

Chapter 1

Introduction

In the ever-evolving landscape of Artificial Intelligence (AI), Natural Language Understanding (NLU) has emerged as a pivotal force, bridging the gap between human communication and machine interpretation. NLU empowers machines to comprehend and decipher the complexities of human language, enabling them to engage in meaningful interactions, extract valuable insights, and perform a variety of sophisticated tasks. This domain encompasses a wide spectrum of capabilities, ranging from fundamental word-level processing to sophisticated analysis of linguistic nuances, sentiment, and intent.

The advent of NLU has ushered in a transformative era, revolutionizing the way we interact with machines. Intelligent systems equipped with NLU capabilities can now engage in near natural conversations, providing personalized assistance through virtual

assistants, powering machine translation tools that break down language barriers, and revolutionizing search engines by understanding the context and intent behind user queries.

At the heart of NLU lies the ability to decode the intricacies of human language, a process that goes beyond merely understanding the literal meaning of words. It involves unravelling the underlying context, intent, and emotions conveyed through language, a task made challenging by the inherent ambiguities, subtleties, and cultural references that permeate human communication.

The rise of Large Language Models (LLMs), including OpenAI's GPT series, Amazon's Titan models, Google AI's LaMDA (now known as Gemini), and DeepSeek etc. has significantly advanced NLU capabilities. These models, trained on massive text and multi-modal datasets, exhibit remarkable prowess in generating human-quality text, translation, crafting creative content, and answering questions in an informative manner. This breakthrough has propelled innovation in a wide range of NLU applications, including machine translation, chatbots, virtual assistants, and content creation.

1.1 Background

Despite significant advancements in NLU, adapting these techniques to low-resource languages remains a considerable challenge. These languages, often characterized by limited digital data, create unique obstacles for training or fine-tuning robust off-the-shelf NLU models. This data scarcity introduces several difficulties, including:

Poor Model Performance: The constrained availability of training data poses a substantial risk to the performance of NLU models. Such models, trained on limited datasets, can struggle to generalize effectively to unseen data. Consequently, this

limitation can manifest in inaccurate predictions and suboptimal performance when confronted with real-world tasks.

Cultural Biases: NLU models inherently mirror the cultural context embedded in their training data. For low-resource languages, this can result into models that perpetuate cultural biases present in the limited data available. The outputs of these models may unwillingly reinforce stereotypes and contribute to discriminatory outcomes.

Gaps in Linguistic Nuances: Low-resource languages exhibit distinct linguistic features, encompassing unique grammatical structures, vocabulary, and idioms. These nuances, integral to effective communication in these languages, pose a challenge for NLU models. The limited exposure to such linguistic diversity during training can result in models struggling to capture and interpret these intricacies accurately, leading to misinterpretations and errors.

Addressing these challenges is crucial for improving the applicability of NLU techniques across diverse linguistic landscapes. Tackling issues such as suboptimal model performance, cultural biases, and gaps in linguistic nuances necessitates innovative strategies that extend beyond traditional NLU methodologies. As we explore the complexities of low-resource languages, it is essential to develop solutions that enhance data availability, reduce biases, and integrate linguistic diversity into NLU models.

1.2 Motivation

The motivation for this research stems from the researcher's prior experience in bootstrapping Machine Translation tools for low-resource languages, involving key processes such as data curation, synthetic data generation, and subsequently the development of both research and production-level Machine Learning (ML) models. A major challenge in this process these days is the heavy reliance on pre-trained Language

Models (LMs), which are primarily trained on data from high-resource languages and inherently influenced by their cultural contexts. These LMs often struggle to generalize effectively to languages with distinct syntactico-semantic structures, socio-cultural nuances, and contextual interpretations, leading to misrepresentation and underperformance in low-resource linguistic settings. Moreover, biases embedded in these models, including gender bias, societal stereotypes, and cultural misrepresentations, pose serious challenges to fairness, inclusivity, and ethical usage. Addressing these issues requires a comprehensive investigation into NLU applications, alongside the development of Bias Identification and Mitigation strategies specifically designed for low-resource languages, ensuring that these models are both linguistically adaptable and socially responsible.

1.3 Research Problem and Objectives

The central focus of this research is bootstrapping NLU applications in low-resource languages, with a specific emphasis on Sentiment Analysis in Hindi. The overarching goal is to develop robust methodologies that transcend linguistic barriers, enabling effective sentiment analysis and subsequently other NLU applications in languages with limited linguistic resources. To achieve this, the research is guided by the following objectives:

Development of an Efficient NLU Framework for Low-Resource Languages: The primary aim of this research is to establish an efficient framework for NLU applications in low-resource languages. With the evolution of technologies ranging from rule-based and statistical models to advanced LLMs, the critical task is to discern the optimal approach for supporting NLU use cases. This involves a comprehensive exploration of various frameworks, assessing their effectiveness in the context of Subjectivity and

Sentiment analysis. The research will provide insights into the strengths and limitations of each framework, demonstrating how to achieve optimal results for specific use cases. This includes the comparative evaluation of rule-based, statistical, and LLM models, considering factors such as accuracy, efficiency, and computational requirements.

Ethical Considerations in Bootstrapping NLU Applications: Given the inherent challenges of limited data points in low-resource languages, ethical considerations play a pivotal role in the bootstrapping process of NLU applications. This objective focuses on evaluating the ethical implications of different NLU models, particularly in the context of potential biases introduced during the bootstrapping phase. The research aims to uncover biases within NLU models and guide practitioners on making data-driven decisions to mitigate such biases, ensuring responsible and ethical NLU applications.

Expanding on this objective further, we aim to provide concrete strategies for mitigating biases identified in NLU models through data-driven decision-making. By evaluating different models and identifying potential sources of bias, the research contributes practical guidelines for practitioners to make informed decisions in the development and deployment of NLU applications in low-resource languages.

Enhancing Generalization Across Low-Resource Languages: Building on the previous objectives, this research objective focuses on enhancing the generalization capabilities of NLU models across similar low-resource languages. By identifying commonalities and unique linguistic features, the research aims to contribute to the development of strategies that can effectively adapt to various low-resource linguistic contexts. This research objective seeks to deepen our understanding of linguistic commonalities and unique features across different low-resource languages. By undertaking a comprehensive examination of linguistic structures, vocabulary, and contextual nuances, in a selected set of similar low-resource languages, the goal is to

enhance the generalization capabilities of NLU models. This involves identifying transferable knowledge and linguistic patterns that transcend specific languages, facilitating the development of strategies with adaptable architectures. The significance lies not only in crafting solutions that cater to individual low-resource languages but also in creating a framework that can intelligently generalize insights, ensuring scalability and effectiveness across diverse linguistic landscapes. Through this objective, the research aims to contribute to the development of NLU models that go beyond mere language-specific optimizations, fostering a more universally applicable solution that embraces the richness of linguistic diversity while maintaining robust performance and impact.

Through these research objectives we aim to address the critical challenges associated with NLU applications in low-resource languages, encompassing technology selection, ethical considerations, bias mitigation, and generalization capabilities. Through a comprehensive exploration of these objectives, this research aspires to contribute actionable insights to the development of responsible and impactful NLU applications in linguistically diverse contexts.

1.3.1 Research Questions

The pursuit of these objectives prompts several key Research Questions (RQs):

1. How can Subjectivity Identification and Sentiment Analysis be effectively implemented in low-resource languages, such as Hindi, leveraging different NLU techniques?
2. To what extent can existing techniques identify and mitigate gender biases in NLU models, fostering fairness in NLU application in low resource languages?

-
3. What are the ethical considerations in the development and deployment of NLU models, particularly in the context of bias detection and mitigation, and how can responsible AI practices be implemented?

1.3.2 Methodology

Addressing the above multifaceted RQs necessitates a comprehensive and multi-dimensional research methodology that integrates insights from relevant literature, leverages state-of-the-art (SoTA) NLU techniques, and employs rigorous empirical evaluations.

1.3.2.1 Sentiment Analysis in Low-Resource Languages

To effectively implement Sentiment Analysis in low-resource languages like Hindi, the research will commence with an extensive review of existing literature on NLU techniques applied to low-resource languages. Lexical Rules (LRs) and linguistic devices will be formalized and adapted for the specific linguistic characteristics of selected low-resource languages. Multiple frameworks, ranging from rule-based methods to LLMs, will be explored to discern the optimal approach. Empirical evaluations will be conducted on diverse datasets from entertainment, lifestyle, politics, sports, and technology domains, with a focus on assessing accuracy and adaptability of proposed model.

1.3.2.2 Bias Identification and Mitigation in NLU Models

This exercise will investigate the identification and mitigation of biases in NLU models. A literature review will provide insights into existing techniques for bias detection and mitigation in the broader context of NLU. This will be followed by an empirical evaluation of these techniques on datasets from low-resource Indian languages. Comparative analyses will gauge the effectiveness of contrastive test sets in identifying

biases, and Prompt Engineering (PE) and In-context Learning (ICL) in mitigating biases, fostering fairness in NLU applications.

1.3.2.3 Ethical Considerations in NLU Model Development and Deployment

To address the ethical considerations, a multi-faceted approach will be adopted. The research will dive into existing literature on ethical considerations in AI and NLU model development, focusing on bias detection, mitigation, and responsible AI practices. Empirical evaluations will be conducted to assess biases within NLU models, particularly in the context of low-resource languages. The development of a novel methodology for ethical evaluation, influenced by template-driven synthetic data generation, will be a focal point. Additionally, guidelines for responsible AI deployment will be proposed based on the research findings.

This research methodology ensures a holistic exploration of the RQs, encompassing theoretical foundations, empirical evaluations, and innovative methodologies. The iterative nature of the methodology allows for continuous refinement and adaptation as the research unfolds, ensuring a robust exploration of the complexities inherent in NLU applications in low-resource languages.

1.3.3 Significance of the Study

In an era characterized by unprecedented technological advancements, this study holds profound significance that transcends the confines of academic inquiry. The ability of AI systems to comprehend and respect linguistic diversity emerges as a pivotal focal point in a world increasingly interconnected through digital interfaces. This research addresses the imperative need for AI systems to navigate low-resource languages, contributing to a more inclusive and culturally aware technological landscape.

The outcomes of this study carry transformative potential across diverse applications, from enhancing sentiment analysis in social media to cultivating culturally sensitive virtual assistants. In a globalized society where communication knows no bounds, the development of NLU models tailored for low-resource languages becomes not merely a technical pursuit but a crucial bridge fostering meaningful engagement across linguistic communities. The study's practical implications extend to empowering users in linguistically diverse regions, facilitating more accurate and contextually relevant interactions with AI systems.

The ethical considerations and guidelines proposed within this research illuminate a path towards the responsible development and deployment of AI technologies. In an environment where biases within NLU models can inadvertently perpetuate stereotypes and discrimination, the study provides a critical examination of these biases. By unravelling the intricacies of bias within NLU models, this research equips developers, policymakers, and practitioners with insights necessary for the creation of fair and inclusive AI systems.

The exploration of biases within NLU models is particularly timely and crucial in addressing ethical concerns surrounding AI applications. Biases have the potential to shape digital narratives, influence decision-making processes, and exacerbate societal inequalities. By scrutinizing biases within NLU models, this study contributes not only to the academic discourse but also to the practical realm of developing technologies that are ethically sound and socially responsible.

Beyond the immediate scope of NLU models, this study ventures into the realm of human-machine dynamics by exploring the relationship between human intelligence and the capabilities of LMs. As society grapples with the implications of AI becoming an integral part of our daily lives, understanding the evolving dynamics between humans and

machines is paramount. The insights derived from this exploration open avenues for shaping the future trajectory of AI research and practice.

In the grand scheme of technological innovation, the impact of this research extends to shaping the landscape of NLU. By fostering inclusivity, the study envisions a future where AI technologies are not only technically advanced but also socially aware and culturally sensitive. The responsible and ethical use of AI technologies emerges as a guiding principle, emphasizing the need for AI to align with human values, diversity, and inclusiveness.

As we embark on this journey of exploration and innovation, the potential of this research to influence the trajectory of AI technologies is palpable. It aspires to contribute to a paradigm shift wherein AI becomes an enabler of positive social change, promoting fairness, inclusivity, and ethical practices. In essence, the significance of this study lies not just in its academic contributions but in its potential to shape the ethical, social, and technological contours of the AI landscape in the years to come.

1.4 Overview of Thesis

The forthcoming chapters of this thesis unfold a comprehensive exploration into the intricacies of NLU applications in low-resource languages, with a specific focus on Subjectivity Identification and Sentiment Analysis in Hindi. Each chapter in this scholarly journey adds a distinctive layer to the overarching narrative, deepening our comprehension of the challenges, opportunities, and ethical dimensions intrinsic to the development and deployment of NLU models.

Chapter 2: Navigating the AI Landscape: Sentiment Analysis and Bias in NLU

This chapter serves as a pedagogical guide, immersing the reader in an in-depth review of existing literature and research in the realm of NLU, particularly focusing on Sentiment

Analysis. By providing a worm's eye view, it imparts foundational insights into NLU techniques, bias detection and mitigation strategies, and the evolving ethical considerations in AI. Synthesizing knowledge from seminal works, this chapter establishes theoretical underpinnings that shape our research, creating a robust foundation for subsequent investigations. The reader gains a comprehensive understanding of the past, present, and future of NLU and bias within the dynamic landscape of AI.

Chapter 3: NLU in low-resource languages: A case of Sentiment Analysis

This chapter details the research methodology, offering a comprehensive overview of the approach employed to address RQs related to Sentiment Analysis in low-resource languages. Drawing insights from relevant literature, the methodology encompasses empirical evaluations, comparative analyses, and innovative frameworks tailored for ethical considerations. The outlined framework provides a robust foundation for exploring Sentiment Analysis. Expanding upon the methodology, this chapter goes into the practical implementation of Sentiment Analysis in low-resource languages, with a specific focus on Hindi and over a dozen other Indian languages. Exploring a spectrum of NLU frameworks, ranging from rule-based models to advanced LLMs, it provides a granular understanding of their strengths, limitations, and optimal applications within the unique challenges posed by low-resource linguistic contexts.

Chapter 4: Responsible AI: Identifying and Mitigating Bias in NLU Models

Unfolding the research methodology, this chapter delineates the approach taken to tackle RQs pertaining to Bias identification and Mitigation in low-resource languages. Rooted in insights from relevant literature, the methodology encompasses a blend of quantitative and qualitative evaluations, comparative analyses, and innovative frameworks tailored for ethical considerations. Through systematic steps, this chapter sheds light on the

exploration of Biases across a spectrum of NLU models, spanning from Rule-based models to advanced LLMs.

This chapter further dives deeper into the identification and mitigation of biases in NLU models, aligning with pertinent RQs. Drawing from existing literature and empirical evaluations, it critically evaluates the effectiveness of techniques such as template-based test-set creation, optimized evaluation, PE, and ICL in promoting fairness within NLU applications in low-resource languages. Importantly, the applicability of these methods extends beyond low-resource languages, encompassing languages and models of all sizes. The chapter concludes by proposing guidelines for responsible AI practices, aligning with RQs and addressing the imperative need for ethical considerations in the AI landscape.

Chapter 5: Conclusion and Future Directions

Chapter 5 combines the key findings drawn from diverse chapters. It not only synthesizes the crux of the research but also accentuates the significance of the undertaken investigations, illuminating their broader implications in the realm of NLU applications within low-resource languages.

Building on the foundational understanding established in Chapter 2, which explored the landscape of AI and its inherent biases, and drawing from the methodological intricacies detailed in Chapter 3, particularly in addressing Sentiment Analysis in low-resource languages, the synthesis in this chapter encapsulates the essence of the research journey. Moreover, Chapter 5 functions as a compass pointing towards the uncharted territories of future research. It unveils unexplored avenues within Bias identification and Mitigation in low-resource languages, echoing chapter 4's focused exploration into gender biases in NLU models. By offering insights and proposing methodological frameworks, this chapter contributes to the ongoing narrative of advancing NLU applications, particularly in linguistic settings with limited resources.