

Guided Projects Artificial Intelligence & Machine Learning

Guided Projects: Natural Language Processing

Machine Translation

Machine translation or MT refers to **fully automated software** that can translate **source content** into target languages. **Humans** may use MT to help them **render text and speech** into another **language**. **nMT** tools are often used to **translate vast amounts** of information involving millions of words that could not possibly be translated the traditional way. It is a sub-field of **computational linguistics** that **investigates** the use of **software to translate text** or speech from **one language to another**.

Question:

Implement a seq2seq model (using LSTM and any deep learning framework of your choice) to perform machine translation from English Language to French Language. The dataset is available here: <http://www.manythings.org/anki/> Download the fr-eng.zip file, which contains the txt file containing the English sentence followed by its French translation (separated by tab). Use only the first few thousands records and not the complete dataset. Make sure that all the necessary preprocessing steps are done prior to feeding the dataset to the model. Use the glove representation for word embedding (as it gives one representation per sentence)