

# Levenshtein Distance & WER Computation

Nuray Taġiyeva

March 7, 2025

# Introduction

- Speech recognition critical in healthcare, legal, and customer service.
- Importance of accuracy: Errors have high costs.
- Word Error Rate (WER): standard metric.
- Levenshtein Distance measures string similarity (insertions, deletions, substitutions).

(Jurafsky & Martin, 2009; MIT NLP Course)

# Theoretical Background

- Naïve Recursive Approach (Brute force):  $O(3^n)$  complexity.
- Dynamic Programming Approach: Efficient polynomial  $O(nm)$  complexity.
- Optimized DP (Rolling Array): Efficient memory use ( $O(\min(n, m))$  space complexity).
- Alternative Metrics: Damerau-Levenshtein, Jaro-Winkler.

# Application in ASR

- Applications: Virtual assistants (Siri, Alexa), Transcription services, Accessibility.
- TED-LIUM dataset for testing.
- Compared ASR Models:
  - Google Speech-to-Text
  - Whisper AI (OpenAI)

# WER Implementation & Visualization

- Implemented using Python (JiWER library).
- WER results:
  - Google STT: 11.34%
  - Whisper AI: 6.02%
- Visualization highlights differences clearly.

# Challenges & Future Improvements

- **Challenges:**

- Scalability issues for large datasets.
- Phonetic error limitations.

- **Proposed Improvements:**

- Phonetic-based metrics.
- Deep learning techniques (wav2vec2, BERT-based corrections).
- Semantic-aware WER computations.

- Jurafsky, D. & Martin, J.H. (2009). *Speech and Language Processing*.
- MIT NLP OpenCourseWare.
- Additional references (See full report and .bib file).