JAYPEE INSTITUTE OF INFORMATION AND TECHNOLOGY

Summer Internship

Generation of Similar Questions and Distractors using Python

Under the provision of : Dr. Suma Dawn

BATCH: B9

SUBMITTED BY:

Aryaman Chaturvedi (20103242)

Table of Contents

Prof. (Dr.) Vikas Saxena, SMIEEE
Ph.D. (CSE), M.E. (CE), B.Tech. (CS&IT)
Head-Department of CSE&IT
Jaypee Institute of Information Technology
A-10, Sector-62, Noida, India, 201307
+91-120-2594273, vikas.saxena@jiit.ac.in

E-Certificate Number: CoE-AIE\SIT-2023\40

To whomsoever it may concern

This is to certify that Mr./Ms. **Aryaman Chaturvedi** (Enrollment no. **20103242**) a student of BTech (CSE) from Jaypee Institute of Information Technology has successfully completed the summer training under the supervision of **Dr. Suma Dawn** from **12th June 2023 to 31st July 2023** (7 weeks) working in a project under Center of Excellence on AI for Education.

We wish him all success in the future.

Prof. Vikas Saxena

PROJECT OBJECTIVE

My project aimed to introduce automation into the process of generating questions, achieved through the utilization of Python programming. The focal point of my project was the conversion of lists to matrices, a topic crucial for data manipulation and analysis across fields like data science, business analytics, and research. The project comprised essential steps aimed at achieving a comprehensive and effective outcome. To commence, the task involved curating a diverse range of questions specifically centered around list-to-matrix conversion. These questions spanned fundamental concepts, intermediate applications, and advanced scenarios. During this curation process, meticulous categorization based on difficulty levels was carried out. This categorization aimed to create a well-balanced question set, challenging learners at various stages of their learning journey.

At the core of the project was the development of a robust system capable of autonomously generating questions using Python. This system was constructed upon a foundation of rules, where each rule represented a distinct guideline for question generation. The innovative aspect of these rules lay in their ability to generate a wide array of questions, contributing to the creation of a diverse question bank. Questions generated under each rule were crafted to ensure uniqueness, eliminating redundancy and elevating the quality of the dataset. These generated questions adhered to the multiple-choice question (MCQ) format, chosen for its effectiveness in assessing knowledge levels and providing instant feedback to learners. The questions were designed to mirror real-world scenarios and problem-solving instances relevant to list-to-matrix conversion, rendering the assessment not only informative but also practical.

The culmination of the project resulted in the creation of a comprehensive MCQ dataset, meticulously structured and organized based on varying difficulty levels. Each question within the dataset originated from a predefined rule, guaranteeing the absence of duplicates and ensuring that each question contributed to a holistic comprehension of the topic. In essence, the project's objective was to harness the capabilities of Python for automating the generation of questions pertaining to the intricacies of list-to-matrix conversion.

Project Description

The inception of this endeavor necessitated a fundamental grasp of Python programming. Python's adaptability and widespread integration across diverse sectors emphasized its significance as a foundational skill. This initial phase of learning not only formed the bedrock for the project but also augmented my coding proficiency. With this solid foundation in place, the project's trajectory seamlessly shifted toward its core theme: the transformation of lists into matrices. This realm, integral to data manipulation and analysis, served as the foundation upon which the project's subsequent stages unfolded.

Data Collection and Templated Question Formation: Curating the Dataset

A pivotal element of the project centered around the meticulous compilation of questions pertaining to the "conversion of lists to matrices." Guided by a methodical template, each question underwent careful crafting to encapsulate the core essence of the topic. This template mandated a well-defined structure for every question, encompassing both the query itself and four potential answer choices. However, the significance of this phase exceeded mere question composition. An essential facet entailed the assessment and classification of questions according to their perceived difficulty levels—ranging from easy to medium to hard. This tiered approach ensured that the final question collection provided a harmonious spectrum of complexity, catering effectively to learners with varying levels of proficiency. To reinforce the authenticity of the question repository, the incorporation of links to the original sources held paramount importance, allowing for proper accreditation of external contributions.

Contributions and Unique Question Generation

Generating the Dataset

Beyond the initial task of curating questions, the project's scope expanded to include the generation of unique, original questions. This significant phase acted as a pivotal bridge connecting theoretical understanding with practical application. It served as an avenue to apply the insights acquired during the extensive data collection process. Consequently, this phase added substantial value to the question repository by introducing a diverse range of question formats, problem-solving scenarios, and perspectives.

Implementation: Rule-Based Logic

The coding phase marked the culmination of the project's diverse journey. Guided by the principle of rule-based logic, the goal was to develop a functional system capable of independently generating questions. Notably, the coding language was limited to Python, highlighting the project's dedication to specific skill proficiency. Rule-based logic emerged as the driving force, empowering the system to follow predefined patterns and rules in crafting questions. This approach capitalized on Python's inherent capabilities to create a sophisticated and efficient solution.

In summary, the project embodied the synergy between Python programming and the intricate realm of "conversion of lists to matrices." The journey showcased a blend of foundational learning, thorough data collection, personal contributions, and code implementation. It encapsulated a process where each phase seamlessly transitioned to the next, resulting in a refined automated question generation system. The success of this project was founded on a commitment to technical proficiency and an innovative approach, affirming the transformative potential of using Python for advanced problem-solving.

Task 1: Creating Question bank from Web Resources

During the initial phase of my project, I embarked on a focused exploration within the expansive domain of Python programming, centering on the subject of "conversion of lists to matrices." This journey was initiated by immersively delving into diverse online platforms and resources. The primary aim was to compile a comprehensive assortment of questions that revolved around the intricacies of handling list-to-matrix conversion in Python. This endeavor demanded a substantial investment of time and effort as I meticulously scrutinized the intricacies of Python programming and its interplay with lists and matrices. This dedicated study not only deepened my comprehension of the subject matter but also equipped me to discern pertinent and perceptive questions within the context of list-to-matrix conversion.

The outcome of my meticulous endeavors materialized as a meticulously curated dataset comprising a total of 100 distinct questions. These questions were judiciously curated from a diverse spectrum of web resources, ensuring a wide array of perspectives and contexts. The dataset was thoughtfully structured, encompassing both the questions themselves and multiple-choice answer options. These answer options were intelligently categorized based on

varying levels of difficulty, providing learners with a systematically progressive journey from foundational concepts to more advanced challenges.

Task 2: Enriching the Question Repository

During Task 2, my focus was on enriching the question repository through the acquisition of an additional set of 100 questions with a central emphasis on the theme of "conversion of lists to matrices." This endeavor spanned a dedicated period of three weeks, during which I diligently gathered questions from diverse sources to ensure a comprehensive array of inquiries. These questions were thoughtfully curated to span a spectrum of difficulty levels, effectively catering to learners at various stages of expertise. Following meticulous selection, these questions seamlessly melded into the existing collection, resulting in a substantial expansion of the repository to a total of 200 questions. This augmentation not only bolstered the repository quantitatively but also enriched its qualitative depth. As a result, learners can now access a more comprehensive and versatile resource that facilitates a profound understanding of the intricacies involved in converting lists to matrices.

Task 3: Rule Formulation

In this task I reached a crucial point where I learned about rule-based logic and how it helps create questions. With the help of mentors and exposure to similar projects, I started learning about the importance of using rules to make questions organized and logical. This made me change how I usually think about programming and focus more on using rules. I spent a good amount of effort understanding this approach and came up with a set of ready-to-use rules.

This experience showed that I'm good at learning new things and adapting. It was like a turning point where I began to see how rules can be used to make questions in a clear way. This shift in thinking helped me realize how rule-based logic can play a big role in making the process of creating questions work really well.

Task 4: Implementation

In this task I moved from theory to practical by starting a coding that focused on using rule-based logic. This phase lasted for two weeks, during which I spent time and effort creating a Python script that would make the predefined rules work. This process involved using Python's built-in abilities to turn the rule-based logic I had thought of into actual working code.

Skills/Technology Developed

- Profound understanding of Python lists and their associated concepts
- Competency in collecting and analyzing datasets
- Proficiency in Python programming for automation and generating questions
- Utilization of Microsoft Excel for managing and modifying datasets
- Enhanced logical skills for crafting multiple-choice questions that assess list-to-matrix conversion

CONCLUSION

In the project, I demonstrated dedication, technical expertise, and collaboration. Each task contributed to an innovative question generation system. I curated questions, applied rule-based logic, and adapted based on feedback. Collaboration with mentors enriched the work. Success relied on adaptability, proactive learning, and a collaborative mindset, showcasing the synergy of skills and contributions.