S., Mamatov, A. A., & Turgunova, N. (2025). Evaluating the effectiveness of text summarization algorithms based on recall-oriented understudy for Gisting evaluation metrics. Fourth International Conference on Digital Technologies, Optics, and Materials Science (DTIEE 2025), 13662, 201–209. https://doi.org/10.1117/12.3072740
14. Niyozmatova, N. A., Mamatov, N. S., Otaxonova, B. I., Samijonov, A. N., & Erejepov, K. K. (2021). Classification Based on Decision Trees and Neural Networks. International Conference on Information Science and Communications Technologies: Applications, Trends and Opportunities, ICISCT 2021. Scopus. https://doi.org/10.1109/ICISCT52966.2021.9670345
15. Zhou, C., Sun, C., Liu, Z., & Lau, F. C. M. (2015). A C-LSTM Neural Network for Text Classification (No. arXiv:1511.08630). arXiv. https://doi.org/10.48550/arXiv.1511.08630
16. Дли, М. И., Dli, M. I., Булыгина, О. В., & O.V., B. (2019). Особенности применения нейро-сетевых моделей для классификации коротких текстовых сообщений. Международный журнал ‘Программные продукты и системы’, 26, 650–654. https://doi.org/10.15827/0236-235X.128.650-654
17. Маматов, Н., Турғунова, Н., Хоитқулов, А., & Алмурадова, Н. (2025). МАТНЛИ МАЪЛУМОТЛАРДАГИ ХАТОЛИКЛАРНИ БАРТАРАФ ЭТИШНИ КВАНТ АЛГОРИТМИ. DIGITAL TRANSFORMATION AND ARTIFICIAL INTELLIGENCE, 3(3), 199-204. DIGITAL TRANSFORMATION AND ARTIFICIAL INTELLIGENCE, 3(3), 199–204.
18. Ниёзматова, Н., & Турғунова, Н. (2025). МАТНЛАРНИ ВЕКТОРЛАШНИ TF_IDF

КВАНТ АЛГОРИТМИ. DIGITAL TRANSFORMATION AND ARTIFICIAL INTELLIGENCE, 3(3), 189–193.

19. Угли, Т. Б. Х., Кизи, Т. Н. М., Кизи, И. М. И., & Кизи, А. Д. А. (2023). АЛГОРИТМЫ ОТДЕЛЕНИЯ ФАКТОВ ОТ ИНФОРМАЦИИ НА УЗБЕКСКОМ ЯЗЫКЕ ИЗ ЗАЯВЛЕНИЙ, ПОЛУЧЕННЫХ ОТ ВИРТУАЛЬНЫХ ПРИЕМНЫХ. Universum: Технические Науки, 12-2 (117), 9–11.