**2**

**An Overview of Natural Language Processing and Neural Networks**

**Learning Objectives**

After completing this chapter, the readers are expected to

* Learn the basics of Natural Language Processing (NLP) necessary for understanding the subsequent chapters.
* Gain knowledge about various semantic and syntactic paradigms in NLP
* Learn the basics of Neural Networks required to understand the subsequent chapters.
* Become familiar with the evaluation metrics employed in neural network modelling. dfgjskdhf

The natural evolution of languages has enabled humans to communicate and share ideas across geopolitical boundaries. However, the concept of a machine processing and understanding these human languages is a complex challenge. Since machines inherently understand numbers and numerical operations, it is necessary to convert natural language into a format that computers can comprehend. This is where NLP comes into play. NLP is an interdisciplinary field within computer science that encompasses techniques to make human language accessible and interpretable by machines. The foundational concepts in NLP draw from a wide range of disciplines, including theoretical computer science, linguistics, statistics, and artificial intelligence.

Statistical methods and linear machine learning models have long been utilised in NLP tasks. However, these traditional approaches depend heavily on manually crafted dictionaries and features, limiting their capacity to capture complex patterns and semantics in language. As the volume of textual data has increased and computational resources have become more accessible, neural architectures have gained prominence in modern NLP techniques. These neural models excel at capturing *latent knowledge* from data, processing inputs of varying lengths, leveraging information from long sequences of text, and autonomously learning features, reducing the need for extensive manual efforts.

This chapter is divided into two parts. In **Part I**, readers are first introduced to the fields of linguistics and NLP. Section 2.1 discusses the goals of computational linguistics and NLP. Section 2.2 describes various tasks in NLP and introduces the NLP pipeline. Section 2.3 explores the linguistic components of language, such as morphology, lexicon, and text normalisation techniques like stemming and lemmatisation. In Section 2.4, we provide an overview of different tokenisation and semantic analysis techniques. Section 2.5 focuses on syntax and grammar-based parsing methods, while Section 2.6 delves into semantics and semantic parsing. Finally, Section 2.7 presents the task of language modelling, emphasising conditional probability and the frequency of co-occurrence.

In **Part II**, we explore neural networks and related concepts to set the foundation for deep learning techniques discussed in later chapters. Section 2.8 introduces the perceptron and its applications in modelling a linear classifier. Section 2.9 presents multilayer perceptrons and popular non-linear activation functions. Section 2.10 covers the gradient-based training process for neural networks and error backpropagation. In Subsection 2.10.3, we discuss the various hyperparameters that influence neural network training. Section 2.11 focuses on the challenges that affect gradient descent. Finally, in Section 2.12, we examine the performance measures commonly used to evaluate deep learning tasks.

**Part I: Natural Language Processing**

Since the advent of computers, researchers have been captivated by the idea of teaching machines to interact like humans. As early as 1963, Joseph Weizenbaum developed ELIZA (Weizenbaum 1983), a rule-based chatbot designed to converse with humans. Fast forward to 2014, Eugene Goostman, a chatbot, passed the *Turing Test*, with human judges unable to discern that Eugene was, in fact, a bot. In 2017, Google revolutionised the field by introducing a new machine translation architecture, now famously known as *Transformers* (Vaswani et al. 2017). More recently, the success and widespread adoption of language models like ChatGPT, and its successors have generated immense interest in language models, accelerating research in both NLP and deep learning. This book is motivated by the need to enhance our understanding of language models and to demystify the concepts from beginner to advanced levels.

Computation aside, a very nuanced problem with language is its ambiguity, contextualisation, and its dynamic