Tokenization & Vocabulary

Week 8 - NLP Course 2025

NLP Course 2025

September 23, 2025

Tokenization & Vocabulary

Week 8: Overview

Learning Objectives

- Understand BPE
- Understand WordPiece
- Apply concepts in practice
- Evaluate different approaches

Topics Covered

- BPE
- WordPiece
- SentencePiece
- Subword tokenization

Introduction

Tokenization & Vocabulary

Content for Tokenization & Vocabulary...

Key Concepts

- Concept 1: BPE
- Concept 2: WordPiece
- Practical applications
- Current research directions

Summary

Key Takeaways

- Main learning point 1
- Main learning point 2
- Main learning point 3

Next week: Continue to next topic