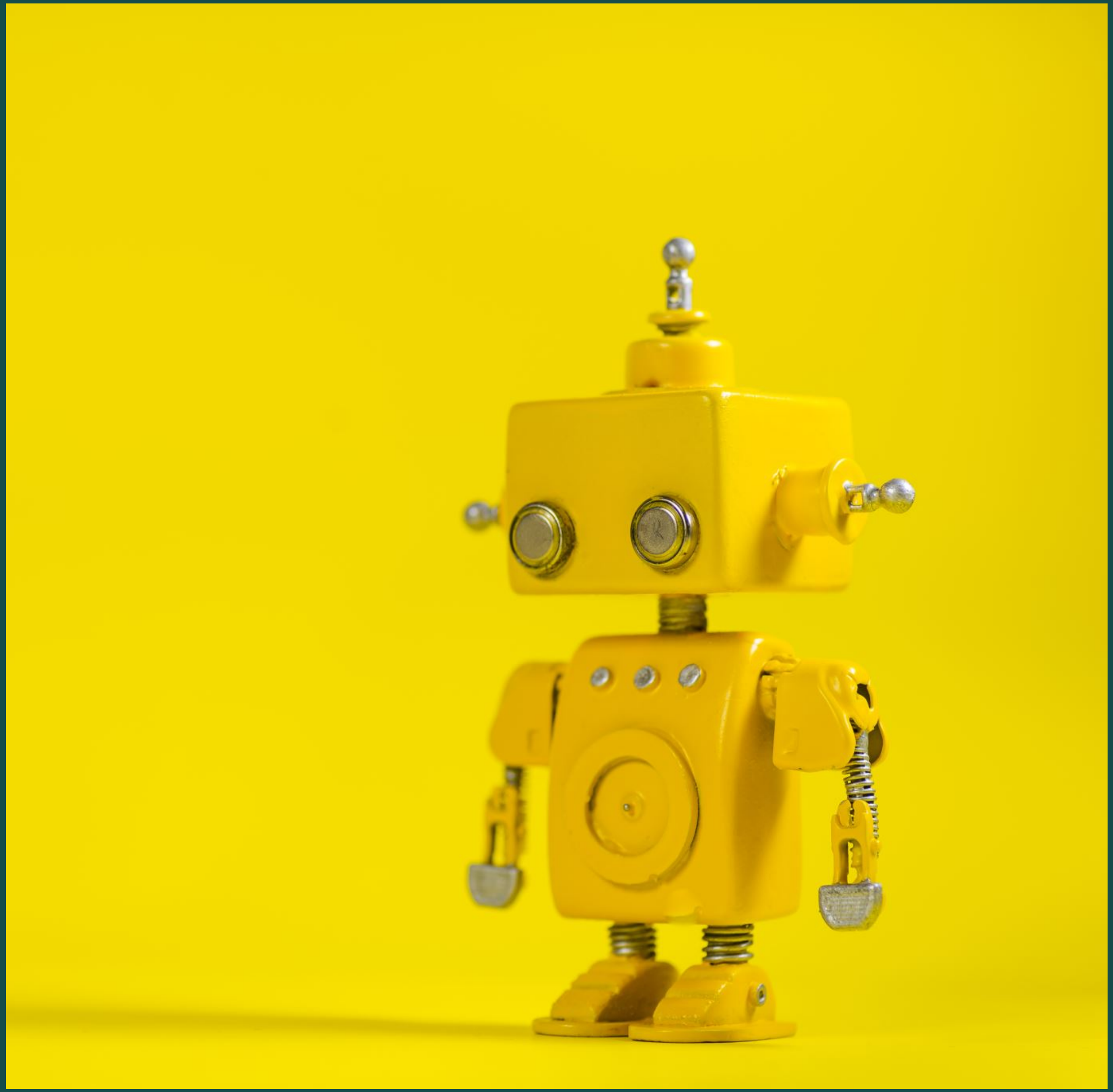
AUTO-CORRECT

Rapid Automated Evaluation of Answer Scripts



Too many papers!

- Correction of papers is a very strenuous and physically tiring task.
- It requires constant concentration and fair hand to mark answers and award a candidate their deserved marks.
- During Examinations, some teachers put in efforts to correct almost a hundred papers everyday.
- This ardent effort sometimes leads to mistakes creeping in, or a non-standardized method of evaluation
- We need to find a way to reduce the pressure on teachers during exam sessions and speed up the process of correction.

Why AutoCorrect?

- The future lies in automation and machine learning
- If answer script evaluation can be automated, it will lead to a standardized as well as a fair method of correction.
- It will also help increase the rate of paper correction and lessen the burden teachers have during the exam season.
- AutoCorrect will certainly go a long way ahead to digitize the entire process of evaluation.

How does AutoCorrect work?

The application has two major sections namely Image Processing and Natural Language Processing.

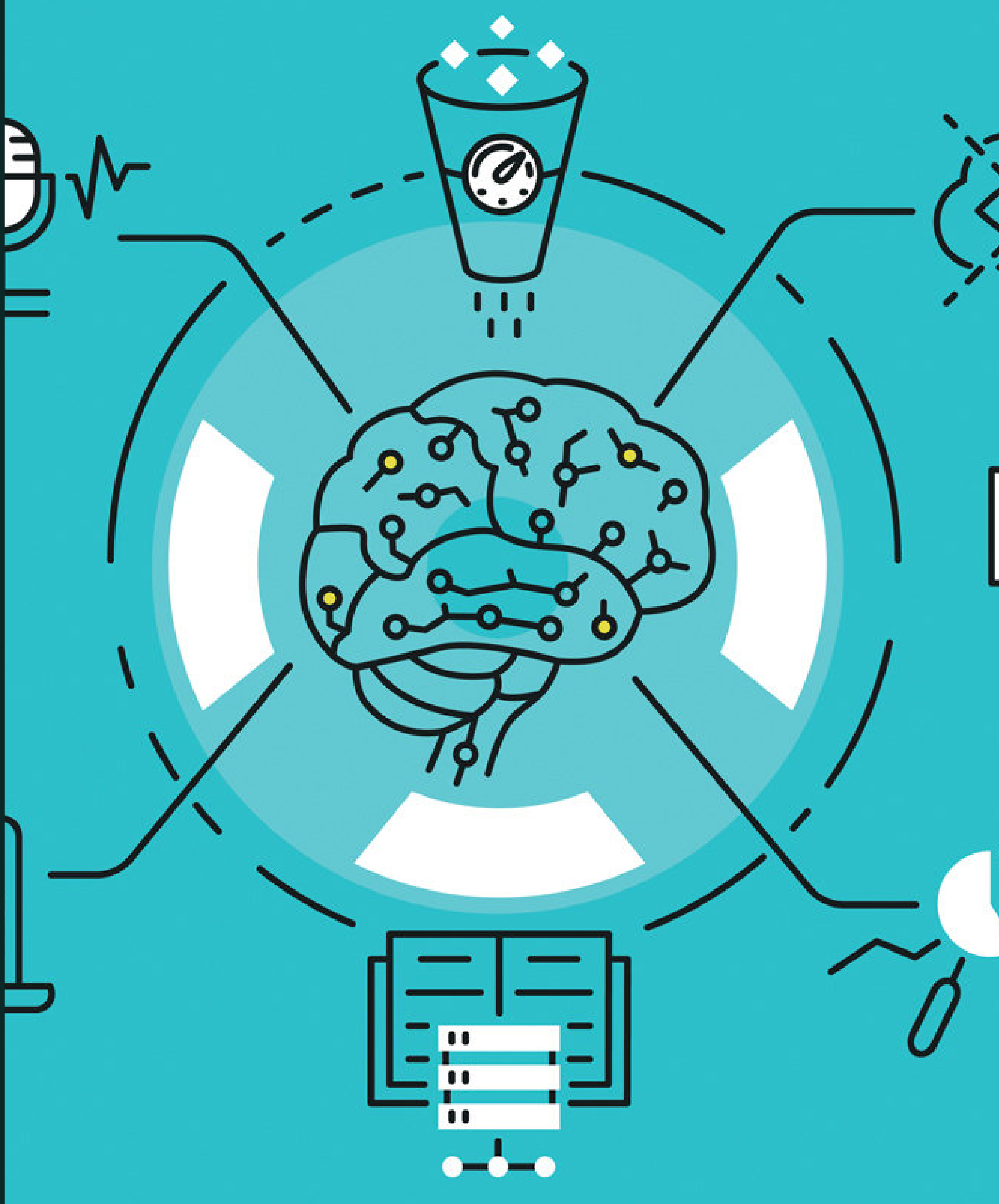


Image Processing

The answer scripts are first scanned and stored as a PDF

These scripts are then processed and converted into text readable formats using OCR.

The OCR is a character level recognition model which identities words and then picks up the characters.

A spell check is then performed to handle misspelt words generated during the OCR.

The answers for different questions are extracted and question-answer pairs are stored separately.

These are then matched with the questions and the key answers present and passed on to the NLP pipeline for evaluation.

Natural Language Processing - I

- The received answer and key answer are first converted into knowledge graphs.
- This is performed to retain concepts as the semantic structure of the answers.
- Each graph is also divided into segments, known as a “sub-graph”.
- These divisions are performed based on concepts – by classifying them into a “main-point” or a “sub-point”.
- A pre-trained embedding model trained on contextual knowledge is also loaded for semantics.

The graphs are then compared at the segment level using cosine similarity.

The main points, sub points and the relation in between are all compared sequentially.

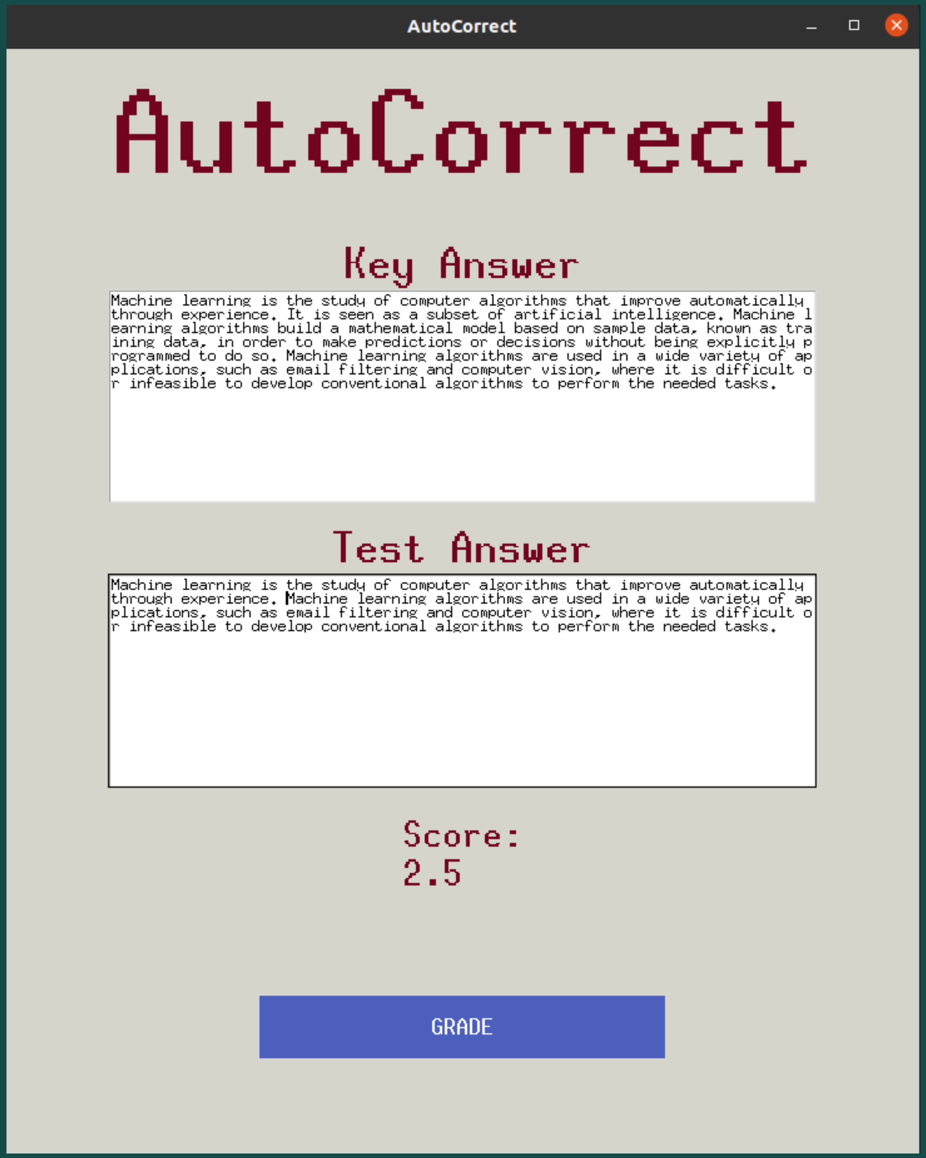
This allows for a check on not only the concept, but also the structure of the concept represented.

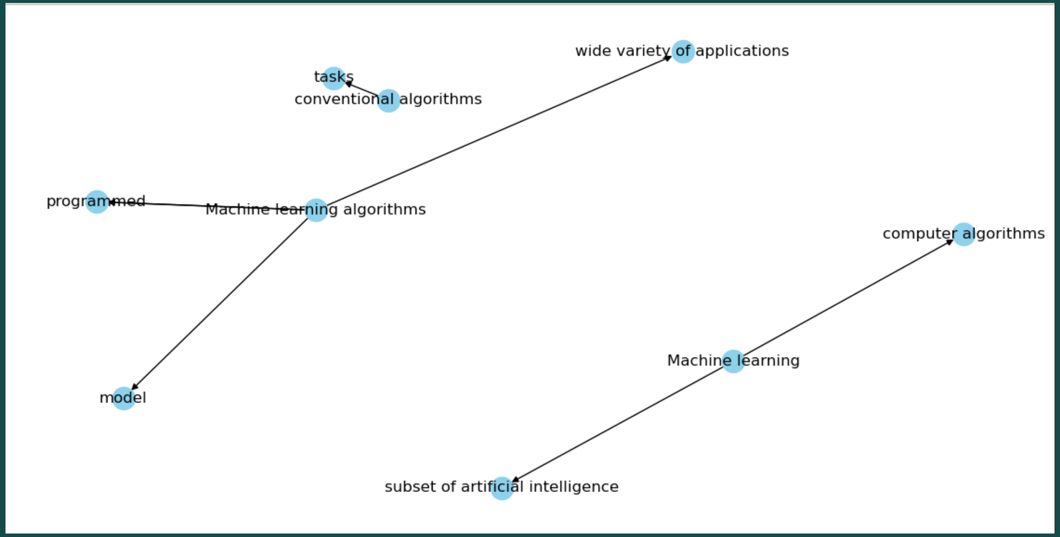
These measures are linearly combined to form the score for each segment

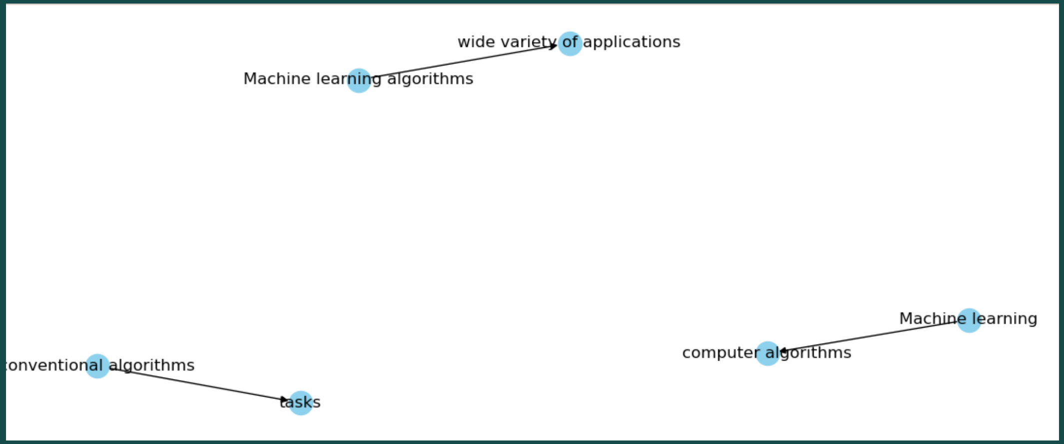
Finally, the scores obtained from each segment is added to form the total allotted marks.

Natural Language Processing - II

Screenshots – Semantic Checking







Screenshots – comparison of Knowledge Graphs

Technology Stack

Python3	Knowledge Graphs	nltk, spaCy & gensim	fastText
Keras & Tensorflow	OpenCV	OCR	Neural Coreferenc e

Thank You