

The background is a dark blue gradient. In the top-left corner, there are two overlapping geometric shapes: a blue parallelogram and a light green parallelogram. In the top-right corner, there is a grey, 3D-rendered circuit board pattern. In the bottom-left, there is a circular inset showing a detailed image of a printed circuit board (PCB) with various electronic components.

Natural Language Processing for EdTech

Taking the best of ML and using it to
empower education software and
professionals

The problem - assessment



Automation

Example question

Here is an example question

Do you...

- ☐ The first answer, which is wrong
- ☐ The second answer, which is correct
- ☐ The third answer, which is wrong
- ☐ The fourth answer, which is wrong

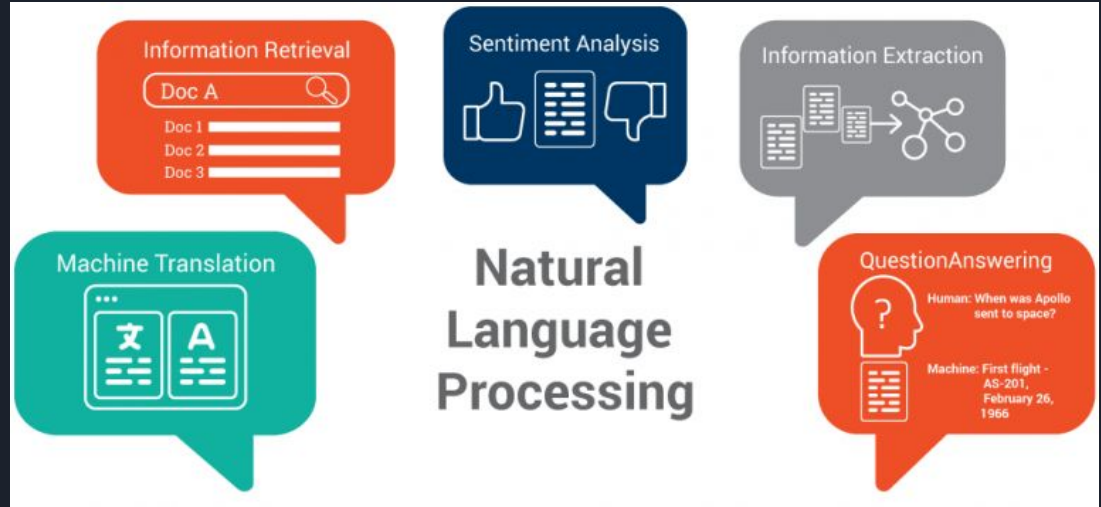
CHECK ANSWER

[< Previous Page](#)

Done!



NLP



Turning text into data and getting insights, e.g. :

- analysing sentiment of a Tweet
- identifying the subject of a text
- translation

The Goal

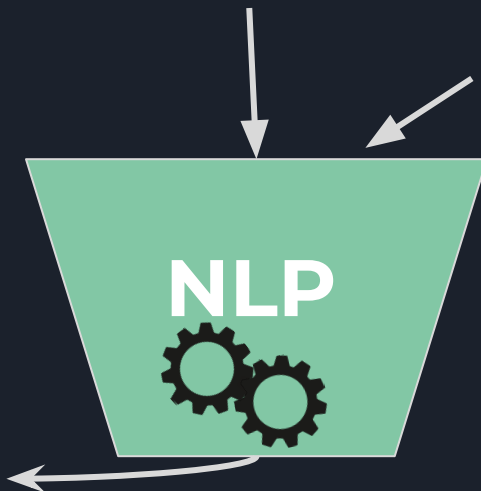
01

Baseline: feed in text, questions and four equally plausible answers come out

The battle of Trafalgar happened in 1805

...

Who was the British
admiral in charge of ... ?

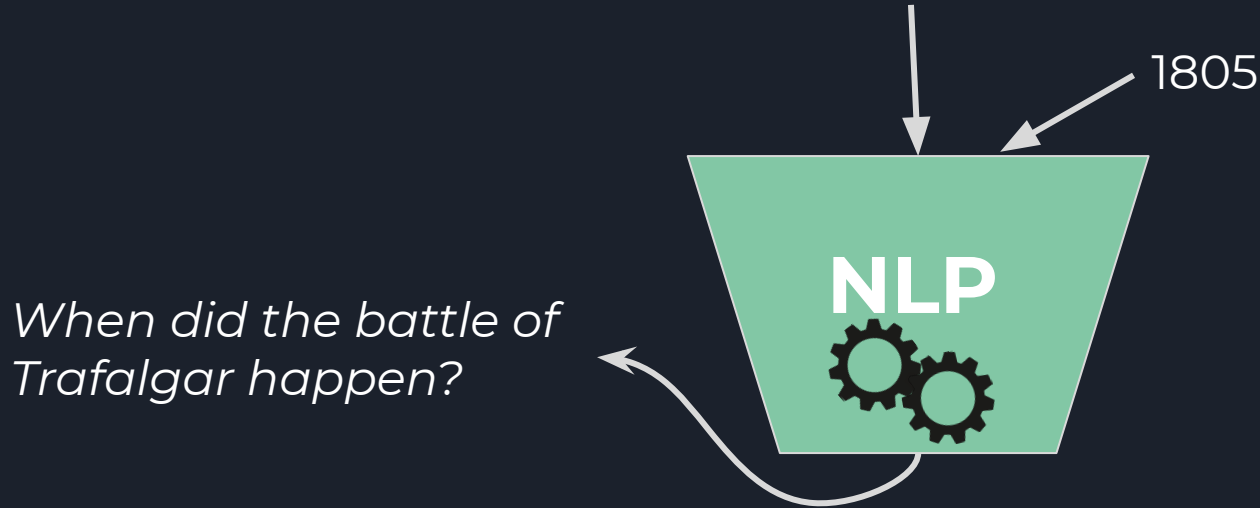


1. Cuthbert Collingwood
2. Horatio Nelson
3. Arthur Wellington
4. King George III

The Goal

02 Even better: takes in text and an **answer**; gives out questions

The battle of Trafalgar happened in 1805 ...



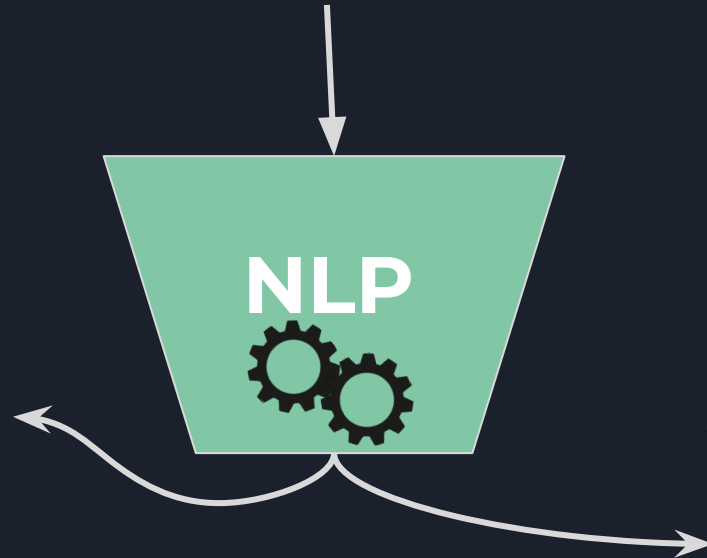
The Goal

03

Just put in **any** text and it spits out questions and four plausible answers.

The battle of Trafalgar happened in 1805

...



1. a. Cadiz
b. Gibraltar
...
2. a. Gravina
b. Villeneuve
...

1. Where was the British fleet coming from?
2. Who led the French force?

What are we working with here ?

SQuAD - Stanford Questions and Answers Dataset

Comprehension questions where the answer is always in the text



Teacher

The Stanford Question Answering Dataset

The role of teacher is often formal and ongoing, carried out at a school or other place of formal education. In many countries, a person who wishes to become a teacher must first obtain specified professional qualifications or credentials from a university or college. These professional qualifications may include the study of pedagogy, the science of teaching. Teachers, like other professionals, may have to continue their education after they qualify, a process known as continuing professional development. Teachers may use a lesson plan to facilitate student learning, providing a course of study which is called the curriculum.

WIKIPEDIA

The Free Encyclopedia

What is a course of study called?

Ground Truth Answers: the curriculum. curriculum curriculum

Prediction: the curriculum

What is another name to describe the science of teaching?

Ground Truth Answers: pedagogy pedagogy formal education

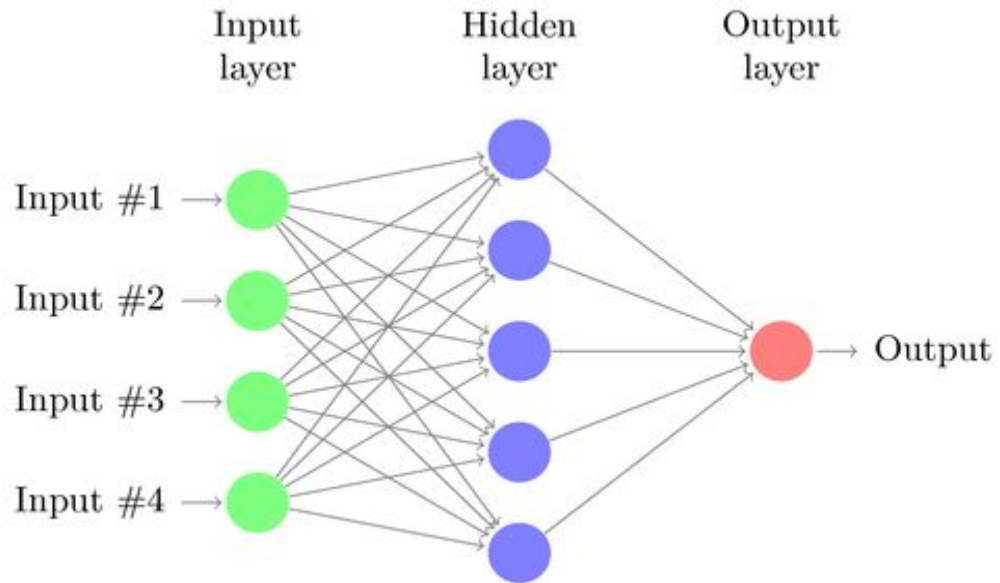
Prediction: pedagogy

Where do most teachers get their credentials from?

Ground Truth Answers: university or college. university university or college

Prediction: university or college

Long term
Short term
Memory
Networks





Development Cycle

1. Architecture

Currently using LSTMs;
2nd stage - Transformers

2. Features to include

GloVe (semantic)
embeddings; POS tags;
stop words



4. Inspect results and detect faults

Overfitting → dropout rate;
Reconsider model limitations,
consult literature

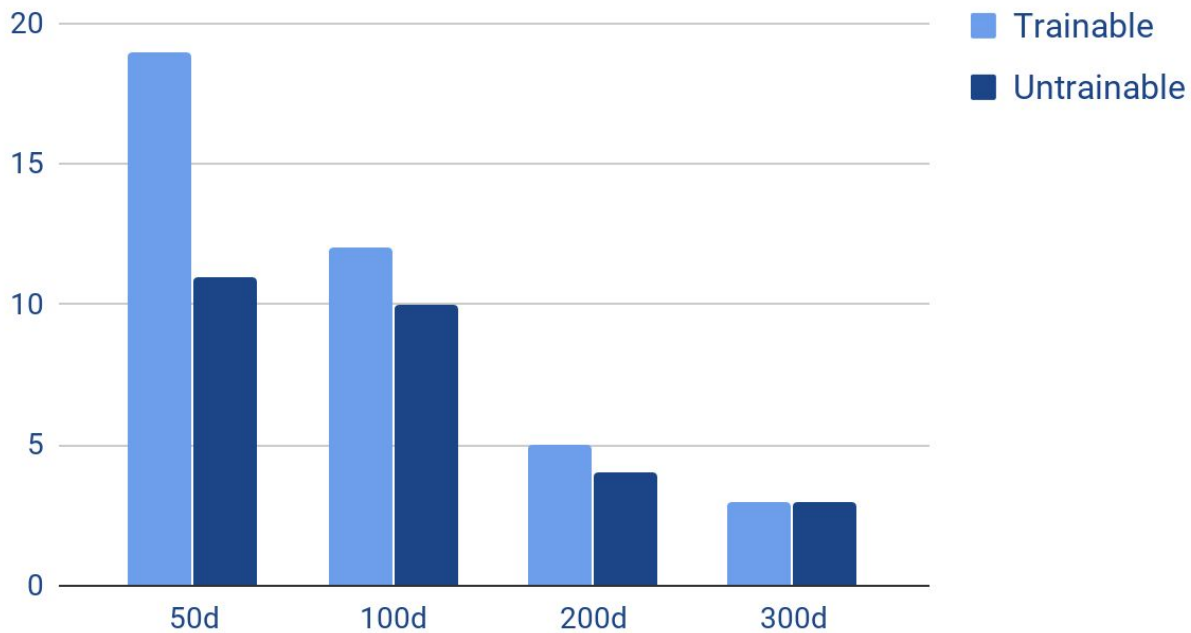
3. Refine hyperparameters

Trainable weights or not;
metrics of success

LSTM so far

Low performance so far. Improved with feature engineering, particularly with pretrained word embeddings

Categorical accuracy % on validation data





Timeline





Thank you!
Any questions?