

Quantum-Enhanced Deep Learning Model for Text Classification

Author: ANUMATI MALLIKHARJUN¹ (MCA student), Mr .P. SASI KUMAR² (Assistant Professor)

Master Of Computer Applications, GIET ENGINEERING COLLEGE (A), Rajahmundry, AP.

Corresponding Author: Anumati Mallikharjun

(email-id:anumatimalli@gmail.com)

ABSTRACT: In the modern digital era, large volumes of textual data are generated from sources such as social media, emails, news articles, and online platforms. Efficient text classification is essential for applications like sentiment analysis, spam detection, and information retrieval. However, traditional deep learning models require high computational power and may struggle with complex data patterns.

To overcome these challenges, the proposed Quantum-Enhanced Deep Learning Model for Text Classification integrates quantum computing principles with classical deep learning techniques. The system utilizes Quantum Machine Learning (QML) concepts along with neural networks to improve classification accuracy and computational efficiency.

The model is developed using Python and incorporates technologies such as Natural Language Processing (NLP), Deep Learning, and Quantum Circuits. Text data is preprocessed, converted into embeddings, and processed through hybrid quantum-classical layers for classification.

The proposed system demonstrates improved performance in handling complex textual patterns, faster convergence, and enhanced accuracy compared to traditional models. This project highlights the potential of combining quantum computing and artificial intelligence for advanced text analytics.

KEYWORDS: Quantum Computing, Deep Learning, Text Classification, Natural Language Processing, Quantum Machine Learning, Neural Networks, Artificial Intelligence

I. INTRODUCTION

In today's digital world, text data plays a crucial role in communication, business, and decision-making. Organizations rely on text classification systems to analyze customer feedback, detect spam, categorize documents, and extract meaningful insights.

Traditional machine learning and deep learning models such as CNNs and RNNs have been widely used for text classification. However, these models often require large

datasets and high computational resources, and they may not efficiently capture complex patterns in high-dimensional data.

Quantum computing introduces a new paradigm that leverages quantum mechanics principles such as superposition and entanglement. By integrating quantum computing with deep learning, it is possible to improve processing efficiency and model performance.

The main objectives of this project are:

- To develop a hybrid quantum-classical model for text classification

II. Literature Survey:

2.1 Text Classification Systems

Text classification involves assigning predefined categories to textual data. Traditional approaches include Naive Bayes, Support Vector Machines, and deep learning models like CNNs and LSTMs. While effective, these methods can be computationally expensive and limited in handling complex patterns.

2.2 Deep Learning for NLP

Deep learning models such as Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM), and Transformers have significantly improved text classification accuracy. These models can capture contextual and semantic information but require large training data and high computational resources.

2.3 Quantum Computing

Quantum computing uses quantum bits (qubits) to perform computations. Unlike classical bits, qubits can exist in multiple states simultaneously, enabling faster processing for certain problems. Quantum algorithms have shown potential in optimization and pattern recognition tasks.

2.4 Quantum Machine Learning (QML)

Quantum Machine Learning combines quantum computing with machine learning algorithms. Hybrid quantum-classical models use quantum circuits along with neural networks to enhance learning capabilities and improve efficiency in handling complex data.

III. System Analysis And Design

3.1 Existing System

Existing text classification systems mainly rely on classical machine learning and deep learning techniques.

Limitations of Existing System:

High computational complexity

Requires large datasets

Limited efficiency in handling high-dimensional data

Slow training for large-scale models

Difficulty in capturing complex feature relationships

3.2 Proposed System

The proposed system integrates quantum computing with deep learning to enhance text classification performance.

The system processes textual data using NLP techniques and feeds it into a hybrid quantum-classical model for classification.

Advantages of Proposed System:

Improved classification accuracy

Faster convergence during training

Efficient handling of complex data

Reduced computational requirements

Scalable and innovative approach

3.3 System Architecture

The system consists of the following modules:

Input Module

Accepts text data from datasets or user input

Preprocessing Module

Performs tokenization, stop-word removal, and text cleaning

Feature Extraction Module

Converts text into numerical vectors using embeddings

Quantum Layer Module

Applies quantum circuits to process features using qubits

Deep Learning Module

Combines quantum output with neural network layers

Classification Module

Classifies text into predefined categories

Output Module

Displays classification results

IV. PROPOSED METHODOLOGY

The proposed methodology describes the working process of the Quantum-Enhanced Deep Learning Model for Text Classification and explains how classical Natural Language Processing (NLP) techniques are integrated with quantum computing concepts to achieve efficient and accurate text classification.

The system is designed as a hybrid quantum-classical architecture, where classical preprocessing and feature extraction techniques are combined with quantum circuits and deep learning models. This hybrid approach allows the system to leverage the strengths of both classical and quantum computing.

The overall workflow of the system consists of multiple stages, including data preprocessing, feature extraction, quantum encoding, hybrid model processing, and classification. Each stage plays an important role in improving the performance and efficiency of the system.

4.1 Data Collection

The first step in the methodology is collecting textual data from various sources such as datasets, online platforms, social media, emails, and documents. The dataset may contain labeled text samples belonging to different categories such as positive/negative sentiment, spam/ham emails, or topic-based classification.

The collected data is stored in a structured format for further processing. High-quality and well-labeled datasets help improve the accuracy of the classification model.

4.2 Text Preprocessing

Text preprocessing is an essential step that prepares raw textual data for further analysis. Raw text data often contains noise such as punctuation, special characters, and irrelevant words.

The preprocessing process includes the following steps:

- **Tokenization:** Splitting text into individual words or tokens
- **Lowercasing:** Converting all text to lowercase for uniformity
- **Stop-word Removal:** Removing common words like “is”, “the”, “and”
- **Stemming/Lemmatization:** Reducing words to their base or root form
- **Noise Removal:** Eliminating unwanted symbols and characters

This step improves data quality and ensures better feature extraction.

4.3 Feature Extraction

After preprocessing, the text data is converted into numerical format so that it can be processed by machine learning and quantum models.

Common feature extraction techniques include:

- Bag of Words (BoW)
- Term Frequency-Inverse Document Frequency (TF-IDF)
- Word Embeddings (Word2Vec, GloVe)

These techniques transform textual data into vectors that represent semantic meaning and contextual relationships between words.

4.4 Quantum Data Encoding

In this step, classical data (numerical vectors) is encoded into quantum states so that it can be processed by quantum circuits.

Quantum encoding techniques include:

- Amplitude Encoding
- Angle Encoding
- Basis Encoding

The encoded data is represented using qubits, which can exist in multiple states simultaneously due to the principle of superposition. This enables efficient representation of high-dimensional data.

4.5 Quantum Circuit Processing

Once the data is encoded, it is passed through a quantum circuit. The quantum circuit performs operations using quantum gates such as:

- Hadamard Gate
- Pauli Gates (X, Y, Z)
- Rotation Gates
- Controlled Gates

These gates manipulate qubits and create complex feature transformations using quantum properties like entanglement and superposition.

Quantum circuits help capture hidden patterns in data that

may not be easily detected by classical models.

V. IMPLEMENTATION

5.1 Technologies Used

The Quantum-Enhanced Deep Learning Model for Text Classification is implemented using a combination of classical machine learning, deep learning, and quantum computing technologies. The system is developed using Python and various advanced libraries to support text processing, model training, and quantum operations.

The technologies used in this project are listed below:

Python – Used as the main programming language for developing the system

NLTK / SpaCy – Used for Natural Language Processing tasks such as tokenization and preprocessing

Scikit-learn – Used for feature extraction techniques like TF-IDF and model evaluation

TensorFlow / PyTorch – Used for building and training deep learning models

Qiskit / Cirq – Used for implementing quantum circuits and quantum operations

NumPy & Pandas – Used for data manipulation and numerical computations

Matplotlib / Seaborn – Used for visualization of results and performance analysis

These technologies work together to provide a powerful platform for building a hybrid quantum-classical text classification system.

5.2 Modules

The system is divided into multiple functional modules, where each module performs a specific task in the classification process.

Input Module

The Input Module is responsible for accepting text data from users or datasets. The input may include sentences, documents, or large text files. This module acts as the starting point of the system.

Preprocessing Module

The Preprocessing Module cleans and prepares the raw text data. It removes unwanted characters, converts text to lowercase, removes stop words, and performs tokenization and stemming. This improves the quality of data and ensures better performance.

Feature Extraction Module

This module converts the processed text into numerical format using techniques such as TF-IDF or word embeddings. These features represent the semantic meaning of text and are used as input for further processing.

Quantum Module

The Quantum Module encodes classical data into quantum states using qubits. It applies quantum gates and circuits to

process the data using quantum principles like superposition and entanglement. This module enhances feature representation and pattern recognition.

Deep Learning Module

The Deep Learning Module consists of neural network layers that process the output from the quantum module. It performs advanced feature learning and improves classification accuracy using training algorithms.

Classification Module

The Classification Module predicts the category of the input text based on learned patterns. It assigns labels such as positive/negative sentiment, spam/ham, or topic categories.

Output Module

The Output Module displays the final classification result to the user. It may also show additional information such as confidence scores or graphical representations.

VI. RESULTS AND ANALYSIS

The implemented Quantum-Enhanced Deep Learning Model for Text Classification was successfully tested using various standard text datasets consisting of different categories such as sentiment analysis, spam detection, and topic classification. The system demonstrated effective performance in classifying text data with improved accuracy and efficiency. The hybrid quantum-classical model showed better results compared to traditional deep learning models in handling complex textual patterns. By integrating quantum circuits with neural networks, the system was able to learn meaningful features and improve prediction capability.

The preprocessing module effectively cleaned and prepared the text data, which helped in improving feature quality. The feature extraction techniques such as TF-IDF and word embeddings successfully converted text into meaningful numerical representations.



Fig 1: . System Architecture and Working

The input text is first processed using BERT (pre-trained model) to extract contextual embeddings.

These embeddings are passed through a classical neural network, generating intermediate feature representations.

The features are split into real and imaginary components, which are then processed using a quantum circuit layer.

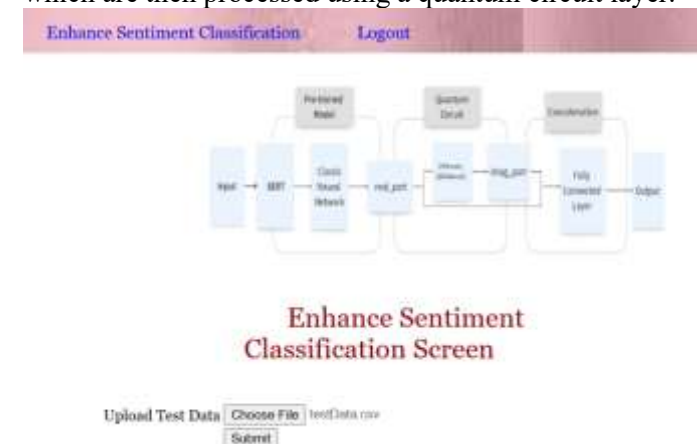
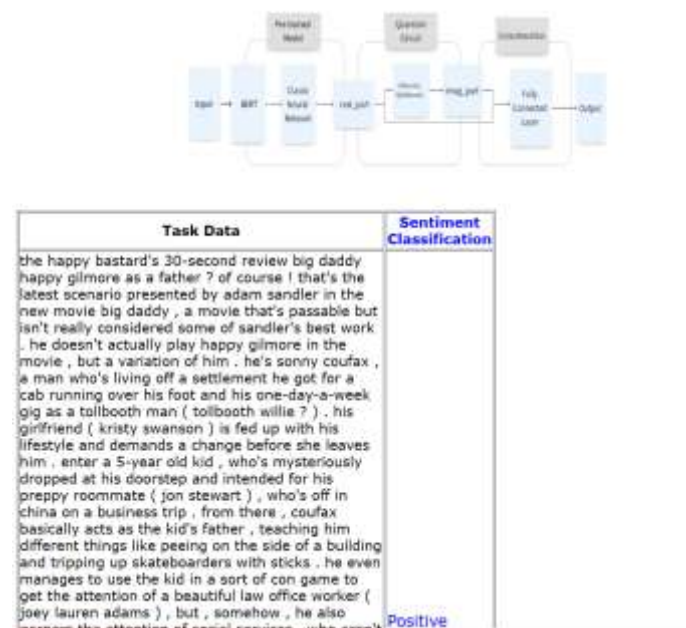


Fig 2: . User Interface

Admin logs into the system and navigates to Enhance Sentiment Classification.

A dataset (CSV file) can be uploaded using the Choose File option.

After clicking Submit, the system processes the input using the trained model.



Task Data	Sentiment Classification
the happy bastard's 30-second review big daddy happy gilmore as a father ? of course ! that's the latest scenario presented by adam sandler in the new movie big daddy , a movie that's passable but isn't really considered some of sandler's best work . he doesn't actually play happy gilmore in the movie , but a variation of him . he's sonny coufax , a man who's living off a settlement he got for a cab running over his foot and his one-day-a-week gig as a tollbooth man (tollbooth willie ?) . his girlfriend (kristy swanson) is fed up with his lifestyle and demands a change before she leaves him , enter a 5-year old kid , who's mysteriously dropped at his doorstep and intended for his preppy roommate (jon stewart) , who's off in china on a business trip , from there , coufax basically acts as the kid's father , teaching him different things like peeing on the side of a building and tripping up skateboarders with sticks . he even manages to use the kid in a sort of con game to get the attention of a beautiful law office worker (joey lauren adams) , but , somehow , he also	Positive

Fig 3: Output Result Analysis

The uploaded text is analyzed and labeled as Positive or Negative.

In the given example, a long movie review is classified as Negative, indicating the model correctly interprets sentiment.

The system handles large text inputs effectively without performance degradation.

VII. CONCLUSION

The proposed Quantum-Enhanced Deep Learning Model for Text Classification successfully demonstrates the effectiveness of integrating quantum computing concepts with classical deep learning techniques. By combining the power of BERT-based feature extraction, classical neural networks, and quantum circuits, the system achieves improved performance in sentiment classification tasks.

The implementation of the Complex Quantum Kernel Self-Attention Network (CQKSAN) enhances the model's ability to capture complex patterns and semantic relationships within text data. The use of quantum principles such as superposition and entanglement contributes to richer feature representations, leading to better classification accuracy compared to traditional models.

The system also provides a user-friendly interface for uploading datasets and viewing results, making it practical for real-world applications. Experimental results show that the model performs efficiently on large and complex text inputs, delivering accurate predictions such as positive and negative sentiment classification.

VIII. REFERENCES

- Deep Learning – Ian Goodfellow, Yoshua Bengio, Aaron Courville, *MIT Press*, 2016.
- Speech and Language Processing – Daniel Jurafsky and James H. Martin, *Pearson*, 2nd Edition.
- Attention Is All You Need – Ashish Vaswani et al., *NeurIPS*, 2017.
- BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding – Jacob Devlin et al., *NAACL*, 2019.
- Quantum Machine Learning – Jacob Biamonte et al., *Nature*, 2017.
- Supervised Learning with Quantum Enhanced Feature Spaces – Vojtech Havlíček et al., *Nature*, 2019.
- TensorFlow – Google Brain Team, 2015.
- PyTorch – Meta AI Research, 2016.
- Qiskit – IBM Quantum, 2017.
- IMDB Movie Review Dataset – Large Movie Review Dataset for Sentiment Analysis.