

Identifying spam

Portfolio assignment 3

MAL2 Autumn 2024

In this assignment, you will implement, train and evaluate **at least two different language models** to classify text messages into *spam* and *not spam* (also known as *ham*). The dataset is obtained from [here](#) and is provided as is.

You will have to think about how to deal with the fact that the texts have different lengths – the simplest solution is to *pad* the texts with padding tokens and then ignore the padding tokens in the Embedding layer with `Embedding(mask_zero=True)`.¹

As stated above, you should construct at least two different language models. The models may differ in any way you find relevant, but some suggestions are:

- GRU vs LSTM vs SimpleRNN cells
- Character vs word embeddings (`TextVectorization(split="character" or "whitespace")`)
- Plain one-hot encoding vs trainable embeddings
- ... and so on

You are to hand in a notebook with

- output (all cells must be run)
- relevant comments describing your approach, experiments, and findings
- a discussion of which metric (accuracy, recall, precision, f1-score, ...) you have chosen to evaluate your model on and why
- your two language models
- at least two interesting figures or animations

¹ https://www.tensorflow.org/guide/keras/understanding_masking_and_padding