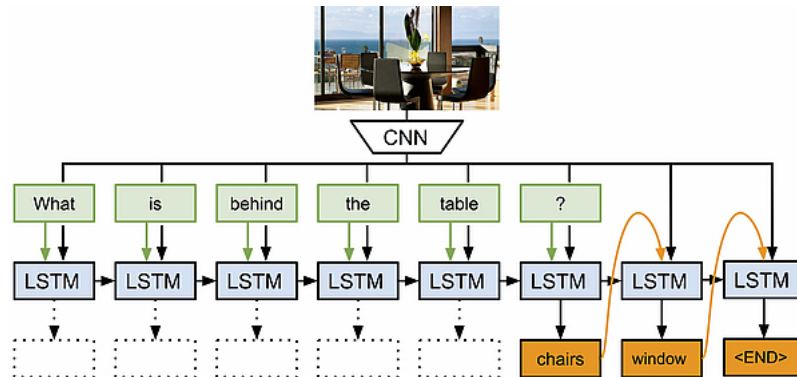# Project Text-Based Information Retrieval: Part 1

Jesse Smits & David Schotmans

March 26, 2019

1 DAQUAR Dataset Modifications

Based on the design of the predictive question-answer model given in the visual-turing-challenge,

we made the decision to convert the training data to fit the sequence prediction, which resulted in the following:

- **Original:**
  Question: what is on the desk in the image3 ?
  Answer: book, scissor, papers, tape, dispenser

- **New:**
  Question: what is on the desk ? *Answer:* book
  Question: what is on the desk ? book *Answer:* scissor
  Question: what is on the desk ? book scissor *Answer:* papers
  Question: what is on the desk ? book scissor papers *Answer:* tape
  Question: what is on the desk ? book scissor papers tape *Answer:* dispenser
  Question: what is on the desk ? book scissor papers tape dispenser *Answer:* <END>

  This way we will continue predicting new possible answers until the *END* symbol is predicted.

2 Encoding & Decoding

To encode our question we used the keras tokenizer from the preprocessing module. After setting up the tokenizer we construct a dense one-hot-encoding of the question. This encoding method results in a rather short 1-dimensional vector (close to length 30). This simple but compact representation has the asset that each encoding corresponds with one unique sentence.
The encoding of the answer is a categorical 1-dimensional vector. This normalized vector with the size of the vocabulary contains the probabilities of the next words. The index of the highest value is decoded with the matching word of the tokenizer.

3 Model training

3.1 Experiments

When training our model, we experimented with the number of LSTM layers, the number of epochs, number of intermediate/hidden nodes and the validation split size. We ended up using 2 LSTM layers with 100 units, a validation split of 20% and 120 epochs.

3.2 Regularization

To increase the validation accuracy of our model, we included a *dropout* layer into the neural network. When using a dropout rate of 0.3, it results in an increase of about 5% in validation accuracy.

4 WUPS evaluation

The neural network is evaluated using the *WUPS* score. By default, the average wups score is calculated over 2000 question-answer test pairs. This results in an average wups of about 0.21.

5 Running the script

Make sure your working directory is set to `part1/`.

Execute the following command to display the help menu: `python3 main.py --help`

You can execute the script with default configuration via: `python3 main.py`

5.1 Dependencies

Run the following command to install all required python3 packages:

```
pip3 install keras tensorflow numpy nltk h5py pickle
```