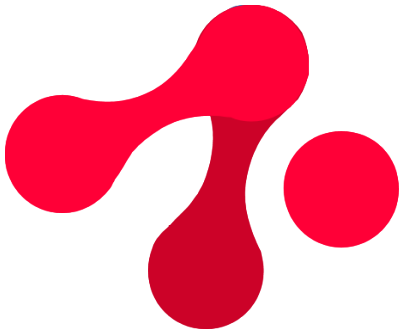

Internship Assignment for NLP

Creation of MCQs with Multiple Correct Answer using generative text models

AKAIKE

11th April 2023

AKAIKE

Objective

The goal of this internship assignment is to test your proficiency in natural language processing. You will be tasked with developing a solution that can automatically generate objective questions with multiple correct answers based on a given chapter from a subject. The generated questions should test the reader's understanding of the chapter and have more than one possible correct answer to increase the complexity and challenge of the questions.The generated questions should not only test the reader's comprehension of the chapter but also encourage them to think beyond the surface level and explore different perspectives and possibilities. Ultimately, the objective of this project is to develop a robust and accurate solution that can aid educators in creating engaging and challenging assessments for their students.

Data

Please download the assignment folder from here. The folder contains PDF files that are chapters from a CBSE text book.

Assumptions

You can assume the following in the image:

- Paragraphs passed will only have one topic in it.

AKAIKE

Example MCAs

Context:
Photosynthesis is a process used by plants and other organisms to convert light energy into chemical energy that, through cellular respiration, can later be released to fuel the organism's activities. Some of this chemical energy is stored in carbohydrate molecules, such as sugars and starches, which are synthesized from carbon dioxide and water – hence the name photosynthesis, from the Greek phōs, "light", and synthesis , "putting together". Most plants, algae, and cyanobacteria perform photosynthesis; such organisms are called photoautotrophs. Photosynthesis is largely responsible for producing and maintaining the oxygen content of the Earth's atmosphere, and supplies most of the energy necessary for life on Earth.

Although photosynthesis is performed differently by different species, the process always begins when energy from light is absorbed by proteins called reaction centers that contain green chlorophyll (and other colored) pigments/chromophores. In plants, these proteins are held inside organelles called chloroplasts, which are most abundant in leaf cells, while in bacteria they are embedded in the plasma membrane. In these light-dependent reactions, some energy is used to strip electrons from suitable substances, such as water, producing oxygen gas. The hydrogen freed by the splitting of water is used in the creation of two further compounds that serve as short-term stores of energy, enabling its transfer to drive other reactions: these compounds are reduced nicotinamide adenine dinucleotide phosphate (NADPH) and adenosine triphosphate (ATP), the "energy currency" of cells.

In plants, algae and cyanobacteria, sugars are synthesized by a subsequent sequence of light-independent reactions called the Calvin cycle. In the Calvin cycle, atmospheric carbon dioxide is incorporated into already existing organic carbon compounds, such as ribulose bisphosphate (RuBP).[5] Using the ATP and NADPH produced by the light-dependent reactions, the resulting compounds are then reduced and removed to form further carbohydrates, such as glucose. In other bacteria, different mechanisms such as the reverse Krebs cycle are used to achieve the same end. The first photosynthetic organisms probably evolved early in the evolutionary history of life and most likely used reducing agents such as hydrogen or hydrogen sulfide, rather than water, as sources of electrons. Cyanobacteria appeared later; the excess oxygen they produced contributed directly to the oxygenation of the Earth, which rendered the evolution of complex life possible. Today, the average rate of energy capture by photosynthesis globally is approximately 130 terawatts, which is about eight times the current power consumption of human civilization. Photosynthetic organisms also convert around 100–115 billion tons (91–104 Pg petagrams, or billion metric tons), of carbon into biomass per year. That plants receive some energy from light – in addition to air, soil, and water – was first discovered in 1779 by Jan Ingenhousz.

MCA Questions:

Submissions

The Submission shall be output image saved in output image path

Q1: Which of the following are examples of light-dependent reactions in photosynthesis?
a. Stripping electrons from suitable substances
b. Reducing and removing compounds
c. Producing glucose
d. Creating NADPH and ATP
Correct Options: (a) & (d)

Q2: Which of the following are sources of electrons used by the first photosynthetic organisms?
a. Oxygen
b. Hydrogen sulfide
c. Water
d. Glucose
Correct Options: (b) & (c)

Q3: Which of the following are byproducts of photosynthesis?
Correct answers:
a. Oxygen
b. Hydrogen
c. NADPH
d. Glucose
Correct Options: (a) & (c)

AKAIKE

Code format

The code submission shall be a .py script that contains the following function. Each function is associated with an objective as explained in the objective section.

```
def get_mca_questions(context: str):
   mca_questions = [“mca1”,”mca2”,”mca3”....]

   return mca_questions
```

The above functions shall have the following behaviour:

1. The functions shall only accept arguments of string data types. If any other data type is passed, it should raise an exception.
2. The functions should return a list of MCA questions as list of strings

Submission guidelines

1. Your code and submission will be subject to automated checks, unit tests and plagiarism checks. Please follow the below guidelines so that it is not rejected.
2. Your code will also be automatically checked for code quality as well. Please follow PEP8 guidelines.
3. Share a writeup discussing the methods and techniques you’ve used in creating your solution.
4. Share the submission as a script (.py) file and include proper documentation that supports your solution.
5. Zip your code and documentation with the following naming convection and upload to the appropriate slot in the google forms **<yourname>_nlp_assignment.zip**

Solution Guidelines

1. Feel free to use any NLP and text processing techniques, libraries, or external data. Please provide due credits to those.
2. Try Large Language Models