Mapping Learning Outcomes to Academic Success: A Novel Dataset for Evaluating Students' Academic Performance

Sapna Rai, Dr. Sunil Sikka, Dr. Anil Kumar

Research Scholar, Amity University Haryana, Gurugram, Haryana, India, sapnasucess@gmail.com

Professor, CSE, Amity University Haryana, Gurugram, Haryana, India, ssikka@ggn.amity.edu

Professor, CSE, Poornima Institute of Engineering and Technology, Jaipur, Rajasthan, India, anilkumar@poornima.org

Abstract

Academic performance evaluation is a critical component of education, shaping student progress, curriculum development and instructional strategies. Traditional assessment methods, primarily reliant on grades and examination scores, often fail to provide a comprehensive measure of students' conceptual understanding and skill proficiency. This study advocates for a paradigm shift towards learning outcome-based assessments, particularly within Computer Science education for Class 11 and 12 students under the Central Board of Secondary Education (CBSE). To facilitate this transition, a novel dataset has been created and systematically mapped to four consolidated key learning outcomes, Computational Thinking and Data Structures, Computer Systems and Cyber Ethics, Programming Concepts and Python Applications and Database Management with SQL Integration. The findings underscore the effectiveness of outcome-based assessments in aligning evaluation methodologies with competency-driven education.

Keywords. Learning outcomes, Computer Science Education, Dataset, Evaluation of academic performance, Students' performance.

1. INTRODUCTION

Academic performance evaluation is a fundamental aspect of education, influencing decisions on student progress, instructional quality, and curriculum planning. Traditional assessment methods, including exam scores, grades, and class participation, have long been used to measure academic success. However, these metrics offer numerical insights but fail to capture a student's mastery of concepts, cognitive skills, or problem-solving abilities. This limitation creates a gap in understanding how well students meet educational objectives.

A shift from traditional assessment to learning outcome-based evaluation is essential in modern education. Traditional performance indicators, such as demographic factors, attendance, socio-economic status, and behavioral data from e-learning platforms, are valuable for predicting academic success. However, they do not fully capture the alignment between assessments and curriculum objectives. Hussain et al. [1] highlight that while attendance and demographic factors can predict performance, they do not accurately assess skill acquisition or conceptual mastery.

Learning outcomes define the specific, measurable competencies students should acquire after completing a course or educational stage. The CBSE has integrated learning outcomes into its curriculum to promote skill-based and competency-driven education. These outcomes provide a structured framework for both teaching and assessment. Hayat et al. [2] emphasize that outcome-based evaluation offers a more detailed understanding of student progress, particularly in higher education, where emotional intelligence and self-efficacy play a crucial role. Similarly, Turjman et al. [3] argue that an outcome-oriented approach allows educators to focus on specific skill sets rather than generalized grading metrics.

Building on this foundation, this study introduces a dataset for evaluating academic performance using learning outcomes in CBSE Class 11 and 12 Computer Science. The paper is structured as follows: Section 2 presents a literature review summarizing relevant research. Section 3 discusses previous datasets and describes the proposed dataset. Finally, Section 4 concludes the study and suggests future research directions for enhancing outcome-based assessments through machine learning

2. LITERATURE REVIEW

This literature review focused on increasing application of diverse datasets and analytical methodologies to evaluate academic performance. Studies such as Hussain et al. [1] have integrated these datasets with academic records to forecast end-semester results, demonstrating how external factors influence student outcomes. Similarly, Kassarnig et al. [4] analyzed datasets incorporating personal and social environment attributes, highlighting the impact of family background, income levels etc. on academic achievement. Internal academic data, including attendance records and assessment scores, also served as a fundamental component in academic performance analysis. Zafari et al. [5] utilized academic scores to classify students into performance categories such as very well, good, medium, and poor. Expanding on this approach, Alsariera et al. [6] and Fahd et al. [7] integrated demographic, internal assessment, and personal attributes for holistic evaluation of student performance.

Chung et al. [8] and Waheed et al. [9] utilized log data to classify students into dropout and non-dropout and at-risk categories. Esparza et al. [10] leveraged student comments to categorize feedback for teachers into positive, neutral, or negative, indirectly evaluating the impact of the teaching environment on student outcomes. Studies have also linked grit and self-efficacy have been identified as significant factors influencing student performance, especially in higher education settings [11]. Brown et al. [12] conducted a five-year longitudinal study examining the relationship between student engagement, learning behaviors, and academic achievement. These comprehensive datasets and analytical approaches highlight the transformative potential of data-driven methodologies in evaluating academic success and optimizing learning experiences.

3. PROPOSED DATASET

The literature review highlights the increasing use of diverse datasets and analytical methodologies to assess academic performance. The study reviewed key attributes into distinct areas, each playing a significant role in understanding student success. The datasets primarily encompass attributes related to student grades, demographic data, attendance records, and standardized test scores [1,5,6,8,13,14,16,17,19,20,21,22], Student behavioral and Engagement data-based studies [4,12,15,26,27,28], Virtual Learning Environments (VLE) [9,10,18,23,24,25], Sentiment-Based Feedback and Psychological Attributes [2,3,11,17]. While existing datasets provide valuable insights, they did not focus on competency-based learning objectives. Comparison between grade-based and outcome-based assessment is presented in Table 1.

Feature	Grade-Based Assessment	Outcome-Based Assessment
Focus	Exam scores, grades, and overall performance	Mastery of skills and competencies
Evaluation Criteria	Based on correct/incorrect answers in standardized tests	Measures conceptual understanding and practical skills
Student Learning Insight	Provides limited insight into skill mastery	Offers a detailed view of strengths and weaknesses
Flexibility	Fixed exams with rigid scoring	Adapts to students' learning progress
Practical Relevance	May not reflect real-world application of knowledge	Aligns with job-market and practical applications

Table 1: Comparison: Grade-Based vs. Outcome-Based Assessments

Our dataset marks a significant step in outcome-based educational evaluation, providing data-driven insights into student performance. It serves as a blueprint for integrating learning outcome assessments into education. Ensuring ethical compliance, it uses publicly available academic questions from CBSE textbooks, sample papers, and online sources without collecting private student data. The dataset is fully anonymized, containing only academic content without personal identifiers. The number of questions extracted based on learning outcomes is shown in Table 2.

LO-1	7035
LO-2	7540
LO-3	8755
LO-4	4635

Table 2: Number of questions for each class

These questions are manually mapped as per the Learning Framework of classes 11 and 12, Computer science by CBSE which are listed in the Figure 1 and Figure 2 respectively and the detailed learning outcomes are consolidated into four key learning outcomes, as shown in Figure 3.

UNIT	LEARNING OUTCOMES
Computer Systems and Organization	1. Develop a basic understanding of computer systems- architecture, and operating systems. 2. Develop a basic understanding of various encoding schemes and number systems along with their applications in digital technology. 3. Develop a basic understanding of new technologies Information technology and how are they affecting the modern lifestyle
Computational Thinking and Programming	4. Appreciate the notion of the algorithm including flowcharts and pseudocode. 5. Explain and use operators, variables and data types including higher-order data structures such as lists, strings, tuples, and dictionaries 6. Develop basic principles of computational thinking
Society, Law and Ethics	7. Explain cyber ethics, cyber safety and cybercrime 8. Understand the value of technology in societies along with consideration of gender and disability issues.

Figure 1: Learning Outcomes by CBSE-NCF for class 11

UNIT	LEARNING OUTCOMES
Computational Thinking and Programming	1 Apply the concept of function 2 Implement search and sort techniques. 3 Create and use Python Libraries 4. Explain and use the concept of exception as well as file handling. 5 Explain the concept of efficiency in algorithms and computing in general. 6. Use basic data structure: Stacks and Queues
Computer Networks	7. Explain computer networks and how it works. 8. Explain the security concerns in networked computers and methods to mitigate them.
Database Management	9. Use Database concepts, SQL to store, retrieve and manipulate data.

Figure 2: Learning Outcomes by CBSE-NCF for class 12

ATTRIBUTES	DESCRIPTION
LO-1	Develop basic computational thinking and appreciate the notion of the algorithm including flowcharts and pseudocode and explain and use operators, variables and data types including higher-order data structures such as lists, stack, queues, strings, tuples, and dictionaries.
LO-2	Develop a basic understanding of computer systems along with architecture, operating systems, encoding schemes, number systems along with their applications in digital technology and how are they affecting the modern lifestyle. Also explain cyber ethics, cyber safety, and cybercrime and understand the value of technology in societies along with consideration of gender and disability issues.
LO-3	Develop basic computational thinking and appreciate the notion of the algorithm including flowcharts and pseudocode and apply the concept of function, Create and use Python libraries and use the concept of exception as well as file handling.
LO-4	Develop basic computational thinking and use Database concepts, SQL to store, retrieve and manipulate data along with connectivity between Python and SQL.

Figure 3: Final combined Learing Outcomes

The dataset comprises 23,464 academic questions with five attributes are manually mapped to the four consolidated learning outcomes. Sample of the dataset is shown in the Figure 4.

1	QUESTION CONTENT	LO 1	LO 2	LO3	LO 4
2	Write the output of-d={0:'zero',1:'one',2:'two',3:'three'}for i in d: print(i, enc	1	0	0	0
3	Find the output of the following Python program: def makenew(mystr): news	1	0	1	0
4	Find the output of the following program def calcresult () : i = 9 while i> 1 :	1	0	1	0
5	Observe the following Python code and find out , which out of the given optic	1	0	1	0
6	Discuss the strategies employed by python for memory allocation?	1	1	0	0
7	Answer the questions (i) and (ii) after going through the following class definit	1	0	1	0
8	Define a class Train in PYTHON with following description: 4 Data Members s	1	0	1	0
9	class ParentClass(objects): def__init__(self) self,x = 1 self.y = 1 def print(self):	1	0	1	0
10	Write a user defined function findname(name) where name is an argument in	1	0	1	0
11	Explain try..except...else ... with the help of user defined function def divide(x	1	0	1	0
12	Write a user defined function arrangelements(X), that accepts a list X of integ	1	0	1	0
13	Consider the following class definition :- 3 class book (): bk = [] def _ init _ (:	1	0	1	0
14	Write a python function generatefibo(n) where n is the limit using a generato	1	0	1	0
15	Evaluate the following postfix using stack & show the content of the stack aft	1	0	0	0
16	Consider the following code : 1 f = open ("mytry", "w+") f.write ("123456789:	1	0	1	0
17	Write a user defined function in Python that displays the number of lines star	1	0	1	0
18	Consider a binary file Employee.dat containing details such as empno:ename	1	0	1	0
19	Differentiate between cardinality and degree of a table with the help of an ex	0	0	0	1
20	Consider the following tables FACULTY and COURSES. Write SQL commands	0	0	0	1
21	State and Verify Absorption law algebraically.	0	1	0	0
22	Draw a logic circuit for the following Boolean expression: A.B+C.D'	0	1	0	0
23	Write the SOP form of a Boolean function F, which is represented in a truth ta	0	1	0	0
24	Obtain a simplified from for a Boolean expression:	0	1	0	0
25	Write any 1 advantage and 1 disadvantage of Bus topology	0	1	0	0
26	i) Suggest a cable layout of connections between the Departments and specif	0	1	0	0

Figure 4: Sample of the Dataset

This dataset will be used to develop a machine learning model for the automatic prediction of learning outcomes based on the question's context.

5. CONCLUSION

This study introduces a novel dataset specifically designed for evaluating academic performance by mapping assessment questions to CBSE-defined learning outcomes for Classes 11 and 12 in Computer Science. Unlike previous research studies, which predominantly relied on traditional academic data or generalized behavioral metrics, this dataset uniquely bridges the gap between granular academic content and predefined educational objectives. The findings demonstrate the power of using a learning outcome-oriented framework to identify key areas where students excel or struggle. The dataset's multi-label and multi-class nature further allows for sophisticated analytical approaches, offering unparalleled insights into how specific learning outcomes influence academic success. This dataset represents a foundational step toward outcome-based education and data-driven academic evaluation. To our knowledge, no prior work has used such a dataset for mapping questions directly to learning outcomes in this context, making it a unique contribution to the field. Future work can expand upon this by integrating this dataset with machine learning algorithms for automated question mapping.

5. REFERENCES

[1] Hussain, Dahan, By-Law, Educational data mining, Indonesian J. Electr. Eng. Comput. Sci., vol. 9, 2018.

[2] Hayat, Saleem, Afzal, Grit self-efficacy achievement, J. Appl. Res. Higher Educ., vol. 12, 2020.

[3] Habeeb, Fadare, Al-Turjman, Artificial Neural Networks, Smart Infrastructures IoT Era, Springer, 2025.

[4] Kassarnig, Mones, Bjerre-Nielsen, Academic performance patterns, EPJ Data Sci., vol. 7, 2018.

[5] Zafari, Sadeghi-Niaraki, Choi, Evaluation high school, Appl. Sci., vol. 11, 2021.

[6] Alsariera, Baashar, Alkawsi, Assessment Machine Learning, Comput. Intell. Neurosci., vol. 2022, 2022.

[7] Fahd, Venkatraman, Miah, Machine Learning Education, Educ. Inf. Technol., pp. 1-33, 2022.

[8] Chung, Lee, Dropout warning systems, Children Youth Serv. Rev., vol. 96, 2019.

[9] Waheed, Hassan, Aljohani, Predicting Academic Performance, Comput. Hum. Behav., vol. 104, 2019.

[10] Gutiérrez, Canul-Reich, Zezzatti, Mining students' comments, Int. J. Comb. Optim. Prob. Inform., vol. 9, 2018.

[11] Faust, Rosendale, Grit self-efficacy predictors, Rev. Educ., vol. 10, 2022.

[12] Brown, Johnson, Williams, Student engagement performance, IEEE Trans. Learn. Technol., vol. 19, 2025.

[13] Baashar et al., Postgraduate performance AI, Alex. Eng. J., vol. 61, 2022.

[14] Alnassar, Blackwell, Homayounvala, Machine Learning Virtual, Int. Conf. Intell. Eng. Manag., 2021.

[15] Madigan, Cross, Smolkowski, Positive behavioral interventions, Educ. Res. Eval., vol. 22, 2016.

[16] Fernandes, Holanda, Victorino, Educational data mining, J. Bus. Res., vol. 94, 2019.

[17] Aziz, Student performance evaluation, Coll. Basic Educ. Res. J., vol. 16, 2020.

[18] Lakkaraju, Aguiar, Shan, Machine Learning Framework, ACM SIGKDD Int. Conf., 2015.
[19] Shanmugarajeshwari, Lawrance, Performance classification techniques, Int. Conf. Comput. Technol., 2016.
[20] Al Mayahi, Al-Bahri, Predicting student success, Int. Congr. Ultra-Mod. Telecomm., 2020.
[21] Motlagh, Amrai, Yazdani, Self-efficacy academic achievement, Procedia-Soc. Behav. Sci., vol. 15, 2011.
[22] Rivas, Fraile, Chamoso, Machine learning techniques, Int. Workshop Learn. Technol., 2019.
[23] Orji, Vassileva, Student engagement performance, Int. Conf. Info. Visualisation, 2020.
[24] Burgos, Campanario, Peña, Tutoring action plan, Comput. Electr. Eng., vol. 66, 2018.
[25] Semerci, Goularas, Flow state learning, J. Educ. Comput. Res., vol. 59, 2021.
[26] Honicke, Broadbent, Academic self-efficacy review, Educ. Res. Rev., vol. 17, 2016.
[27] Adams, Blair, Time management behaviors, Sage Open, vol. 9, 2019.
[28] Ardura, Galán, Learning approaches self-efficacy, Int. J. Sci. Educ., vol. 41, 2019.