Unveiling The Virtual Classroom: An In-Depth Analysis of The Online Education System

1. INTRODUCTION

1.1 Overview

Online learning platforms proliferated during the COVID-19 pandemic. These online learning platforms became the only source for teaching and learning. Every individual and educational institute has completely depended on these platforms. Usually, in regular classroom teaching, students' behavior and learning progress are monitored. But, in the case of online learning platforms, especially when there is a huge number of participants, require the help of artificial intelligence (AI), data analytics, etc. to monitor the learners' behavior and progress.

1.2 Purpose

Developing a comprehensive understanding of the online education ecosystem, this project pursues to address various aspects of virtual classrooms. By precisely evaluating its strengths, weaknesses, opportunities, and challenges, this project aims to offer key insights to educational institutions, policymakers, and online platforms. The objective is to enhance the efficiency and inclusivity of online education.

To achieve this objective, several business requirements need to be considered:

- Research and data collection
- Active participation of stakeholders
- Technology infrastructure
- Expertise
- Data privacy
- Effective Analytics
- Documentation
- Adaptability and scalability
- Extensibility

2. LITERATURE SURVEY

To better engage students and to connect with their instructors artificial intelligence provides good support [1]. Similarly, information and communication technologies have prompted a different learning style on the other side of the classroom [2]-[3]. The advancement of computer technology and education technology (EdTech) products showcased significant effects on school learning [4]. The examination of several education policy issues and complexities for the delivery of EdTech plays a key role [5]. In addition to the specific EdTech products, YouTube also has a significant impact on the learning process in terms of self-directed learning. It provides formal and informal education, but we do not the reliability of the content [6]-[7]. The courses available in online learning platforms are referred to as massive open online courses (MOOCs). There is a dire need for proper learning analytics of MOOCs [8]. To authenticate and proctor the online learning students, a biometric-based technology is implemented [9]. During the pandemic, AI and machine learning have attempted to solve human problems and also supported teachers to assess their students [10]. The importance of e-learning has been emphasized by conducting a comprehensive bibliometric analysis [11]. Discussed the barriers such as connectivity, device access, etc. while integrating technology with special education [12].

The abovementioned works discussed the importance of online learning during a disease outbreak, the implementation of EdTech products, the importance of learning analytics, and the barriers to integrating technology into special education.

The factors that influence the student while adopting the technology were discussed in [13]. Further, the effect on the student's long-term learning by leveraging computer technology in the classroom was discussed in [14]. The online learning behavior in terms of experience, engagement, and the pattern of K12 students in China during COVID-19 were discussed in [15]. A report was prepared based on the experiences of computer science students with emergency remote teaching [16]. To monitor the students' online learning behavior, an enhanced extended nearest neighbor technique was implemented [17]. The identification of actual online learning behavior of college students based on head gesture recognition was discussed in [18]. The learning analytics was implemented to observe the change in the learning behavior of students in special education [19]. Furthermore, the student behavior analysis was conducted based on self-organizing maps neural networks through the clustering of user settings [20]. The identification of the multidimensional engagement of students in the online learning platform using multi-channel data was discussed [21]. A study was conducted to observe the variances when the students are engaged in traditional learning systems and learning management systems [22]. In addition, early prediction of the students at risk in the e-learning platforms using Hidden Markov Models was discussed in [23]. Besides, the early prediction of student behavior and the support of pedagogical intervention were discussed [24]-[25]. A model was developed to analyze the student learning experiences [26]. As well, the key concepts of building an intelligent education system were discussed in [27].

2.1 Existing problem

From these literature works it is observed that there is a dire need to develop an intelligent virtual learning platform that engages students and analyses their learning behavior during the learning. Further, they emphasized the new pedagogical intervention and practices.

2.2 Proposed solution

A web-based and unified business intelligence application i.e., IBM Cognos analytics is used to extract insights such as learning behavior, progress, etc. in online education platforms.

3. THEORETICAL ANALYSIS

3.1 Block diagram

To achieve the project deliverables, the following steps shown in Fig. 1 need to be performed.

- Create "smartinternz" account as a student with the institution email ID
- Create "IBM Cloud" account as a student with the institution email ID to access the free IBM Cloud account for one year
- Attend the Bootcamp
- Take the courses and earn badges and public URLs
- Share the course badges on social networks
- Start the project implementation using IBM Cognos Analytics

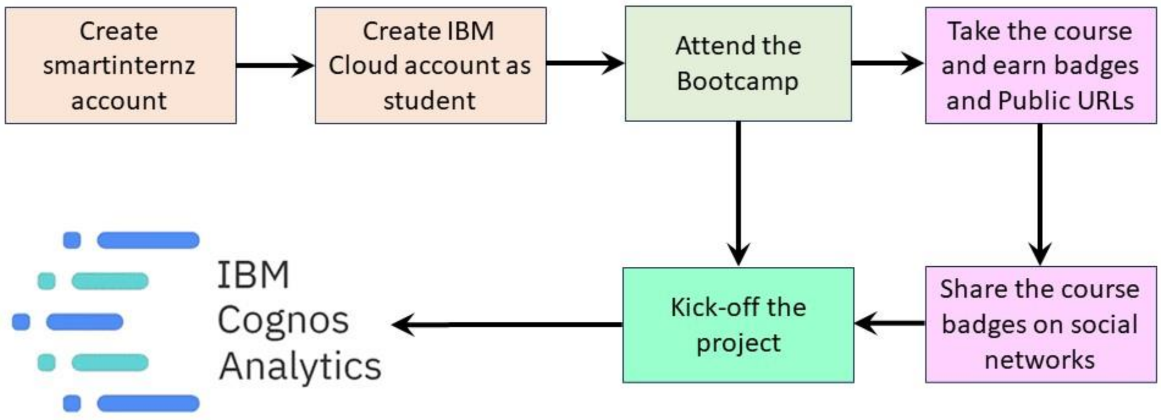


Fig. 1 Block diagram of the project

3.2 Hardware / Software Designing

Hardware Requirements

- **RAM:** 4GB RAM minimum
- **Processor:** Intel i3 minimum
- **Hard disk storage:** 50GB minimum
- **Networking:** High-speed Internet Access

Software Requirements

- IBM Cognos Analytics
- Web Page Templates
- Python (latest version)
- Visual Studio Code
- Zotero for reference management

4. EXPERIMENTAL INVESTIGATIONS

Name	Description
Age(Years) by Your level of satisfaction in Online Education	▪ The total number of results for Age(Years), across all Your level of satisfaction in Online Education, is over a thousand. ▪ Average is the most frequently occurring category of Your level of satisfaction in Online Education with a count of 541 items with Age(Years) values (52.4 % of the total).
Internet facility in your locality by Your level of satisfaction in Online Education	▪ Across all values of Your level of satisfaction in Online Education, the sum of Internet facility in your locality is over 3500. ▪ Internet facility in your locality ranges from 823, when Your level of satisfaction in Online Education is Bad, to nearly two thousand, when Your level of satisfaction in Online Education is Average. ▪ Internet facility in your locality is unusually high when Your level of satisfaction in Online Education is Average.
Performance in online by Level of Education	▪ Over all values of Level of Education, the sum of Performance in online is nearly seven thousand. ▪ Performance in online ranges from 562, when Level of Education is School, to nearly 5500, when Level of Education is Under Graduate.

	▪ Performance in online is unusually high when Level of Education is Under Graduate.
Time spent on social media (Hours) by Device type used to attend classes	▪ Across all device type used to attend classes, the sum of Time spent on social media (Hours) is over 2500. ▪ Time spent on social media (Hours) ranges from 71, when Device type used to attend classes is Desktop, to over 1500, when Device type used to attend classes is Laptop. ▪ Time spent on social media (Hours) is unusually high when Device type used to attend classes is Laptop.
Engaged in group studies? colored by Engaged in group studies? sized by Performance in online	▪ Over all values of Engaged in group studies? and Engaged in group studies?, the sum of Performance in online is nearly seven thousand. ▪ The summed values of Performance in online range from nearly three thousand to over four thousand. ▪ Performance in online is unusually high when the combination of Engaged in group studies? and Engaged in group studies? is No and No. ▪ For Performance in online, the most significant value of Engaged in group studies? is No, whose respective Performance in online values add up to over four thousand, or 58.1 % of the total.
Average marks scored before pandemic in traditional classroom sized by Average marks scored before pandemic in traditional classroom	▪ The total number of results for Average marks scored before pandemic in traditional classroom, across all average marks scored before pandemic in traditional classrooms, is over a thousand. ▪ The counts are unusually high when the values of Average marks scored before pandemic in traditional classroom are 81-90 and 71-80. ▪ 81-90 (33.2 %) and 71-80 (30.3 %) are the most frequently occurring categories of Average marks scored before pandemic in traditional classroom with a combined count of 656 items with Average marks scored before pandemic in traditional classroom values (63.5 % of the total).
Age(Years) by Study time (Hours) colored by Age(Years)	▪ The overall number of results for Age(Years) is over a thousand. ▪ 4 is the most frequently occurring category of Study time (Hours) with a count of 213 items with Age(Years) values (20.6 % of the total).
Average marks scored before pandemic in traditional classroom by Level of Education	▪ The total number of results for Average marks scored before pandemic in traditional classroom, across all Level of Education, is over a thousand. ▪ Under Graduate is the most frequently occurring category of Level of Education with a count of 817 items with Average marks scored before pandemic in traditional classroom values (79.1 % of the total).
Performance in online by Device type used to attend classes colored by Gender	▪ Performance in online is unusually high when Device type used to attend classes is Laptop. ▪ Over all device type used to attend classes and genders, the sum of Performance in online is nearly seven thousand. ▪ The summed values of Performance in online range from 53 to over 2500.

	▪ Performance in online is unusually high when the combination of Device type used to attend classes and Gender is Laptop and Male. ▪ For Performance in online, the most significant value of Device type used to attend classes is Laptop, whose respective Performance in online values add up to nearly 4500, or 64.3 % of the total. ▪ For Performance in online, the most significant value of Gender is Male, whose respective Performance in online values add up to over four thousand, or 57.9 % of the total.
Study time (Hours) by Level of Education	▪ The total number of results for Study time (Hours), across all Level of Education, is over a thousand.
Study time (Hours) for Have separate room for studying?	▪ The total number of results for Study time (Hours), across all Have separate room for studying?, is 672.
Clearing doubts with faculties in online mode and Performance in online for Study time (Hours) colored by Time spent on social media (Hours)	▪ Over all values of Study time (Hours), the sum of Performance in online is nearly seven thousand. ▪ Performance in online ranges from 82, when Study time (Hours) is 9, to almost 1500, when Study time (Hours) is 4. ▪ Over all values of Study time (Hours), the sum of Clearing doubts with faculties in online mode is nearly three thousand. ▪ Clearing doubts with faculties in online mode ranges from 39, when Study time (Hours) is 9, to 648, when Study time (Hours) is 4. ▪ Clearing doubts with faculties in online mode is unusually high when Study time (Hours) is 4.
Performance in online by Study time (Hours)	▪ Over all values of Study time (Hours), the sum of Performance in online is nearly seven thousand. ▪ Performance in online ranges from 82, when Study time (Hours) is 9, to almost 1500, when Study time (Hours) is 4. ▪ Performance in online is unusually high when Study time (Hours) is 4.

5. FLOWCHART

The following Fig. 2 showcases the high-level flow of activities in the project. The project starts with the definition of the problem statement. The immediate activity is the collection of project requirements in terms of business, hardware, and software according to the problem statement. Once the requirements are finalized, collect the dataset and integrate it with IBM Cognos on the IBM cloud platform. The preprocessing of the data is necessary to ensure the correctness of the data. Once the data are pre-processed, the descriptive analysis of the data can be done by creating different kinds of visualizations. The pinning of visualizations is required to retain them in the exploration and to access them whenever required. Once all the visualizations are pinned, then need to create a dashboard, story, and report for a better and quicker understanding of the insights of online education. The deliverables are to be verified for correctness and quality in the performance testing. Dashboard, story, and report are to be integrated with the web using Flask to access the project deliverables. The final step in the project is the preparation of project documentation and demonstration video.

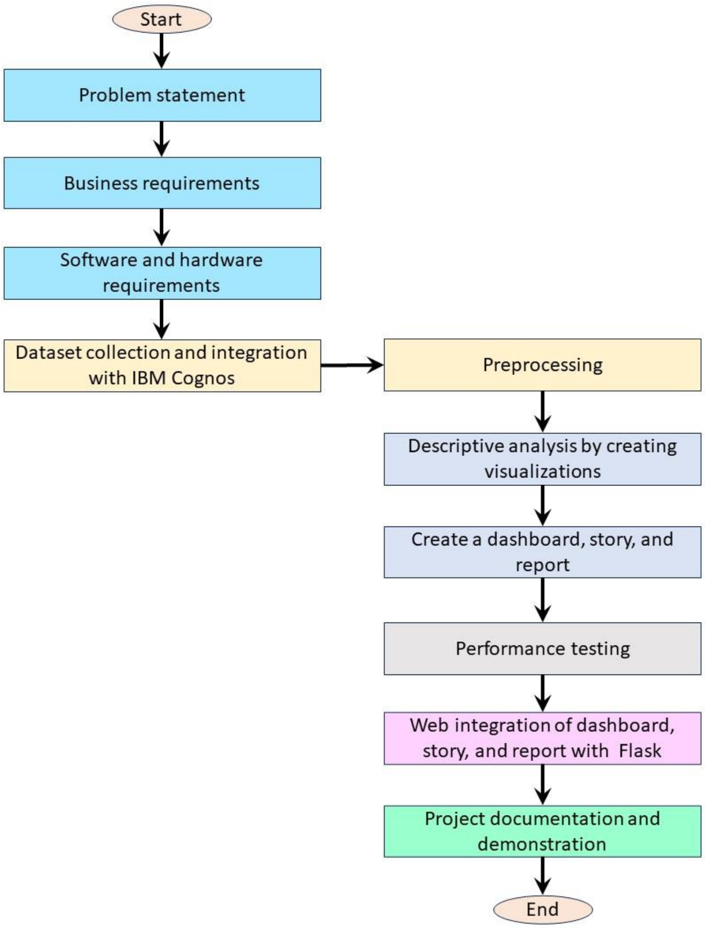


Fig. 2 Flow of activities in the project

6. RESULT

The following are the experimental results in terms of various kinds of visualizations that provide insights into online education.

1. Column Chart: Age(Years) by Your level of satisfaction in Online Education

Fig. 3 Column chart

1. **Bar Chart: Internet facility in your locality by Your level of satisfaction in online Education**

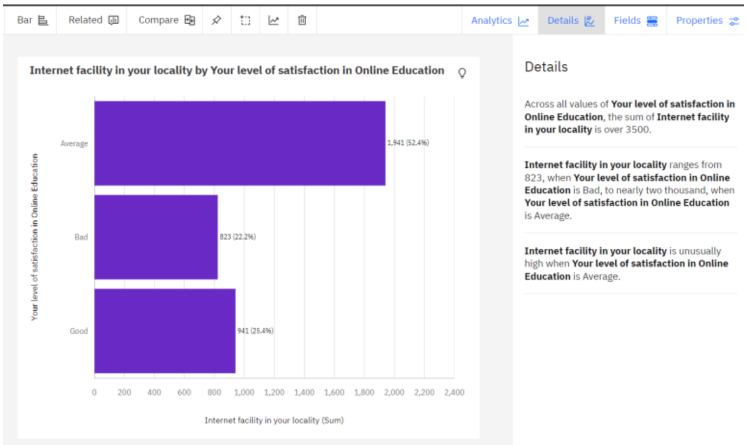


Fig. 4 Bar chart

2. **Bar chart: Performance in online by Level of Education**

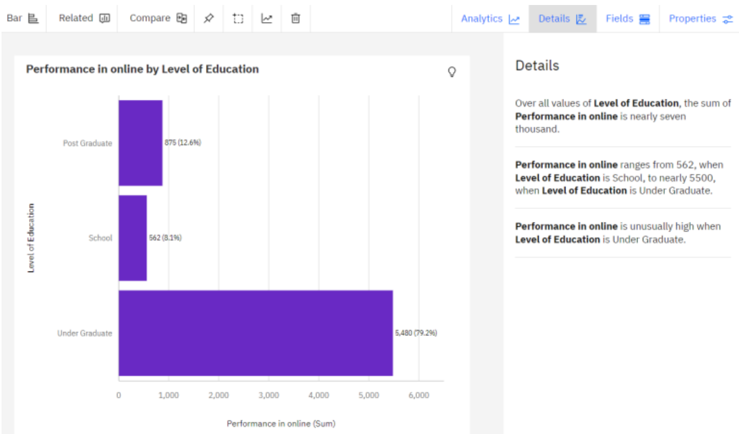


Fig. 5 Bar chart

3. **Pie Chart: Time spent on social media (Hours) by Device type used to attend classes**

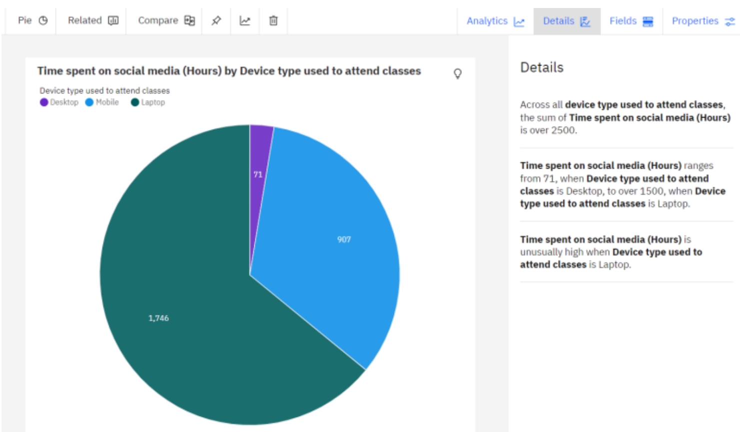


Fig. 6 Pie chart

4. **Packed bubbles: Engaged in group studies? colored by Engaged in group studies? sized by Performance in online**

Fig. 7 Packed bubble chart

5. **Wordcloud: Average marks scored before pandemic in traditional classroom**

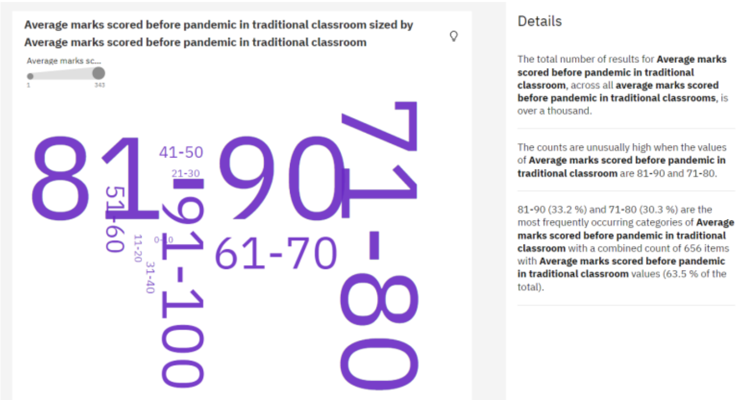
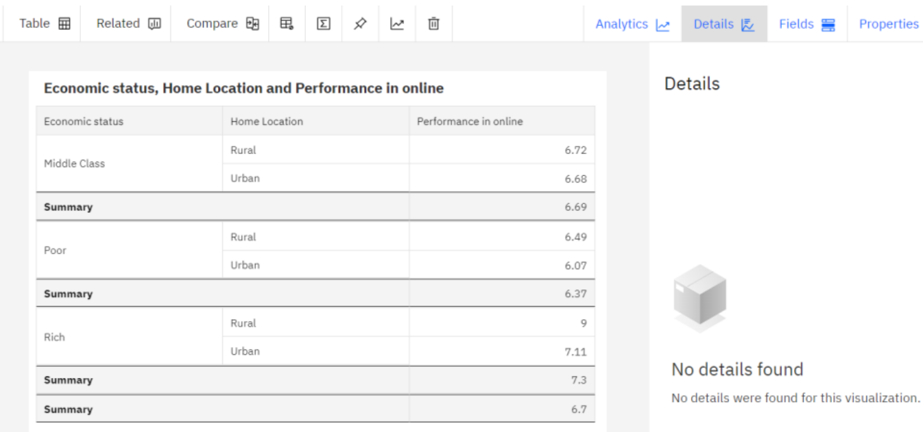


Fig. 8 Wordcloud

6. **Table: Economic status, Home Location and Performance in online**



Economic status, Home Location and Performance in online

Economic status	Home Location	Performance in online
Middle Class	Rural	6.72
	Urban	6.68
Summary		6.69
Poor	Rural	6.49
	Urban	6.07
Summary		6.37
Rich	Rural	9
	Urban	7.11
Summary		7.3
Summary		6.7

Fig. 9 Representation of data in a table

7. **Radial Chart: Age(Years) by Study time (Hours) colored by Age(Years)**

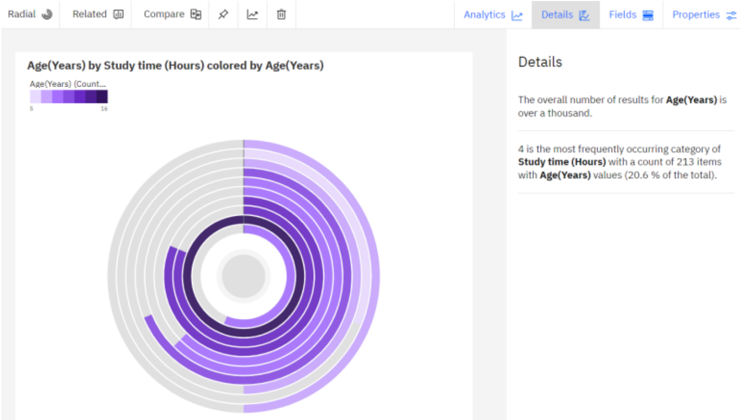


Fig. 10 Radial Chart

8. **Line Chart: Performance in online by study time(hours)**

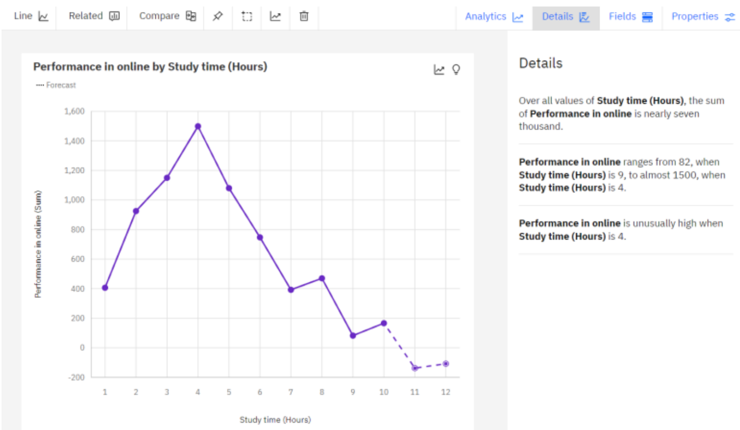


Fig. 11 Line Chart

9. **Line Chart: Performance in online by sleep time(hours)**

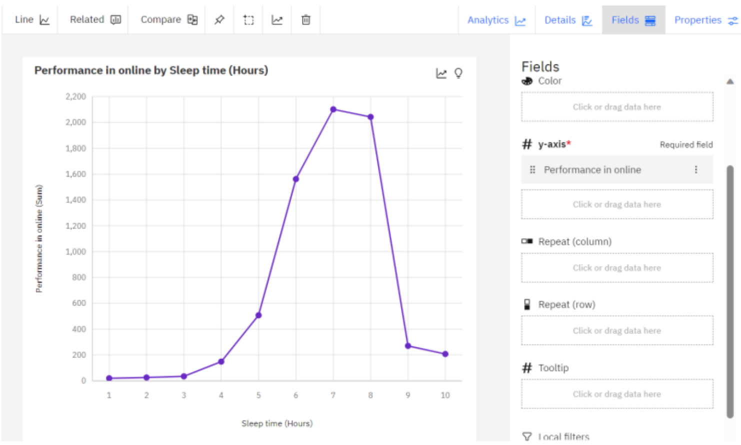


Fig. 12 Line Chart

10. Stacked Bar Chart: Performance in online by Device type used to attend classes colored by Gender

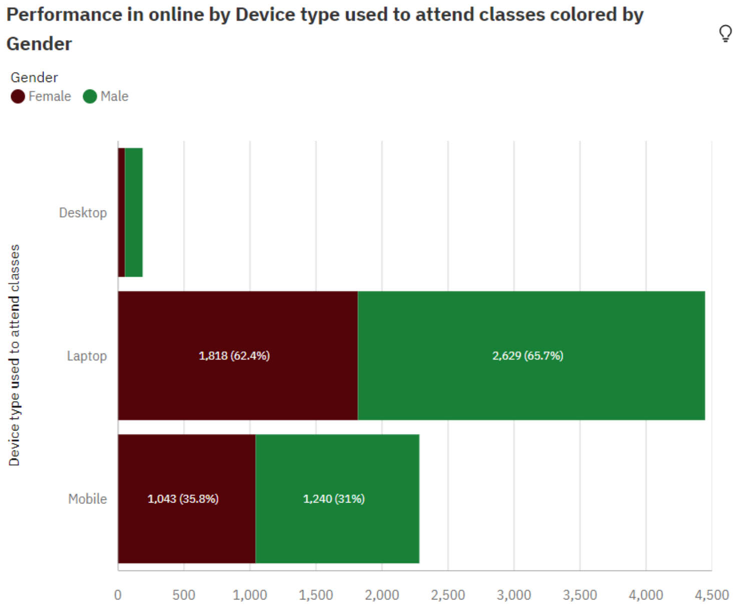


Fig. 13 Stacked bar chart

11. Dashboard

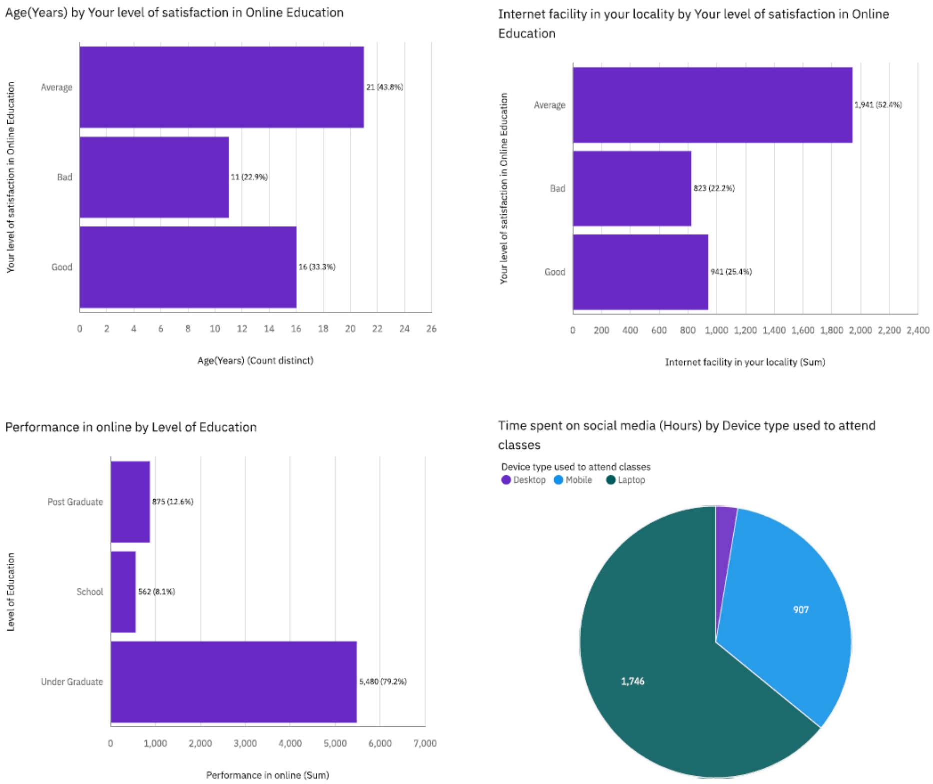
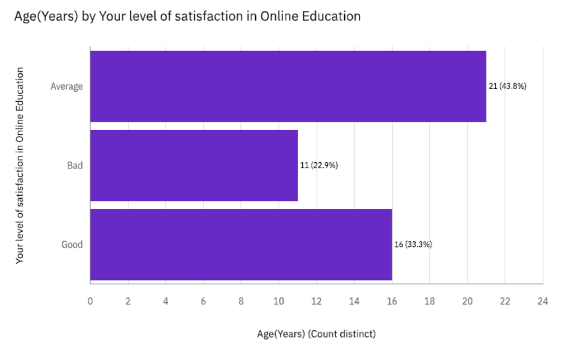


Fig. 14 Dashboard

12. Story

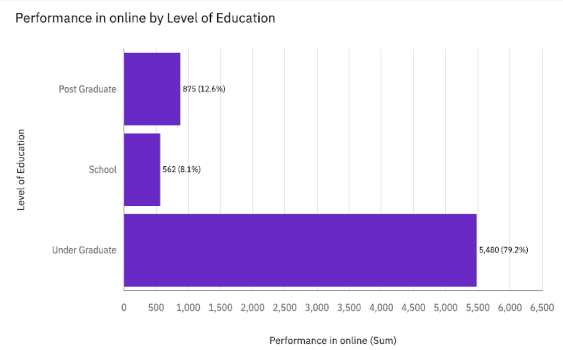
Age (Years) by Your level of satisfaction in Online Education

- This is the horizontal bar chart.
- In this bar chart, it is observed that the level of satisfaction "Average" has the highest value (43.8%) at the age of 21.
- Similarly, the level of satisfaction "Bad" has the lowest value (22.9%) at the age of 11.



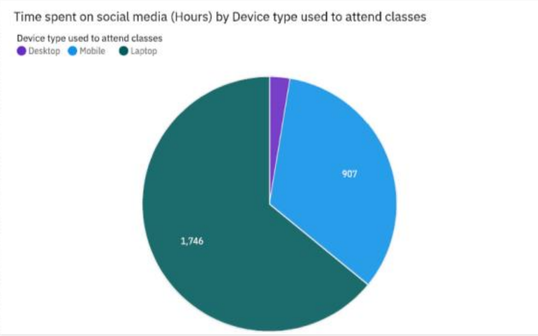
Performance in online by Level of Education

- This is a horizontal bar chart.
- In this bar chart, it is observed that the Level of Education "Under Graduate" has the highest value (5,480) at Performance in Online.
- Similarly, it is observed that the Level of Education "School" has the lowest value (562) at Performance in Online.



Time spent on social media (Hours) by Device type used to attend classes

- This is a pie chart.
- From this, it is observed that more time is spent on social media using the device type "Laptop".



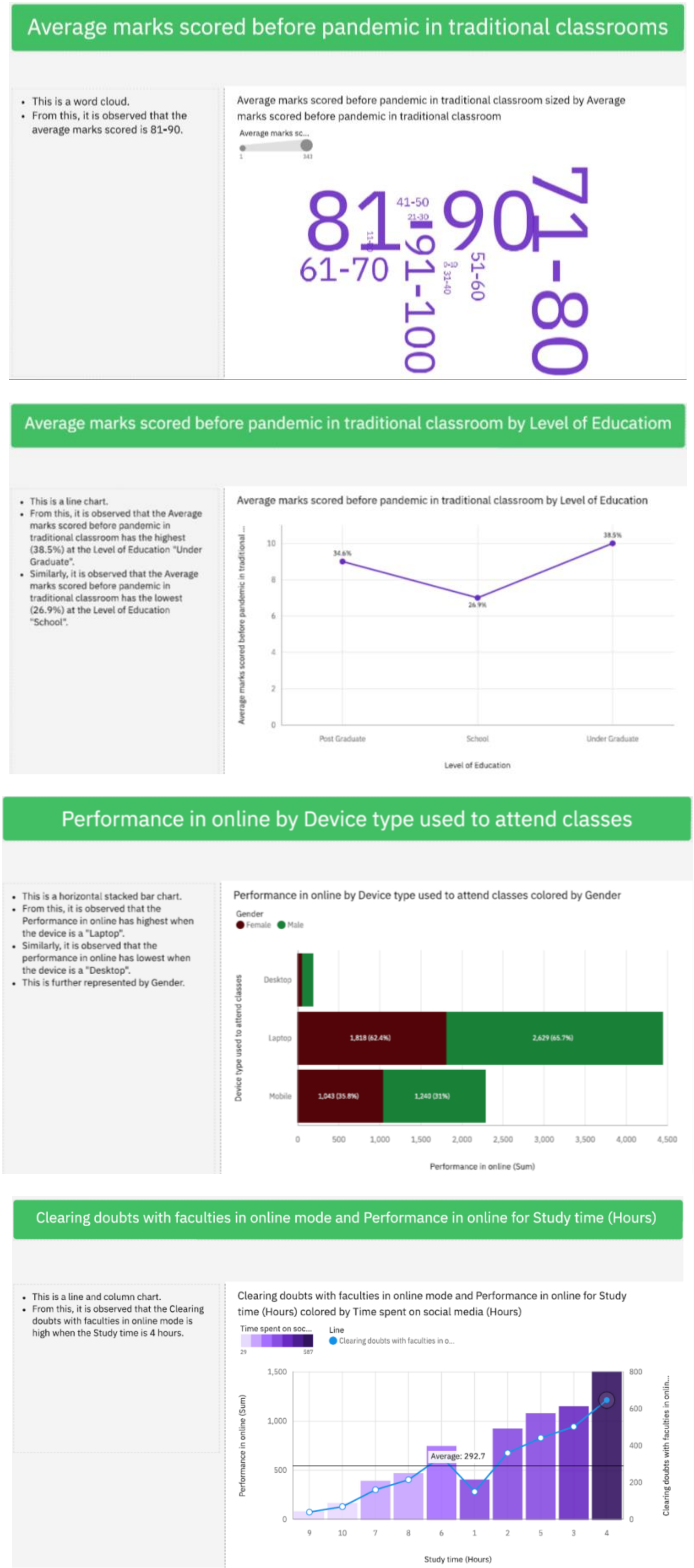

Average marks scored before pandemic in traditional classrooms
This is a word cloud.
From this, it is observed that the average marks scored is 81-90.
Average marks scored before pandemic in traditional classroom sized by Average marks scored before pandemic in traditional classroom
81-90
71-80
61-70
91-100
51-60
41-50
Average marks scored before pandemic in traditional classroom by Level of Education
This is a line chart.
From this, it is observed that the Average marks scored before pandemic in traditional classroom has the highest (38.5%) at the Level of Education "Under Graduate".
Similarly, it is observed that the Average marks scored before pandemic in traditional classroom has the lowest (26.9%) at the Level of Education "School".
Average marks scored before pandemic in traditional classroom by Level of Education
34.6%
26.9%
38.5%
Post Graduate
School
Under Graduate
Level of Education
Performance in online by Device type used to attend classes
This is a horizontal stacked bar chart.
From this, it is observed that the Performance in online has highest when the device is a "Laptop".
Similarly, it is observed that the performance in online has lowest when the device is a "Desktop".
This is further represented by Gender.
Performance in online by Device type used to attend classes colored by Gender
Gender
Female
Male
Desktop
Laptop
Mobile
Device type used to attend classes
Performance in online (Sum)
Clearing doubts with faculties in online mode and Performance in online for Study time (Hours)
This is a line and column chart.
From this, it is observed that the Clearing doubts with faculties in online mode is high when the Study time is 4 hours.
Clearing doubts with faculties in online mode and Performance in online for Study time (Hours) colored by Time spent on social media (Hours)
Average: 292.7
Performance in online (Sum)
Study time (Hours)

Fig. 15 Story

13. Report

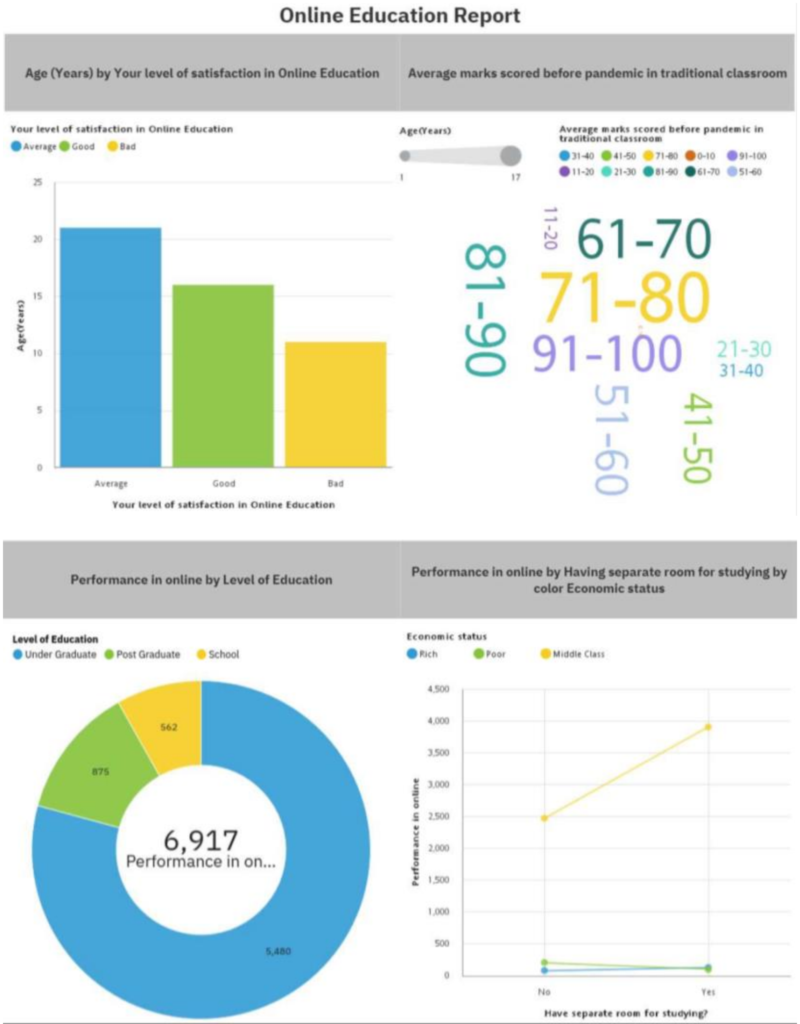


Fig. 16 Report

14. Web Integration using Flask

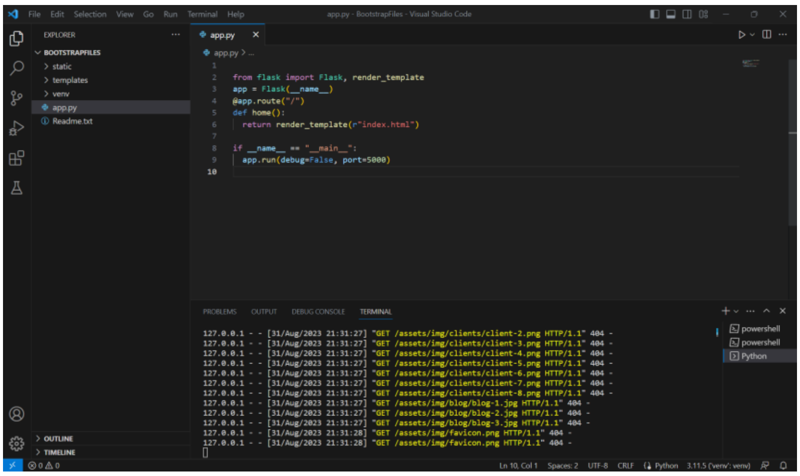


Fig. 17 Web integration with flask

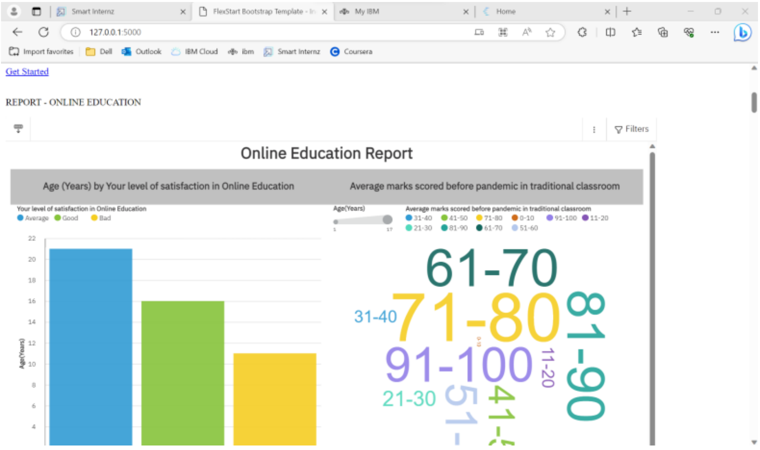


Fig. 18 Running on a local host after web integration

7. ADVANTAGES & DISADVANTAGES

Advantages

- Enhanced online learning
- Informed decision-making
- Improvement in curriculum design
- Advancement in research

Disadvantages

- Difficulty of adapting to technology integration
- Limited generalization of insights
- Analysis of different learning styles of learners
- Loss of on campus study experience

8. APPLICATIONS

The applications of this project are as follows.

- Integration of technology in educational institutions
- Enables educational institutes to understand student learning behavior, design the curriculum, and policy-making
- Implementation of innovative pedagogical strategies
- Helps educators and researchers
- Helps in finding ways to keep the students motivated and engaged in online learning

9. CONCLUSION

The key findings and insights in online education are the impact on education, motivation, and learning behavior. It would tie together the various aspects explored throughout the analysis and highlight their implications for the future of education. Finally, a clear understanding of the insights gained from the analysis and inspire them to take action in advancing the field of online education.

10. FUTURE SCOPE

There are several paths for further exploration and development that can contribute to the continuous growth and development of online education. Some potential areas of future scope are as follows.

- Innovation in pedagogical practices
- Long-term analysis of online education

- Implementation of effective assessment methods
- Robust learning analytics
- Ethical practices in learning

11. BIBLIOGRAPHY

[1]. P.-H. Lin, A. Wooders, J. T.-Y. Wang, and W. M. Yuan, "Artificial intelligence, the missing piece of online education?," *IEEE Eng. Manag. Rev.*, vol. 46, no. 3, pp. 25–28, Sep. 2018, doi: 10.1109/EMR.2018.2868068.

[2]. C. Gupta, V. Gupta, and A. Stachowiak, "Adoption of ICT-based teaching in engineering: An extended technology acceptance model perspective," *IEEE Access*, vol. 9, pp. 58652–58666, 2021, doi: 10.1109/ACCESS.2021.3072580.

[3]. M. Hernandez-de-Menendez and R. Morales-Menendez, "Technological innovations and practices in engineering education: a review," *Int J Interact Des Manuf*, vol. 13, no. 2, pp. 713–728, Jun. 2019, doi: 10.1007/s12008-019-00550-1.

[4]. J. R. Morrison, S. M. Ross, and A. C. K. Cheung, "From the market to the classroom: how ed-tech products are procured by school districts interacting with vendors," *Education Tech Research Dev*, vol. 67, no. 2, pp. 389–421, Apr. 2019, doi: 10.1007/s11423-019-09649-4.

[5]. F. Peruzzo, S. J. Ball, and E. Grimaldi, "Peopling the crowded education state: Heterarchical spaces, EdTech markets and new modes of governing during the COVID-19 pandemic," *International Journal of Educational Research*, vol. 114, p. 102006, 2022, doi: 10.1016/j.ijer.2022.102006.

[6]. F. Mohamed and A. Shoufan, "Choosing YouTube videos for self-directed learning," *IEEE Access*, vol. 10, pp. 51155–51166, 2022, doi: 10.1109/ACCESS.2022.3174368.

[7]. A. Shoufan and F. Mohamed, "YouTube and education: A scoping review," *IEEE Access*, vol. 10, pp. 125576–125599, 2022, doi: 10.1109/ACCESS.2022.3225419.

[8]. G. Alexandron, L. Y. Yoo, J. A. Ruipérez-Valiente, S. Lee, and D. E. Pritchard, "Are MOOC learning analytics results trustworthy? With fake learners, they might not be!," *Int J Artif Intell Educ*, vol. 29, no. 4, pp. 484–506, Dec. 2019, doi: 10.1007/s40593-019-00183-1.

[9]. M. Labayen, R. Vea, J. Florez, N. Aginako, and B. Sierra, "Online student authentication and proctoring system based on multimodal biometrics technology," *IEEE Access*, vol. 9, pp. 72398–72411, 2021, doi: 10.1109/ACCESS.2021.3079375.

[10]. A. Ara Shaikh, A. Kumar, K. Jani, S. Mitra, D. A. García-Tadeo, and A. Devarajan, "The role of machine learning and artificial intelligence for making a digital classroom and its sustainable impact on education during Covid-19," *Materials Today: Proceedings*, vol. 56, pp. 3211–3215, 2022, doi: 10.1016/j.matpr.2021.09.368.

[11]. M. A. Fauzi, "E-learning in higher education institutions during COVID-19 pandemic: current and future trends through bibliometric analysis," *Heliyon*, vol. 8, no. 5, p. e09433, May 2022, doi: 10.1016/j.heliyon.2022.e09433.

[12]. A. C. Starks and S. M. Reich, ""What about special ed?": Barriers and enablers for teaching with technology in special education," *Computers & Education*, vol. 193, p. 104665, Feb. 2023, doi: 10.1016/j.compedu.2022.104665.

[13]. H. P. VanDerSchaaf, T. U. Daim, and N. A. Basoglu, "Factors influencing student information technology adoption," *IEEE Trans. Eng. Manage.*, vol. 70, no. 2, pp. 631–643, Feb. 2023, doi: 10.1109/TEM.2021.3053966.

[14]. N. Bianchi, Y. Lu, and H. Song, "The effect of computer-assisted learning on students' long-term development," Journal of Development Economics, vol. 158, p. 102919, Sep. 2022, doi: https://doi.org/10.1016/j.jdeveco.2022.102919.

[15]. M. Zuo, Y. Ma, Y. Hu, and H. Luo, "K-12 students' online learning experiences during COVID-19: Lessons from China," *Front Educ China*, vol. 16, no. 1, pp. 1–30, Mar. 2021, doi: 10.1007/s11516-021-0001-8.

[16]. G. Toti and M. A. Alipour, “Computer science students’ perceptions of emergency remote teaching: An experience report,” *SN COMPUT. SCI.*, vol. 2, no. 5, p. 378, Sep. 2021, doi: 10.1007/s42979-021-00733-2.

[17]. S. Lai, B. Sun, F. Wu, and R. Xiao, “Automatic personality identification using students’ online learning behavior,” *IEEE Trans. Learning Technol.*, vol. 13, no. 1, pp. 26–37, Jan. 2020, doi: 10.1109/TLT.2019.2924223.

[18]. L. Qianqian *et al.*, “Research on behavior analysis of real-time online teaching for college students based on head gesture recognition,” *IEEE Access*, vol. 10, pp. 81476–81491, 2022, doi: 10.1109/ACCESS.2022.3192349.

[19]. R. Y.-Y. Chan, C. M. V. Wong, and Y. N. Yum, “Predicting behavior change in students with special education needs using multimodal learning analytics,” *IEEE Access*, vol. 11, pp. 63238–63251, 2023, doi: 10.1109/ACCESS.2023.3288695.

[20]. S. Delgado, F. Moran, J. C. S. Jose, and D. Burgos, “Analysis of students’ behavior through user clustering in online learning settings, based on self organizing maps neural networks,” *IEEE Access*, vol. 9, pp. 132592–132608, 2021, doi: 10.1109/ACCESS.2021.3115024.

[21]. J. Yue *et al.*, “Recognizing multidimensional engagement of e-learners based on multi-channel data in e-learning environment,” *IEEE Access*, vol. 7, pp. 149554–149567, 2019, doi: 10.1109/ACCESS.2019.2947091.

[22]. A. M. Villa, C. P. Molina, and M. Castro, “Students’ behavior when connecting to the LMS: A case study at UNED,” *IEEE R. Iberoamericana Tecnologias Aprendizaje*, vol. 14, no. 3, pp. 87–94, Aug. 2019, doi: 10.1109/RITA.2019.2942253.

[23]. A. Gupta, D. Garg, and P. Kumar, “Mining sequential learning trajectories with Hidden Markov Models for early prediction of at-risk students in e-learning environments,” *IEEE Trans. Learning Technol.*, vol. 15, no. 6, pp. 783–797, Dec. 2022, doi: 10.1109/TLT.2022.3197486.

[24]. H. Wan, Z. Zhong, L. Tang, and X. Gao, “Pedagogical interventions in SPOCs: Learning behavior dashboards and knowledge tracing support exercise recommendation,” *IEEE Trans. Learning Technol.*, vol. 16, no. 3, pp. 431–442, Jun. 2023, doi: 10.1109/TLT.2023.3242712.

[25]. H. Wan, K. Liu, Q. Yu, and X. Gao, “Pedagogical intervention practices: improving learning engagement based on early prediction,” *IEEE Trans. Learning Technol.*, vol. 12, no. 2, pp. 278–289, Apr. 2019, doi: 10.1109/TLT.2019.2911284.

[26]. S. Joshi and S. K. Sharma, “A model for good learning environment using learning data analytics,” *Procedia Computer Science*, vol. 194, pp. 156–164, 2021, doi: 10.1016/j.procs.2021.10.069.

[27]. V. Terzieva, S. Ilchev, K. Todorova, and R. Andreev, “Towards a design of an intelligent educational system,” *IFAC-PapersOnLine*, vol. 54, no. 13, pp. 363–368, 2021, doi: 10.1016/j.ifacol.2021.10.474.

APPENDIX

Web Integration using Flask

app.py

```
from flask import flask, render_template
app = flask(__name__)
@app.route("/")
def home():
 return render_template("index.html")
if __name__ == "__main__":
  app.run(debug=False, port=5000)
```