Education and Information Technologies
https://doi.org/10.1007/s10639-023-12045-4

Improvised progressive model based on automatic calibration of difficulty level: A practical solution of competitive-based examination

Aditya Shah[1] · Ajay Devmane[1] · Mehul Ranka[1] · Prathamesh Churi[1]

Received: 14 February 2023 / Accepted: 10 July 2023

Abstract
Online learning has grown due to the advancement of technology and flexibility. Online examinations measure students' knowledge and skills. Traditional question papers include inconsistent difficulty levels, arbitrary question allocations, and poor grading. The suggested model calibrates question paper difficulty based on student performance to improve understanding assessment. The suggested student assessment system paradigm involves determining difficulty, creating the exam, and assessing the student. Based on the previously established relationship between question difficulty and right responses, questions are computed and then divided into difficulty categories. This model improves testing by adapting to the student's ability in real-time. This method ensures that all students are graded uniformly and fairly using pre-determined questions and criteria. The methodology can also cut exam creation and administration time, freeing up teachers and administrators to focus on other assessment tasks. It considers more evidence, learner-centered assessment can help employers evaluate candidates more accurately and meaningfully. It might boost academic productivity by letting assessors quickly write high-quality papers and save up time for deeper investigation and experimentation. This may accelerate scientific progress. Automatic paper generation raises ethical questions about research validity and reliability.

Keywords Progressive Model · Competitive-Based Examination · Difficulty Level · Difficulty based test · Online assessment · Question Paper Generation

Abbreviations

AQBDMS	Adaptive Question Bank Development and Management System
AQG	Automatic Question Generator
AQP	Automatic Question Production
BN	Bayesian Network
CSP	Constraint Satisfaction Problem

Extended author information available on the last page of the article

Published online: 09 August 2023

DR	Derived Relationships
GMAT	Graduate Management Admission Test
GRE	Graduate Record Examination
IRT	Item Response Theory
KNN	K-Nearest Neighbour
LMKT	Language Models for Deep Knowledge Tracing
QPD	Question Paper Designer
QR	Quantitative Relationships
QUESTOURnament	A tool to find the real difficulty of Question
RCM	Regularised Competition Model
SR	Semantic Relationships

1 Introduction

Online learning has become increasingly popular in recent years, with the advancement of technology and the need for flexibility in education. Online courses offer students the opportunity to study from anywhere at any time, making education accessible to a wider range of individuals (Chakraborty et al., 2021). Online assessments, a critical component of online learning, are used to evaluate students' knowledge and skills. However, online assessments come with their own set of challenges, particularly when it comes to ensuring fair evaluation (Elzainy et al., 2020; Holden et al., 2021). Moreover, the lack of standardization in online assessment systems can lead to inconsistencies in the questions and difficulty levels (Martin et al., 2019). This can create an unequal playing field for students, as some may have access to easier or harder questions based on the specific platform or assessment they are using.

Another point of view is that a question paper represents the student's comprehension of the subject, and its accuracy is critical for an appropriate assessment of the student. Competitive exams such as the GRE, GMAT, or equivalent exams necessitate a complete evaluation system and assess a student's knowledge based on how well the student does on a specific question (Al-Maqbali & Hussain, 2022; Guangul et al., 2020).

Numerous drawbacks of the conventional approach of question paper generating include variable degrees of difficulty, arbitrary question selection, and inaccurate marking. These restrictions may make it more difficult to accurately evaluate the student, which may result in erroneous judgements about their potential for further study or work. The suggested model offers a novel method for creating question papers that adjusts the level of difficulty based on student performance in order to address these problems. The model offers a more constant level of difficulty across the segments, lowering the possibility of inaccurate marking and enabling a more precise evaluation of the student's comprehension.

We created a novel model for the creation of question papers with calibrated difficulty levels after taking into consideration the necessity of the problem in the evaluation of an assessment system. The model described in (Singh et al., 2021) is the basis for our model. The model was created with competitive exams in mind, taking into account both the student's level of comprehension and test performance. The following objectives are the focus of our research:

- To introduce a new and improvised model for question paper generation in competitive exams. The model is a revised version of model published in Springer education and Information Technologies Journal (Singh et al., 2021).
- A designed model calibrates the difficulty level of questions based on the student's performance and ability to answer a particular question.
- We present a system that can provide a consistent level of difficulty across segments.
- To reduce the chances of manual incorrect marking by evaluators and provide a more accurate assessment of the student.

Teachers, assessment specialists, academics, and technology experts who are interested in the creation and implementation of online learning and assessment systems make up the target audience for this research study. The paper is also aimed at students and other people who are searching for a just and efficient grading system. The important takeaways from this paper include a thorough understanding of the present issues with the online assessment system and how our suggested model resolves these issues by adjusting the level of difficulty of questions in accordance with the proficiency of the student. By guaranteeing that the questions are suited to the student's abilities rather than offering a standard collection of questions with the same level of difficulty for all students, our methodology offers a solution for fair evaluation.

The proposed model offers several explanations. First, it will accurately test students' understanding. Second, the measured difficulty of the question paper will provide a positive experience for the student, helping them to feel more secure and confident when completing the test. As such, the model can effectively and accurately assess a large student population, providing a valuable tool for competitive exam designers. Additionally, the model will enable questionnaires to be tailored to individual students, allowing for a more personalized experience.

The practical implications of this research are significant as it can be adopted by various educational institutions, organizations, and examination boards to conduct online assessments, especially for competitive exams like GRE and GMAT. With the increasing popularity of online learning, it is crucial to have a fair and effective evaluation system in place. Our model ensures that the evaluation is not just based on luck but on the student's actual understanding and abilities. This will result in a more accurate evaluation of the student's abilities and lead to a better educational experience.

This paper's Introduction is followed by a discussion of the Search Criteria, which provides a brief overview of the sources from which we have drawn our works. Our third section is devoted to Related Work, in which we investigated several publications to determine where they fall short and how they should be implemented. The fourth section of the paper explains why we developed a scenario for the reader to understand what sort of thinking existed and why we should build this model. The fifth section of this work is devoted to the implementation of our Question paper generating model, which we prefer to refer to as the Progressive model. The final two sections are the Results and Discussion, in in which we describe the model and analyse probable repercussions.

2 Search criteria

For the literature review we adopted criteria of Systematic literature review (SLR). We exclusively searched on the Google scholar platform based on keywords mentioned in Fig. 1. A total of 289 papers were searched out of which we have taken papers published in Scopus and/or web of science and/or ERIC indexed journals. A few papers which were not indexed but considered to be published in reputed publications were also brought into the literature survey. The papers published in public repositories such ArXiV, SSRN were also taken into consideration. A total of 98 papers were filtered which were majorly based on publishers like IEEE, Elsevier, Springer, MDPI, ACM Transactions etc. Papers published in conferences with either higher citations and/or reputed flagship conference were also taken into consideration. Finally, a total number of 40 papers were taken into consideration. Figure 1 and Table 1 depict the search methodology of our extensive research work.

3 Related works

There is no research where a user's genuine capabilities are known, faculty or the invigilator can choose the user's difficulty level, difficulty-based questions are generated, and the questions are generated randomly throughout the test. To overcome these challenges, the study describes a progressive approach to paper creation, in which the model gauges the complexity of the questions based on the user's multiple-choice response selections.

Search performance is inversely related to question difficulty for a given topic, according to research by the authors of (Singh et al., 2021), and can be used as a measure of question difficulty when computed using standard points of reference. This was done by considering precision, recall, and a Meadows composite effectiveness score generated from 50 questions. Each measure was classified using non-parametric discriminant analysis based on its two closest neighbours, and then each measure was encoded into a spatial representation of inter-question similarity. The architects for incorporating a computerised adaptive testing module into an adaptive learning system, where the test items are created using a domain knowledge tree, are presented by the authors of (Chen & Zhang, 2008). The content, suitable response, level of difficulty, discrimination, and guessing parameter for each test item in the item bank are all listed. Test items are chosen based on a learner's aptitude and graded dynamically based on their response in computerised adaptive testing. In this iterative method, the learner's skill level is initially measured using an item response model.

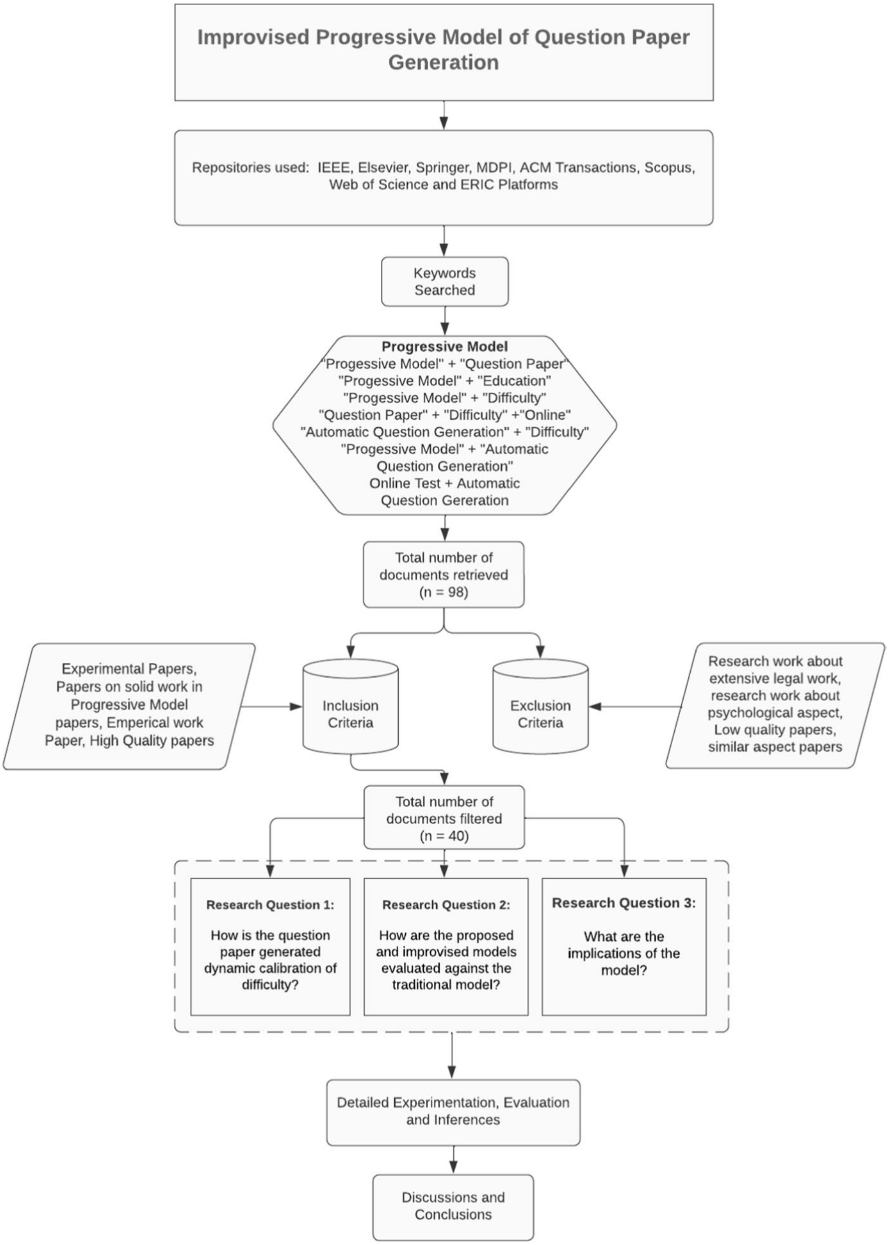

Fig. 1 Search Criteria

A method to select test questions that adapt to individual needs has been presented by (Vozár & Bieliková, 2008) combining three approaches—the first selecting the most appropriate learning topic based on course structure, the second using item response theory to select the k-best questions with the right level of difficulty,

Table 1 Research questions

Research Question	Question Statements	Response
RQ1	How is the question paper generated dynamic calibration of difficulty?	This is discussed in Sect. 5 of the Proposed Architecture
RQ2	How are the proposed and improvised models evaluated against the traditional model?	This is presented in Sect. 6 which discusses the Results and Outcomes of the model
RQ3	What are the implications of the model?	This is discussed in Sect. 7, which focuses on the Model's Discussions and Conclusions

and the third based on prioritising questions based on usage history. The necessary conditions are shown in the prerequisite graph. The relationship between a student's accuracy and knowledge is revealed by the IRT model. By avoiding repeatedly asking the same question on a certain topic, history-based question selection also aims to raise the quality of question selection. As opposed to this paper (Sun et al., 2009), they created a relational data model that is comparable to data structures to explain the constraints of the test paper. Other data structures exist that preserve corresponding constraint factors in addition to those already listed, such as Chapters, Difficulty, Content Level, and Knowledge Point. There are five additional qualities displayed for each question, in addition to the content and response: the question type, difficulty level, knowledge point, reference score value, and content level. The phrases total score, difficulty level, duplication level, and grade for core and keystone questions, question kinds, and chapters must be specified by the evaluator.

An algorithm to generate questions that are tailored to students' knowledge and misconceptions is presented in Paper (Khodeir et al., 2014). By considering students' knowledge and controlling the difficulty level of the generated questions by adapting generated questions, students' actual student model is estimated. Bayesian Network (BN) knowledge representation for probabilistic domains is used in this. The student knowledge estimator, evidence selector, question assessment module, question difficulty level estimator, and question selector are the main components of the question generation process. By applying established parameters for measuring the difficulty levels, the authors of (Singhal et al., 2016) have proposed a difficulty model for generating questions across formal domains in accordance with the difficulty level provided by the user. The elements that determine a question's difficulty can be divided into three groups: those that affect semantic relationships (SR), quantitative connections (QR), and relationships generated from meta information (DR).

By using students freely constructed one- to three-sentence answers to determine the empirically observed difficulty of short-answer questions, the authors in (Pado, 2017) empirically evaluated the predictive power of both Bloom dimensions. They then used logistic regression models to predict the best feature set using the bloom heuristic and the student average variation heuristic. depending on the Cognitive Process and Knowledge component, the query difficulty. To maximise performance throughout the learning process, (Comas-Lopez et al., 2018) presents a sequential adaptive exam paradigm in which students are given questions divided into groups of 4 based on the difficulty. Students must properly answer a predetermined number of questions in order to advance; else, they will be placed back on the previous level. Each set of questions contains parameters that can be used to forecast and correct ineffective teaching and learning strategies, such as the required number of attempts. This method has allowed us to demonstrate that 75% of students' grades increase when using our technique, which is described in this article, as opposed to standard adaptive testing.

(Sullivan, 2001) offered a novel and ground-breaking mobile-based progressive approach for determining the difficulty level of multiple-choice questions based on test-takers' inputs and adjusting it to the student's skills. This method is based on the idea that if a student answers more and more tough questions in a set, they will be given more difficult questions in consecutive sets. The number of incorrect responses and the level of difficulty are significantly associated, as evidenced by the fact that a question's difficulty is determined before it is used to generate the question paper. (Perez et al., 2012) explores if evaluators are better at determining the complexity of a problem than pupils believe. To assist them, they developed an automatic classification expert system that evaluates the level of complexity of the questions using the challenging online learning platform QUESTOURnament. Several evolutionary algorithms and fuzzy logic methodologies were used to create the expert system. This method investigates the level of difficulty as perceived by the pupils.

(Purohit et al., 2012) propose an Adaptive Question Bank Development and Management System (AQBDMS), which intelligently selects questions from a large database (question bank) and models the question using inputs or parameters supplied by the question paper designer (QPD). This question bank employs a concept map created using a visual technique and hierarchical knowledge of a certain topic to ensure that questions are modelled as closely as possible in accordance with components such as Bloom's Taxonomy, difficulty level, grading system, and so on. The evaluation of the developed question model will provide feedback on the students' overall understanding of the topic. An innovative approach to measuring question difficulty utilised by (Wang et al., 2014) is the Regularised Competition Model (RCM), which integrates question user comparisons and question textual descriptions into a cohesive framework. They solve data scarcity and cold starts by leveraging textual data. They also forecast the difficulty levels of the new questions using an unsupervised machine learning technique known as the K-Nearest Neighbour (KNN) approach based on textual similarities.

(Mishra & Jain, 2016) classified the QAs based on a variety of criteria, including the application domain (Restricted or Open-domain QAs), the type of analysis performed on the questions and source documents (Morphological, Semantic), the different types of questions asked (Hypothetical, Factoid, or Casual questions), the data attributes (language, size), the data surveyed in a data frame (structured or unstructured), and the types of description QAS. (Sunil, 2020) demonstrated a Smart Paper Generator, a piece of software that allows instructors to write test questions fast and automatically. A randomization method will produce the question paper from the question bank that the evaluator must upload for each topic.

Controlled sequence generation was studied by (Srivastava & Goodman, 2021). They improvised the Language Models for Deep Knowledge Tracing (LMKT). This model forecasts a learner's chance of properly answering a question and generalises to barriers not encountered during training. After defining the aim and data with LMKT, we train a model to generate questions depending on student and target difficulty. The results show that they created new, precisely calibrated language translation questions for second language learners from a recognised online education platform. The project proposed by (Dwivedi et al., 2019) develops and deploys an automated system to find and create engineering exam questions. A productive dynamic process creates the question paper. The difficulty index of inputted question items allows the question paper to be adjusted to the testing need, i.e., basic to advanced. The assessment module creates reaction cues for subjective questions and password-protected, expert-validated answer keys for objective questions. The strategy produces hard, original, and fast question sets.

(Patil et al., 2022) Created a system that creates test questions for various assessments and exams based on the input of a complicated Marathi sentence. The two main components of the workload in our AQG system are pre-processing and AQG. The Pre-processing module is divided into three sections: I stop word elimination, tokenization, and POS tagging. The Automatic Question Production (AQG) Engine performs three subtasks: (1) question word selection, (2) clause selection, and (3) question production. (Shah et al., 2022) says that they are replacing the traditional method of question paper generation with an autonomous question paper generator technology, which could save time. Automatic test paper creation is a method for solving the constraint satisfaction problem (CSP) that involves selecting questions from a question bank and automatically producing various types of papers that match instructional requirements. The system also employs security controls to prevent question recurrence, ensuring the test's security and objectivity (Ragasudha & Saravanan, 2022). "Bloom's taxonomy must be used to create a database of questions for automatic question paper generation." In order to safely communicate the question paper to the authorised person, it was encrypted and decrypted using cryptography. The keyword matching method of evaluation is utilised to evaluate the response automatically.

Although each of the aforementioned ways is effective and efficient in its own right, each has weaknesses that can be addressed. The writers of paper (Singh et al., 2021) have merely proposed a hypothetical progressive paradigm for question paper creation; no actual implementation of this concept has occurred. Paper (Vozár & Bieliková, 2008) evaluation, which only had a small number of participants, should have included a larger group of people. For the technique presented in paper (Sun et al., 2009) it would be ideal to have a large database of questions because the number of questions may impact how effective the algorithm is. It might not be as efficient with a smaller database of questions. Paper (Khodeir et al., 2014) generates questions based on simulations of students' knowledge, and the model was evaluated on fake students rather than actual students because the replies of real students can vary much more than those of the 50 fictitious students. Before a question is produced in (Singhal et al., 2016) the user must provide his or her difficulty preferences. This is a time-consuming trial-and-error method in which the user must modify the difficulty options until the desired question is generated. Because the difficulty can only be assessed after grading the students, the (Pado, 2017) approach is difficult to apply in many Short Answer Grading circumstances. Although Paper (Comas-Lopez et al., 2018) is a very useful model, analysis shows that using the programme continuously does not result in proper knowledge assimilation and is not always the optimum technique during autonomous learning. The model in (Sullivan, 2001) creates a graphical representation in the form of a question map, but it makes no attempt to analyse the map. Because it is based on the results of the questions' searches, it may produce false results owing to incorrect search results.

According to (Perez et al., 2012), confusing issue statements make students nervous, so evaluators should rephrase that question. It could also be attributed to a major misunderstanding that was previously missed. (Purohit et al., 2012) suggests that, while this application is hypothetical and yet under development, its success is expected to considerably aid the organisation in successful question modelling and evaluation. (Wang et al., 2014) Graph storage and processing RCM's bottleneck is Laplacian. RCM has difficulty functioning in non-technical contexts. (Mishra & Jain, 2016) QAs were unable to understand any natural language question or its representation; they were also unable to comprehend knowledge derived from documents (structured, semi-structured, unstructured, and semantic web) and search for relevant, accurate, and succinct solutions that could meet users' informational needs. Future model improvement will necessitate more comprehensive student data to improve the efficacy of our LMKT model, as (Srivastava & Goodman, 2021) identifies the availability of training data and the computational restrictions of large LMs as drawbacks of their technique. (Dwivedi et al., 2019) conclude that, because subjective question items lack established answer patterns, answer signals with their analytical breakdown must be inserted into the system for it to recognise specific patterns in writing material. Due to a lack of coreference resolution, the suggested technique in (Patil et al., 2022) generates several erroneous questions, and several of these questions have been semantically misconstrued during manual review. (Shah et al., 2022) had specific questions that may be repeated in many question papers due to partiality in a system caused by the professional's personal contact with them. As a result, there is no guarantee that the question paper was generated at random. Other potential issues with this strategy include a lack of storage capacity, ineffective document transmission, and editing challenges. (Ragasudha & Saravanan, 2022) discusses how long processing took because the answer sheet had to be scanned, uploaded, and then analysed based on the keyword. This is the main issue with the current system. In general, related research using machine learning classification methods has been done to estimate the complexity of problems.

4 Scenario and need for progressive model

Up till today, we still use the typical traditional form of assessment in which the evaluator makes a Question Paper for students and assesses them on the pre-defined set of questions which is analogous to the famous quote by Albert Einstein "Everybody is a genius. But if you judge a fish by its ability to climb a tree, it will live its whole life believing that it is stupid". This points out to an observation that each student has different calibre and learning capabilities and judging them on a single set of Question Paper is not fair. This is due to the fact that all of the questions are the same, as is the structure in which they have been presented. These papers rely on the memory power of an individual and not the understanding of the student. Hence, a new and dynamic approach is proposed which helps to evaluate the students according to the knowledge they possess which is depicted in the use case below –

Once upon a time, there were three best friends named Rohit, Karan, and Akash who were classmates in the 10th grade. Rohit was a topper, Karan was an average student, and Akash was a backbencher. They all attended the Batra Institute in Kota, which is renowned for producing the most IITians. They used to sit in the same row, and their seats were adjacent to each other. It was the day of their JEE MCQ exam, and they all gathered in the examination hall with their respective belongings and now they have started to access the students through Progressive approach of Questions which is new for all of them.

Rohit, who was the topper of the class, was very confident and cool. He had already revised the entire syllabus many times, and he was sure that he would score an A+. Karan, who was an average student; he had revised the entire syllabus, but he was not confident about his answers. Akash, who was a backbencher, was the most nervous of all and was sure that he would fail the exams. The examination started, and the students were busy writing their answers.

Rohit felt confident and calm as he began attempting the paper; he completed the common portion for everyone with flying colours and moved on to the next section with the difficulty level equal to the average of the previous section, which is high because he satisfactorily completed it. From now on, he receives questions based on his capacity to solve them; if he correctly answers the question, the difficulty increases, and his calibre and knowledge are assessed. If he is unable to solve a particular question, the difficulty level of the succeeding question remains the same, giving him another opportunity to progress at a higher difficulty level.

Karan seemed a little worried and unsure about his exam preparation. As a result, he made mistakes in answering because he was unsure of the proper option, therefore he solved the common portion like a regular ordinary student who knows the answer but hesitates due to lack of confidence. So, after the common section, he received an average level difficulty question, and he is also assessed for his knowledge, and the difficulty level varies as he answers some questions correctly, so the difficulty level increases, giving him a chance to score higher, and if he consequently incorrectly solves it two times, so the difficulty level decreases. He completed his exam in this manner.

Akash was the most worried since he expected to fail because he had not studied, yet he was an excellent listener and hence understood and remembered the concept through the notion in his subconscious memory. He began with the common section, which had questions of all levels, and because he did not pay attention in class, he missed many of them and scored lower in the common section, resulting in the following section's difficulty being below average. But these were the questions he knew and correctly answered, and the difficulty level began to rise, and he also used his subconscious memory to tackle out many questions and at least reach a decent average difficulty level questions up to which he can solve and above which requires a little practise. So, he was quite certain that he would not receive good marks, but he would not fail the exam.

After the exams, all the students went home, and they all eagerly waited for their results. The day of the results arrived, and the students gathered in the school assembly to check their results. Rohit was the first one to see his result. He had scored an A+. He was thrilled, and his parents were proud of him. Karan was next, and he had scored a B+. He was happy that he had passed, and his parents were proud of him. Akash was the last one to see his result, and he was nervous. To his surprise, he had scored a B+. He was overjoyed, and his parents were proud of him. The three friends were happy that they had passed their exams, and they celebrated their success together.

From the above scenario we can conclude that there is a need of a new and advanced evaluation & assessing technique such as a Progressive Question Papers so as to assess each candidate on the basis of their learning and not the memory power possessed by them as for example Akash would have studied the last night of exam and memorized all the concept then come to exam hall and had solved the paper and then forget about it but as seen he understood the concept an knows how to go about solving a question which helped him fetch good marks in exam; which ensures fair evaluation of each and every student.

5 Proposed architecture of progressive model

5.1 Data collected

The information was gathered utilising a variety of tools, including Sanfoundry.com, 11th Standard Textbooks, two to three other websites, and some questions that the papers' writers themselves had created. 200 Questions in all were added to the database, which was kept in a CSV file, and the questions were fetched from it when the code was performed. Each difficulty level had exactly 40 questions, which explains how there could be 5 difficulties (1 being the lowest & 5 being the highest) and 200 questions overall.

5.2 Proposed algorithm

In the Proposed Algorithm, adaptive testing has been implemented by first receiving two inputs from the user, X and Q, where X represents the evaluation's difficulty level and Q is the number of questions the evaluator is prepared to ask. The model will properly distribute the questions into a total of N sections based on the evaluator's preference, with the first Y sections serving to assess the students' expectations during the Preliminary Stage and Section (Y + 1) being generated by calculating the average difficulty level of the answers scored in the Y sections. The result will be computed up to the Y sections, and the level of difficulty will be determined according to Table 1. Then, following the $(Y+1)^{th}$ part, the model generates a difficulty level based on the same question number solved correctly or incorrectly; the difficulty would increase or decrease respectively based on the answer to question number (M), which is the number of questions in the preceding section. In this manner, a paper is generated automatically using the Progressive model, and a fair analysis is performed for the students (Figs. 2, 3 and 4).

5.3 Architectural diagram (Fig. 5)

5.3.1 Described model

The Model consists of a total of 10 parts, each of which contains 5 questions, for a total of 50 questions in the Test. All different types of questions with varying degrees of difficulty are included in the first three sections, from 1 to 3. After

```
Adaptive Testing
Input: X, Q
Data: Global C, Global LQ
            // Q = Number of Questions
            // C = List of Correctly Solved
            // LQ = Level Presently Question on
            // MS = Marks Scored
            // Y = Pre-Progressive Section Number
            // DF = Difficulty of Y th Section
            // M = Total number of questions in each Section (Here M = 5)
if Q%M ≠ 0 then
    Q ← Next multiple of M;
```

Fig. 2 Proposed Algorithm

Sect. 3 is finished, it will determine the average difficulty level for Sect. 4 by considering the correct answers the user provided in Sects. 1 through 3. Additionally, that challenge would be outlined in Sect. 4, after which the Progressive model would begin to take effect.

5.3.2 Testing user's ability

This model's primary goal is to identify the user's capabilities so that they can feel secure and in control while completing the tasks and taking the exam. The paper's model does this by first randomly assigning three questions of each difficulty level to each of the three parts. When a user answers every question, the organisation can compare the user's responses to the test questions' difficulty levels to determine where the user actually stands.

5.3.3 Parameter selection for difficulty calculation

The difficulty of the Sect. 4 questions is determined by averaging the first three sections. (User's Ability & its Perception of Difficulty) The average is determined by the right answers provided by the user in accordance with the kind of questions' difficulty level; the higher the degree of difficulty, the more likely the next section's questions will be of a higher difficulty.

```
end if
N ← Q/M;
        // N = Total Number of Sets
for i ← 0 to Y-1 do
    for j ← 0 to M do
        res ← response of X[j];
                // res either has response has Right or Wrong
        if res == "Right" then
            MS ← MS + LQ[j]
        end if
    end for
end for
DF ← MS/ ((Y-1) * (M*(M+1)/2))
C ← [0] * M;
LQ ← [DF] * M;
for i ← Y-1 do
    for j ← 0 to M do
        res ← response of X[j];
                // res either has response has Right or Wrong
        if res == "Right" then
            MS ← MS + LQ[j]
            C[j] ← C[j] + 1
            res == "Wrong"
            C[j] ← C[j] - 1;
        end if
    end for
end for
```

Fig. 3 Proposed Algorithm

```
for i ← Y to N do
    Construct Question set with Levels in LQ;
    responses ← Take responses;
    Call Check-for-Set(responses);
    End of set;
end for
for i ← Y to N do
    for j ← 0 to M do
        res ← response of X[j];
                // res either has response has Right or Wrong
        if res == "Right" then
            if LQ[j] < M and C[j] == 1 then
                LQ[j] ← LQ[j] +1;
                C[j] ← 0;
                C[j] ← C[j] + 1
                res == "Wrong"
                if C[j] == -2 then
                    if LQ[j] > 1 then
                        LQ[j] ← LQ[j] -1;
                        C[j] ← 0;
                    end if

                else
                    C[j] ← C[j] - 1;
                end if
            end if
            MS ← MS + LQ[j]
```

Fig. 4 Proposed Algorithm

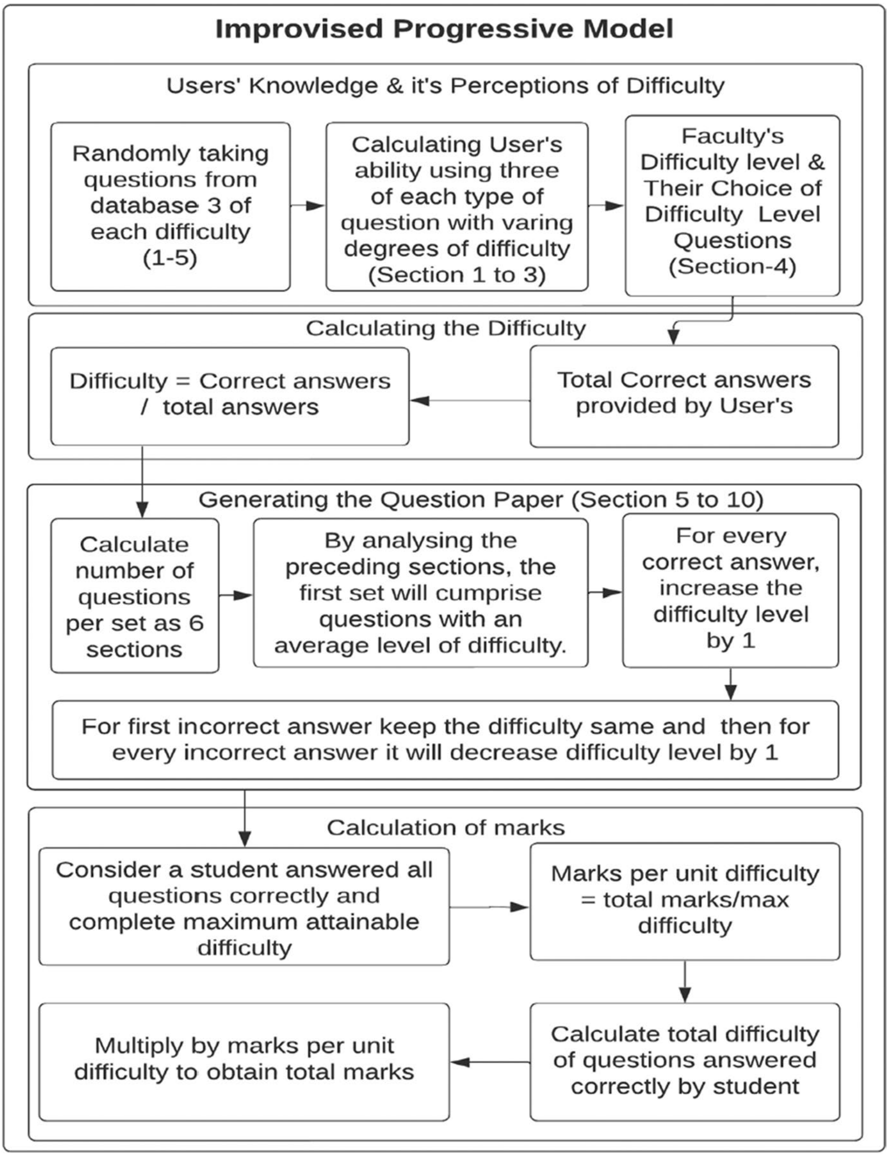

Fig. 5 Improvised Progressive Model Flow Chart

5.3.4 Proposed progressive model

The three primary processes in the suggested model for the student assessment system are determining the difficulty, producing the exam, and assessing the student. The difficulty of a question is computed first, and the questions are then classified into different difficulty levels based on the previously established association between a question's difficulty and the correct answers. The answers are then used to grade the learner after the questions have been utilised to generate the questions based on the algorithm.

Explaining the model using the case of a test-taker who completed it and correctly responded to every query up until the third degree of difficulty in the first three portions. As a result, all of the questions in Sect. 4 would be of level 3 difficulty, with the User's average difficulty being three. The Progressive model will now use its algorithm, which is the user's, to answer the questions in the following part. For example, the response to question number 1 will only affect the remaining question 1 of every subsequent section, whether the answer is correct or incorrect. The other number sequences in any segment will not be impacted. The difficulty level would rise if the response was accurate, stay the same if the response was incorrect again, and fall if the response was incorrect, respectively. If the answer is incorrect twice in a row, the difficulty level will not decrease. This is the most advanced model that can be used to help the company select the best and ideal candidate from the competition.

5.3.5 Computing the difficulty

Our goal is to calculate the difficulty of a multiple-choice question that n students have attempted. In a perfect world, where time, age, gender, and other restrictions are unaffected, the difficulty (d) of any question can be expressed as the proportion of right answers (c) to all responses (n) as shown and stated in formula (1).

$$Difficulty\ (d) = \frac{number\ of\ correct\ responses\ (c)}{total\ number\ of\ responses\ (c)} \tag{1}$$

Knowing the user's ability to solve all problems once and the faculty's difficulty level of expectation, the model can compute and provide an average difficulty level of correct responses to the number of questions answered by the user in terms of computing its score using the given formula.

$$Difficulty\ level\ for\ Initial\ Stage\ of\ Progressive\ Model = \frac{Correct\ Responses\ *\ Difficulty\ Level}{45} \tag{2}$$

When a user provides all correct responses for a specific difficulty, the set may be one or more in terms of correct responses for the entire difficulty indicated which is shown in formula (2) how the difficulty is being calculated for the next section which is Progressive Model with Initial stage. Here, we provide a variety of instances and ensure that no incorrect distribution of difficulty occurs incorporating our proposed approach, and that consistency is maintained for each and every user and organisation that employs our model.

As noticed in the evaluation of students, many students cheat and receive high grades even though they have not studied. This is because all questions are identical, as is the format in which they are presented. By using our model, we can generate unique questions for each student, and if any student receives the same question, it will be in a different question number, making it difficult for students to compare their papers even if they wish to do so. This will be a truly fair system for those who work hard for the evaluation. When the model poses varying questions, the question

Table 2 Ranges for determining difficulty levels of questions

Difficulty Levels	Range for Each Level
Difficulty Level 1	0.00 < = x < 0.20
Difficulty Level 2	0.20 < = x < 0.40
Difficulty Level 3	0.40 < = x < 0.60
Difficulty Level 4	0.60 < = x < 0.80
Difficulty Level 5	0.80 < = x < 1.00

arises, how would we take care of the biasness between the students. Here, difficulty plays a major role. Thus, the questions feed in the database are the one which are ranked by the specialists, so once the difficulty has been accurately categorised in the database, it will produce unique questions with the same level of difficulty, so that each student will be treated equally if the questions are unique.

The difficulty levels range from 1 to 5, with 1 being the easiest and 5 being the most challenging. When determining and calculating the average of the difficulty based on the marks earned by the user in the first four sections, where each difficulty has been assigned some fixed ratios ranging from 0–1. The fixed ratio of marks obtained by the User to the total marks of the four sections has been distributed, as shown in Table 2. The table shows that if the ratio calculated using the above-mentioned difficulty formula falls in the range 0–0.2, then difficulty level 2 falls in the range 0–0.2, and so on for each of the five difficulty levels, a 0.2 or 20% range is provided.

5.3.6 Evaluating the user

$$Marks\ Scored\ by\ the\ User = \frac{(Questions\ Correct)\ (Their\ Difficulty\ Level)}{195} \times 100 \tag{3}$$

The formula (3) stated above will be used to determine the user's final test scores. The test was divided into 10 sections, the first three of which tested the user's ability to solve problems mentally, with a possible score of 45. Then the progressive model is used, where the user can earn a maximum of 150 marks by passing each of the six sections one at a time, and the average difficulty is determined based on the four sections. The final step is to determine the overall marks based on the User's responses. By multiplying the correct responses in accordance with the difficulty level of the questions, which is also noted by the system, the system calculates the correct answers to the questions. The user's final grade is determined by scaling their total points earned down to 100 using the mentioned formula.

5.3.7 Generating the question paper

The most important phase in this progressive strategy is the creation of test questions. A student will receive more difficult questions in the following set if they have answered more questions and questions of higher complexity in a set, and vice versa. We add up all the questions (Q) and all the marks.

6 Results and analysis

6.1 Simulation analysis of proposed progressive model

This section presents the analysis of the initial Proposed Progressive Model shown in Tables 3, 4, 5 and 6. The Proposed Progressive Model consisted of the first 3 sections containing questions of all difficulty levels for assessing the level of the student. The fourth section was based on a difficulty level that was given determined by the evaluator. This section was added for the purpose of getting a sense of the difficulty level of the questions that the evaluator expected from the students. From the fifth section onwards, questions were generated dynamically based on the student's responses.

This, however, posed some inconsistencies while calculating and determining the difficulty for the 'Progressive Sections' ranging from Sects. 6th to 10th. Table 7 demonstrates one of the multiple cases in which inconsistency occurred while calculating the difficulty for Sect. 5 when the evaluator's/mentor's difficulty expectation was taken into consideration. The arrows on the right of the image indicate the incorrect determination of the difficulty levels. For instance, when the student correctly answered all the questions of the first section till difficulty level 3 (row 4), and the evaluator gave an input of DF0 for the 'Expectation Section', the difficulty calculated by our model for Sect. 5 was inconsistent with the rest of the model. The model assigned a level of DF2 for the 'Progressive Section' whereas the expected level was DF3. Due to these inconsistencies, the 'Expectation Section' which considered the evaluator's input was removed, and the model was developed without it.

Table 3 Preliminary Sections

Preliminary Sects. (1st—3rd)					
Sections	Q1	Q2	Q3	Q4	Q5
Difficulty Levels	DF1	DF2	DF3	DF4	DF5
Questions	3	3	3	3	3

Table 4 Expectation sections

Expectation Sects. (4th)					
Sections	Q1	Q2	Q3	Q4	Q5
Difficulty Levels	Assigned by the Faculty (e.g., DF4)				
Questions	1	1	1	1	1

Table 5 Progressive section initial section

Progressive Section Initial Sect. (5th)					
Sections	Q1	Q2	Q3	Q4	Q5
Difficulty Levels	Average of the above 4 Sections				
Questions	1	1	1	1	1

Table 6 Progressive sections

Progressive Sects. (6th-10th)					
Sections	Q1	Q2	Q3	Q4	Q5
Difficulty Levels	Based on previous Responses				
Questions	5	5	5	5	5

Table 7 Expected vs assigned difficulties of section in proposed progressive model

DF1	DF2	DF3	DF4	DF5	Section 4	Sum	Total marks	Fraction	Expected	Assigned	
0	0	0	0	0	0	0	0	0.00	0	0	
3	0	0	0	0	0	3	3	0.05	1	1	
3	3	0	0	0	0	6	9	0.15	2	1	←
3	3	3	0	0	0	9	18	0.30	3	2	←
3	3	3	3	0	0	12	30	0.50	4	3	←
3	3	3	3	3	0	15	45	0.75	5	4	←

6.2 Result and analysis of improvised progressive model

This section presents the results and outcomes of the built progressive model. As per the model built, the first 3 sections contain questions of all difficulty levels for assessing the level of the student. From the fourth section onwards, questions are generated dynamically based on the student's responses. For research purposes, the paper pattern used in the model is shown in Table 8.

The factors given above try to mimic real-life test patterns and can be altered according to the needs of the test maker. The diagram below represents the broad cases that could take place while the student is giving the test (Fig. 6).

6.2.1 Best case –(Fig. 7)

In the best case, all the questions answered by the student would be correct. This means that he will attain the maximum possible difficulty achievable. Hence, the marks obtained by the student would be 100% of the maximum marks which is 220. It is observed that since the student gives all correct answers in the first 3 sections, his difficulty is the highest in the consecutive sections and remains the same throughout.

Table 8 Details of the Improvised Progressive Model used in the Paper

No of sections	10
No. of questions in each section	5
Difficulty Levels	5 (DF1, DF2, DF3, DF4, DF5)
Total Questions	50

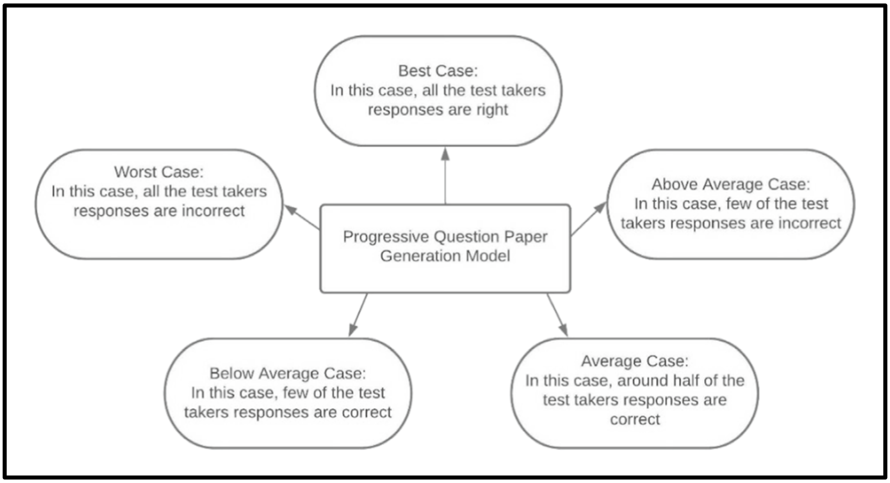

Fig. 6 Cases of Improvised Progressive Model

6.2.2 Above average case –(Fig. 8)

In the above average case, more than half of the questions answered by the student would be correct. Hence, the marks obtained by the student would be between 50 to 100% of the maximum marks. Our model is consistent with this as shown in the above case. The student scored 165 out of a cumulative difficulty of 220 which is 75.00%. It is observed that the scores 33/45 in the first 3 sections, due to which his next section starts with D4. The student answers most of the questions correctly and when he answers an incorrect question, due to the design of the model, he is given a chance to score on the question of the same difficulty again.

6.2.3 Average case –(Fig. 9)

In the average case, around half of the questions answered by the student would be correct. Hence, the marks obtained by the student would be around 50% of the maximum marks. Our model is consistent with this as shown in the above case. The student scored 110 out of a cumulative difficulty of 220 which is 50.00%. It is observed that the scores 22/45 in the first 3 sections, due to which his next section starts with D3. The rate of decrease in difficulty is lower than the rate of increase in difficulty due to which the sections in the end are of higher difficulty.

	Q1	Q2	Q3	Q4	Q5
SET 1	QD1	QD1	QD1	QD2	QD2
✓/ ×	✓	✓	✓	✓	✓
Diff.	1	1	1	2	2
SET 2	QD2	QD3	QD3	QD4	QD4
✓/ ×	✓	✓	✓	✓	✓
Diff.	2	3	3	3	4
SET 3	QD4	QD4	QD5	QD5	QD5
✓/ ×	✓	✓	✓	✓	✓
Diff.	4	4	5	5	5
Diff. of SET4	$= \frac{\text{Correctly Solved Difficulty}}{\text{Total Difficulty till SET 3}} = \frac{45}{45} = 1 \Rightarrow$ D5				
SET 4	QD5	QD5	QD5	QD5	QD5
✓/ ×	✓	✓	✓	✓	✓
Diff.	5	5	5	5	5
SET 5	QD5	QD5	QD5	QD5	QD5
✓/ ×	✓	✓	✓	✓	✓
Diff.	5	5	5	5	5
SET 6	QD5	QD5	QD5	QD5	QD5
✓/ ×	✓	✓	✓	✓	✓
Diff.	5	5	5	5	5
SET 7	QD5	QD5	QD5	QD5	QD5
✓/ ×	✓	✓	✓	✓	✓
Diff.	5	5	5	5	5
SET 8	QD5	QD5	QD5	QD5	QD5
✓/ ×	✓	✓	✓	✓	✓
Diff.	5	5	5	5	5
SET 9	QD5	QD5	QD5	QD5	QD5
✓/ ×	✓	✓	✓	✓	✓
Diff.	5	5	5	5	5
SET 10	QD5	QD5	QD5	QD5	QD5
✓/ ×	✓	✓	✓	✓	✓
Diff.	5	5	5	5	5

Fig. 7 Best Case Scenario

	Q1	Q2	Q3	Q4	Q5
SET 1	QD1	QD1	QD1	QD2	QD2
✓/ ×	✓	✓	✓	×	✓
Diff.	1	1	1	2	2
SET 2	QD2	QD3	QD3	QD4	QD4
✓/ ×	✓	✓	✓	✓	✓
Diff.	2	3	3	3	4
SET 3	QD4	QD4	QD5	QD5	QD5
✓/ ×	✓	✓	×	✓	×
Diff.	4	4	5	5	5
Diff. of SET4	$= \frac{\text{Corr. Solved Diff.}}{\text{Ttl Diff. till SET 3}} = \frac{33}{45} = 0.7333 => D4$				
SET 4	QD4	QD4	QD4	QD4	QD4
✓/ ×	✓	×	✓	✓	✓
Diff.	4	4	4	4	4
SET 5	QD5	QD4	QD5	QD5	QD5
✓/ ×	✓	✓	✓	×	✓
Diff.	5	4	5	5	5
SET 6	QD5	QD5	QD5	QD5	QD5
✓/ ×	✓	✓	×	×	✓
Diff.	5	5	5	5	5
SET 7	QD5	QD5	QD5	QD4	QD5
✓/ ×	✓	✓	×	✓	✓
Diff.	5	5	5	4	5
SET 8	QD5	QD5	QD4	QD5	QD5
✓/ ×	✓	×	✓	✓	✓
Diff.	5	5	4	5	5
SET 9	QD5	QD5	QD5	QD5	QD5
✓/ ×	✓	×	✓	✓	✓
Diff.	5	5	5	5	5
SET 10	QD5	QD4	QD5	QD5	QD5
✓/ ×	✓	✓	✓	✓	✓
Diff.	5	5	5	5	5

Diff. Increases
Diff. Decreases
Diff. Remains Same

Fig. 8 Above Average Case Scenario

6.2.4 Below average case –(Fig. 10)

In the below average case, less than half of the questions answered by the student would be correct. Hence, the marks obtained by the student would be between 00 to 50% of the maximum marks. Our model is consistent with this as shown in the above case. The student scored 55 out of a cumulative difficulty of 220 which is 25.00%. It is observed that the scores 15/45 in the first 3 sections, due to which his next section starts with D2. The student answers a few of the questions correctly and gets a chance to score more with every correct answer.

	Q1	Q2	Q3	Q4	Q5
SET 1	QD1	QD1	QD1	QD2	QD2
✓/✗	✓	✓	✓	✓	✓
Diff.	1	1	1	2	2
SET 2	QD2	QD3	QD3	QD4	QD4
✓/✗	✓	✓	✓	✓	✗
Diff.	2	3	3	3	4
SET 3	QD4	QD4	QD5	QD5	QD5
✓/✗	✗	✓	✗	✗	✗
Diff.	4	4	5	5	5
Diff. of SET4	$= \frac{\text{Corr. Solved Diff.}}{\text{Ttl Diff. till SET 3}} = \frac{22}{45} = 0.4889 \Rightarrow$ D3				
SET 4	QD3	QD3	QD3	QD3	QD3
✓/✗	✓	✓	✓	✓	✓
Diff.	3	3	3	3	3
SET 5	QD4	QD4	QD4	QD4	QD4
✓/✗	✓	✗	✓	✓	✓
Diff.	4	4	4	4	4
SET 6	QD5	QD4	QD5	QD5	QD5
✓/✗	✓	✓	✗	✗	✓
Diff.	5	4	5	5	5
SET 7	QD5	QD5	QD5	QD5	QD5
✓/✗	✗	✓	✗	✗	✗
Diff.	5	5	5	5	5
SET 8	QD5	QD5	QD4	QD4	QD5
✓/✗	✓	✗	✓	✓	✓
Diff.	5	5	4	4	5
SET 9	QD5	QD5	QD5	QD5	QD5
✓/✗	✓	✗	✓	✗	✓
Diff.	5	5	5	5	5
SET 10	QD5	QD4	QD5	QD5	QD5
✓/✗	✗	✗	✗	✓	✗
Diff.	5	4	5	5	5

Diff. Increases
Diff. Decreases
Diff. Remains Same

Fig. 9 Average Case Scenario

6.2.5 Worst case –(Fig. 11)

In the worst case, all the questions answered by the student would be incorrect. This means that he will attain the minimum possible difficulty achievable. Hence, the marks obtained by the student would be 0% of the maximum marks which is 0. It is observed that since the student gives all incorrect answers in the first 3 sections, his difficulty is the lowest in the consecutive sections and remains the same throughout.

	Q1	Q2	Q3	Q4	Q5
SET 1	QD1	QD1	QD1	QD2	QD2
✓/ ×	✓	×	✓	✓	✓
Diff.	1	1	1	2	2
SET 2	QD2	QD3	QD3	QD4	QD4
✓/ ×	✓	×	×	✓	×
Diff.	2	3	3	3	4
SET 3	QD4	QD4	QD5	QD5	QD5
✓/ ×	×	✓	×	×	×
Diff.	4	4	5	5	5
Diff. of SET4	$= \frac{\text{Corr. Solved Diff.}}{\text{Ttl Diff. till SET 3}} = \frac{15}{45} = 0.3 \Rightarrow D2$				
SET 4	QD2	QD2	QD2	QD2	QD2
✓/ ×	✓	×	×	✓	✓
Diff.	2	2	2	2	2
SET 5	QD3	QD2	QD2	QD3	QD3
✓/ ×	✓	×	✓	×	×
Diff.	3	2	2	3	3
SET 6	QD4	QD1	QD3	QD3	QD3
✓/ ×	×	×	×	✓	✓
Diff.	4	1	3	3	3
SET 7	QD4	QD1	QD3	QD4	QD4
✓/ ×	×	✓	×	×	✓
Diff.	4	1	3	4	4
SET 8	QD3	QD2	QD2	QD4	QD5
✓/ ×	×	×	×	×	✓
Diff.	3	2	2	4	5
SET 9	QD3	QD2	QD2	QD3	QD5
✓/ ×	✓	×	✓	✓	×
Diff.	3	2	2	3	5
SET 10	QD4	QD1	QD3	QD4	QD5
✓/ ×	✓	✓	×	×	×
Diff.	4	1	3	4	5

Diff. Increases
Diff. Decreases
Diff. Remains Same

Fig. 10 Below Average Case Scenario

6.3 Comparison with basic progressive model

This section presents the differences between the Basic Progressive Model presented in (Singh et al., 2021) and the Improvised Progressive Model that was inspired by it (Fig. 12).

It is inspired in the sense of determining and calculating the difficulty of the questions of the next sections, but with some minor changes. The Basic Progressive Model in (Singh et al., 2021) consisted of the first section to contain one question each from all difficulties. Consequently, based on the users' responses, the difficulties of the questions changed in the following manner:

	Q1	Q2	Q3	Q4	Q5
SET 1	QD1	QD1	QD1	QD2	QD2
✓/ ✗	✗	✗	✗	✗	✗
Diff.	1	1	1	2	2
SET 2	QD2	QD3	QD3	QD4	QD4
✓/ ✗	✗	✗	✗	✗	✗
Diff.	2	3	3	3	4
SET 3	QD4	QD4	QD5	QD5	QD5
✓/ ✗	✗	✗	✗	✗	✗
Diff.	4	4	5	5	5
Diff. of SET4	$= \frac{\text{Correctly Solved Difficulty}}{\text{Total Difficulty till SET 3}} = \frac{0}{45} = 0 => \text{D1}$				
SET 4	QD1	QD1	QD1	QD1	QD1
✓/ ✗	✗	✗	✗	✗	✗
Diff.	1	1	1	1	1
SET 5	QD1	QD1	QD1	QD1	QD1
✓/ ✗	✗	✗	✗	✗	✗
Diff.	1	1	1	1	1
SET 6	QD1	QD1	QD1	QD1	QD1
✓/ ✗	✗	✗	✗	✗	✗
Diff.	1	1	1	1	1
SET 7	QD1	QD1	QD1	QD1	QD1
✓/ ✗	✗	✗	✗	✗	✗
Diff.	1	1	1	1	1
SET 8	QD1	QD1	QD1	QD1	QD1
✓/ ✗	✗	✗	✗	✗	✗
Diff.	1	1	1	1	1
SET 9	QD1	QD1	QD1	QD1	QD1
✓/ ✗	✗	✗	✗	✗	✗
Diff.	1	1	1	1	1
SET 10	QD1	QD1	QD1	QD1	QD1
✓/ ✗	✗	✗	✗	✗	✗
Diff.	1	1	1	1	1

Fig. 11 Worst Case Scenario

- If the user gave a correct answer, the corresponding question of the next section had a question of a higher difficulty. The difficulty was increased by 1 each time the user gave the correct answer.
- If the user gave the incorrect answer, the corresponding question of the next section had a question of the same difficulty. Furthermore, if the user gave a consequent incorrect answer, the difficulties of the questions of the corresponding sections decreased each time by 1 till the user gives a correct answer.

Unlike the model in (Singh et al., 2021), the proposed model here increases the number of sections containing questions of all difficulties. The initial sections will contain more than one question from each difficulty. This will allow to account for any of the mistakes by giving them another chance to attempt the question of the said difficulty, which will give a better understanding of the student's knowledge level. Based on the student's responses, the difficulties of the questions change in the following manner:

- If the user gives a correct answer, the corresponding question of the next section will a question of a higher difficulty. The difficulty is increased by 1 each time the user gives a correct answer.
- If the user gives the incorrect answer, the corresponding question in the next section will have a question of the same difficulty. Furthermore, if the user gave a consequent incorrect answer, the difficulty of the questions of the next section is decreased by 1. This cycle then continues till the end of the test.

Another difference between the two models is in the way the overall test is generated. In the model proposed in the (Singh et al., 2021), the algorithm for increase

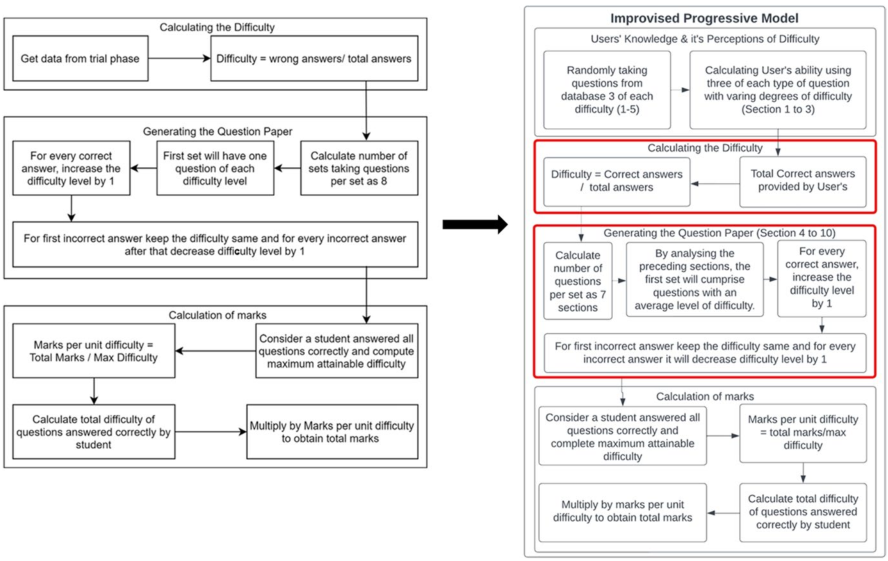

Fig. 12 Difference between Basic Progressive Model and Improvised Progressive Model

and decrease of the difficulties of the questions is used from the initial section of the test. This creates a unique question paper for each, and every user based on their responses. In the Improvised Proposed model, the initial sections are set to be identical for all the users. The number of identical sections depends on the organisation/team conducting the test. After the set of identical sections, the knowledge level of the student is determined and from there on the algorithm for increase and decrease of the difficulties of the questions is used for dynamic question paper generation.

As far as we have observed, all of the past research has focused only on a single aspect such as either predicting question difficulty, determining the best balancing parameters for classifying questions, generating questions using an already trained model, or employing paper evaluating methods such as Genetic Algorithms and Fuzzy Logic, used by (Perez et al., 2012) which investigates evaluator' ability to categorise questions by difficulty level, comparing it to students' perceptions and the measures obtained by an expert system of automatically classifying the questions using QUESTOURnament, which evaluates the level of complexity of the question, and it demonstrates that students' perceptions will outperform the evaluators perceptions. Due to these same inconsistencies, the 'Expectation Section' that considered the evaluator's input was eliminated, and our model was developed without it as mentioned in Sect. 6.1. Next is (Purohit et al., 2012) 's an Adaptive Question Bank Development and Management System (AQBDMS) that selects questions from a rich database (question bank) intelligently and represents the question model using inputs or parameters supplied by the question paper designer (QPD). This question bank employs a concept map built using a visual technique and hierarchical knowledge of a certain topic to ensure that questions are modelled as closely as possible in accordance with components such as Bloom's Taxonomy, difficulty level, grading system, and so on.

Automatic test paper creation (Shah et al., 2022) and AQG (Patil et al., 2022) method attempted to solve the constraint satisfaction problem (CSP), which involves selecting questions from a question bank and automatically producing various types of papers that match instructional requirements, saving time using various NLP techniques, but the system also needed to employ security controls to prevent question recurrence, ensuring the test's security and objectivity, so Bloom's taxonomy (Ragasudha & Saravanan, 2022) must be used to create a database of questions. Cryptography was used to encrypt and decrypt the question paper to safely transfer it to the permitted individual. To evaluate the response automatically, the keyword matching method is used. Because the answer sheet had to be scanned, uploaded, and then analysed based on the keyword, the processing time was quite slow. And, because we were looking for objective questions, incorporating Bloom Taxonomy's keyword matching was pointless by wasting time on all the processes.

But our goal was to come up with a complete package of system consisting of automatic question paper generator and an evaluation system in an integrated system which could be used to evaluate the candidates through an adaptive test (Refer Table 9). This would be a fully automated system in which the candidates' perspective of a question paper was taken into consideration by evaluating them in the Preliminary Sections and to make the test adapt to the difficulty level and understanding of the candidate. Next in the Progressive Section we evaluate the candidate then and there before they reach the next set of Questions which depends on the responses

Table 9 Comparative Analysis of Progressive Model with related work

Citation	Method Used	Generation of Automatic Question Paper	Automatic Calibration of Difficulty Level	Calculation of Score
(Perez et al., 2012)	Genetic algorithms and Fuzzy Logic	x	✓	x
(Purohit et al., 2012)	Adaptive Question Bank Development and Management System (AQBDMS)	✓	✓	x
(Wang et al., 2014)	K-Nearest Neighbor	x	✓	✓
(Mishra & Jain, 2016)	Logical Queries	✓	x	x
. (Sunil, 2020)	Smart Paper Generator	✓	x	x
(Srivastava & Goodman, 2021)	Language Models Knowledge Tracing (LM-KT) Model	✓	x	✓
(Dwivedi et al., 2019)	Automated system for Retrieving and Designing exam	✓	x	x
(Patil et al., 2022)	Automatic Question Generation (AQG) Framework	✓	x	x
(Shah et al., 2022)	Automatic question paper generator system	✓	x	x
(Ragasudha & Saravanan, 2022)	Bloom's taxonomy	✓	x	✓
	Our Model	✓	✓	✓

of the previous set of Questions and hence the difficulty increases and decreases in accordance with the thorough approach described in Sect. 5.3 and the final result or score is described on the basis of the responses which are correctly answered by the candidate and the difficulty of the question.

This change in the generation of questions with a set of identical sections allows a hybrid of static and dynamic method to assess the student. It assesses the student's initially on common ground from the group of student's and later proceeds in a dynamic fashion. The static part of the test allows for a more customised test paper which allows for better assessment and evaluation of the student.

7 Discussion and conclusion

Learning via the internet is becoming increasingly popular because it allows students to learn anything, at any time, from any location by utilising web resources. In any learning system, assessment is essential (Das et al., 2021). An assessment system can help learners identify self-learning gaps and progress in their learning.

According to the results, we were able to create two types of models. On the one hand, we used the Evaluator's expectation model, but this created bias because the student answered to each difficulty level questions correctly, including the evaluator's difficulty level questions, so the results were not accurate. So, by changing the model, we only kept the Preliminary stage while removing the Evaluator's expectation model, where our model produces more accurate results than expected, and some cases that we demonstrated while explaining the model. This demonstrates how accurate the results are while only using the Preliminary Answers provided by the students to judge them for the subsequent evaluation process. For fair results for everyone, it judges how you answer, and the marks the student will receive will be proper marks that he deserves, and by this we can say that the model will give chances to those who deserve them, while not being generous to those who do not learn for the evaluation (Jopp & Cohen, 2022). When compared to other processes, our Progressive Question Generation Model will give students more flexibility to double their motivation and scoring rate, because they will receive questions based on the difficulty they have scored in previous sections, and our model will take the average of the questions answered correctly, so that the model can know your ability to further classify the problems in the evaluation. Other evaluation processes use the same paper with the same questions and format, so they judge each student using the same paper. However, with these models, they will have more flexibility to respond based on their expected difficulty level (Goss, 2022). We used the Paper (Singh et al., 2021), which we modified and incorporated into our model. In our model, the User's knowledge in the test is determined first, and then subsequent questions are given based on that knowledge (Facilitating an Online and Sustainable Learning Environment for Cloud Computing Using an Action Research Methodology, 2023; Computer Science Education, 2022 and Tsunemoto et al., 2022).

7.1 The applicability of the progressive model

The model is a Question Generating model used in the Graduate Record Examination (GRE) and the Graduate Management Admission Test (GMAT) (Wendler & Bridgeman, 2014). This model uses adaptive technology to determine a student's ability level in real-time and adjusts the difficulty of each subsequent question based on the student's previous answers (Huang et al., 2009). The progressive Model aims to provide a more accurate and reliable assessment of a student's abilities, tailored to the experience for the individual. The model works by starting with an average difficulty question, then varying the complexity of the subsequent questions depending on the student's answer to the first question. If the student answers the previous question correctly, the next question will be more difficult. If the student answers incorrectly, the next question will be of lower difficulty. This process continues until completion, providing an accurate and individualized assessment of the student's abilities.

In comparison to typically standardised testing methodologies, the Progressive Model has various advantages. For one thing, it makes testing more efficient and effective because the test adapts to the student's abilities in real-time. Furthermore, it delivers a more accurate score that more accurately reflects the student's ability and assists admissions committees in making more informed selections. Our Progressive Model is primarily concerned with learner-centered assessment. The emphasis on individualized learning distinguishes learner-centered assessment from more traditional approaches to evaluation. The emphasis in learner-centered assessment is on the learner's growth and understanding rather than the evaluator's capacity to assess their knowledge. This method has been demonstrated to be more effective in encouraging deeper learning, more student involvement, and a better knowledge of the subject matter.

The Progressive Model is thus a sophisticated and effective scoring system used in standardized tests such as the GRE and GMAT. It provides a more efficient and effective testing experience by tailoring the test experience to the individual and providing a more accurate score, reducing test anxiety, and assisting admissions committees in making more informed decisions. In contrast to our approach, which is based on the student's preferences and test-taking abilities, the competition's model provides a random or fixed difficulty level input.

The model also not only assists in the proper evaluation of students but also determines the topics in which the students excel by not only giving different types of difficulties but also considering different types of subjects such as logical analysis, technical, mathematical, social, general, history, and many others. It also aids in Placement Activities and industry certifications because the emphasis is on the individual learner's abilities and needs rather than the exam itself. The assessment process would be tailored to the students' specific learning style and goals, considering their strengths, weaknesses, and any special needs (Starkweather et al., 2017).

The coronavirus pandemic has forced thousands of students and educators to reevaluate how the education system keeps everchanging and the necessity of adapting efficiently to the changing pace of times. Many schools went completely online with little or no notice, and teachers were working hard to ensure that no compromise takes place in their teaching with the newly presented mode of teaching (Gill et al. 2022). With its aim to improve social justice by re-designing the curriculum of Queen Mary University of London's Cloud

Computing course (Gill et al., 2023) used the blended learning technique for cloud computing to foster prior experience by incorporating active learning activities such as interactive video, Mentimeter quiz, flipped classroom, group-based projects, and lab. The progressive model of question paper generation can be used by the students to create quizzes for their classmates. This will ensure that the teaching methodology proposed in (Gill et al., 2023) which includes research knowledge and teamwork skills would be fulfilled. The model is progressive and adapts based on the test taker's ability. This will be helpful for the student taking the quiz which would be an indication of the areas in which he lacks, and additionally will also give him an idea of how different students think while creating a quiz. This will give an all-around inclusive experience (Gill et al., 2023) to the student. Furthermore, the student who creates hard quizzes for his teammates would mean he understands the concepts of the subject well. This will provide an excellent opportunity and platform for students to test themselves in such a curriculum while catalyzing professional development. It also increases engagement which is important for the development of students (Gill et al. 2022). The progressive model of question paper generation falls very well in place for the proposed model in (Gill et al., 2023) where it can be used to evaluate the students in the output and outcomes phases of the research plan pathway. It would help the teachers to evaluate how well students have learned using virtual laboratories of cloud computing to see where they were succeeding and where they were failing. This will help to develop unique and interesting types of teaching methods based on the students' performances. The use of the progressive model of question paper would also ensure that no misconduct (Gill et al. 2022) takes place during the tests as the papers generated would be based on the student's answers and understanding, making it less likely for students to have the same questions in their tests (Curriculum Redesign for Cloud Computing to Enhance Social Justice and InterculturalDevelopment in Higher Education, 2023).

7.2 Contribution to the literature

The model can be highly regarded as a set of standard tests used by full-time employees of the organisation to evaluate candidates in accordance with their preferences and get an edge in hiring the best candidate using this approach (Fung et al., 2022). They don't directly add to the body of literature. However, the tests evaluate verbal and mathematical abilities, which are crucial for success in many areas, including literature. As a result, they may indirectly contribute to the study and appreciation of literature by assisting in the identification of people who have the aptitude to pursue such studies. These exams are made to gauge a candidate's critical thinking, problem-solving, and analytical writing abilities, all of which are thought to be necessary for success in the real world. An individual who performs well on these tests is more likely to succeed in a job and enjoy the environment there because he possesses the information and skills necessary for success (Menon & Suresh, 2022). A high score in our model can also make someone more competitive on the job market and boost their chances of admission to a top-ranked school. The exams will be widely acknowledged by academic institutions and businesses across the world, making them an important resource for individuals interested in pursuing post-secondary education or a career in any major companies (Nieminen, 2022).

7.3 Practical implications

The Progressive Model has practical implications in the assessment process by providing a standardized and efficient way to create and administer tests and exams (Hardigan et al., 2001; Yudono et al., 2022). This approach helps to ensure that all learners are evaluated consistently and fairly, based on a pre-determined set of questions and criteria. Question paper generation can also help to reduce the time and effort required to create and administer exams, freeing up evaluators and administrators to focus on other important aspects of the assessment process.

The Model provides several benefits to students, including fair and consistent assessments, reduced stress and anxiety, improved learning outcomes, and increased efficiency. Additionally, question paper generation can motivate students to focus on their learning and improve their outcomes and free up eevaluators to focus on other important aspects of the assessment process such as providing feedback.

The Progressive Model has many practical implications, such as how a learner-centred approach to assessment can provide several practical benefits in both the admissions process and improving job prospects. A learner-centred approach to admissions can provide a more comprehensive and meaningful evaluation of a candidate's abilities and potential for success in a specific programme or course (Yudono et al., 2022). This method can provide practical, real-world evidence of a candidate's abilities while also considering their individual learning needs and styles. In terms of job prospects, a learner-centred approach to assessment can provide employers with a more accurate and meaningful evaluation of a candidate's abilities because it considers a broader range of evidence, such as projects, self-assessments, and performance tasks, in addition to traditional methods such as exams and quizzes (Ranjan et al., 2022). This ensures that candidates are evaluated fairly and on the basis of their unique strengths and abilities.

7.4 Social implications

The automatic generation of question papers has potential social implications that are both positive and negative. On one hand, it has the potential to increase productivity and efficiency in academia by allowing evaluators to quickly generate high-quality papers and free up time for more in-depth analysis and experimentation. This could lead to a higher rate of scientific discovery and advancement.

On the other hand, automatic paper generation raises ethical concerns about the validity and reliability of the research being produced. If the generated papers contain false or misleading information, this could have serious consequences for the scientific community and society at large (Jaiswal et al., 2022). Additionally, the widespread use of automatic paper generation could lead to a reduction in the diversity of ideas and perspectives in the scientific community, as well as a decrease in the importance placed on original thought and creativity.

Furthermore, the commercialization of automatic paper generation tools could further exacerbate these negative social implications. For example, companies could use these tools to generate low-quality or biased research to support their products or services, leading to the dissemination of false or misleading information.

7.5 Future scope & limitations

As technology continues to advance, the Progressive model is expected to expand its test formats to include computer-based and online options. This will make the exams more accessible to a wider range of individuals and provide greater flexibility in terms of when and where the exams can be taken. In addition to being used for graduate school admission, the exams may become more widely used for job placement purposes, as employers seek to assess the aptitude and skills of job candidates. The increased use of these exams for job placement is due to the growing recognition of their value in evaluating an individual's knowledge, skills, and abilities. Furthermore, as the workforce continues to evolve, there may be a greater emphasis on evaluating individuals' soft skills, such as communication, leadership, and teamwork, in addition to their technical knowledge. This will provide a more comprehensive evaluation of an individual's abilities, leading to better outcomes and improved success in both education and employment.

While the Progressive model has several benefits, it is not without its drawbacks. One concern is that the exams may not accurately reflect a students' real-world skills and abilities, as they focus primarily on aptitude and knowledge rather than practical application. Another issue, the cost of taking the exams can be a barrier for some students', particularly those from low-income backgrounds. There is also a tendency for students to focus too much on test-taking strategies rather than actual knowledge and skills, leading to a narrow emphasis on exam performance rather than overall development. The high stakes of the exams can also cause significant stress and anxiety for students', potentially impacting their performance and overall well-being. It is important to consider these limitations and address them to ensure that the exams provide an accurate and fair evaluation of a student's abilities and potential for success.

This conversation highlights ChatGPT's impact on our model which lacks consistency across various subject areas as discussed in (Gill et al., 2023). However, its use raises several problems, including the creation of misleading or erroneous information and the threat it poses to academic legitimacy. Since ChatGPT was trained on a sizable amount of data, there are significant ethical questions about whether it may be biased or inaccurate. If this is the case, it might seriously undermine the quality of questions generated with it & the ranking of difficulty as well as the validity of the data used to train the system (ChatGPT). The findings of this study urge that educational organisations should immediately adjust their policies and practises for combating plagiarism and move to include artificial intelligence (AI) into their teaching and assessment processes rather than be wary of it. Teachers need to be trained on how to use ChatGPT effectively and need to be made aware of ChatGPT's capabilities, limitations, and potential effects on their academics.

ChatGPT could be used to accomplish various tasks in our presented model. It could be used to provide personalized learning experiences by analyzing student learning and answering patterns and providing a detailed analysis of the areas in which they lack and excel. On the other hand, it can be used to produce the incorrect options of the MCQ-based test questions which could be close to or far off from the correct answer depending on the degree of variability provided by the questions paper generator. Furthermore, it can be used to change the wording of the questions and choices presented to the students so that if in case a question is asked to two or more students, they have a paraphrased version of the question and not identical questions. Among many applications of ChatGPT, these seem to be the most relevant and efficient use of the chatbot.

In conclusion, the Progressive Model have the potential to provide numerous benefits, it's important to acknowledge and address their limitations to ensure that the exams are fair and equitable for all individuals.

Acknowledgements The authors would like to say thank you to the anonymous reviewers and respected editors for taking valuable time to go through the manuscript.

The Introduction and Discussion sections have been reworded using CHATGPT.

The authors would like to express a sincere thanks to Cerebranium and it's founder Omkar Pimple for pioneering initial version of Progressive Model (https://cerebranium.com). We also would like to thank Rishabh Singh who played pivotal role in structuring initial version of progressive model.

Funding Authors of this paper confirm that there is no funding received for this research work.

Data availability Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors of this research study declare that there is NO conflict of interest.

References

Al-Maqbali, A. H., & Hussain, R. M. R. (2022). The impact of online assessment challenges on assessment principles during COVID-19 in Oman. *Journal of University Teaching and Learning Practice,19*(2), 73–92. https://doi.org/10.53761/1.19.2.6

Chakraborty, P., Mittal, P., Gupta, M. S., Yadav, S., & Arora, A. (2021). Opinion of students on online education during the COVID-19 pandemic. *Human Behavior and Emerging Technologies,3*(3), 357–365. https://doi.org/10.1002/hbe2.240

Chen, S., & Zhang, J. (2008). Ability Assessment based on CAT in Adaptive Learning System. *2008 International Workshop on Education Technology and Training and 2008 International Workshop on Geoscience and Remote Sensing, ETT and GRS 2008,1*, 88–91. https://doi.org/10.1109/ETTandGRS.2008.151

Comas-Lopez, M., Molins-Ruano, P., Atrio, S., & Sacha, G. M. (2018). Sequential adaptive tests to improve the assimilation of contents during learning. *International Symposium on Computers in Education (SIIE),2018*, 1–5. https://doi.org/10.1109/SIIE.2018.8586732

Curriculum Redesign for Cloud Computing to Enhance Social Justice and Intercultural Development in Higher Education. In Handbook of Research on Fostering Social Justice Through Intercultural and Multilingual Communication. edited by Meletiadou, Eleni, 62–80. Hershey, PA: IGI Global, 2023. https://doi.org/10.4018/978-1-6684-5083-3.ch004

Das, B., Majumder, M., Phadikar, S., & Sekh, A. A. (2021). Automatic question generation and answer assessment: A survey. *Research and Practice in Technology Enhanced Learning,16*(1), 5. https://doi.org/10.1186/s41039-021-00151-1

Dwivedi, P., Shankar, R. T., Meghana, B., Sushaini, H., Sudeep, B. R., & Pooja, M. R. (2019). Designing an adaptive question Bank and question paper generation management system. In *International Conference on Artificial Intelligence and Data Engineering* (pp. 965-973). Singapore: Springer Nature Singapore.

Elzainy, A., El Sadik, A., & Al Abdulmonem, W. (2020). Experience of e-learning and online assessment during the COVID-19 pandemic at the College of Medicine, Qassim University. *Journal of Taibah University Medical Sciences,15*(6), 456–462. https://doi.org/10.1016/j.jtumed.2020.09.005

Facilitating an Online and Sustainable Learning Environment for Cloud Computing Using an Action Research Methodology. In Handbook of Research on Implications of Sustainable Development in Higher Education. edited by Meletiadou, Eleni, 43–70. Hershey: IGI Global, 2023

Fung, C. Y., Su, S. I., Perry, E. J., & Garcia, M. B. (2022). Development of a socioeconomic inclusive assessment framework for online learning in higher education. In *Socioeconomic Inclusion During an Era of Online Education* (pp. 23–46). IGI Global. https://doi.org/10.4018/978-1-6684-4364-4.ch002

Gill, S. S., Xu, M., Patros, P., Wu, H., Kaur, R., Kaur, K., Fuller, S., Singh, M., Arora, P., Parlikad, A. K., Stankovski, V., Abraham, A., Ghosh, S. K., Lutfiyya, H., Kanhere, S. S., Bahsoon, R., Rana, O., Dustdar, S., Sakellariou, R., … Buyya, R. (2023). Transformative effects of ChatGPT on modern education: Emerging Era of AI Chatbots. *Internet of Things and Cyber-Physical Systems*. https://doi.org/10.1016/j.iotcps.2023.06.002

Goss, H. (2022). Student Learning Outcomes Assessment in Higher Education and in Academic Libraries: A Review of the Literature. *The Journal of Academic Librarianship,48*(2), 102485. https://doi.org/10.1016/j.acalib.2021.102485

Goldenson, R. P., Avery, L. L., Gill, R. R., & Durfee, S. M. (2022). The virtual homeroom: utility and benefits of small group online learning in the COVID-19 era. *Current Problems in Diagnostic Radiology, 51*(2), 152–154.

Guangul, F. M., Suhail, A. H., Khalit, M. I., & Khidhir, B. A. (2020). Challenges of remote assessment in higher education in the context of COVID-19: A case study of Middle East College. *Educational Assessment, Evaluation and Accountability,32*(4), 519–535. https://doi.org/10.1007/s11092-020-09340-w

Hardigan, P., Lai, L., Arneson, D., & Robeson, A. (2001). Significance of academic merit, test scores, interviews and the admissions process: A case study. *American Journal of Pharmaceutical Education*, *65(1), 40-43.*

Holden, O. L., Norris, M. E., & Kuhlmeier, V. A. (2021). Academic Integrity in Online Assessment: A Research Review. In *Frontiers in Education* (Vol. 6). Frontiers Media S.A. https://doi.org/10.3389/feduc.2021.639814

How Covid-19 Changed Computer Science Education, ITNOW, Volume 64, Issue 2, Summer 2022, Pages 60–61, https://doi.org/10.1093/itnow/bwac062

Huang, Y. M., Lin, Y. T., & Cheng, S. C. (2009). An adaptive testing system for supporting versatile educational assessment. *Computers and Education,52*(1), 53–67. https://doi.org/10.1016/j.compedu.2008.06.007

Jaiswal, A., Varma, A., & Pereira, V. E. (2022). An Empirical Investigation of Academic Integrity in Assessments- Evidence from an Emerging Country. *Academy of Management Proceedings*, *2022*(1). https://doi.org/10.5465/AMBPP.2022.11364abstract

Jopp, R., & Cohen, J. (2022). Choose your own assessment–assessment choice for students in online higher education. *Teaching in Higher Education,27*(6), 738–755. https://doi.org/10.1080/13562517.2020.1742680

Khodeir, N., Wanas, N., Darwish, N., & Hegazy, N. (2014). Bayesian based adaptive question generation technique. *Journal of Electrical Systems and Information Technology,1*(1), 10–16. https://doi.org/10.1016/j.jesit.2014.03.007

Martin, F., Ritzhaupt, A., Kumar, S., & Budhrani, K. (2019). Award-winning faculty online teaching practices: Course design, assessment and evaluation, and facilitation. *Internet and Higher Education,42*, 34–43. https://doi.org/10.1016/j.iheduc.2019.04.001

Menon, S., & Suresh, M. (2022). Development of assessment framework for environmental sustainability in higher education institutions. *International Journal of Sustainability in Higher Education,23*(7), 1445–1468. https://doi.org/10.1108/IJSHE-07-2021-0310

Mishra, A., & Jain, S. K. (2016). A survey on question answering systems with classification. *Journal of King Saud University-Computer and Information Sciences,28*(3), 345–361.

Nieminen, J. H. (2022). Assessment for Inclusion: Rethinking inclusive assessment in higher education. *Teaching in Higher Education*. https://doi.org/10.1080/13562517.2021.2021395

Pado, U. (2017). Question Difficulty – How to Estimate Without Norming, How to Use for Automated Grading. *Proceedings of the 12th Workshop on Innovative Use of NLP for Building Educational Applications*, 1–10. https://doi.org/10.18653/v1/W17-5001

Patil, P. M., Bhavsar, R. P., Pawar, B. V. (2022). Automatic question generation for subordinate conjunctions of Marathi. *2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS)*, *1*, 169–173.

Perez, E. V., Santos, L. M. R., Perez, M. J. V., de Castro Fernandez, J. P., & Martin, R. G. (2012). Automatic classification of question difficulty level: Teachers' estimation vs. students' perception. *Proceedings - Frontiers in Education Conference, FIE*. https://doi.org/10.1109/FIE.2012.6462398

Purohit, V. K., Kumar, A., Jabeen, A., Srivastava, S., Goudar, R. H., Shivanagowda, & Rao, S. (2012). Design of adaptive question bank development and management system. *2012 2nd IEEE International Conference on Parallel, Distributed and Grid Computing*, 256–261. https://doi.org/10.1109/PDGC.2012.6449828

Ragasudha, R., & Saravanan, M. (2022). Secure Automatic Question Paper Generation with the Subjective Answer Evaluation System. *International Conference on Smart Technologies and Systems for next Generation Computing (ICSTSN),2022*, 1–5.

Ranjan, A., Gujar, S., & Ramani, V. (2022). Dynamic matching in campus placements: The benefits and affordability of the dream option. *IIMB Management Review,34*(3), 262–277. https://doi.org/10.1016/j.iimb.2022.08.001

Shah, P., Faquih, U., Devkar, R., & Shahare, Y. (2022). An Intelligent Question Paper Generator using Randomized Algorithm. *International Journal of Engineering Research & Technology (IJERT), 11*(04). www.ijert.org

Singh, R., Timbadia, D., Kapoor, V., Reddy, R., Churi, P., & Pimple, O. (2021). Question paper generation through progressive model and difficulty calculation on the Promexa Mobile Application. *Education and Information Technologies,26*(4), 4151–4179. https://doi.org/10.1007/s10639-021-10461-y

Singhal, R., Goyal, S., & Henz, M. (2016). User-Defined Difficulty Levels for Automated Question Generation; User-Defined Difficulty Levels for Automated Question Generation. *2016 IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI)*. https://doi.org/10.1109/ICTAI.2016.126

Srivastava, M.,& Goodman, Noah. (2021) Question generation for adaptive education. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing* (vol 2: Short papers), pp 692–701. Online. Association for Computational Linguistics. https://doi.org/10.1007/s10639-023-12045-4

Starkweather, A., Sargent, L., Nye, C., Albrecht, T., Cloutier, R., & Foster, A. (2017). Progressive Assessment and Competency Evaluation Framework for Integrating Simulation in Nurse Practitioner Education. *Journal for Nurse Practitioners,13*(7), e301–e310. https://doi.org/10.1016/j.nurpra.2017.04.012

Sullivan, T. (2001). Locating question difficulty through explorations in question space. *Proceedings of the ACM International Conference on Digital Libraries*, 251–252. https://doi.org/10.1145/379437.379669

Sun, C., Liu, B., Jin, Z., & Wang, C. (2009). An adaptive algorithm for generating question paper with multi-factors. *Proceedings of 2009 4th International Conference on Computer Science and Education, ICCSE 2009*, 1048–1051. https://doi.org/10.1109/ICCSE.2009.5228540

Sunil, A. (2020). Smart Paper Generator. *International Journal of Progressive Research in Science and Engineering,1*(9), 17–19.

Tsunemoto, A., Trofimovich, P., Blanchet, J., Bertrand, J., & Kennedy, S. (2022). Effects of benchmarking and peer-assessment on French learners' self-assessments of accentedness, comprehensibility, and fluency. *Foreign Language Annals, 55*(1), 135–154. https://doi.org/10.1111/flan.12571

Vozár, O., & Bieliková, M. (2008). Adaptive test question selection for web-based educational system. *Proceedings - 3rd International Workshop on Semantic Media Adaptation and Personalization. SMAP,2008*, 164–169. https://doi.org/10.1109/SMAP.2008.15

Wang, Q., Liu, J., Wang, B., & Guo, L. (2014). A regularized competition model for question difficulty estimation in community question answering services. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1115–1126.

Wendler, C., & Bridgeman, B. (2014). *The Research Foundation for the GRE revised General Test: A Compendium of Studies*. www.ets.org/gre/compendium

Yudono, M. A. S., Faris, R. M., Wibowo, A. de, Sidik, M., Sembiring, F., & Aji, S. F. (2022). Fuzzy Decision Support System for ABC University Student Admission Selection. *Proceedings of the International Conference on Economics, Management and Accounting (ICEMAC 2021)*, 230–237. https://doi.org/10.2991/aebmr.k.220204.024

Authors and Affiliations

Aditya Shah[1] · Ajay Devmane[1] · Mehul Ranka[1] · Prathamesh Churi[1]

✉ Ajay Devmane
ajay.devmane@gmail.com

Aditya Shah
adityashah.001.as@gmail.com

Mehul Ranka
mehul.ranka05@gmail.com

Prathamesh Churi
Prathamesh.churi@ieee.org; Prathamesh.churi@gmail.com

[1] B. Tech- Computer Science and Business Systems (CSBS), Computer Engineering Department, Mukesh Patel School of Technology Management & Engineering, NMIMS University, Mumbai, India