UniTs - University of Trieste

Faculty of Data Science and Artificial Intelligence

Department of mathematics informatics and geosciences

Natural Language Processing

Lecturer:
Prof. Alberto Cazzaniga

Author:
Christian Faccio

September 30, 2025

github.com/christianfaccio christianfaccio@outlook.it

Preface

As a student of Data Science and Artificial Intelligence, I've created these notes while attending the **Natural Language Processing** course.

The course provides a comprehensive introduction to the field of natural language processing, covering both theoretical concepts and practical applications. The notes encompass a variety of topics, including:

- ML and DL Fundamentals
- Tokenization and Text Preprocessing
- Word Embeddings
- RNN for NLP
- Transformers and Attention Mechanisms
- Language Models (e.g., BERT, GPT)
- Understanding LLMs
- Visutal Language Models

While these notes were primarily created for my personal study, they may serve as a valuable resource for fellow students and professionals interested in natural language processing.

Contents

Bibliography

[1] David Goldberg. “What every computer scientist should know about floating-point arithmetic”. In: *ACM computing surveys (CSUR)* 23.1 (1991), pp. 5–48.

[2] *High Performance Computing — digital-strategy.ec.europa.eu*. https://digital-strategy.ec.europa.eu/en/policies/high-performance-computing.