|  |
| --- |
| TOI Applied Mathematics |
| Deep Learning Minor |
| Assignment 3 |
| Inholland, 1 februari 2024 |

in tijden van social media.

Lectoraat Jeugd en Samenleving, 2020

in tijden van social media.

Lectoraat Jeugd en Samenleving, 2020

in tijden van social media.

Lectoraat Jeugd en Samenleving, 2020



# Learning Objectives

* Prepare text data for deep learning using Python
* Train recurrent deep learning models using Tensorflow/Keras
* Train a transformer using Tensorflow/Keras

# Steps of the assignment

The ‘TED-Parallel-Corpus’ Data Set (<https://github.com/ajinkyakulkarni14/TED-Multilingual-Parallel-Corpus/tree/master> contains sentences from TED talks translated to a number of different languages. The goal of this assignment is to use this data set to train a model for automatic translation from one language to another.

# Step 1: Data preparation

Download the data set.

Choose two languages from the available languages: one source language and one target language. The sentences in the remaining languages can be discarded.

Apply relevant data preprocessing steps to create a suitable trainings set consisting of sentences from the two selected languages.

Step 2: RNN

Implement an RNN encoder-decoder for machine translation using Keras/TensorFlow.

Train the RNN encoder-decoder on your training set.

Step 3: Transformer

Implement a transformer encoder-decoder for machine translation using Keras/Tensorflow.

Train the transformer encoder-decoder on your training set.

Step 4: Sample results

Create a small test set of new sentences (not from the corpus). Design them in such a way that they will challenge the translation models in various ways.

Try the two translation models on the test sentences.

Make an informal, qualitative comparison between the translations of the two models.

# Requirements

This assignment needs to be done individually. You need to hand-in:

* Your code in github. Code needs to be of high quality and well-documented.
* A report of at most 6 pages describing at least the following:
* an introduction describing the context and goal of the task;
* the preprocessing methods used;
* a brief explanation of your implementations;
* outcomes of the qualitative evaluation;
* conclusions;
* discussion;
* references.

# Assessment

## Assessment criteria

|  |
| --- |
| 1. Report quality |
| All parts of the work are clearly described.  The paper has a logical structure.  The paper is of appropriate length.  Language is unambiguous and easy to read.  Correct use of spelling and grammar.  Layout is appropriate.  Figures and tables are used appropriately.  References are in accordance with APA style. |
| 2. Preprocessing |
| The data set is correctly loaded.  Necessary preprocessing steps are identified.  Necessary preprocessing steps are correctly applied.  The final data set has a suitable format. |
| 3. RNN |
| An RNN encoder-decoder is implemented correctly.  An RNN encoder-decoder is correctly trained on the data set. |
| 4. Transformer |
| A transformer encoder-decoder is implemented correctly.  A transformer encoder-decoder is correctly trained on the data set. |
| 5. Qualitative evaluation |
| Suitable test sentences are created.  The RNN encoder-decoder is correctly tested on the test sentences.  The transformer encoder-decoder is correctly tested on the test sentences.  The results of the tests are effectively qualitatively analysed and compared. |

## Rubrics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Criterion | 0 poor | 1 insufficient | 2 sufficient | 3 good | 4 excellent |
| 1 Report quality | The report is incomplete or the structure or language is so poor that it is not understandable. | The report does not describe all parts of the work or the structure, language or formatting contains mistakes that make the report hard to read. | The report describes all parts of the work. The structure, language or formatting contains mistakes, but these do not hamper  readability. | The report clearly describes all parts of the work. The structure, language and formatting are of high quality. | The report clearly describes all parts of the work. The structure, language and formatting are professional. |
| 2 Preprocessing | The data set is not correctly loaded or essential preprocessing steps are missing. No suitable training set is created. | The data set is correctly loaded. The most important preprocessing steps are identified and applied, but some steps are missing or suboptimal. The final training set is not optimal for training the models. Preprocessing code is of low quality. | The data set is correctly loaded. Essential preprocessing steps are identified and applied. The final data set has a suitable format. Code quality is sufficient. | The data set is correctly loaded. Useful preprocessing steps are identified and applied. Choices are motivated. The final data set has a suitable format. Code is efficient and of high quality. | The data set is correctly loaded. Useful preprocessing steps are identified and applied. Advanced preprocessing, like improved sentence tokenization is applied. The final data set has a suitable format. Code is professional. |
| 3 RNN | No working RNN encoder-decoder is implemented or the model is not successfully trained. | An RNN encoder-decoder is implemented and trained, but the implementation or training contains flaws. | An RNN encoder-decoder is implemented correctly using the right elements from Keras/ TensorFlow. The model is correctly trained on the data set. Code is of sufficient quality. | An RNN encoder-decoder is implemented correctly using the right elements from Keras/ TensorFlow. The model is correctly trained on the data set. Some hyperparameter training has been done or training is effectively  monitored and steered. Code is of high quality. | An RNN encoder-decoder is implemented correctly using the right elements from Keras/ TensorFlow. The model is correctly trained on the data set. The model is fully optimized using hyperparameter tuning. Training is correctly  monitored and steered. Code is professional. |
| 4 Transformer | No working Transformer encoder-decoder is implemented or the model is not successfully trained. | A transformer encoder-decoder is implemented and trained, but the implementation or training contains flaws. | A transformer encoder-decoder is implemented correctly using the right elements from Keras/ TensorFlow. The model is correctly trained on the data set. Code is of sufficient quality. | A transformer encoder-decoder is implemented correctly using the right elements from Keras/ TensorFlow. The model is correctly trained on the data set. Some hyperparameter training has been done or training is effectively  monitored and steered. Code is of high quality. | A transformer encoder-decoder is implemented correctly using the right elements from Keras/ TensorFlow. The model is correctly trained on the data set. The model is fully optimized using hyperparameter tuning. Training is correctly  monitored and steered. Code is professional. |
| 5 Qualitative evaluation | No useful test sentences are created or  the two models are not successfully tested on the test sentences. | Some test sentences are created and  the two models are tested on the test sentences, but the test sentences or tests contain flaws.  No or insufficient relevant differences between the models are identified. | Some suitable test sentences are created.  The two models are correctly tested on the test sentences.  Through qualitative analysis some relevant differences between the models are identified. | A well-chosen number of test sentences is created. The choice for the sentences is motivated.  The two models are correctly tested on the test sentences.  Through qualitative analysis  various relevant differences between the models are identified. Implications of the differences are discussed. | A well-chosen number test sentences is created.  Sentences are purposely constructed to reveal differences between the models.  The two models are correctly tested on the test sentences.  Through qualitative analysis the important differences between the models are identified. Implications of the differences are discussed in detail. |

# Grading

The grade for this assignment is computed as follows: *grade = 1+0.45\*points.* To get a passing grade at most one criterion can be assessed as *insufficient* or *poor*. If these condition is not met but the score is still 10 or higher, the grade will be 5.0.