ALGERIAN DEMOCRATIC AND POPULAR REPUBLIC

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

KASDI MERBAH UNIVERSITY OUARGLA

FACULTY OF NEW INFORMATION AND COMMUNICATION TECHNOLOGIES

DEPARTMENT OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY

THESIS SUBMITTED IN CANDIDACY FOR A MASTER DEGREE IN COMPUTER SCIENCE,OPTION

ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

PRESENTED By **Oumaima Zitouni**.

TITLE

User Intention Prediction Using Text Messages

Evaluation Date : 23 /6/2024

Before the Jury:

Dr.YACINE KHALDI	SUPERVISOR	UKM OUARGLA
Dr.YOUCEFA.M	PRESIDENT	UKM OUARGLA
Dr.BILAL KHALDI	EXAMINER	UKM OUARGLA

ACADEMIC YEAR: **2023/2024**

ACKNOWLEDGMENT

First and foremost, I am grateful for the blessings and guidance bestowed upon me by Allah,

I extend my deepest thanks to my family, especially my mother, whose unwavering love and support have been my foundation throughout this thesis journey, her encouragement has meant everything to me,and my father has been a constant source of inspiration, for which I am deeply grateful.

I am also immensely thankful for the support of my splings and my friends , who stood by me during the challenging times and celebrated the successes with me,Their friendship has been a source of strength and joy.

I express sincere appreciation to Dr. Yacine Khaledi, my dedicated supervisor, for his expert guidance and unwavering encouragement.

A special thank you to the esteemed members of the jury for their time and consideration in evaluating my thesis, I extend my gratitude to all my teachers for their invaluable contributions to my academic growth.

Thank you all for being part of this meaningful journey.

Abstract

Recently, social media has become an important source of information, where user-generated content can be analyzed, making it a significant research area in natural language processing (NLP) and machine learning (ML).

These fields can be utilized to understand user intent based on text, although analyzing intent from text is a challenging task due to the informal nature of text and its context.

The main contribution of this work is to harness the power of NLP and ML to detect patterns in textual data that predict user intent, provide insights into user behavior, and enhance predictive models. The proposed method combines two models using machine learning: the Bidirectional Encoder Representations from Transformers (BERT) model and the Long Short-Term Memory (LSTM) model applied to textual data. Through the innovative application of these models, our study demonstrated the feasibility of accurately classifying users' intentions regarding locations, achieving high accuracy in our dataset.

Our results show the model's efficiency in identifying dialogues where the user does not discuss a location (class "0") with high accuracy and recall. However, this highlights the need to improve the detection of positive cases where user intent includes location (class "1"), especially with regard to recall.

The discrepancy in performance between class "0" and class "1" indicates that while the model can effectively differentiate non-location dialogues, it struggles with the nuances of location intent.

The importance of our work lies not only in the impressive accuracy achieved but also in the model's contribution to predictive capabilities in the field of NLP.

Our approach provides a foundation for developing intelligent systems capable of understanding and predicting user intentions based on textual data, with potential applications in virtual assistants, targeted advertising, and location-based services.

Keywords: Intent Prediction, Destination prediction, Machine Learning, Naturel Language Processing, LSTM, BERT.

Résumé

Récemment, les médias sociaux sont devenus une source d'information cruciale, où le contenu généré par les utilisateurs peut être analysé et transformé en un domaine de recherche clé dans le traitement du langage naturel (NLP) et l'apprentissage automatique (ML).

Ces domaines permettent de comprendre les intentions des utilisateurs à partir du texte, bien que l'analyse des intentions basée sur le texte soit une tâche complexe en raison de la nature informelle du texte et de son contexte.

La principale contribution de ce travail réside dans l'exploitation de la puissance du NLP et de l'apprentissage automatique pour détecter des motifs dans les données textuelles qui prédisent les intentions des utilisateurs, fournissent des informations sur leur comportement et améliorent les modèles prédictifs.

La méthode proposée combine deux modèles d'apprentissage automatique : le modèle Bidirectional Encoder Representations from Transformers (BERT) et le modèle Long Short-Term Memory (LSTM) appliqués aux données textuelles.

Grâce à l'application innovante de ces modèles, notre étude a démontré la faisabilité de classifier avec précision les intentions des utilisateurs concernant les emplacements, atteignant ainsi une grande précision dans notre ensemble de données. Nos résultats montrent l'efficacité du modèle à identifier les dialogues dans lesquels l'utilisateur ne discute pas d'un emplacement (classe « 0 ») avec une grande précision et un rappel élevé.

Cependant, cela met en évidence la nécessité d'améliorer la détection des cas positifs où l'intention de l'utilisateur inclut la localisation (classe « 1 »), notamment en ce qui concerne le rappel.

L'écart de performance entre la classe « 0 » et la classe « 1 » indique que, bien que le modèle puisse différencier efficacement les dialogues ne concernant pas les emplacements, il éprouve des difficultés avec les nuances des intentions relatives aux emplacements.

L'importance de notre travail réside non seulement dans la précision impressionnante obtenue, mais aussi dans la contribution du modèle aux capacités prédictives dans le domaine du NLP.

Notre approche constitue une base pour développer des systèmes intelligents capables de comprendre et de prédire les intentions des utilisateurs à partir de données textuelles, avec des applications potentielles dans les assistants virtuels, la publicité ciblée et les services géolocalisés.

Mots-clés: Classification des intentions, Prédiction de destination, Apprentissage automatique, Traitement du langage naturel, LSTM, BERT.

ملخص

في الآونة الأخيرة، أصبح التواصل الاجتماعي مصدراً مهماً للمعلومات، حيث يمكن تحليل المحتوى الذي ينشئه المستخدمون ليصبح مجالاً بحثياً مهماً في معالجة اللغة الطبيعية (NLP) والتعلم الآلي (ML).

يمكن استخدام هذين المجالين لفهم نوايا المستخدم استناداً إلى النص، ويعد تحليل النوايا بناءً على النص مهمة صعبة للغاية بسبب الطبيعة غير الرسمية النص والسياق.

تتمثل المساهمة الرئيسية لهذا العمل في تسخير قوة البرمجة اللغوية العصبية والتعلم الآلي للكشف عن الأنماط في البيانات النصية التي تتنبأ بنية المستخدم، وتقديم رؤى حول سلوكه، وتعزيز النماذج التنبؤية.

تجمع الطريقة المقترحة بين نموذجين باستخدام التعلم الآلي: نموذج تمثيلات التشفير ثنائية الاتجاه من المحولات (BERT) ونموذج الذاكرة طويلة قصيرة المدى (LSTM) على البيانات النصية، ومن خلال التطبيق المبتكر لهذه النماذج، أظهرت دراستنا جدوى التصنيف الدقيق لنوايا المستخدمين فيما يتعلق بالمواقع، مما يحقق دقة عالية في مجموعة البيانات الخاصة بنا.

وتظهر نتائجنا كفاءة النموذج في تحديد الحوارات التي لا يناقش فيها المستخدم موقعاً (فئة "0") بدقة واستدعاء عاليين، ومع ذلك، يشير ذلك إلى الحاجة إلى تحسين اكتشاف الحالات الإيجابية التي تتضمن فيها نية المستخدم الموقع (الفئة "1")، خاصة فيما يتعلق بالاستدعاء.

يشير التباين في الأداء بين الفئة "0" والفئة "1" إلى أنه في حين أن النموذج يمكنه التمييز بين الحوار غير المتعلق بالموقع بشكل فعال، إلا أنه يعاني من الفروق الدقيقة في نية الموقع، لا تكمن أهمية عملنا في الدقة المبهرة التي تم تحقيقها فحسب، بل أيضًا في مساهمة النموذج في القدرات التنبؤية في مجال البرمجة اللغوية العصبية.

يوفر نهجنا أساسًا لتطوير أنظمة ذكية قادرة على فهم نوايا المستخدم والتنبؤ بها استنادًا إلى البيانات النصية، مع تطبيقات محتملة في المساعدين الافتراضيين، والإعلانات المستهدفة، والخدمات القائمة على الموقع.

الكلمات المفتاحية: التواصل الاجتماعي، معالجة اللغة الطبيعية NLP التعلم الاليML، نوايا المستخدم بيانات النصية, نموذج BERT, نموذج LSTM، الأنظمة الذكية، المساعدين الافتراضيين، الإعلانات المستهدفة.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER1: GENERAL INTRODUCTION

1- Introduction

The field of natural language processing (NLP) has revolutionized human-computer interaction by enabling innovative approaches to understanding textual data. Our study focuses on predicting users' destinations based on text messages, a topic that sits at the intersection of linguistics, artificial intelligence, and data analytics.

By harnessing the power of NLP and machine learning, our research aims to uncover patterns in textual data that can predict user intent, offer insights into user behavior, and improve predictive models. This investigation explores various methodologies and algorithms to enhance the accuracy and efficiency of intent classification, advancing towards more intuitive and responsive digital environments.

2- Problematic

The challenge addressed by this research is to accurately distinguish whether dialogues refer to the user's current location or a future destination using NLP techniques. This involves parsing nuanced language patterns to infer user intent effectively.

3- Motivations

The motivation behind this study lies in the growing importance of understanding user intentions and behavior through textual analysis. Predicting user destinations from dialogues can significantly improve the design of personalized services and recommendation systems.

4- Objectives of the Study

This study aims to distinguish whether dialogue refers to the user's current location or a future destination,through the use of advanced NLP models such as Long Short-Term Memory (LSTM) and Bidirectional Encoding Representations from Transformers (BERT) networks, the research first classifies dialogues based on the importance of location before delving deeper into the nuances of presence versus intent

5- Scope and Limitations

The study leverages a unique dataset of 9,508 lines extracted from TV series subtitles, providing a specific context for location-based intent detection. Limitations include dataset size, generalizability to other dialogue types, and potential biases in the extracted data.

6- Thesis structure

Our work is organized into four chapters that collectively explore the intersection of natural language processing (NLP) and user behavior analysis. The "General Introduction" sets the stage by providing an overview of our research topic. The subsequent chapter conducts a "literature review", surveying existing research in NLP and user intention detection. Following this, "the proposed model "chapter details our methodology and design, focusing on innovative techniques like LSTM and BERT. Finally, "the experimental analysis and discussion "chapter presents empirical findings derived from applying our NLP models to a unique dataset.

CHAPTER2 : LITERATURE REVIEW

1. Introduction :

Natural Language Processing (NLP) stands at the forefront of artificial intelligence, serving as a bridge between human communication and computer understanding. With the exponential growth in digital communication, NLP has become essential for extracting meaningful information from vast amounts of unstructured text. From sentiment analysis and machine translation to chatbots and voice recognition, NLP applications are transforming industries by enabling machines to understand, interpret, and respond to human language in a valuable way.

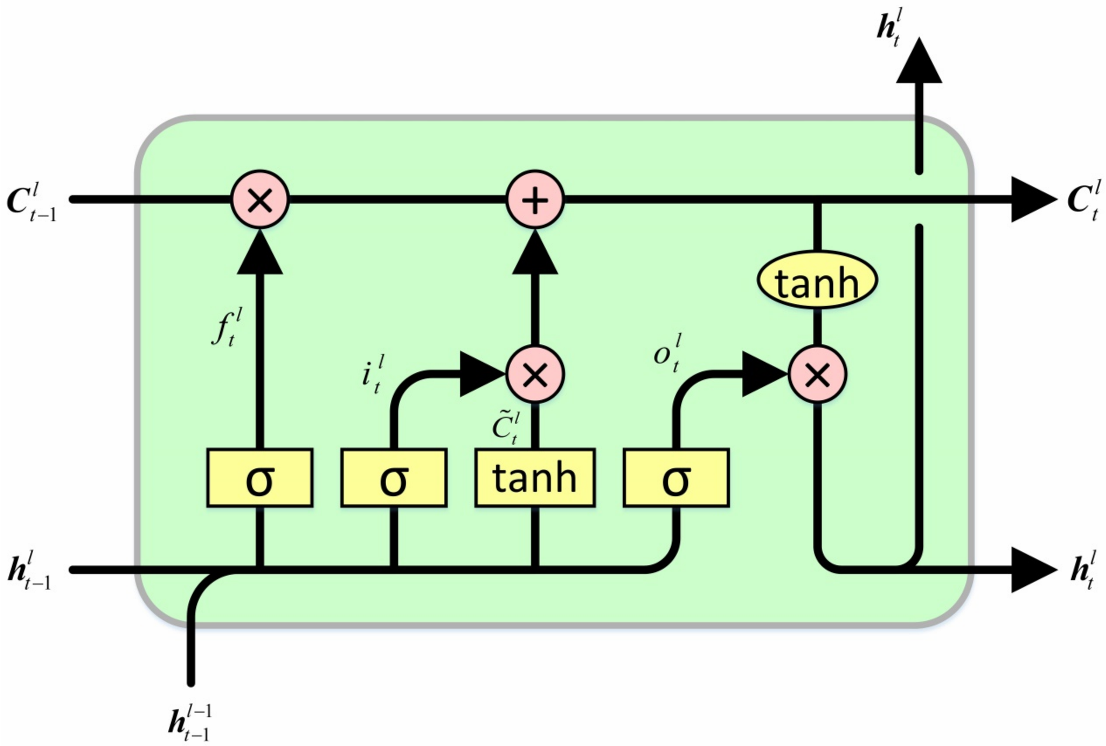
One of the key areas where NLP has shown significant impact is in user intent classification. This involves deciphering the underlying intent behind a user's input, which is crucial for enhancing user experience in various applications such as virtual assistants, customer service automation, and personalized recommendations. Understanding user intent allows systems to provide more accurate and contextually relevant responses, thereby improving interaction efficiency and satisfaction.

This section reviews the current literature on user intent classification using NLP, exploring the methodologies, models, and challenges that researchers and practitioners face. We will examine the evolution of NLP techniques from traditional machine learning algorithms to advanced deep learning models, highlighting the progress and ongoing research in this dynamic field. By synthesizing existing studies, we aim to provide a comprehensive understanding of the state-of-the-art in user intent classification and identify potential avenues for future research.

In recent years, the development of sophisticated models has revolutionized NLP. The most prominent models include the Long Short-Term Memory (LSTM) network and the Bidirectional Encoder Representations from Transformers (BERT). These models have set new benchmarks in various NLP tasks due to their unique architectures and learning capabilities , so in this work we mainly focuses on those two models in user intent classification.

2. Long Short-Term Memory (LSTM) (1)

LSTM is a type of recurrent neural network (RNN) that was proposed by Sepp Hochreiter and Jürgen Schmidhuber in their seminal 1997 paper. Traditional RNNs suffer from the problem of vanishing and exploding gradients, which makes them ineffective at capturing long-term dependencies in sequences. LSTM addresses this issue through its unique architecture, which includes memory cells and gating mechanisms. These gates—input, output, and forget gates—allow LSTM networks to selectively retain and forget information, making them adept at modeling long-range dependencies.



The diagram the current input and previous hidden state should be added to the cell state, and the output gate controls what part of the cell state should be used to produce the current hidden state. These gates, along with the cell state (Ct-1) and hidden state (ht-1), allow the LSTM to selectively remember and update its internal memory as it processes the input sequence.

The cell state (Ct-1) acts as a memory that carries information from the previous time step, while the hidden state (ht-1) represents the output of the LSTM layer at the previous time step. These states are critical for the LSTM's ability to maintain and update its memory, enabling it to capture long-term dependencies in the input data. The tanh activation function is applied to the cell state and hidden state to introduce non-linearity and enable the LSTM to learn complex patterns in the input data. Additionally, the sigmoid activation function is used to compute the gate values (ft, it, ot), which control the flow of information within the LSTM cell. (1)

As the input sequence passes through the stacked LSTM layers, the model learns to extract relevant features and patterns from the text. The final LSTM output is a probability binary value, which determines if the user is going into a destination or not. This values ranges from 0 to 1, if the output is under the threshold 0.5 then the classification result is false, if not then the classification result is True. (1)

2.1 Benefits of LSTM

1. **Long-term dependency learning:** LSTMs can capture dependencies over long sequences, which is crucial for tasks like language modeling and machine translation.

2. **Robustness to vanishing gradients:** The gating mechanisms help mitigate the vanishing gradient problem, ensuring more stable and efficient training.

3. **Versatility:** LSTMs have been successfully applied to a variety of tasks, including speech recognition, time series prediction, and text generation.

2.2 Drawbacks of LSTM

1. **Complexity:** The architecture of LSTM is more complex compared to traditional RNNs, leading to increased computational costs.

2. **Training time:** Due to their complexity, LSTMs often require more time to train, which can be a limitation for real-time applications.

3. **Overfitting:** LSTMs are prone to overfitting, especially when dealing with small datasets, necessitating careful regularization and tuning.

3. Bidirectional Encoder Representations from Transformers (2)

Bidirectional Encoder Representations from Transformers (BERT), introduced by Google in 2018 in paper called “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”, represents a major advancement in NLP through its innovative use of the transformer architecture. Transformers, proposed by Vaswani et al. in 2017, rely entirely on attention mechanisms to process sequences of data. BERT leverages the transformer architecture by employing bidirectional training, allowing it to consider both the left and right context of a word simultaneously. This bidirectional nature enables BERT to generate more nuanced and contextually relevant representations.

3.1 History and Architecture

BERT's architecture consists of multiple layers of transformers, each layer containing self-attention mechanisms and feedforward neural networks. The model is pre-trained on vast amounts of text data using a masked language modeling objective, where random words in a sentence are masked, and the model learns to predict them. This pre-training step allows BERT to capture deep contextual information. BERT can then be fine-tuned on specific NLP tasks, such as user intent classification, by adding a simple classification layer on top of the pre-trained model. (2)

3.2 Benefits of BERT

1. **Contextual Understanding:** BERT's bidirectional training allows it to understand the context of words more effectively than unidirectional models.
2. **Transfer Learning:** Pre-trained BERT models can be fine-tuned on various tasks with relatively small amounts of task-specific data, making it versatile and efficient.
3. **Performance:** BERT has achieved state-of-the-art results on numerous benchmarks, including the GLUE (General Language Understanding Evaluation) tasks and SQuAD (Stanford Question Answering Dataset).

3.3 Drawbacks of BERT

1. **Resource Intensive:** BERT requires significant computational resources for both pre-training and fine-tuning, which can be a barrier for smaller organizations.
2. **Inference Time:** The large size of BERT models can lead to slower inference times, making them less suitable for real-time applications.
3. **Complexity:** Fine-tuning BERT requires careful parameter tuning and understanding of its architecture, which can be challenging for practitioners.

4. Deep Learning Approaches to User Intent Classification :

(2) **:**

Understanding user intent is a critical factor in enhancing the adoption and utilization of conversational AI systems like chatbots. As large language models (LLMs) like ChatGPT become increasingly prevalent, the ability to accurately recognize and respond to user intents is essential for improving user satisfaction and trust.

Recent studies have examined the performance of state-of-the-art LLMs, such as GPT-4 Turbo, in understanding and responding to user intents. The findings suggest that while these advanced models excel at processing diverse data and generating coherent responses, they still face challenges in accurately recognizing specific user intents, particularly for less common or complex query types.

Researchers have proposed various approaches to address this issue, including the development of intent taxonomies and techniques for intent-based prompt reformulation (3).These strategies aim to improve the models' ability to understand the user's underlying goals and provide more relevant and satisfactory responses.

The literature highlights the importance of considering user intent as a key factor in the design and evaluation of conversational AI systems. As LLMs continue to evolve, further research is needed to explore effective methods for enhancing intent recognition and user satisfaction in human-AI interactions.

(4)**:**

This paper proposes a novel two-stage approach for user feedback-based online learning for intent classification. The method uses a multi-armed contextual bandit framework that leverages a pre-trained RoBERTa text encoder to extract features from utterances and learn joint representations of text and intent. The training has two stages: 1) Offline pretraining, where the bandit policy is trained on a small labeled dataset using an offline policy learning approach to address distribution shift, and 2) Online fine-tuning, where the pretrained policy is further fine-tuned using the REINFORCE algorithm with user feedback, without requiring true intent labels, and a sliding window strategy is used to handle distributional shift in the online setting. This two-stage approach allows the method to efficiently adapt to dynamic user preferences and data distributions, achieving competitive performance compared to supervised benchmarks trained on much larger labeled datasets. Extensive experiments on benchmark datasets show the proposed method significantly outperforms alternatives that omit either the offline pretraining or online fine-tuning stages. In summary, the paper presents a novel online learning framework for intent classification that effectively leverages both static and dynamic data sources to address the challenges of distribution shift and label scarcity in real-world settings.

(5) :

The paper presents a systematic literature review on user intent modeling for conversational recommender systems (CRS). User intent modeling is a crucial process in natural language processing that aims to identify the underlying purpose behind a user's request, enabling personalized responses. The review covers the context of user intent modeling, which has diverse practical applications in domains like e-commerce, healthcare, education, and virtual assistants, involving various machine learning models like SVMs, LDA, Naive Bayes, BERT, Word2Vec, and MLPs. The problem addressed is that selecting the most suitable intent modeling approach for CRS can be challenging due to the wide array of available models and lack of a clear classification scheme. The authors conducted a systematic literature review and developed a decision model to assist researchers in selecting appropriate intent modeling frameworks. The results analyzed 59 distinct intent modeling approaches and identified 74 commonly used features, providing insights into model combinations, trends, quality concerns, evaluation measures, and frequently used datasets. The study offers practical insights and a comprehensive understanding of user intent modeling, empowering the development of more effective and personalized conversational recommender systems, and the proposed decision model can help researchers perform a systematic and efficient assessment of fitting intent modeling frameworks.

(6) :

The paper discusses using large language models (LLMs) to generate, validate, and apply user intent taxonomies for analyzing search and chat log data. Extracting user intents from log data is challenging due to the fluidity of intents and lack of sufficient context, especially for emerging search modalities like AI-driven chat. Traditional qualitative methods are time-consuming, while existing quantitative methods may not capture the nuances.

The authors propose a new methodology that combines LLM-generated taxonomies with human expert verification and curation to leverage the strengths of both, aiming to reduce human effort while ensuring validity. This integrative methodology for generating, validating, and applying purpose-driven taxonomies could be useful beyond just user intent analysis.

The authors demonstrate the effectiveness of their approach by applying it to search and chat logs from the Microsoft Bing search engine, uncovering new insights about user intents. The key research questions addressed are: (1) Can LLMs reliably generate user intent taxonomies? (2) Can an LLM correctly apply a taxonomy to annotate logs? (3) How do LLM-based approaches compare to human annotation?

Overall, the paper presents a novel framework for leveraging LLMs and human expertise to tackle the challenge of understanding user intents at scale, with implications for improving information access systems.

(7) :

The study is based on the Technology Acceptance Model (TAM) and employs statistical analysis to examine three review types: regular reviews (type A), comparative reviews (type B), and suggestive reviews (type C).

Existing opinion mining studies have focused on only two types of reviews - regular and comparative. This research aims to address the gap in determining useful review types from the perspective of both customers and designers. By using samples of users (N=400) and designers (N=106), the study examines the three review types in terms of their perceived usefulness, perceived ease of use, and behavioral intention.

The results indicate that a positive perception of suggestive reviews (type C) can improve users' decision making and business intelligence. The study suggests that type C reviews could be considered a new useful review type in addition to the existing types A and B. To establish the validity of these findings, the paper employs regression, correlation, and mediation analysis to quantitatively analyze the survey data.

(8) :

The paper discusses the problem of question-answering (QA), which is a complex task in natural language processing (NLP) and artificial intelligence (AI).

Since finding a generic answer to any question is still a distant goal, the authors propose to break down the problem into simpler parts. Specifically, they focus on intent classification for providing an answer given a question.

The authors explain that traditional feedforward neural networks are not well-suited for modeling the dynamics of language, as they cannot effectively capture the temporal dependencies between words. This led to the development of recurrent neural networks (RNNs), which can maintain an internal state and pass information through time. However, RNNs suffer from the vanishing gradient problem, limiting their ability to retain long-term memory.

To address this, the authors introduce the Long Short-Term Memory (LSTM) architecture. Unlike a classic RNN, the LSTM has a more complex cell structure with specialized "gates" that regulate the flow of information in the network. The forget gate decides what information to remove from the cell state, the input gate determines what new information to add, and the output gate controls what information to output. This allows the LSTM to effectively model long-range dependencies in language.

The authors then present their predictive model for intent classification, which uses an LSTM network. They discuss the training and evaluation of this model, and provide an example of how it can be used within a basic prototype responder system. The intent classification approach can be useful as a building block for more advanced question-answering systems, as well as for enhancing the capabilities of chatbot-based dialogue interfaces.

(9) **:**

The research paper examines the use of recurrent neural network (RNN) and long short-term memory (LSTM) models for the important task of utterance classification. Utterance classification is a critical pre-processing step for many speech understanding and dialog systems. In multi-user settings, a system first needs to identify if an utterance is even directed at the system, before then determining the intent of the user's input.

The authors propose RNN and LSTM models for both of these subtasks - addressee detection and intent classification. They compare the performance of these neural network approaches to more traditional baselines, such as n-gram language models, feedforward neural network language models, and boosting classifiers.

A key challenge the paper aims to address is the limitations of n-gram based models. N-gram approaches suffer from two fundamental issues - the limited temporal scope of n-grams, and the sparseness of n-gram distributions, which requires large amounts of training data for good generalization. Longer n-grams can capture more contextual information, but exacerbate the data sparsity problem. The recurrent neural network architectures proposed in this work hold the promise of being able to better model long-term dependencies in the utterances.

To further address the data sparsity issue, the paper also investigates using character n-gram based word input encoding, rather than standard one-hot word vectors. This character-based encoding is shown to outperform the one-hot approach, likely by better handling rare and out-of-vocabulary words.

The experimental results demonstrate the effectiveness of the RNN and LSTM models. They achieve over 30% relative reduction in equal error rate compared to boosting classifier baselines on an utterance intent classification task. For an addressee detection task, the neural network models provide over 3.9% absolute reduction in equal error rate compared to a maximum entropy language model baseline.

An interesting finding is that the RNN architecture works best for short utterances, while the LSTM excels on longer utterances. This is likely due to the LSTM's superior ability to model long-range dependencies, which become more important as utterance length increases. Overall, this work presents a compelling approach using recurrent neural network models to significantly improve upon traditional n-gram based methods for the critical tasks of utterance classification and intent determination. The results highlight the potential of these neural architectures to better capture the contextual information present in spoken language interactions.

(11)**:**

This paper introduces ILLUMINER, a new way to classify user intents and fill in slots using instruction-tuned large language models (Instruct-LLMs). Instead of traditional methods that rely on large datasets, ILLUMINER frames these tasks as language generation, utilizing Instruct-LLMs' ability to understand and follow instructions.

The paper makes several key contributions: it introduces a more efficient single-prompt approach for slot filling, evaluates various Instruct-LLMs on popular datasets, compares different techniques for fine-tuning LLMs, and conducts extensive analyses to understand the impact of model size, training data, and other factors. It also explores how well these methods work across different datasets and languages.

The authors conclude that ILLUMINER, especially when using LoRA fine-tuned FLAN-T5, outperforms other methods, especially when dealing with limited data. This approach holds significant promise for task-oriented dialogue systems, offering improved performance while requiring less data and computational resources.

5. Conclusion

Despite the extensive research that has been conducted on the problem of user intent classification, accurately identifying user goals and preferences remains a significant challenge, especially in the face of dynamic and evolving user behaviors and data distributions. Traditional static classification models often struggle to keep pace with changes in user communication patterns and preferences over time.

The literature highlights the crucial importance of continuous model training and adaptation for improving the performance of conversational AI systems and enhancing overall user satisfaction. As user needs and goals shift, the ability to rapidly update and fine-tune intent classification models is essential for maintaining high accuracy and providing a seamless user experience.

CHAPTER3: Proposed Method For Destination Prediction

1. Introduction

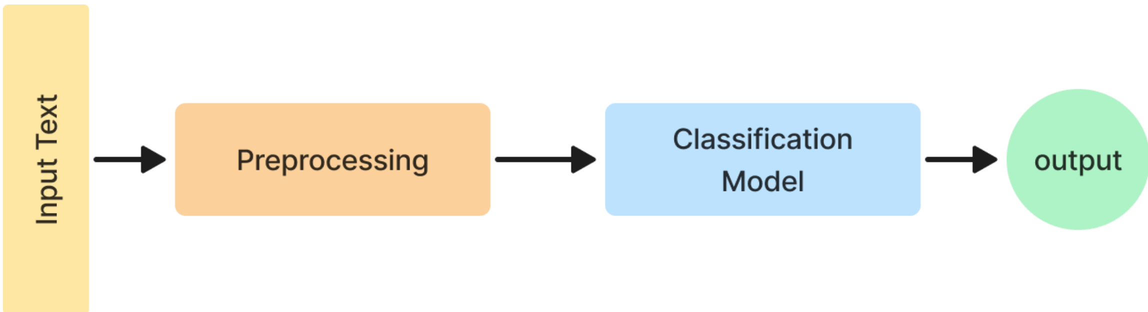
This chapter presents a detailed analysis of the two models employed to tackle the destination binary classification task in this research. The primary objective is to evaluate the efficacy of a BERT (Bidirectional Encoder Representations from Transformers) architecture and an LSTM (Long Short-Term Memory) network in accurately predicting the destination of a given input.

The chapter begins by outlining the general framework and high-level architecture that serves as the foundation for both the BERT-based and LSTM-based models. This includes a comprehensive overview of the data preprocessing steps, along with a detailed explanation of the overall model design. Subsequently, the chapter delves into the specific characteristics, model structures, and training procedures employed for each modeling approach.

By providing this detailed exposition, the chapter aims to offer a transparent and reproducible account of the research methodology. This will enable readers to gain a thorough understanding of the underlying principles, appreciate the nuances of the proposed approaches, and potentially replicate or build upon the findings in future studies.

2. General Pipeline of Text Classification

The general text classification pipeline, as illustrated in Figure 1, adheres to a



straightforward three-step process:

a) Preprocessing: This initial step involves preparing the input data for the subsequent modeling phase. This may include tasks such as:

- **Data Cleaning**: Addressing issues like missing values, incorrect data entries, and inconsistent formatting to ensure data integrity.

- **Feature extraction** : Identifying and extracting relevant features from the raw data, which are essential for the classification model to learn meaningful patterns.
- **Data Transformation**: Applying necessary transformations to the data, such as normalization or scaling, to optimize the model's performance.

b) Classification Model: This is the core of the framework, where the machine learning or deep learning model is applied to the preprocessed data. In this research, we explore two distinct approaches:

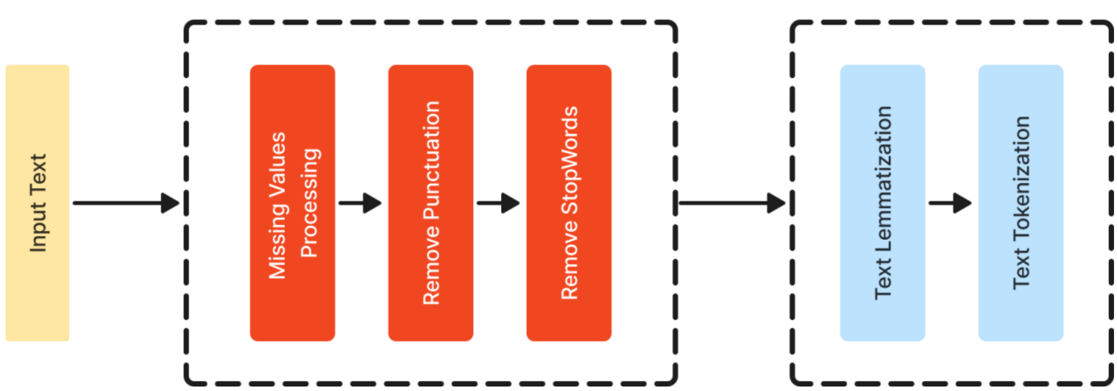
- **BERT-based Model**: Utilizing the pre-trained BERT architecture, which has demonstrated exceptional performance in various natural language processing tasks.
- **LSTM-based Model**: Implementing a recurrent neural network (RNN) architecture specifically designed to handle sequential data, such as text.

c) Output: The final step involves the trained model generating the binary classification prediction based on the input data. This output can be further analyzed, used for decision-making, or integrated into a larger system, depending on the specific application.

3. Preprocessing

Data preprocessing is the crucial first step in any data analysis project. It involves cleaning, refining, and organizing raw data to make it suitable for machine learning models. This ensures the models can learn accurate patterns and achieve optimal performance. As shown in Figure 2 preprocessing in details includes the following two main phases:

a) Data Cleaning:



- **Handling Missing Values**: Techniques like imputation, deletion, or mean/median substitution are employed to address missing values. The chosen method depends on the nature of the missing data and the specific requirements of the model. In this work, we opt for removing data samples containing missing values to avoid potential issues during model training.

- **Removing Punctuation**: Unnecessary punctuation marks are removed from the data to simplify the textual input and reduce noise. This step helps the model focus on the essential information within the text.
- **Removing Stopwords**: Common, non-informative words (e.g., "the," "a," "and") are identified and removed to concentrate on more meaningful features. This process enhances the model's ability to learn from the most relevant terms in the text.
- These techniques were implemented using the nltk library in Python, which provides comprehensive tools for natural language processing tasks, including punctuation removal, stopword lists for various languages, and other effective functionalities.

b) Data Transformation:

- **Lemmatization**: Words in the textual data are transformed into their base or dictionary forms, reducing the dimensionality of the input features. Lemmatization ensures that words with different inflections (e.g., "run," "runs," "running") are treated as the same underlying term, improving the model's understanding of the language.
- **Tokenization**: The cleaned and lemmatized text is then broken down into smaller, meaningful units or tokens that can be effectively processed by the classification model. This step creates a sequence of tokens that the model can utilize to learn patterns and make predictions. Additionally, padding is applied to the text, ensuring that the input to the model has a fixed length. This is achieved by adding empty tokens to shorter sequences and truncating longer sequences to the specified maximum length.

By applying these data preprocessing techniques systematically, the raw input data is transformed into a clean, consistent, and well-structured format that can be efficiently utilized by the subsequent Classification Model. This comprehensive preprocessing enhances the overall accuracy and performance of the system by ensuring the model has access to high-quality, relevant input data.

4. LSTM-based Model

The Long Short-Term Memory (LSTM) model we propose is designed to effectively capture the sequential dependencies inherent in user input data. Our model architecture consists of a single LSTM layer followed by a dense layer, tailored to classify user intents with high accuracy.

Model Architecture

1. **LSTM Layer**: The core of the model features an LSTM layer with 100 LSTM cells. This layer is responsible for processing the input sequences and capturing temporal dependencies within the data.

2. **Dense Layer**: The output from the LSTM layer is fed into a dense layer comprising 100 neurons. This layer utilizes a sigmoid activation function to perform the final classification of user intents.

The LSTM layer's ability to maintain long-term dependencies and the dense layer's classification capability combine to make this model both powerful and efficient for user intent classification tasks.

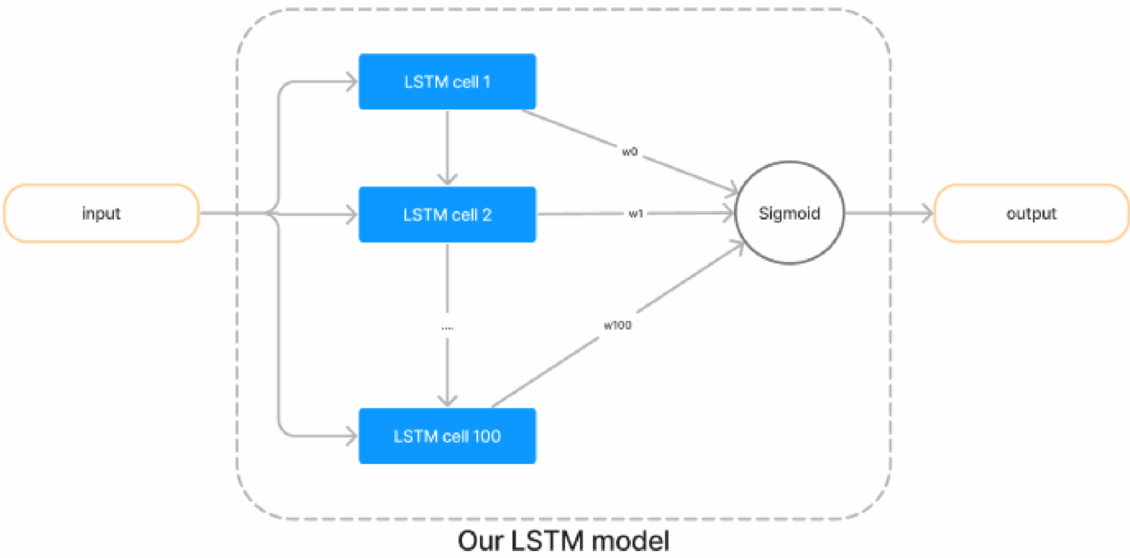


Figure 4 : Proposed LSTM model architecture

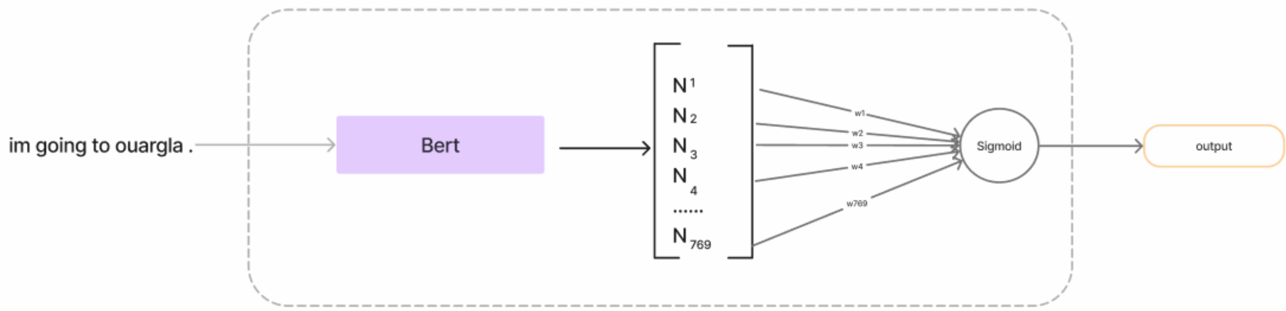
5. BERT-based Model

Our proposed BERT model leverages the power of Bidirectional Encoder Representations from Transformers (BERT) to enhance the understanding of user intents. This model integrates a pre-trained BERT with a simple yet effective classification layer.

Model Architecture :

1. **BERT Layer**: We utilize a pre-trained BERT model to capture complex language patterns and contextual information from the input text. BERT's ability to understand bidirectional context makes it ideal for this task.
2. **Dense Layer**: Following the BERT layer, we add a dense layer with 769 weights, employing a sigmoid activation function. This layer serves as the classification head, converting BERT's contextual embeddings into user intent predictions.

   By leveraging BERT's sophisticated language understanding and a straightforward classification layer, this model aims to achieve robust performance in user intent classification.



These proposed models, with their distinct architectures, provide two complementary approaches to tackle the problem of user intent classification, offering a balance between simplicity and effectiveness.

6. Conclusion

This chapter presented a comprehensive analysis of two deep learning architectures, LSTM and BERT, for addressing the destination binary classification task. The LSTM-based approach leveraged the model's ability to capture complex patterns in sequential data, while the BERT-based approach utilized the pre-trained BERT language model to learn contextual representations of the input features. Together, these models provide a complementary set of deep learning tools to effectively tackle the binary classification objectives of this research.

The chapter provided a detailed explanation of the general modeling framework, data preprocessing techniques, model architectures, training procedures, and a comparative analysis of the two approaches. This comprehensive exposition enables readers to understand the underlying principles, appreciate the nuances of the proposed methods, and potentially replicate or build upon the findings in future studies.

CHAPTER4: Experimental Results and Analysis

1. Introduction

In this chapter, we detail the experiments conducted to compare the performance of BERT and LSTM models for the task of user intent classification. Understanding user intent is crucial for a wide range of applications, including customer service automation, search engines, and virtual assistants. Accurately classifying user intent enables these systems to provide more relevant and efficient responses, thereby improving user experience.

To evaluate the effectiveness of BERT and LSTM in this context, we designed a series of experiments using a publicly available dataset. Both models were trained and tested on the same dataset, allowing for a fair comparison of their capabilities. We meticulously preprocessed the data to ensure consistency and reliability in the experimental results.

The chapter is organized as follows: we begin by describing the dataset and its characteristics, followed by the experimental setup, including hardware and software configurations. Next, we outline the training procedures and present the quantitative and qualitative results of our experiments. We also discuss the implications of these results, compare them with existing literature, and highlight any limitations of our study. Finally, we suggest potential avenues for future research in the field of user intent classification.

By conducting these experiments, we aim to provide a comprehensive analysis of how BERT and LSTM perform in user intent classification tasks, offering valuable insights for researchers and practitioners looking to implement these models in real-world applications

2. Dataset

2.1 Dataset Definition

The dataset under consideration consists of subtitle extracts from over 300 different TV series, encompassing approximately 24,000 lines. This dataset is curated for the purpose of exploring the sequential nature of subtitle lines, particularly focusing on the context provided by the preceding and succeeding lines of dialogue. Each line in the dataset is presented along with its immediate previous and next lines, allowing for a comprehensive analysis of conversational flow within the subtitles.

2.2 Composition and Structure

- **Total Lines:** Approximately 24,000
- **Total Series:** Over 300 TV series
- **Data Format:** Each line is accompanied by its preceding and succeeding lines.
- **Positive Samples:** 1,600 lines (indicated as True)
- **Negative Samples:** 22,400 lines (indicated as False)

3.3 Example Entries

The dataset is structured to facilitate binary classification, where each line is labeled as either a positive sample (True) or a negative sample (False). Below are examples illustrating this structure:

- **Positive Sample:**
    - Previous Line: "Let's meet today at the harbor"
    - Current Line: "I'm waiting for you in front of Starbucks" (True)
    - Next Line: "How are you doing?"
- **Negative Sample:**

- Previous Line: "How are you doing?"
- Current Line: "I've already been to London" (False)
- Next Line: "What are your plans for the weekend?"

3.4 Annotation and Labeling

The dataset includes 1,600 positive samples, wherein the dialogue lines are contextually relevant and sequentially coherent within the subtitles. The remaining 22,400 lines are labeled as negative samples, where the lines do not form a logically consistent sequence. This labeling aids in training models to discern meaningful conversational continuity from arbitrary or contextually irrelevant sequences. This dataset is particularly useful for applications in natural language processing (NLP), especially for tasks related to dialogue generation, conversational AI, and sequence prediction. Researchers can leverage the contextual information provided by the preceding and succeeding lines to enhance the performance of models tasked with understanding and generating human-like conversations.

By analyzing this dataset, one can gain insights into the dynamics of conversational flow in scripted television dialogues, providing a rich resource for advancing the field of NLP and related technologies.

3.5 Train/test Data Split

Table 1 illustrates the train/test data distribution.

Table 1 : Train/test Data Split

Label	train	test	Total
True	1306	317	1623
False	6300	1585	7885

3.6 DataProccessing and Transformations

a) Adding label field

To enhance the dataset, we added a labels field that consolidates the information from user_is_there and user_is_intending_to_go_there into a single categorical label. This new field helps in simplifying the intent classification process by creating a unified indicator of user intent. The labeling was done based on the following criteria:

- If the user is both at a place and intending to go to a place, the label is set to '1'.
- If the user is neither at a place nor intending to go to a place, the label is set to '0'.
- If the user is not at a place but is intending to go there, the label is set to '1'.
- If the user is at a place but not intending to go to any other place, the label is set to '1'.
- If only the user's presence is recorded as true, the label is '1'.
- If only the user's intention to go is recorded as true, the label is '1'.
- If either the user's presence or intention data is missing (None), the label is set to '0'.

This transformation enables us to create a more streamlined binary classification for our intent detection models, facilitating a clearer distinction between different user intents in our analysis.

b) Cleaning the text :

in this step we used nltk library to remove special characters and special words then we lemmatize the result.

c) Text Embedding :

at the end of processing we have main data for BERT and the lemmatized data for LSTM, before feeding it to the models we used custom tokenizer for LSTM model fitted on the Data . but for BERT we used the pretrained BERT Tokenizer .

3. Experimental Setup

The experiments were conducted in the Google Colab environment, which provides access to powerful computational resources including GPU acceleration, facilitating efficient model training and evaluation.

The experiments were configured with the following parameters:

- **Batch Size**: We used a batch size of 32. This batch size strikes a balance between training speed and stability, allowing for efficient gradient updates while avoiding excessive memory usage.

- **Epochs**: The models were trained for 20 epochs. This duration was chosen based on preliminary tests to ensure sufficient learning without overfitting.
- **Learning Rate** : we used 0.001 as learning rate

The combination of these hyperparameters was selected based on a series of preliminary experiments aimed at optimizing the performance of our models. Each model's performance was monitored using standard evaluation metrics, ensuring that the chosen parameters provided the best trade-off between training time and model accuracy.

The entire setup aimed to leverage the robustness of the Adam Optimizer and the relevance of Binary Cross-Entropy loss for binary classification, providing a reliable framework for training our NLP models to detect user intent accurately.

4. Model Training :

For BERT models we used pretrained bert-base-uncased and its pretrained tokenizer and finetune it on our data. while for LTSM we train directly on ourdata.

Training setup :

- **Loss Function**: We used Binary Cross-Entropy as the loss function. This is appropriate for our binary classification task, as it measures the performance of a classification model whose output is a probability value between 0 and 1. Binary Cross-Entropy loss helps in optimizing the model by minimizing the difference between the predicted and actual labels.

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

- **Optimizer**: For both models, we employed the Adam Optimizer, known for its computational efficiency and low memory requirements, The Adam Optimizer is well-suited for handling sparse gradients on noisy problems, making it an ideal choice for our NLP tasks.

5. Performance metrics:

To evaluate the performance of our models, we used several classification metrics:

- **Accuracy**: Measures the proportion of correct predictions out of the total predictions. It is a useful metric for getting a quick overview of the model's performance.
- **Precision**: Indicates the proportion of true positive predictions out of all positive predictions made by the model. It is crucial when the cost of false positives is high.
- **Recall**: Reflects the proportion of true positive predictions out of all actual positives. It is important when the cost of false negatives is high.
- **F1 Score**: The harmonic mean of precision and recall, providing a balanced measure when there is an uneven class distribution.
- **ROC-AUC (Receiver Operating Characteristic - Area Under the Curve)**: Evaluates the model's ability to distinguish between classes. A higher AUC indicates better performance in distinguishing between positive and negative classes.

These metrics provide a comprehensive assessment of our models' performance, helping us understand their strengths and weaknesses in different aspects of binary classification.

6. Results

This section presents the performance metrics of the two models, LSTM and BERT, on our binary classification task. The models were evaluated using various metrics, including accuracy, F1-score, precision, and recall, for both classes (False and True) as well as overall metrics.

6.1 Model Performance

The detailed performance metrics are summarized in Table 2.

Table 2: Performance Metrics for LSTM and BERT Models

Model	**LSTM**				**Bert**			
Class	**Accuracy**	**F1-score**	**Precision**	**Recall**	**Accuracy**	**F1-score**	**Precision**	**Recall**
False	0.97	0.91	0.85	0.97	0.94	0.93	0.92	0.94

True	0.14	0.23	0.53	0.14	0.59	0.63	0.69	0.59
total	0.84	0.57	0.69	0.56	0.89	0.79	0.81	0.77

6.2 Model Complexity

In addition to performance metrics, the complexity of each model is also presented in terms of the number of trainable parameters and the model size in Table 3.

Table 3 : Model Complexity Metrics

Models	Trainable Parameters Number	Model Size
BERT	109,483,009	417.64 MB
LSTM	66,689	260.50 KB

6.3 Discussion

The results demonstrate that the BERT model outperforms the LSTM model in almost all performance metrics across both classes. Specifically, the BERT model achieves a significantly higher F1-score and accuracy for the True class compared to the LSTM model. This indicates that BERT is better at correctly identifying instances of the True class, which is often the more challenging class in binary classification tasks.

For the False class, both models perform well, but BERT still shows a slight edge in terms of precision and recall. The overall metrics (accuracy, F1-score, precision, and recall) further highlight BERT's superior performance, with an overall accuracy of 0.89 compared to LSTM's 0.84, and an F1-score of 0.79 compared to LSTM's 0.57.

However, the complexity and resource requirements of BERT are significantly higher, with over 109 million trainable parameters and a model size of 417.64 MB, compared to LSTM's 66,689 parameters and 260.50 KB model size. This indicates a trade-off between model performance and computational efficiency.

In summary, while the BERT model provides superior classification performance, the LSTM model remains a viable option when computational resources are limited. The choice between these models should therefore be guided by the specific constraints and requirements of the application.

6.4 Analysis of Confusion Matrices

The confusion matrices for the LSTM and BERT models (figure 06 and figure 07) provide a detailed breakdown of the models' classification performance on the binary user intent classification task. Here's a detailed discussion:

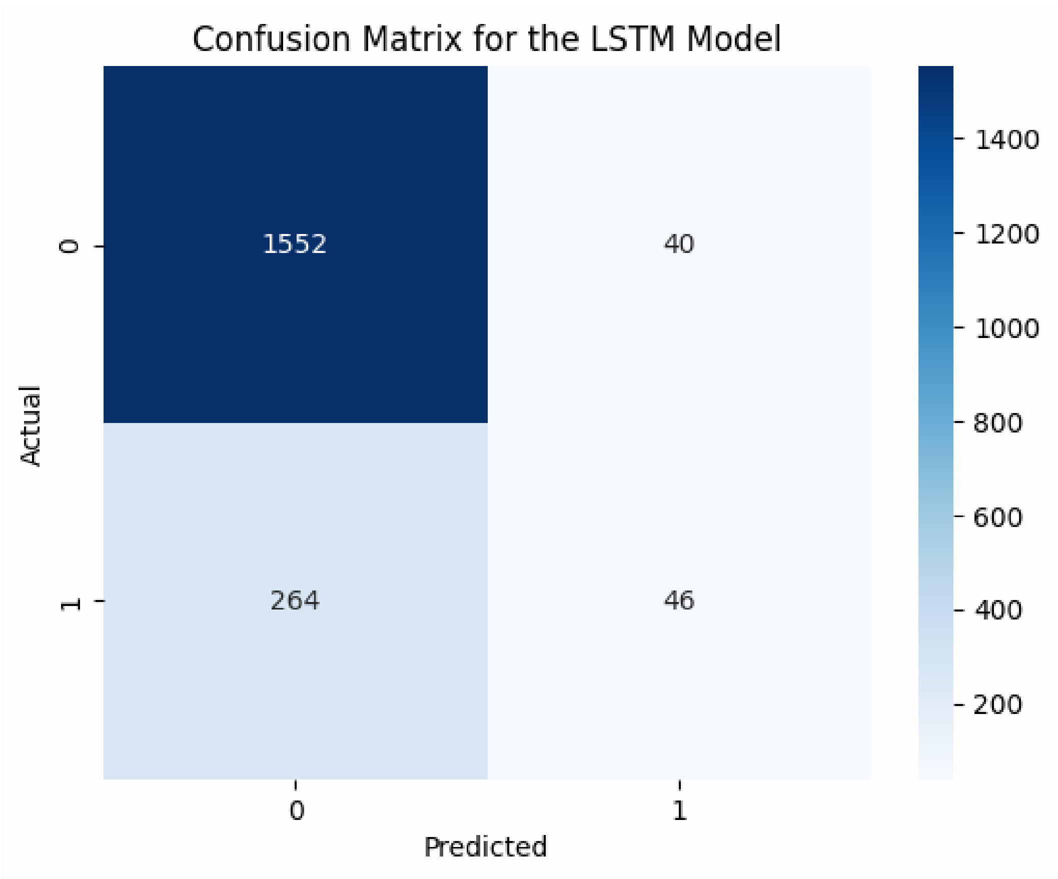


Figure 6 : Confusion matrix for the LSTM model

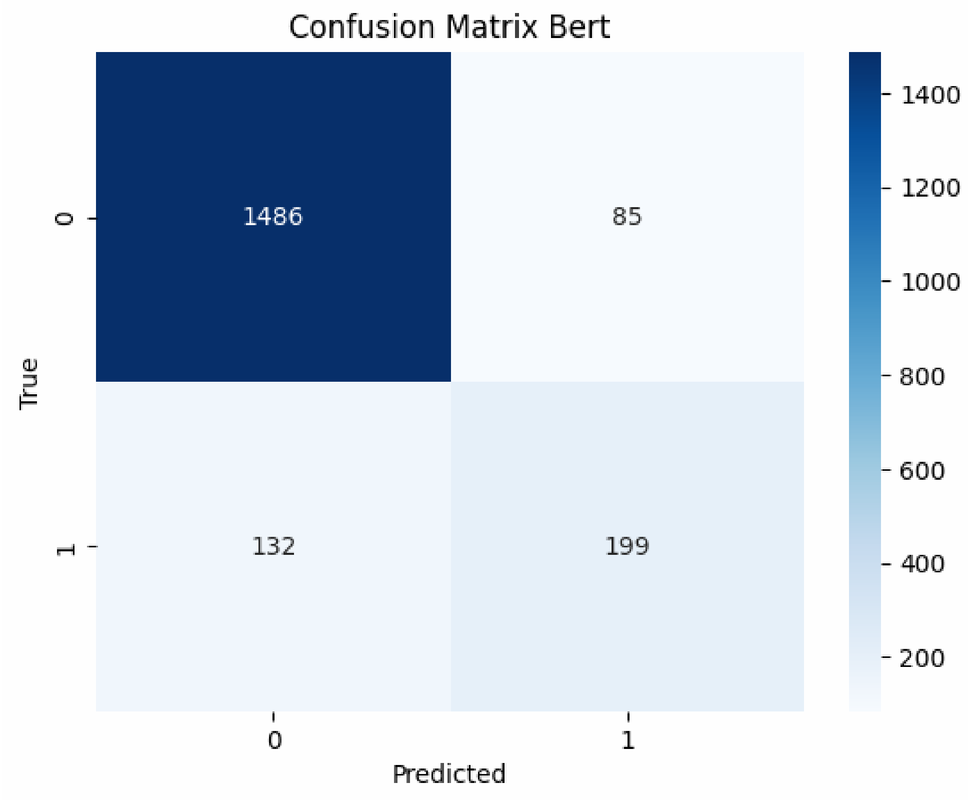


Figure 7 : Confusion matrix for the BERT model

LSTM Model:

- **Strengths**: LSTM shines in its ability to identify negative cases (class 0). Its high true negative rate signifies a commendable proficiency in correctly filtering out irrelevant or unwanted instances. This makes LSTM a valuable tool when dealing with tasks where accurately classifying negative examples is paramount.
- **Weaknesses**: LSTM stumbles when it comes to recognizing positive cases (class 1). The low recall and F1-score paint a picture of a model struggling to capture a significant portion of these crucial instances. This translates to a high number of false negatives – positive cases that are mistakenly classified as negative. In user intent classification, this could lead to misinterpreting user requests and hindering the overall effectiveness of the system.

BERT Model:

- **Strengths**: The BERT model demonstrates better balance with a higher true positive rate, leading to higher recall and F1-score. Its precision is also substantially better than the LSTM model.

- **Weaknesses**: While BERT's false positive rate is higher than LSTM's, the overall improvement in true positive identification outweighs this drawback.

Conclusion

- **Overall Performance**: The BERT model outperforms the LSTM model across all key metrics (accuracy, precision, recall, and F1-score). This aligns with the observations from the ROC curve analysis, further confirming BERT’s superiority in this classification task.
- **Model Selection**: Based on these confusion matrices and calculated metrics, the BERT model is clearly the better choice for the user intent classification task.

6.5 Roc Curve

The ROC (Receiver Operating Characteristic) curve presented in the Figure 7 is a graphical plot that illustrates the performance of two binary classifiers (BERT and LSTM) for user intent classification. Here’s a detailed analysis of the ROC curve:

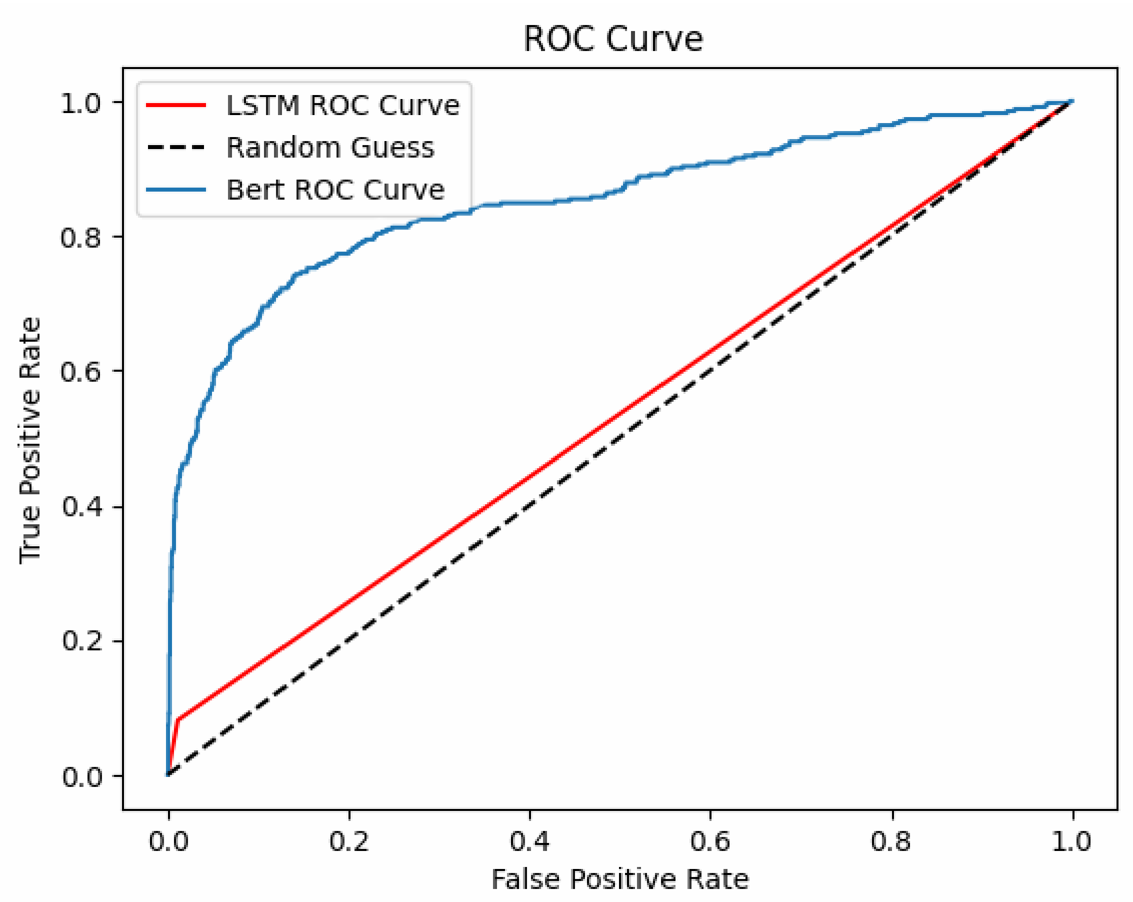


Figure 8 : The ROC curve for both Bert and LSTM

ROC Curve Components

1. **True Positive Rate (TPR)**: Also known as recall or sensitivity, this is plotted on the Y-axis. It represents the proportion of actual positives correctly identified by the model.
2. **False Positive Rate (FPR)**: This is plotted on the X-axis. It represents the proportion of actual negatives incorrectly identified as positives by the model.
3. **Diagonal Line**: The dashed line represents random guessing. Any point along this line indicates that the classifier is no better than random guessing.

Analysis of the Curve

1. **BERT Model (Blue Curve)**
   - The BERT model's ROC curve is significantly above the diagonal line, indicating that it performs much better than random guessing.
   - The curve shows a high true positive rate with a relatively low false positive rate across most of the threshold range, suggesting that BERT is a strong classifier for this task.
   - The area under the ROC curve (AUC) for BERT would likely be quite high, indicating excellent performance.
2. **LSTM Model (Red Curve)**
   - The LSTM model's ROC curve is also above the diagonal line, indicating it performs better than random guessing but not as well as the BERT model.
   - The LSTM curve lies below the BERT curve, indicating that at any given false positive rate, the true positive rate for the LSTM is lower than that for the BERT model. This suggests that the BERT model is superior to the LSTM model for this particular classification task.
   - The area under the ROC curve (AUC) for LSTM would be lower than that of BERT, suggesting comparatively poorer performance.

Conclusion

- **Model Performance**: BERT outperforms LSTM in the binary user intent classification task, as indicated by its ROC curve which lies above that of the LSTM model.
- **Classifier Quality**: Both models perform better than random guessing, but the BERT model demonstrates a much higher capability in distinguishing between the two classes.
- **Decision Making**: If one were to choose between these two models based on this ROC curve, BERT would be the preferred model due to its superior performance as reflected in the ROC curve.

6.8 Acurracy Curve Analysis

BERT Model:

- The accuracy plot for the BERT model shows a clear distinction between train accuracy and test accuracy over 10 epochs.
- **Train Accuracy:** It starts around 0.86 and increases steadily, reaching approximately 0.96 by the 10th epoch.
- **Test Accuracy:** It starts around 0.87, peaks at approximately the 3rd epoch at around 0.90, and then exhibits a slight decrease and stabilization around 0.88-0.89.

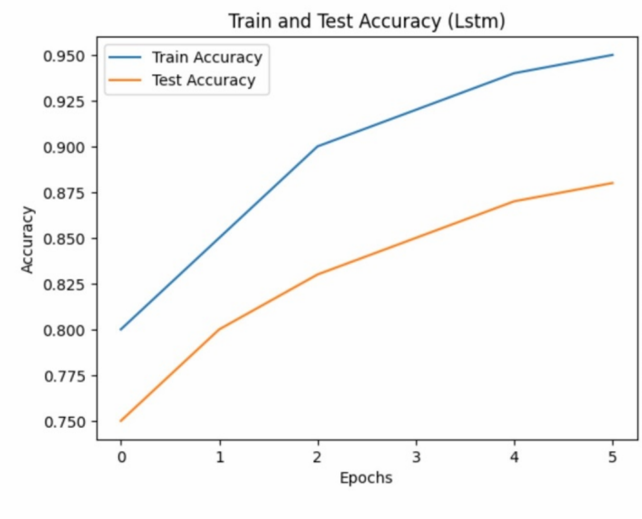
LSTM Model:

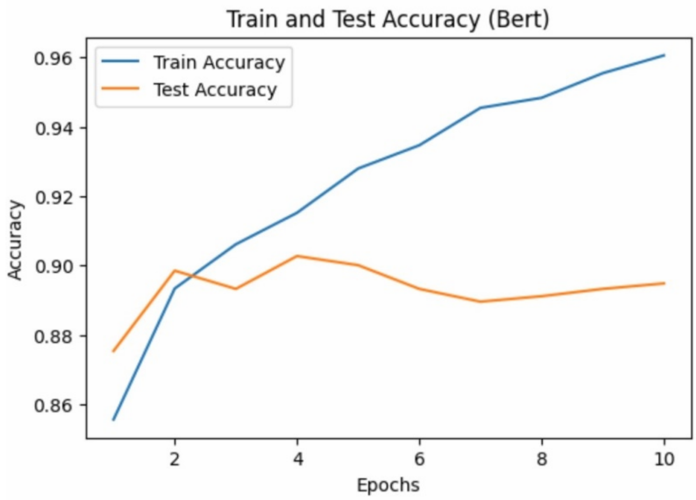
- The LSTM model's accuracy plot is depicted for 5 epochs.
- **Train Accuracy:** It shows a steady increase from around 0.75 to approximately 0.95 by the 5th epoch.
- **Test Accuracy:** It starts at 0.75, steadily increasing to about 0.88 by the end of the 5th epoch.

Interpretation:

- Both models show a significant increase in training accuracy over epochs, indicating effective learning.

- The BERT model achieves higher train accuracy compared to the LSTM model, but the test accuracy does not improve beyond the initial epochs, suggesting possible overfitting.
- The LSTM model shows a more consistent improvement in both train and test accuracy, with a narrower gap between them, indicating better generalization compared to the BERT model.





Loss Curve Analysis:

BERT Model:

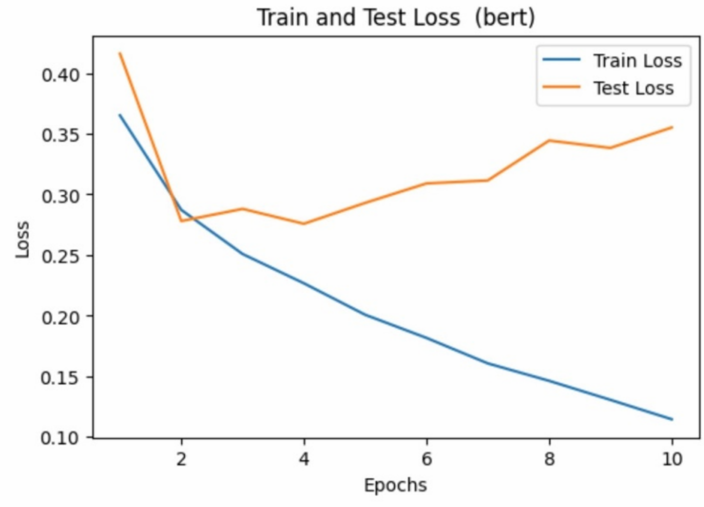
- The loss plot for the BERT model shows a significant decrease in train loss, from around 0.35 to 0.10, indicating effective learning.
- The test loss, however, decreases initially and then increases slightly, stabilizing around 0.35, which suggests overfitting after the initial epochs.

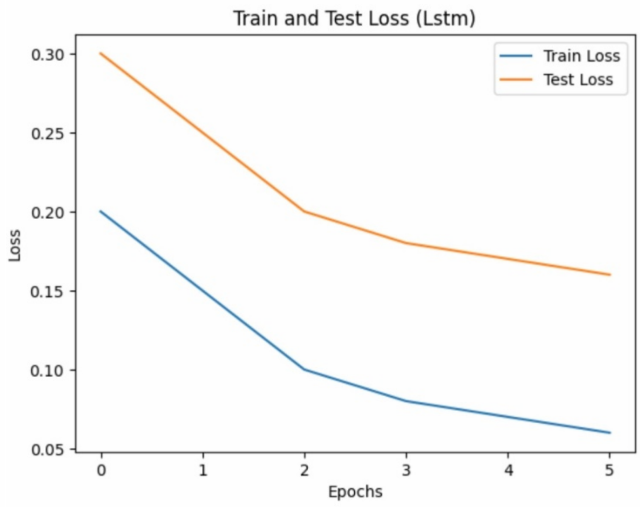
LSTM Model:

- The loss plot for the LSTM model shows a consistent decrease in both train and test loss over 5 epochs.
- **Train Loss:** Decreases from around 0.30 to approximately 0.05.
- **Test Loss:** Decreases from around 0.30 to about 0.20, showing continuous improvement without signs of overfitting.

Interpretation:

- The BERT model shows signs of overfitting, as evidenced by the divergence of train and test loss after the initial epochs.
- The LSTM model demonstrates better generalization, with both train and test loss decreasing consistently.





- **BERT Model:** Exhibits high training accuracy and low training loss, but the test performance suggests overfitting, meaning the model fits the training data very well but does not generalize effectively to unseen data.
- **LSTM Model:** Shows more balanced performance with continuous improvement in both accuracy and loss for train and test data, indicating better generalization capabilities.

7. Conclusion

Our experiments shed light on the strengths and weaknesses of BERT and LSTM models for user intent classification. While BERT emerged victorious, its victory dance comes with a hefty price tag – increased computational demands due to its larger size. This trade-off between performance and efficiency presents a double-edged sword for developers.

For applications where top-notch classification accuracy reigns supreme and resources are abundant, BERT stands tall as the champion. Its ability to excel, particularly in identifying the often-tricky positive class, makes it a compelling choice. However, for scenarios where computational resources are limited or the primary focus lies in efficiently filtering out irrelevant negative cases, LSTM remains a strong contender. Its lower complexity makes it a more lightweight option for resource-constrained environments.

The future holds exciting possibilities for user intent classification. Advancements in creating efficient BERT implementations or even hybrid approaches that merge the strengths of both models could unlock even better solutions. By staying abreast of these developments, researchers and practitioners can ensure their systems wield the most effective tools for deciphering user intent, ultimately leading to more intuitive and user-friendly experiences.

General Conclusion

The use of social media creates a huge set of data that serves researchers in the field of natural language processing (NLP) and serves as a source of data for building intelligent systems capable of analyzing different parts of human lives, such as behavior and feelings. The goal of our research was to harness the potential of NLP and advanced machine learning techniques, namely LSTM and BERT models, to predict the user's current or intended location from dialogues within subtitles of TV series.

Through the innovative application of these models, our study demonstrated the feasibility of accurately classifying user intentions regarding sites, achieving high accuracy on our dataset. Our results demonstrate the efficiency of the model in identifying dialogues in which the user does not discuss a site, category “0”, with high accuracy and recall. ,However, it indicates a need for improvement in detecting positive ,instances where user intent involves location ,category “1,” particularly in terms of recall. The disparity in performance between the “O” category and the “l” category indicates that while the model can differentiate non-site dialogue effectively, it struggles with the nuances of site intent.

The importance of our work lies not only in the impressive accuracy achieved, but also in the model's contribution to predictive capabilities in the field of NLP.

Our approach provides a foundation for developing intelligent systems capable of understanding and predicting user intentions based on textual data, with potential applications in virtual assistants, targeted advertising, and location-based services.

To address this, future work could involve class rebalancing to provide a more evenly distributed training set, which may help improve the model's sensitivity to location-related dialogues. Moreover, enhancing the feature extraction process or improving the model structure can provide deeper contextual understanding, thus improving the classification of site intent dialogues

References

1. *Long short-term memory.* **Hochreiter, S., & Schmidhuber, J"urgen.** s.l. : 1735–1780., 1997, Neural Computation, p. 9(8).

2. **Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova.** *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.* Minneapolis, Minnesota : Association for Computational Linguistics, 2019.

3. *User Intent Recognition and Satisfaction with Large Language Models: A User Study with ChatGPT.* **Anna Bodonhelyi, Efe Bozkir, Shuo Yang, Enkelejda Kasneci, Gjergji Kasneci.** 2024, arXiv 2402.02136.

4. *Training language models to follow instructions with human feedback.* **Long Ouyang et al.** 2022, arXiv, p. 2203.02155.

5. *User Feedback-based Online Learning for Intent Classification.* **Kaan Gönç et al.** October, 2023. INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION, Paris, France, .

6. *Understanding User Intent Modeling for Conversational Recommender Systems: A Systematic Literature Review.* **Siamak Farshidi et al.** s.l. : arXiv:2308.08496v1, 2023.

7. *Using Large Language Models to Generate, Validate, and Apply User Intent Taxonomies.* **Chirag Shah et al.** s.l. : arXiv:2309.13063v3, 2024.

8. *A Preliminary Investigation of User Perception and Behavioral Intention for Different Review Types: Customers and Designers Perspective.* **Atika Qazi et al.** s.l. : Hindawi Publishing Corporation e Scientific World Journal, 2014.

9. *Intent Classification in Question-Answering Using LSTM Architectures.* **Giovanni Di Gennaro et al.** s.l. : arXiv:2001.09330v1, 2020.

10. *Recurrent Neural Network and LSTM Models for Lexical Utterance Classification.* **Suman and Andreas.** s.l. : INTERSPEECH, 2015.

11. *ILLUMINER: Instruction-tuned Large Language Models as Few-shot Intent Classifier and Slot Filler.* **Bhat, Paramita Mirza and Viju Sudhi and Soumya Ranjan Sahoo and Sinchana Ramakanth.** s.l. : arXiv, 2024. 2403.17536.

12. *Development of Social Media Analytics System for Emergency Event Detection and Crisis Management.* **Khatoon et al.** s.l. : Computers, Materials and Continua, 2021, Computers, Materials and Continua, pp. 3079-3100.