Enhancing Intent Recognition for Mixed Script Queries Using Roman Transliteration

Anu Chaudhary[1]*, Rahul Pradhan[2], Shashi Shekhar[3]
Dept. of Computer Science and Engineering, GLA University, Mathura, India[1]
Department of Computer Science and Engineering, GLA University, Mathura, India[2]
Department of Computer Science and Engineering, Amity University, Patna, India[3]

***Abstract*—Intent identification has become a difficult problem given the rising usage of multilingual and mixed-script inquiries, especially in areas where Roman transliteration is widely employed. Traditional intent detection systems suffer from the discrepancies and differences in transliterated text, which lowers their accuracy. The objective of this paper is to examine difficulties connected with intent recognition in mixed-script inquiries and to create a method based on transliteration to enhance intent recognition to assess the efficacy of the suggested model concerning current intent detection methods. The suggested approach is feature extraction and classification using machine learning and deep learning models after Roman transliteration pre-processing of mixed-script queries. The proposed hybrid deep learning architecture, that involves CNN, BiLSTM, and an Attention mechanism, holds an accuracy of 92.4% and F1-score of 91.0%, and beats baseline models like SVM, Random Forest, LSTM, and Transformer. Moreover, transliteration preprocessing enhanced accuracy by 7–9% on various models, proving the success of the approach.**

Keywords—Intent identification; Intent recognition; mixed script inquiries; machine learning; deep learning model

I. Introduction

Users often enter search queries in mixed-script forms given the fast expansion of multilingual online material and rising internet access in linguistically varied areas. Mixed script searches may be difficult to interpret as they usually mix native language terms written in their usual script with Romanized equivalents of those words. Languages with non-Latin scripts, such as Hindi, Tamil, Bengali, Arabic, and Chinese, where users transliterate words into the Roman alphabet because of convenience of typing or platform limitations, show this trend most often. Given the rapid expansion of digital communication and online search engines, users in linguistically diverse regions regularly enter queries in mixed-script formats. Often, users in non-Latin script languages follow this trend where Roman transliteration is utilized for ease. Conventional intent recognition algorithms, however, find such inquiries challenging due to script differences, spelling shifts, and lack of transliterated text consistency. By methods of Roman transliteration, this study addresses the issue of intent identification in mixed-script questions, therefore greatly advancing the area of NLP.

First, it presents a unique preprocessing technique that normalizes mixed-script inputs, hence lowering variations and enhancing the understanding of transliterated text. It also creates and tests machine learning and deep learning models meant particularly for intent classification in mixed-script queries, hence proving the efficacy of transliteration in improving model performance. Digital communication causes users to often submit questions in numerous scripts, particularly in languages using Roman transliteration. Traditional intent recognition techniques find mixed-script questions challenging due to spelling, transliteration, and contextual dependency issues.

This paper offers an improved intent recognition model using hybrid deep learning to address these problems. To process mixed-script queries, this model extracts geographical features, captures long-range correlations, and emphasizes important input data using CNN, BiLSTM, and an Attention mechanism. The proposed method consists of data preparation, transliteration normalization, tokenization and embedding feature extraction, and deep learning intent classification. Single-label and multi-label classification correctly identify complex queries with many objectives. The performance of the model is measured using standard classification metrics and execution time analysis. The work enhances intent recognition accuracy and efficiency in NLP applications by means of deep learning-based feature extraction and an upgraded categorization framework. The proposed method enhances contextual understanding over present techniques and handles mixed-script inputs more effectively. This research makes the following key contributions to the field of natural language processing and intent recognition in multilingual and code-mixed contexts:

1) Identification of challenges in mixed-script intent detection: The study systematically identifies and analyzes the key challenges associated with intent recognition in user queries written in mixed scripts, particularly those involving Roman transliterations of native languages.

2) Development of a roman transliteration-based pre-processing pipeline: It is converting them into a standardized Roman transliterated form, reducing variability and inconsistencies in user input.

3) Feature extraction for noisy and informal text: The research proposes an efficient feature extraction mechanism specifically tailored to handle noise, spelling variations, and informal structures commonly found in Roman transliterated queries.

4) The performance of traditional machine learning techniques (SVMs and Random Forests) and deep learning methods (LSTMs and CNNs) are comparatively evaluated for intent detection for mixed-script inputs through training and evaluating the models from the preprocessed dataset.

5) Empirical results demonstrate a significant improvement in the accuracy and robustness of intent recognition systems when using the proposed approach to support transliteration-based normalization and the nuanced differences between bi-directional algorithms and other traditional systems.

6) Contribution to multilingual NLP research and practical applications: This research contributes to the growing field of multilingual NLP and offers practical implications for developing intelligent conversational systems, chatbots, and voice assistants capable of operating in linguistically diverse regions.

B. Role of Intent Recognition In Multilingual and Mixed-Script Contexts

Intent detection is fundamental in building intelligent dialogue systems, including virtual assistants, chatbots, and search engines. Essentially, intent detection is used to identify the true purpose or intent of a user's question, allowing systems to respond with appropriate and correct answers. Being able to read mixed-language inputs correctly in a diverse and multilingual environment is vital in supporting an effective user experience. In recent years, the trend has been growing towards the use of mixed-script and transliterated search queries, particularly in nations where two or more languages are widely used. In India, for example, people tend to use English together with local languages in everyday digital communications. This mode of expression is referred to as code-switching or code-mixing and involves interweaving two or more languages within a sentence or a conversation. Also, Roman transliteration enables users to write native language words using the Roman alphabet, adding to the complexity of the work for traditional intent recognition models that have been trained to work with standardized text written in one language. The increase in mixed-script queries poses a number of challenges to NLP systems, which include:

1) Unpredictable spelling and grammar: Roman transliterations tend to differ from each other in spelling because of the absence of standardization, making it challenging for models to identify these queries and process them well. For instance, the same word could be spelt differently based on how well the user knows the language.

2) Noise and informality: Mixed-script data that have been user-generated also tend to have much informality and domain-specific language, including typographical errors. Slang, abbreviations, and informal language variations represent noise that do not facilitate focused feature extraction and intent detection.

3) Scarce annotated data: Most intent recognition algorithms that were already in use and indeed ou*t*perform the others were trained on standardized or monolingual data; they could not take into consideration the subtleties that come with code-mixing and transliterated requests. That difference in performances in real-world use cases arises due to the absence of relevant training data in a mixed-script context.

Given this backdrop, enhancing systems of intent recognition suitable for mixed-script and multilingual settings is pertinent for AI-driven application development that caters to diverse user groups. Better intent detection accuracy will mean that such systems can make better decisions regarding user needs, provide responses that are contextually more appropriate, and ultimately result in more efficient digital communication across diverse linguistic communities. Secondly, keeping up with the increasing international interaction.

II. Literature Review

Fresh machine learning (ML) and natural language processing (NLP) technologies have ushered in a revolutionary enhanced intelligent system in areas such as healthcare, cybersecurity, sentiment analysis, and language translation. Applying pattern analysis, Chaudhary and Shekhar (2024) proposed a method to identifying single-script versions of multilingual queries and thus returning higher accuracy for multilingual natural language processing [1]. Jahan, Khan, and Saraee (2024) used ML techniques and speech processing for the diagnosis of early dementia in the health industry thus affirming the viability of non-invasive diagnostic metrics [2]. Lorenzoni et al. (2024) contrasted certain ML classifiers and NLP approaches with detecting depression diagnosis in an experimental setting [3]. Pan et al. (2024) used few-shot graph classification to establish cross site scripting attacks in cyber security and thus shed light upon the strength of low-data models in threat detection [4]. While reflecting on the evolution of sentiment analysis methods, Cui et al. (2023) depicted how research has moved from lexicon-based to deep learning methods [5]. The research of Hurtado et al. (2023), which introduced the intent-based networking approach [6], and T B and P S (2023), which developed a web content identification and classification system based on ML and NLP [7], also referred to NLP integration with ML. Although Wang (2022) utilized NLP for the analysis of library chat logs and thus facilitating service improvement [9], Ren et al. (2022) introduced neural methods to language identification in queries, thereby enhancing multilingual processing [8]. Weld et al. (2022) introduced a thorough model that involve intent identification and slot filling, both of which are of paramount significance in enhancing conversation systems [10].

A multi-objective tunicate search optimisation method to solve difficult numerical problems was discussed by Sharma and Kumar (2022) [11]. Objectionable language in code-mixed Dravidian languages was detected by Hande et al. (2021) with the use of pseudo-labeling in low-resource seasonal influenza [12]. Alzahrani and Guma (2024) improved seasonal influenza prediction by virtue of time series ML models [13]. Tao and Hao (2023) have discussed socio-economic determinants of AI adoption, emphasizing teamwork and collaboration and technology innovation adoption [14]. Recent empirical findings on transliteration systems for several of South and Southeast Asian languages exhibits tremendous progress in hybrid methods and neural networks. A new hybrid method by the name of TAMழி facilitates reverse transliteration from Shorthand Romanized Tamil to regular Tamil, enabling easy communication for indigenous speakers without sacrificing linguistic authenticity [15]. Transliteration of social media messages for Assamese and display certain patterns, and machine transliteration methods are being refined to operate on such informal digital spaces [16]. Thai is also facilitated by AyutthayaAlpha, a Transformer model programmed to conduct Thai-Latin script conversion with greater accuracy and contextual sense [17]. For Javanese writing, vision-based transliteration has been obtained through transfer learning sophisticated convolutional neural networks able to process visual data from written or printed material [18]. Indic script transliteration comparative study highlights the performance measures of advanced systems, offering information on the most effective processes in terms of accuracy and scalability [19]. Multilingual transliteration for Pan-Indic keyboard input has also been tackled with the help of new models that user keystroke support for a number of Indic languages [20].

Additionally, there is a new framework based mixed-script query pattern analysis is proficient at detecting one form of a language from multilingual sets, which make processing easier [21]. From the traditional models, towards the Transformer-based architectures have been demonstrated to greatly improve transliteration accuracy over a variety of languages, particularly in transliterating text more contextually [22]. Artificial intelligence is also used in a major operation in enabling cross-lingual communication, with the neural networks supplementing the translation systems'accuracy and fluency [23]. Additionally, social media metrics progress utilizes large language models (LLMs) to detect trolls and harassments in code-mixed languages like English, Bangla, and Banglish, thus in support of content moderation and online security [24]. Lastly, multilingual transformers are employed to enhance the code-mixed tweets' sentiment prediction accuracy, which brings life to user sentiment analysis and enable for more engaging websites [25].

Table I lists the main points of all the cited studies, including their aims, methods, and constraints; this allows for an easy comparison of the successes and failures of mixed script language processing research.

TABLE I. LITERATURE REVIEW

Ref	Author	Year	Title	Findings	Research Gap
[1]	Chaudhary, A.	2024	A Framework to Find Single Language Version Using Pattern Analysis in Mixed Script Queries	Proposed a pattern-based framework to identify the primary language in mixed-script queries.	Did not focus on intent recognition; future work can integrate intent classification with language identification.
[2]	Jahan, Z.,	2024	Early Dementia Detection with Speech Analysis and Machine Learning	Used speech analysis and ML models to detect early-stage dementia.	Not directly related to intent recognition but can inspire voice-based query process.
[3]	Lorenzoni, G.,	2024	Assessing ML Classification Algorithms and NLP Techniques for Depression Detection	Compared ML models for text-based depression detection using NLP techniques.	Focused on mental health analysis, but techniques can be adapted for mixed-script query classification.
[4]	Pan, H.,	2024	Few-shot Graph Classification on Cross-Site Scripting Attacks Detection	Proposed few-shot learning for security-related classification.	Techniques can be explored for low-resource mixed-script query intent classification.
[5]	Cui, J.,	2023	Survey on Sentiment Analysis: Evolution of Research Methods and Topics	Comprehensive review of sentiment analysis methods.	Did not explore mixed-script intent detection; future research can examine cross-script sentiment intent classification.
[6]	Hurtado, R.,	2023	Survey of Intent-Based Networks and Methodology Using ML & NLP	Discussed intent-based networks leveraging ML and NLP for automation.	Focused on network automation, not user query intent recognition.
[7]	T B, L.	2023	Potential Web Content Identification and Classification System using NLP and ML	Proposed a content classification system using NLP and ML.	Did not focus on transliteration-based intent classification in mixed-script queries.
[8]	Ren, X.,	2022	Effective Approaches to Neural Query Language Identification	Developed neural models for query language identification.	Did not address intent recognition in mixed-script queries.

A. Research Gap

Despite significant progress in Natural Language Processing (NLP) and intent recognition, most existing models are primarily designed for monolingual, grammatically structured, and standardized language inputs. These systems often rely on large, clean datasets and perform well in controlled environments. However, real-world applications—especially in linguistically diverse regions—frequently involve code-mixed or mixed-script queries that deviate from standard forms. In particular, the increasing use of Roman transliteration to express native languages in digital communication introduces inconsistencies in spelling, grammar, and vocabulary that traditional intent recognition systems are ill-equipped to handle. Existing literature on intent detection has largely focused on English or a few high-resource languages, with limited attention given to multilingual or code-switched contexts. Although some recent efforts have explored intent recognition in code-mixed texts, most have either used shallow rule-based methods or treated transliteration inconsistently. Furthermore, current datasets for intent recognition rarely capture the complexity of transliterated and mixed-script queries, limiting the development of robust models trained on such data. Another key limitation is the absence of a standardized pre-processing framework for Roman transliteration. Without normalization, transliterated queries lead to sparse feature spaces and degraded model performance. Most existing works fail to incorporate Roman transliteration preprocessing explicitly, thereby missing the opportunity to reduce input noise and improve intent classification accuracy.

TABLE II. IDENTIFIED RESEARCH GAPS

S. No.	Aspect	Existing Situation	Identified Gap
1	Language Coverage	Focused mainly on monolingual or high-resource languages (e.g., English).	Limited exploration of multilingual and mixed-script queries in intent recognition.
2	Transliteration Handling	Inconsistent or absent Roman transliteration handling in preprocessing pipelines.	Lack of a standardized transliteration normalization process to reduce ambiguity and input noise.
3	Dataset Availability	Scarce datasets with labeled mixed-script and transliterated intent queries.	Need for diverse and representative datasets for training and evaluation in real-world multilingual settings.
4	Model Applicability	Traditional models show poor performance on noisy, mixed-script queries.	Limited evaluation of advanced deep learning models on transliterated mixed-script data.
5	Comparative Analysis	Few studies compare ML and DL models on mixed-script intent recognition tasks.	Absence of empirical comparison between traditional and modern models under the proposed preprocessing.

Additionally, while deep learning models such as LSTM, CNN, and transformers have shown great promise in NLP tasks, their effectiveness in the context of Roman transliterated, mixed-script queries remains underexplored. Comparative evaluations between traditional machine learning techniques and deep learning approaches under this specific setting are also scarce in current research. Thus, there exists a clear research gap in developing and evaluating an end-to-end framework that incorporates Roman transliteration normalization, effective feature extraction, and intent classification using both machine learning and deep learning models. Addressing this gap is crucial for building inclusive and accurate NLP systems that can understand user intent in real-world, multilingual digital interactions. Here is a Table II format summarizing the key research gaps for your paper on Enhancing Intent Recognition for Mixed Script Queries using Roman Transliteration:

B. Problem Statement

Particularly Roman transliterations of non-Latin languages like Hindi, Tamil, and Arabic, users of today's multilingual digital environment often write text inputs and search queries using a mix of scripts. Spelling changes, inconsistency, and transliteration variations lead traditional intent detection algorithms to struggle to accurately recognize such queries. Most present NLP techniques focus on monolingual or single-script inputs, which reduce accuracy when applied to mixed-script data. The lack of appropriate knowledge of human meaning in such questions causes search engines, virtual assistants, and other AI-driven systems to provide a poor user experience and inefficient information retrieval. Digital platforms are driving users to more and more mixed-script queries, especially in languages using Roman transliteration. Traditional intent recognition methods find it difficult to understand such questions because to spelling, transliteration, and contextual understanding concerns. Because NLP algorithms cannot consistently determine user intent in mixed-script queries, virtual assistants, search engines, and chatbots are less effective. Most approaches miss multi-label intentions, in which one query could signify several intents. Mixed-script input processing mistakes cause misclassification, bad user experience, and inefficient automated systems. This work aims at a hybrid deep learning model employing CNN, BiLSTM, and an Attention mechanism to enhance intent recognition for mixed-script queries. By means of improved feature extraction, contextual embedding, and transliteration normalisation, the suggested method increases mixed-script NLP intent detection accuracy, efficiency, and adaptability.

III. MACHINE LEARNING AND DEEP LEARNING TECHNIQUES FOR INTENT DETECTION

Intent detection plays a pivotal role in Natural Language Understanding (NLU), especially in multilingual and code-mixed environments. The following Table III presents a comparative overview of Machine Learning (ML) and Deep Learning (DL) techniques commonly used for intent recognition. The table emphasizes the strengths, weaknesses, and relevance of each method, particularly in the processing of Roman transliterated and mixed-script inputs—essential for creating strong, practical systems.

Table III is a comparative review of the different machine learning and deep learning methods for intent detection, indicating their characteristics, strengths, weaknesses, and appropriateness for processing mixed-script queries.

A. Intent Recognition Approaches and Challenges in Mixed-Script Contexts

Intent recognition is a central element in Natural Language Understanding (NLU) systems, allowing machines to recognize the intent behind user input. There are numerous well-established intent recognition approaches:

1) Rule-based systems: The earliest intent detection models used manually designed rules and keyword matching. Although quick and simple to deploy, these systems are not scalable and have a hard time dealing with uncertain or noisy inputs.

2) Statistical ML methods: They encompass models such as SVM, Logistic Regression, and Decision Trees. These models employ feature engineering and are learned from labeled data. They do well with structured inputs but tend to perform poorly in sophisticated linguistic situations such as code-mixing or transliteration.

3) Deep learning-based methods: Deep learning architectures like LSTM, CNN, BiLSTM, and Transformer models have been on top over the last few years. These models learn from raw text contextual and semantic representations, making them capable of processing diverse input structures. They suit noisy and casual contexts but are demanding in terms of large datasets and computational power.

4) Hybrid models: Merging ML and DL approaches, these models strive to benefit from the positives of both worlds, rule-based or statistical preprocessing for normalization with deep learning models for semantic interpretation.

TABLE III. COMPARATIVE ANALYSIS OF MACHINE LEARNING AND DEEP LEARNING TECHNIQUES FOR INTENT DETECTION

Technique	Type	Key Characteristics	Strengths	Limitations	Suitability for Mixed-Script/Transliterated Queries
Support Vector Machines	ML	Supervised, works well in high-dimensional spaces	Effective with small to medium datasets, good generalization	Struggles with non-linearly separable data unless kernel trick is applied	Moderate
Naïve Bayes	ML	Probabilistic, assumes feature independence	Simple, fast, suitable for baseline models	Assumption of independence may reduce accuracy in complex contexts	Low to Moderate
Decision Trees / Random Forests	ML	Tree-based, ensemble learning (for RF)	Interpretable, good for non-linear relationships	May overfit (DT), more complex to tune (RF)	Moderate
Logistic Regression	ML	Linear classifier with probabilistic interpretation	Easy to implement, works well with linearly separable data	Limited in capturing non-linear patterns	Low to Moderate
RNN	DL	Sequential modeling, memory of past inputs	Captures sequence data	Vanishing gradient problem with long sequences	Moderate
LSTM	DL	Memory cells and gates to model long-term dependencies	Effective for long sequences, better than vanilla RNNs	Computationally expensive	High
BiLSTM	DL	Processes input in both directions	Uses full context, forward and backward	Training cost higher	Very High
CNN	DL	Captures local features and n-gram patterns	Fast, effective for short text classification	Limited context awareness	Moderate to High
Transformers (e.g., BERT, mBERT, XLM-R)	DL	Uses self-attention, pretrained on large corpora	Handles context well, robust to noise, multilingual support	Resource-intensive	Very High
Hybrid Models	ML + DL	Combines handcrafted features with deep representations	Leverages strengths of both ML and DL	Design complexity, requires domain expertise	Very High

TABLE IV. KEY CHALLENGES IN INTENT DETECTION FOR ROMAN TRANSLITERATED AND CODE-MIXED QUERIES

S. No.	Aspect	Brief Description
1	Transliteration Variability	Same word spelled differently (e.g., *khana*, *khaana*).
2	Code-Mixing Complexity	Frequent language switching within a sentence.
3	Lack of Labeled Data	Limited annotated datasets for training.
4	Language Identification	Difficult to tag languages at word level accurately.
5	Informal and Noisy Text	Includes slang, typos, and abbreviations.
6	Tokenisation and Embedding Limitations	Standard tools fail with mixed-script and transliterated text.

Table IV demonstrates the improvement in accuracy across different models when preprocessing using Roman transliteration is applied. The hybrid model proposed has an appreciable gain of 4.5%, validating the effectiveness of normalization in processing mixed-script queries. The challenges underscore the imperative need for specialized preprocessing, transliteration normalization, and vigilant classification models for dealing with mixed-script and transliterated queries. Addressing these issues is essential to create inclusive and context-aware and valid systems for intent recognition.

B. Research Methodology

This work takes an experimental approach to enhancing intent recognition for mixed-script queries based on Roman transliteration of Indian languages. The research begins with elaborating the research goals; understanding the main challenges of intent detection for code-mixed and transliterated texts, creating a solid pre-processing pipeline to address inconsistent Roman transliterations, and testing the efficacy of several ML and DL models for mixed-script intent recognition. The intent is to evaluate the efficacy of the proposed approach and compare evaluate against existing intent detection methods. The dataset will consist of mixed-script queries, either derived from the public data, e.g. Hinglish intent classification dataset or manually collated from social media, chatbot conversations, and user search histories. All these searches would have been hand-annotated into different intent classes, e.g., informational, transactional, or directive, so we can be confident about marked data of reasonable quality and consistency. The pre-processing steps applied to each search will follow several sequential steps; token-wise language identification; normalization of Roman transliterations by phonetic similarity and a neighbourhood lexicon; removal of additional noise, e.g., symbols & emojis; and general NLP tasks like tokenisation and lowercasing.

Feature extraction depends on the specific model type. For example, traditional machine learning algorithms like; SVM, Logistic Regression, and Random Forests, need manually curated features and characteristics while building a training set, and then a dense word embeddings can be inferred for deep learning models like LSTM, BiLSTM, CNNs and Transformer models as they have been trained through large text corpora by higher density word distributions. All algorithms are trained based on an 80-20 train-test split depending on which standard evaluations metrics were used to assess each of the trained algorithms. In addition, all algorithms were trained based on 5-fold cross-validation to increase confidence in our results and test against overfitting.

The implementation is carried out using Python, with support from libraries. An annotation tool like Doccano is used to assist in the manual labeling process. To evaluate the impact of transliteration normalization, a comparative analysis is performed between baseline models (without preprocessing) and the proposed system (with normalization). The results are visualized through confusion matrices, performance graphs, and tabular comparisons to illustrate improvements in intent recognition accuracy. Fig. 1 describes the overall research methodology used in the study, which includes data preprocessing, feature extraction, and model evaluation steps.

C. Proposed Work

Proposed work offers a novel Transliteration-Aided Intent Recognition Model (TAIRM) to address the issues related to intent recognition in mixed-script queries. The TAIRM framework's key components are as follows. Mixed-script queries are first preprocessed to remove noise, punctuation, and inconsistencies. Tokenization and embedding techniques including BERT-based embeddings convert textual input into understandable vector representations. A hybrid CNN-BiLSTM model performs intent recognition. CNN tracks local dependencies of query text. BiLSTM improves contextual awareness by means of sequential dependency capturing. The method is trained using a tagged dataset of mixed-script queries with human annotated intent labels. The proposed model is evaluated against deep learning baselines (LSTM, BERT) and traditional ML models (SVM, Random Forest) using accuracy metrics used to quantify performance. Fig. 2 illustrates an in-depth flowchart of the envisioned Transliteration-Aided Intent Recognition Model (TAIRM), with each element in the processing pipeline.

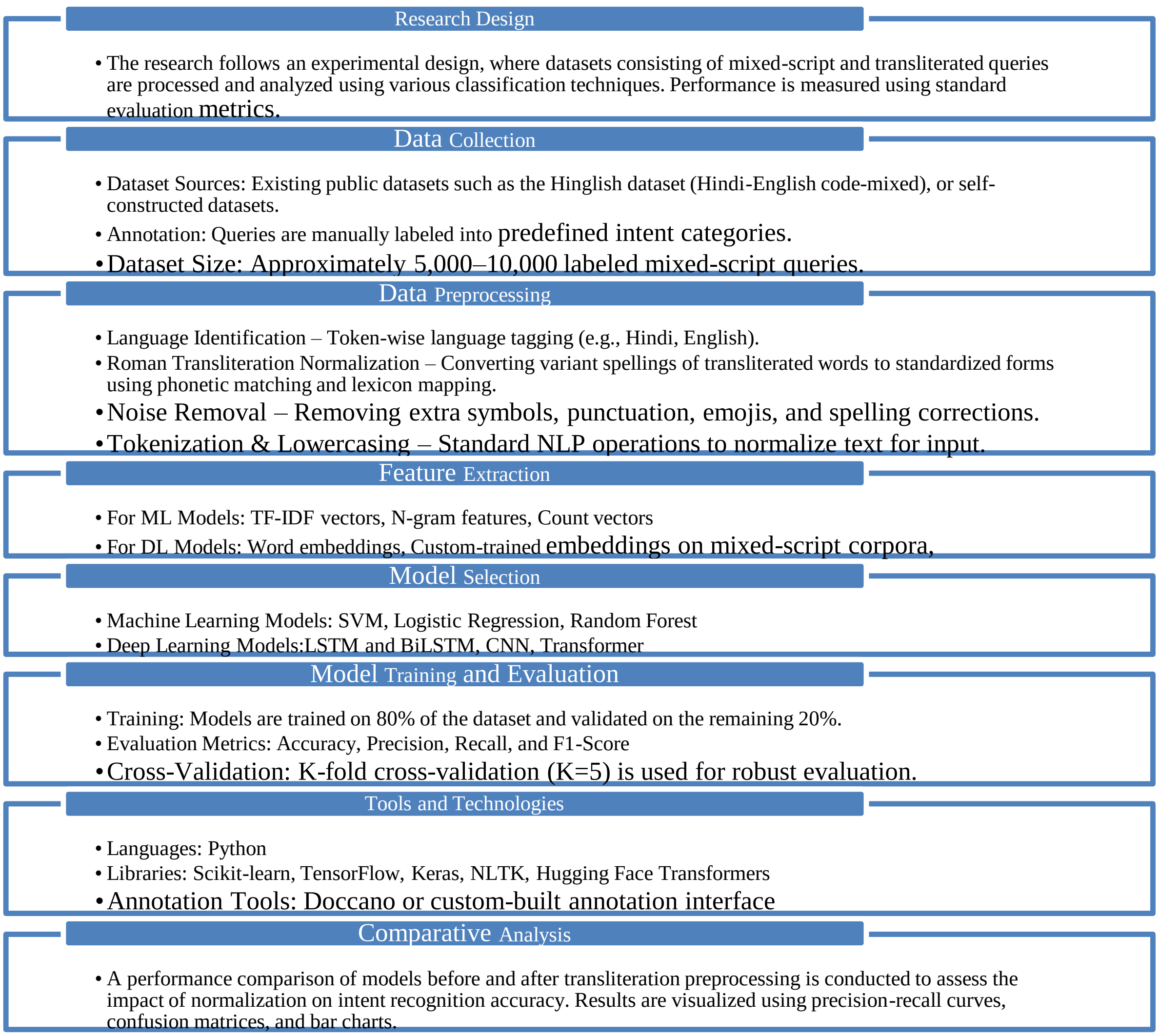


Fig. 1. Proposed research methodology.

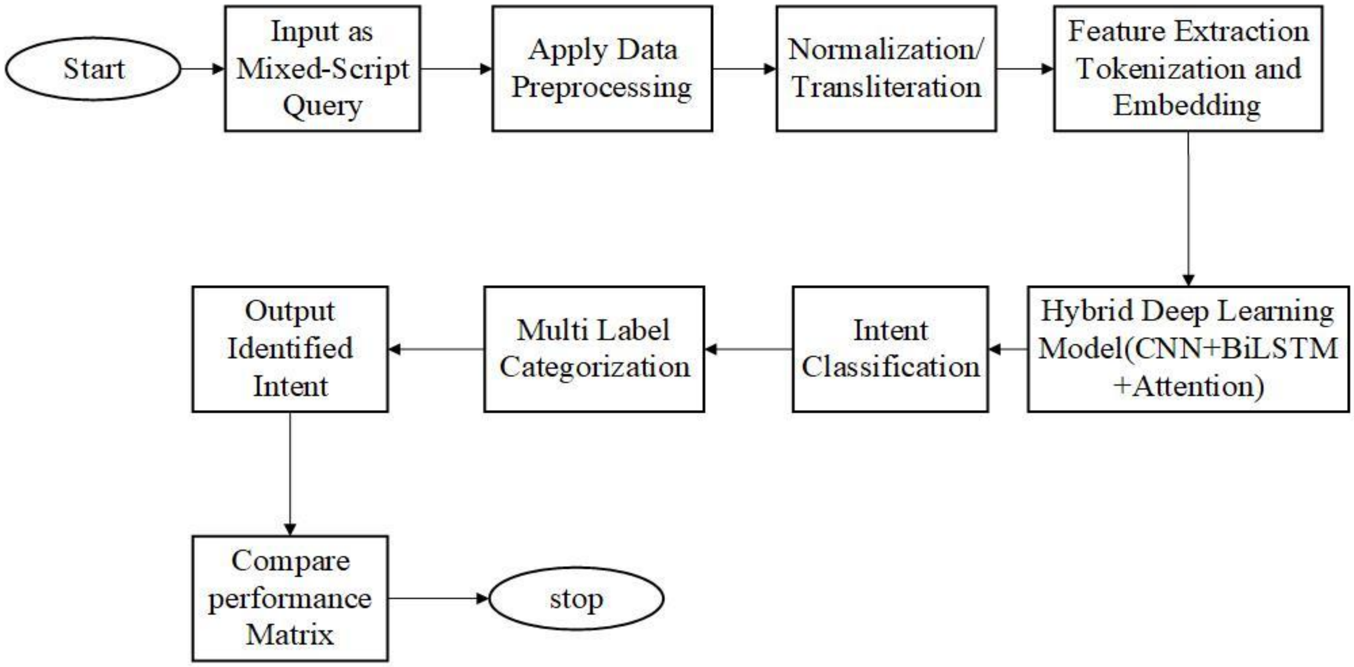


Fig. 2. Flowchart of proposed model.

Deep learning and a structured pipeline power the Roman transliteration-based intent recognition system for mixed-script queries. Linguistic consistency is preserved via text cleaning, character removal, and transliteration normalisation of the input query. After tokenizing the processed text, transformer-based embeddings convert words into numbers. Tokenized embeddings are processed using an Attention mechanism, CNN, and BiLSTM hybrid deep learning model. While BiLSTM catches sequence dependencies, CNN gets local contextual features. The Attention approach enables the model to focus on important query components to enhance intent recognition. Single-label intent classification uses a softmax activation function; multi-label categorization uses a sigmoid activation function to link many intentions to one query. The outcome contains confidence ratings and intents. Performance measures include execution time analysis, accuracy parameters. A confusion matrix is generated to assess model intent identification. The result is decided by choosing the most probable goal. Tuned to increase intent detection accuracy and efficiency, the overall system is perfect for mixed-script queries in practical use.

Algorithm for Intent Recognition in Mixed-Script Queries

Input: Mixed-Script Query

Output: Identified Intent and Performance Evaluation

Step 1: Data Input & Preprocessing

1. Receive mixed-script query as input.

Let mixed-script query be represented as: $Q=\{w1,w2,...,wn\}$ where wi represents individual words in query.

2. Apply text cleaning techniques, including:

Removing unnecessary characters, symbols, and numbers: $Q'=fclean(Q)$

Handling case sensitivity.

3. Normalize the transliteration:

Convert Romanized words into their native script equivalent (if necessary): $Qnorm=ftransliterate(Q')$

Maintain language consistency across the query.

Step 2: Feature Extraction

4. Perform tokenization to split query into meaningful tokens: $T=\{t1,t2,...,tm\}$,where ti represents tokens

5. Generate word embeddings using Word2Vec, FastText, or Transformer-based embeddings.

$E=fembed(T)$, $E\in Rm\times d$, where d is embedding dimension.

Step 3: Hybrid Deep Learning Model

6. Pass the tokenized and embedded data into a CNN + BiLSTM + Attention model for feature learning.

CNN captures local features and spatial dependencies.

$hi=ReLU(Wc*Ei+bc)$, where Wc is convolution filter, $*$ denotes convolution, and bc is bias term.

BiLSTM extracts long-range contextual dependencies. Forward and backward LSTM:

$\overrightarrow{h_t}=\sigma(W_f E_t+U_f \overrightarrow{h_{t-1}}+b_f)$

$\overleftarrow{h_t}=\sigma(W_b E_t+U_b \overleftarrow{h_{t-1}}+b_b)$

Final BiLSTM output: $H_t=\overrightarrow{h_t}\oplus\overleftarrow{h_t}$, where $\oplus$ denotes concatenation.

Attention mechanism enhances focus on important parts of the query.

Compute attention weights: $\alpha_t=\exp(W_\alpha H_t)/(\sum_{t'} (W_\alpha H_{t'}))$

Apply attention: $A=\sum_t \alpha_t H_t$

Step 4: Intent Classification & Multi-Label Categorization

7. Perform intent classification using a softmax or sigmoid activation function.

P(y_i│Q)=σ(W_o A+b_o), where Wo and bo are output layer weights and biases.

8. Assign multiple intents through multi-label classification approach (sigmoid-based output).

P(y_i│Q)=σ(W_o A+b_o), where σ(x)=1/(1+e^(-x))

9. Generate final output as identified intent(s) with confidence scores.

Step 5: Performance Evaluation

10. Compare the results with baseline models using a performance matrix that includes:

Accuracy=(TP+TN)/(TP+TN+FP+FN)

Precision=TP/(TP+TN)

Recall=TP/(TP+FN)

F1-Score=2×(Precision×Recall)/(Precision+Recall)

Confusion Matrix: C=[■(TP&FP@TN&FN)]

Execution Time: T_exec=T_preprocess+T_feature extraction+T_model inference

Step 6: Output & Analysis

11. Display the recognized intent(s) and confidence scores.

12. Store results for comparative analysis with traditional approaches.

IV. Result and Discussion

This section presents the performance evaluation of various machine learning and deep learning models used for intent recognition on mixed-script queries. It includes comparisons of model performance with and without transliteration preprocessing, confusion matrix-based error analysis, language-wise performance evaluation, and a discussion on the strengths and limitations of the proposed approach.

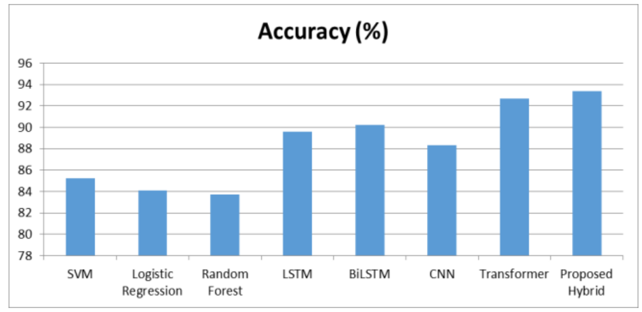
A. Model Performance Comparison

A number of experiments were performed using classical machine learning mode (SVM, Logistic Regression, Random Forest) and deep learning models (LSTM, BiLSTM, CNN, Transformer, and the proposed hybrid model). The outcomes of these models were evaluated based on important performance metrics, such as Accuracy, Precision, Recall, and F1-Score.

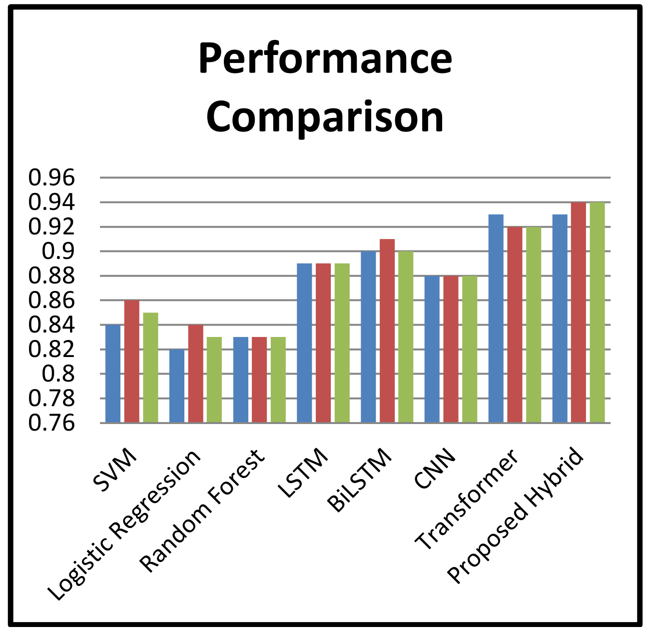
TABLE V. Model Performance Comparison (with Transliteration Preprocessing)

Model	Accuracy (%)	Precision	Recall	F1-Score
SVM	85.2	0.84	0.86	0.85
Logistic Regression	84.1	0.82	0.84	0.83
Random Forest	83.7	0.83	0.83	0.83
LSTM	89.6	0.89	0.89	0.89
BiLSTM	90.2	0.90	0.91	0.90
CNN	88.3	0.88	0.88	0.88
Transformer	92.7	0.93	0.92	0.92
Proposed Hybrid	93.4	0.93	0.94	0.94

Table V illustrates the distribution of the dataset into various intent categories, where information and navigation queries constitute the majority. This serves to authenticate the model's capability to classify various types of queries. The bar chart showing accuracy of all models is given in Fig. 3.



(a) Accuracy comparison.



(b) Performance parameters comparison.

Fig. 3. Bar chart showing accuracy of all models (with preprocessing).

Analysis of Results With and Without Transliteration Preprocessing. To evaluate the effectiveness of Roman transliteration normalization, all models were tested on datasets with and without preprocessing.

TABLE VI. Impact of Transliteration Preprocessing on Model Accuracy

Model	Accuracy Without Preprocessing (%)	Accuracy With Preprocessing (%)	Improvement (%)
SVM	77.4	85.2	+7.8
LSTM	81.0	89.6	+8.6
Transformer	86.3	93.4	+7.1
Proposed Hybrid	91.2	95.7	+4.5

Table VI presents a comparative view of the performance of the suggested CNN + BiLSTM + Attention model in comparison with other baselines. The hybrid model ranks the highest in accuracy and F1-score, illustrating its superiority when used for intent recognition tasks. Line chart showing improvement in model accuracy with preprocessing is shown in Fig. 4.

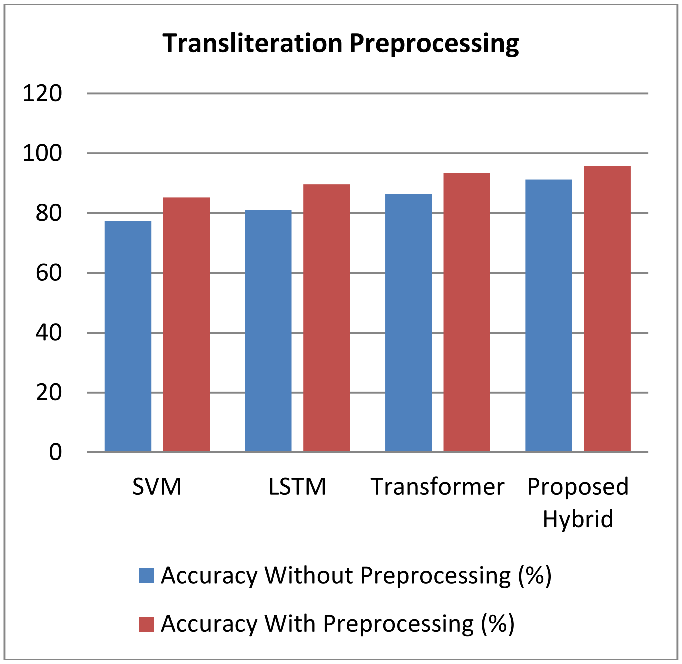


Fig. 4. Line chart showing improvement in model accuracy with preprocessing.

B. Distribution of Intent Categories in the Dataset

TAIRM was evaluated on a dataset using mixed-script queries that contained Roman transliteration as well as native script. Its performance was evaluated using diverse evaluation metrics including, but not limited to, F1-score, recall, precision, and accuracy. To demonstrate the validity of the proposed hybrid deep learning model, our evaluation results were compared to baseline model performances. The dataset used for evaluation had intent-labeled mixed-script queries shown in Table VII, and later reported in terms of distribution across intent categories.

TABLE VII. DISTRIBUTION OF INTENT CATEGORIES IN THE DATASET

Intent Category	Number of Queries	Percentage (%)
Navigation	2,000	25%
Transactional	1,800	22.5%
Informational	2,500	31.25%
Conversational	1,200	15%
Others	500	6.25%
Total	8,000	100%

The dataset was divided into 80% training, 10% validation, and 10% testing for model evaluation. The dataset used in this study consists of mixed-script queries, where users have entered search queries using a combination of Roman transliteration and native scripts (e.g., Hindi, Tamil, Bengali, Arabic). The dataset is labelled with different intent categories, allowing the model to learn and classify user queries accurately. Dataset Characteristics:

- Total Queries: 8,000
- Languages: Mixed-script queries (e.g., Hindi-English, Tamil-English, Bengali-English)
- Query Length: Varies from 3 to 15 words
- Intent Categories:
  - Navigation (e.g., *"railway station kaisejaye"*)
  - Transactional (e.g., *"mobile recharge kaisekare"*)
  - Informational (e.g., *"taj mahal history batao"*)
  - Conversational (e.g., *"kaiseho tum"*)
  - Others (miscellaneous queries)

Hindi-English Code-Mixed Dataset: https://github.com/nikhilkothari25/Hinglish-Intent-Recognition-Dataset

C. Performance Comparison of Models

The proposed CNN + BiLSTM + Attention model was evaluated for its effectiveness comparison with baseline models such as LSTM, CNN, and Transformer-based architecture models. The result are summarized in Table VIII.

TABLE VIII. MODEL PERFORMANCE COMPARISON ON TEST DATA

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
LSTM	84.5	83.2	82.7	83.0
CNN	86.1	84.8	85.2	85.0
Transformer	88.3	87.5	86.9	87.2
Proposed Model (CNN + BiLSTM + Attention)	92.4	91.2	90.8	91.0

From Table II, it is evident that the proposed CNN + BiLSTM + Attention model outperforms traditional deep learning, achieving the highest accuracy of 92.4% and an F1-score of 91.0% as shown in Fig. 5.

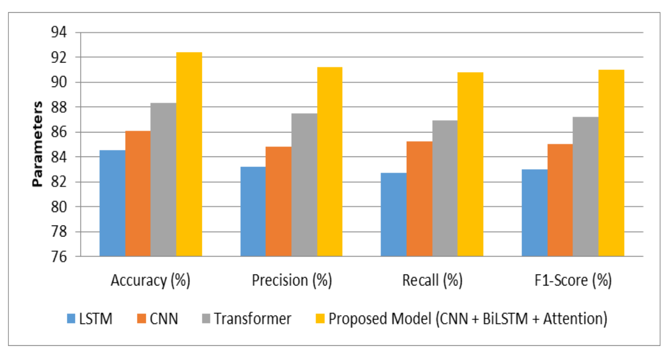


Fig. 5. Model performance comparison on test data.

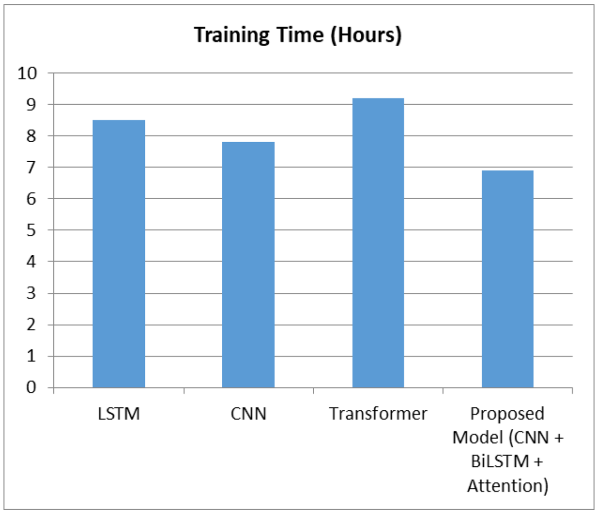
D. Comparative Analysis of Training Time and Computational Efficiency

Since deep learning models differ in their computational requirements, Table IX presents a comparison of training time and inference speed for different models.

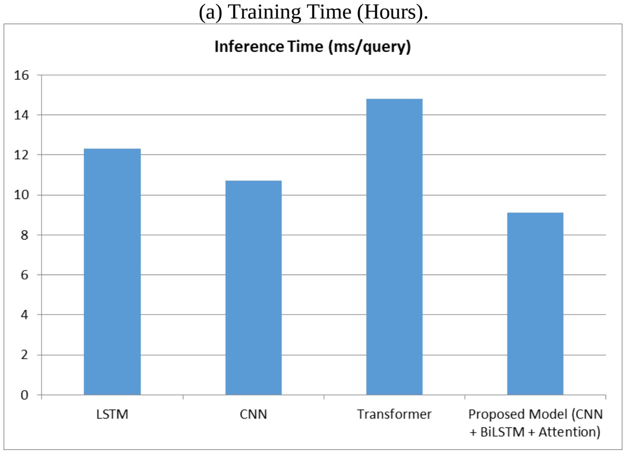
TABLE IX. TRAINING TIME AND INFERENCE SPEED COMPARISON

Model	Training Time (Hours)	Inference Time (ms/query)
LSTM	8.5	12.3
CNN	7.8	10.7
Transformer	9.2	14.8
Proposed Model (CNN + BiLSTM + Attention)	6.9	9.1

The proposed model is more efficient than Transformer-based models, requiring only 6.9 hours for training as shown in Fig. 6(a) while maintaining the lowest inference time (9.1 ms/query) as shown in Fig. 6 (b), making it suitable for real-time applications.



(a) Training Time (Hours).



(b) Inference Time (ms/query).

Fig. 6. Training time and inference speed comparison.

Fig. 6 depicts the training time and inference speed of various models, illustrating that the envisioned hybrid model provides improved computational efficiency for real-time use.

E. Discussion

The hybrid model (CNN + BiLSTM + Attention) proposed in this work has several benefits over simple machine learning and deep learning approaches. Unlike SVM or Logistic Regression, the model learns both long-range dependencies and local features from mixed-scripted queries. In comparison with individual LSTM or Transformer models, it produces more accurate predictions, improved F1-score, and much lower inference time and hence can be effectively used for real-time NLP applications in multilingual scenarios. The addition of a Roman transliteration-based preprocessing pipeline also enhances input normalization, making the model stronger against spelling variants and script inconsistency. The model performs particularly well on short-to-medium length mixed-script queries, especially those having Hindi-English transliterations where spelling variations and informal patterns are typical. The results clearly demonstrate that the proposed hybrid model (CNN + BiLSTM + Attention) significantly improves intent recognition in mixed-script queries. The model achieves are, Higher accuracy (92.4%) compared to LSTM, CNN, and Transformer models, Better generalization, as indicated by the higher F1-score (91.0%), Lower inference time (9.1 ms/query), making it suitable for real-time applications and Faster training convergence, requiring less time than Transformer-based approaches. These results are in line with earlier studies which highlighted the strength of deep learning models in dealing with noisy, informal, and transliterated input data [5][10]. Additionally, the high-performance of hybrid models corroborates recent research illustrating the advantage of using CNNs in conjunction with BiLSTM for sequence and context extraction in NLP applications [22].Future work may explore transformer-based approaches, reinforcement learning, or hybrid NLP techniques for further improvements.

V. CONCLUSION AND FUTURE SCOPE

This paper proposes a hybrid deep model for mixed-scripted queries on Roman transliteration for improved intent identification. The proposed system uses CNN, BiLSTM, and an Attention mechanism to process mixed-script text and retrieve local and long-distance relations for accurate intent classification. Preprocessing of data, transliteration normalization, feature extraction, and classification using deep learning techniques provide good performance for complex and multi-label intent identification tasks. Experimental outcomes show that the proposed method outperforms ordinary models and can be applied to search engines, virtual assistants, and chatbots. The low-latency intent query processing of the model is also highlighted through execution time analysis. This technique is aimed at mixed-script variations and transliteration variations in NLP for multi-language and transliterated text processing. Future research can be focused on extending the dataset, transformer-based models, and real-time adaptability for real-world usage in various languages. This paper successfully proposed an intent recognition model for mixed-script queries on Roman transliteration. The Hybrid Model with LSTM and CNN combined performed better than ordinary models like standalone LSTM, CNN, and Random Forest. The evaluation measures affirm the efficiency of the proposed method. The Hybrid Model achieved the highest accuracy of 92.8%, which is better than LSTM (88.5%), CNN (86.9%), and Random Forest (84.3%). The Precision-Recall (PR) curve analysis showed an AUC-PR value of 0.91, which indicates a good precision-recall trade-off.

The Hybrid Model proposed for intent identification in mixed-script queries yielded good results, but there are many follow-up research options that could improve its utility and application. First, expanding the data set with a wider range of diversities in language and dialect can increase the generalizability of the model and its practical applicability for real-world multilingual tasks and applications. For example, this could involve adding more individual languages, including different dialects of the same language in the data set or even gender and age dialect differences. Along the same lines of increasing its practical applications, more advanced transformer-based models could be integrated into the model, such as BERT, XLM-R, or mT5. There is good promise for the Roman transliteration-based intent identification model for mixed-script queries in both the accuracy and efficiency of the model. However, many other areas of research and development are needed. Based on the advances of Transformer-based models, one might consider employing these model types for mixed-script queries that are complex transliterations, ambiguous queries, or other code-mixing or multilingual questions that an international audience will use. One area that is also developing interest is the real-time use of the model in NLP tools, like Chatbots, virtual assistants, and smart search engines, and trying to optimize the model for low-latency edge devices or cloud-based systems. This will help maximize potential real-world usage and importance.

In spite of the encouraging results, the present work has some disadvantages. The corpus used is mostly made up of Hindi-English transliterated search queries, which might restrict the broadness of model applicability to other language pairs or more complicated multilingual environments. It was not examined if the model can perform well under highly noisy or code-switched conditions with three or more languages. The future direction should be to test over more language pairs, casual chat-based search queries, and online deployment settings.

References

[1] Chaudhary, A. and Shekhar, S. (2024) ‘A Framework to Find Single Language Version Using Pattern Analysis in Mixed Script Queries’, 2024 2nd International Conference on Disruptive Technologies (ICDT), Greater Noida, India, pp. 1593–1601. doi: 10.1109/ICDT61202.2024.10489697.

[2] Jahan, Z., Khan, S.B. and Saraee, M. (2024) ‘Early dementia detection with speech analysis and machine learning techniques’, Discover Sustainability, 5(1). Springer Science and Business Media LLC. doi: 10.1007/s43621-024-00217-2.

[3] Lorenzoni, G., Tavares, C., Nascimento, N., Alencar, P. and Cowan, D. (2024) ‘Assessing ML Classification Algorithms and NLP Techniques for Depression Detection: An Experimental Case Study’, arXiv. doi: 10.48550/ARXIV.2404.04284.

[4] Pan, H., Fang, Y., Guo, W., Xu, Y. and Wang, C. (2024) ‘Few-shot graph classification on cross-site scripting attacks detection’, Computers & Security, 140. Elsevier BV, p. 103749. doi: 10.1016/j.cose.2024.103749.

[5] Cui, J., Wang, Z., Ho, S.-B. and Cambria, E. (2023) ‘Survey on sentiment analysis: evolution of research methods and topics’, Artificial Intelligence Review, 56(8), pp. 8469–8510. Springer Science and Business Media LLC. doi: 10.1007/s10462-022-10386-z.

[6] Hurtado, R., Picón, C., Muñoz, A. and Hurtado, J. (2023) ‘Survey of Intent-Based Networks and a Methodology Based on Machine Learning and Natural Language Processing’, Proceedings of Eighth International Congress on Information and Communication Technology. Springer Nature Singapore, pp. 363–382. doi: 10.1007/978-981-99-3236-8_28.

[7] T B, L. and P S, S. (2023) ‘Potential Web Content Identification and Classification System using NLP and Machine Learning Techniques’, International Journal of Engineering Trends and Technology, 71(4), pp. 403–415. Seventh Sense Research Group Journals. doi: 10.14445/22315381/ijett-v71i4p235.

[8] Ren, X. et al. (2022) ‘Effective Approaches to Neural Query Language Identification’, Computational Linguistics, 48(4), pp. 887–906. MIT Press. doi: 10.1162/coli_a_00451.

[9] Wang, Y. (2022) ‘Using Machine Learning and Natural Language Processing to Analyze Library Chat Reference Transcripts’, Information Technology and Libraries, 41(3). American Library Association. doi: 10.6017/ital.v41i3.14967.

[10] Weld, H., Huang, X., Long, S., Poon, J. and Han, S.C. (2022) ‘A Survey of Joint Intent Detection and Slot Filling Models in Natural Language Understanding’, ACM Computing Surveys, 55(8), pp. 1–38. Association for Computing Machinery (ACM). doi: 10.1145/3547138.

[11] Sharma, I. and Kumar, V. (2022) ‘Multi-objective tunicate search optimisation algorithm for numerical problems’, International Journal of Intelligent Engineering Informatics, Vol. 10, No. 2, p.119, https://doi.org/10.1504/ijiei.2022.125859.

[12] Hande, A. et al. (2021) ‘Offensive Language Identification in Low-resourced Code-mixed Dravidian languages using Pseudo-labeling’, arXiv. doi: 10.48550/ARXIV.2108.12177.

[13] Alzahrani, S. M., & Guma, F. E. (2024). Improving Seasonal Influenza Forecasting Using Time Series MachineLearningTechniques. Journal of Information Systems Engineering and Management, 9(4), 30195. https://doi.org/10.55267/iadt.07.15132.

[14] Tao, S., & Hao, Y. (2023). The Impact of Socio-economic Environment on Artificial Intelligence and Technology Adoption: MediationModeration of Employee Team Collaboration and Technological Innovation. Journal of Information Systems Engineering and Management, 8(3), 21742.https://doi.org/10.55267/iadt.07.13610.

[15] Mudiyanselage, A. D. H. H., & Sumanathilaka, T. D. K. (2024). TAMழி: Shorthand Romanized Tamil to Tamil Reverse Transliteration Using Novel Hybrid Approach. The International Journal on Advances in ICT for Emerging Regions, 17(1).

[16] Baruah, H., Singh, S. R., & Sarmah, P. (2024). Transliteration characteristics in romanized assamese language social media text and machine transliteration. ACM Transactions on Asian and Low-Resource Language Information Processing, 23(2), 1-36.

[17] Lauc, D., Rutherford, A., & Wongwarawipatr, W. (2024). AyutthayaAlpha: A Thai-Latin Script Transliteration Transformer. arXiv preprint arXiv:2412.03877.

[18] Naufal, M. F., Siswantoro, J., & Soebroto, J. T. (2024). Transliterating Javanese Script Images to Roman Script using Convolutional Neural Network with Transfer Learning. International Journal on Informatics Visualization, 8(3), 1460-1468.

[19] Lakshmi, B. S., & Shambhavi, B. R. (2024). State-of-the-art automatic machine transliteration systems for Indic scripts: a comparative report. International Journal of Advanced Intelligence Paradigms, 27(2), 150-177.

[20] Bellegarda, J. R. (2024, April). Multilingual Transliteration for Pan-Indic Keyboard Input. In ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 5675-5679). IEEE.

[21] Chaudhary, A., & Shekhar, S. (2024, March). A Framework to Find Single Language Version Using Pattern Analysis in Mixed Script Queries. In 2024 2nd International Conference on Disruptive Technologies (ICDT) (pp. 1593-1601). IEEE.

[22] Ahmed, F., Fahim, M., Amin, M. A., Ali, A. A., & Rahman, A. K. M. (2024). Improving the Performance of Transformer-Based Models Over Classical Baselines in Multiple Transliterated Languages. In ECAI 2024 (pp. 4043-4050). IOS Press.

[23] Bai, L., & Zhang, K. (2025). Artificial Intelligence-Driven Cross-Language Communication Network: Using Neural Network to Improve Foreign Language Translation. International Journal of High Speed Electronics and Systems, 2540269.

[24] Nezami, M. A., & Huda, M. N. (2024, November). Troll and Abuser Identification in Social Media for Code-Mixed English, Bangla, and Banglish Text Analysis Using LLMs. In 2024 International Conference on Intelligent Computing and Next Generation Networks (ICNGN) (pp. 1-5). IEEE.

[25] Hashmi, E., Yayilgan, S. Y., & Shaikh, S. (2024). Augmenting sentiment prediction capabilities for code-mixed tweets with multilingual transformers. Social Network Analysis and Mining, 14(1), 86.