

Abstractive Text Summarization Datasets, Models, and Tokenization Approaches for Turkish and Hungarian

Batuhan Baykara, Tunga Güngör

Boğaziçi University, Computer Engineering, Istanbul, Turkey

- Text summarization
 - Extractive text summarization
 - Abstractive text summarization
- This work is related to abstractive text summarization
- Contributions:
 - Two large-scale publicly available summarization datasets for Turkish and Hungarian
 - Strong baselines for both datasets
 - Comparing pointer-generator model (commonly-used baseline model for summarization) with BERT-based models
 - Two morphological tokenization methods
 - SeparateSuffix
 - CombinedSuffix

