

ENGLISH TO ARABIC TRANSLATOR

By:

Sara Almutairi

DATASET

 English →

 Dataset Summary

 Sentence Pairs

 Language Pair

 Type

 Use Case

Arabic Translator

FeatureDetail

12,561

English → Arabic

Parallel Text Corpus

NMT (Seq2Seq Model)



TOOLS & LIBRARIES USED

Data Handling & Preprocessing:

- pandas, numpy – Data manipulation and analysis
- re, unicodedata.normalize – Text cleaning and normalization
- Tokenizer, pad_sequences – Text tokenization and sequence padding

Deep Learning (TensorFlow / Keras):

- Embedding, LSTM, Dense, Dropout – Model layers
- Input, Model – Model architecture
- EarlyStopping – Training optimization (prevent overfitting)
- Custom Attention Layer – For sequence-to-sequence modeling with context

User Interface:

- gradio – Simple and interactive web-based interface for testing the translator

Thank you.

