

Predicting primary and middle-school students' preferences for online learning with machine learning

**Authors:**

V. Selvakumar¹ 
Tilak Pakki Venkata² 
Teja Pakki Venkata² 
Shubham Singh² 

Affiliations:

¹Department of Mathematics and Statistics, Faculty of Statistics, Bhavan's Vivekananda College of Science, Humanities and Commerce, Secunderabad, India

²Department of Mathematics and Statistics, Bhavan's Vivekananda College of Science, Humanities and Commerce, Secunderabad, India

Corresponding author:

Selvakumar,
drselva2022@gmail.com

Dates:

Received: 02 Feb. 2023
Accepted: 11 July 2023
Published: 31 Aug. 2023

How to cite this article:

Selvakumar, V., Venkata, T.P., Venkata, T.P. & Singh, S., 2023, 'Predicting primary and middle-school students' preferences for online learning with machine learning', *South African Journal of Childhood Education* 13(1), a1324. <https://doi.org/10.4102/sajce.v13i1.1324>

Copyright:

© 2023. The Authors.
Licensee: AOSIS. This work is licensed under the Creative Commons Attribution License.

Read online:

Scan this QR code with your smart phone or mobile device to read online.

Background: The COVID-19 pandemic has brought attention to student psychological wellness. Because of isolation, lack of socialisation and intellectual and physical development from excessive media use, primary and secondary school students are at high risk for health problems.

Aim: This study aimed to identify the most effective machine learning model for predicting the offline and online instructional strategies students would choose during a pandemic.

Setting: The study was carried out at a number of primary and middle schools in Hyderabad, India.

Methods: We evaluated the data using machine learning methods such as logistic regression, K-nearest neighbour (KNN), decision trees, bagging and boosting using the Python programming language.

Results: In this study, 414 instances were collected from different schools. Exploratory data analysis showed that few students chose online courses. According to the research, very few students choose online classes, and the majority of students favoured offline classes over online because of physical and mental health difficulties; online education effects include a lack of social and peer relationships that affects young children psychologically, and they may not be disciplined enough to resist internet diversions. Smartphones, laptops, etc., affect their vision, causing headaches and impaired eyesight.

Conclusion: The KNN was the most accurate machine learning algorithm, with 92.13% accuracy to fits the data to identify the preferences of online education.

Contribution: This article examined the perspectives of primary and middle-school children on online education. Most students in this survey also reported experiencing mental or physical health issues that made online education difficult for them. Machine learning algorithms were applied to identify the most effective model for predicting students' online and offline study preferences. This machine learning method will help schools improve their course delivery methods, allowing students to continue their studies without interruption.

Keywords: online education; classification; logistics regression; bagging and boosting; K-nearest neighbours classifier.

Introduction

In recent years, advancements in technology have revolutionised the field of education, offering students a range of learning methods to choose from. However, COVID-19 pandemic has further accelerated the adoption of online learning as schools and educational institutions have had to adapt to remote teaching and learning environments. This unprecedented situation has brought to light the importance of understanding student preferences regarding learning methods, particularly the shift towards online modes of instruction. The pandemic has posed unique challenges for students, with varying access levels to technology and resources at home. Some students have embraced online learning as a convenient and flexible option, while others have faced difficulties adjusting to the virtual classroom environment.

Additionally, the impact of prolonged online learning on children's well-being and academic performance has been a subject of concern. Furthermore, the shift to online learning has also necessitated changes in assessment methods, with many educational institutions opting for

online exams. Understanding student preferences regarding online or offline modes of examination is crucial in ensuring fair and effective evaluation processes. This study aims to identify the most effective machine learning model for predicting student's preferred learning methods, explicitly focusing on the dichotomy between online and offline learning modes. This research becomes especially relevant in the context of the COVID-19 pandemic, where traditional in-person learning has been disrupted, and online learning has become a prominent alternative. By leveraging the power of machine learning algorithms, this research aims to uncover valuable insights into the factors that influence student preferences for online or offline learning. The analysis will encompass 414 online and offline students, utilising Google Forms and a questionnaire and direct contact with students to determine the impact of the lockdown on the children's education, health, social life, attention and other areas.

This study seeks to bridge the gap between student preferences and educational offerings by identifying the best machine-learning model for predicting whether students prefer online or offline learning methods. In the context of the COVID-19 pandemic, understanding student preferences becomes even more crucial as the education landscape continues to evolve.

Reviews

A study informs that the government imposed a lockdown because of COVID-19, which negatively impacted several sectors, including income, education and so forth (Zhou & Kan 2021). According to a study, students in lockdown accepted online learning, and most desired to continue taking classes online even in an epidemic (Zheng et al. 2021). The study says that online education is a relatively new teaching method and to increase students' interest in it. Some tactics, such as encouraging and assisting students, can be used (Shuanglan Luo & Xueqin Huang 2012). A study briefed that the majority of students preferred to complete their coursework online to fit their schedules, but many curriculum streams are practical in nature and cannot be completed online (Muthuprasad et al. 2021). A study shows that online education is a good option and helpful in pandemic situations (Allo 2020). A study suggests that the best way to teach students during a lockdown is an online mode, which allows teachers to teach everything from home using videos, PDFs and other tools (Jena 2020). One more study says that online instruction can be carried out even in the most challenging circumstances because the notes, resources and recordings of classes can be consulted at any time and in the comfort of one's home (Ayebi-Arthur 2017).

A study notifies that various online platforms offer materials like notes, PDFs, etc., and various exercises and assignments where students can independently understand the language (English) (Anggaira & Aryanti 2021). Another study brief that students used to experiment with various teaching methods from various institutions during lockdown to learn more about the subject. Social media is the most significant

diversion in online learning (Affendy, Ayi & Basir 2022). A study informs that students will show more interest when online courses are highly interactive with teachers and students, such as through discussions, quizzes and other activities (Arbaugh 2000).

A study describes that, although the students dislike online classes, they have accepted the pandemic situation and kept taking them despite the struggles (Valizadeh & Soltanpour 2021). A study specifies that students supported web-based learning during the lockdown, and most of them all came to the class using their mobile devices. When comparing interest in online activities by gender, female students demonstrated greater interest than male students (Vinayachandra et al. 2021). A study explains that it is not like every student is fit for online education. Teachers should identify their students' strengths and weaknesses regarding online learning strategies and create courses that will enable every student to grasp the material (Kauffman 2015). A study describes that if the instructions and guidelines for the system are clear, the student will be able to understand the online classes' course (Tetyana Rios, Marquita Elliott & Jean Mandernach 2018). Another study says that more students use mobile data to listen to online classes. Students were engaged in various academic activities like assignments and tests after each class. Future class information has been updated on social media platforms (Omar et al. 2021).

According to one survey, all resources, instruction, tests and exams are given online, which is the more efficient format for the students. Because of the pandemic, there was no longer any transportation between homes and schools, allowing students to spend more time on their studies (Darius, Gundabattini & Solomon 2021). A study specifies that students who participate actively in online classes interact with their classmates, demonstrating the strong correlation between students in online education (Fatima Afzal & Lynn Crawford 2022). A study indicates that mothers in lockdown dedicated more time to household and children's educational tasks than other family members (Xue & McMunn 2021). A study explains that students during the pandemic are less dependent on teachers and more dependent on family members and technology, so these online classes are bridging the knowledge gap for the students (Goudeau et al. 2021). A study specifies that students can easily understand concepts using various tools during online lessons. This study also shows that teachers and students can be encouraged to use online tools even after the COVID-19 situation (Pokhrel & Chhetri 2021).

According to a study summary, online education teachers and school administrators will encounter numerous difficulties (Kitahara & Westfall 2007). Another study suggests that students are uncomfortable with the new online teaching methods, and the benefits of taking classes online are not as great as their drawbacks (Nguyen 2022). A study explains that online classes were very helpful in the corona situation, but there were various challenges, such as poor internet, power outages, etc. (Safura 2021;

Alaidaros et al. 2021). Another study suggests that urban students find online learning much simpler than rural students because urban areas are more developed and students have better technology accessibility (Bast 2021). A study shows that female students use their phones more frequently than male students. But even so, mobile devices have negatively influenced students, causing male students to lose focus and self-control while causing female students' behaviour changes (Nayak 2018).

Another study notifies that COVID-19 is the reason that many students have left school (Kuhfeld et al. 2020). A study informs that students' communication skills have decreased because of online classes. Students are no longer confident and have trouble understanding materials found online (Alawamleh, Al-Twait & Al-Saht 2020). A survey indicates that teachers' e-learning teaching was below expectations. The students were content with the online classes but were not happy with how well they learned the material (Syauqi, Munadi & Triyono 2020). A study informs that students are not satisfied with the tests and exams that were given online, even though they received good grades and excellent test results (Hussain et al. 2020). Also, another study demonstrates that online classes are not interactive, and taking them for longer than 10 h can have serious health effects (Chandra et al. 2021). A study reviews that online learning has a detrimental impact on students, with depression and anxiety being the main effects. The government sector can act to provide students with the necessary equipment and internet access at a low cost during the pandemic (Agung & Surtikanti 2020).

Materials and methods

The data for the proposed study was gathered in two ways: (1) online using Google Forms and (2) offline using a questionnaire and direct contact with the students. The survey was done with safety and preventative measures in mind. The collected data were processed for analysis using Python. We employed plots and graphs for data visualisation. Machine learning methods were used to get insights from the investigation.

Statistical analysis

Data preprocessing is a phase in the data extraction and evaluation process that converts raw data into a format that computers and machine learning algorithms can interpret and analyse. The collected data consist of categorical variables. In the first step, before analysing the data, we converted categorical data into numeric data in the data preprocessing step. We only considered around 26 of the 28 characteristics in our data, and we did the following transformations:

- Two-variable columns have been replaced with 0s and 1s.
- Columns with five variables, such as excellent, very good, good, fair and poor, have been reduced to three variables, better, good and bad.

- By using the dummification technique, the columns with three variables have been transformed into 0s and 1s.
- According to the replies in the columns having check boxes, students who have picked one choice are placed in one category, while students who have selected several options are placed in a different category and converted to 0s and 1s.

After this procedure, the number of columns increased to 75; thus, we selected 56 of those 75 columns that were considered for analysis.

The data was analysed using Python, which is a programming language that can be used for powerful statistical analysis. The essential questions were considered and analysed using machine learning techniques as per the data collected using a questionnaire. The machine learning techniques that were implemented were:

- Logistic regression
- K-nearest neighbours (KNN)
- K-nearest neighbours using cross-validation
- Decision tree
- Bagging
- Adaptive boosting
- Gradient boosting
- Extreme gradient boosting

We used the metric accuracy to compare the machine learning models that we have implemented. The higher the accuracy, the better the model. Accuracy is the ratio of correct predictions to the total number of predictions.

Results

According to the data, male responses comprised 50.24% of the total responses, while female responses comprised 49.76%. The students come from various communities: 25.12% are from rural regions, 76.46 are from urban areas and the remaining are from peri-urban areas. When we look at responses by class, class 1 had a response rate of 2.42%, class 2 of 5.31%, class 3 of 6.52%, class 4 of 6.52%, class 5 of 4.83, class 6 of 24.64%, class 7 of 25.60% and class 8 of 24.15%. Considering the parent's income, 39.13% of kids said their parents' income was sufficient, 15.46% reported that it was insufficient and 45.41% reported that it was just partially enough.

From Figure 1, we can examine the correlation between the gender of the pupils and the frequency with which they desired that the school perform its classes. Male students who wanted daily lessons were somewhat more numerous than female students. The kids were all equally interested in the morning classes.

We can see that most male and female students wanted lessons to be held every day. While classes that met once a week or twice a week had similarly less interest from both males and females.

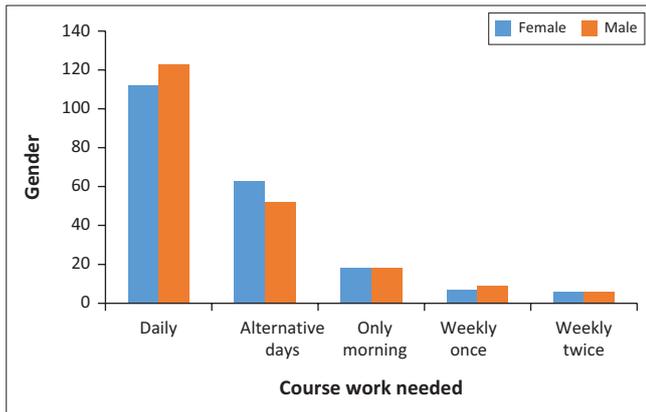


FIGURE 1: Plot of gender versus how often classes required.

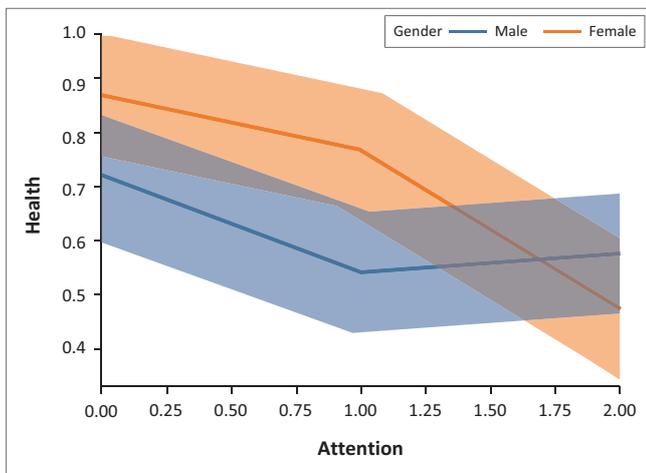


FIGURE 2: Plot of attention versus health.

In the plot (Figure 2) in the X axis:

- 0-Able to pay attention to classes
- 1-partially able to pay attention to classes
- 2-Not able to pay attention to classes

In the plot (Figure 2) in the Y axis:

- 0-No health issues
- 1-facing health issues

The graph in Figure 2 displays the interaction between students of both genders. According to their health and desire to attend online classes, students with excellent health could pay attention in class, but those in poor health could pay less attention. When we look at both genders, we can find that females in excellent health were better able to pay attention than males, but men in poor health were still able to do so.

In Figure 3, we can see that the majority of students have opted for offline classes, whereas a minority of students have opted for online classes.

In K-nearest neighbours (KNN), we are trying to determine the optimal number of neighbours, which would help us get results with the best accuracy possible, so we have implemented cross-validation in KNN to determine the optimal number of neighbours, leading to good accuracy.

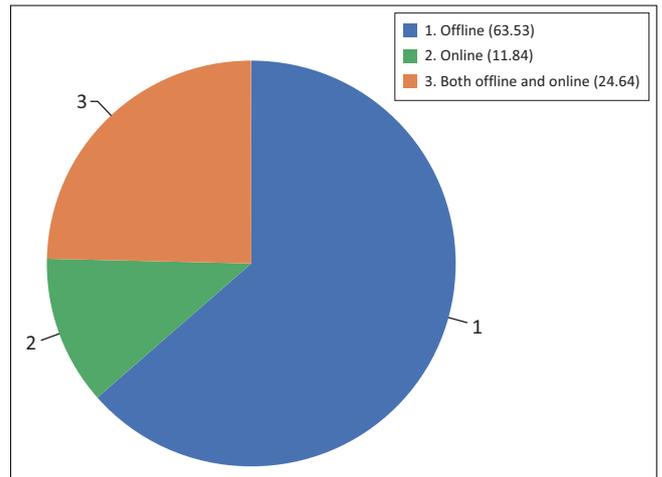


FIGURE 3: Plot of the desired mode of classes.

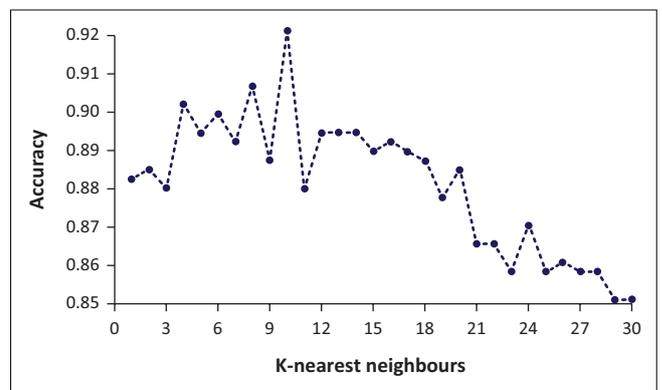


FIGURE 4: K-nearest neighbours with cross-validation.

TABLE 1: Comparing different machine learning algorithms.

Machine learning model	Accuracy %
Logistic regression	69.6
K-nearest neighbours	70.4
K-nearest neighbour with cross-validation	92.1
Bagging	67.2
Adaboosting	73.6
Gradient boosting	73.6
X Gradient boosting	72.8
Decision tree	67.2

The plot in Figure 4 is used to show the relationship between the number of neighbours and the accuracy. We see that we got the highest accuracy when the value of k was 10, where accuracy was 92%, whereas we got the lowest accuracy when the value of k was 30.

Table 1 shows that we have the following percentage accuracies after implementing all machine learning algorithms. We can see that K-nearest neighbour with cross-validation has the highest accuracy, 92.1%, and among all machine learning algorithms, the lowest accuracy we got was for decision tree and bagging, which is 67.2%.

Conclusion

We can observe that survey participation was equally split between the males and females. According to students'

opinions, the majority of them preferred daily offline education. Owing to health and concentration concerns, the majority of students chose offline classes; nevertheless, because of the lockdown, offline instruction was not possible. Additionally, students encountered significant difficulties when taking online exams since they lacked comprehensive notes and resources. The course's material has been condensed, which has greatly aided students' learning but has also lowered their understanding of how to grade their classes. Teachers also praised students for their excellent study habits throughout that difficult time. Our dataset's best-performing machine learning model was KNN with cross-validation.

By utilising cutting-edge technologies like machine learning, we can comprehend the conduct and circumstances of the children. Although every student has tried their hardest in COVID, online education is still new to them. Because of the lengthy online sessions held in one home, all students were stressed. As a result, the majority of students have chosen the offline option, and they are pleased that the schools have reopened.

Acknowledgements

Competing interests

The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

Authors' contributions

For this study, all the listed authors have made various contributions to the development of this article. Tilak P.V. and Teja P.V. collected data from different schools. The second, third and fourth authors derived machine learning models and predicted the model's accuracy. V.S. conceived the study and were in charge of the overall direction and providing some significant suggestions that helped improve the overall quality of the article. Finally, V.S. and Tilak P.V. wrote the manuscript with input from all authors. All authors discussed the results and contributed to the final manuscript.

Ethical considerations

This article followed all ethical standards for research without direct contact with human or animal subjects.

Funding information

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data availability

Data sharing is applicable to this article from the corresponding author, V.S., upon request in the future.

Disclaimer

The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of any affiliated agency of the authors.

References

- Affendy, N.A.R., Ayi, T.P. & Basir, M.S.S.M., 2022, 'A Study on Students' Online Learning Performance and Satisfaction During the Pandemic: The Case for TVET Education Institution in Borneo Malaysia,' in *2022 XII International Conference on Virtual Campus (JICV)*, Arequipa, Peru, 2022, pp. 1–4. <https://doi.org/10.1109/JICV56113.2022.9934419>
- Agung, A.S.N. & Surtikanti, M.W., 2020, 'Students perception of online learning during COVID-19 pandemic: A case study on the English students of STKIP Pamane Talino', *Soshum: Jurnal Sosial dan Humaniora* 10(2), 225–235. <https://doi.org/10.31940/soshum.v10i2.1316>
- Alaidaros, H., Kherd, A. & Al-Aidroos, H.A., 2021, 'Students' perception of online learning during COVID-19 pandemic at Al-Ahgaff University, Yemen: A Survey', in *2021 International Conference of Modern Trends in Information and Communication Technology Industry (MTICTI)*, Sana'a, Yemen, pp. 1–5. <https://doi.org/10.1109/MTICTI53925.2021.9664781>
- Alawamleh, M., Al-Twait, L.M. & Al-Saht, G.R., 2020, 'The effect of online learning on communication between instructors and students during COVID-19 pandemic', *Asian Education and Development Studies* 11(2), 380–400. <https://doi.org/10.1108/AEDS-06-2020-0131>
- Anggaira, A.S. & Ariyanti, N., 2021, 'Students perception towards Online-Digital Media in English learning during COVID-19 pandemic', in *ICLIQE 2021: Proceeding of The 5th International Conference on Learning Innovation and Quality Education*, Universitas Sebelas Maret, Surakarta Indonesia, September 4, 2021, pp. 1–4.
- Arbaugh, J.B., 2000, 'How classroom environment and student engagement affect learning in internet-based MBA courses', *Business Communication Quarterly* 63(4), 9–26. <https://doi.org/10.1177/108056990006300402>
- Ayebi-Arthur K., 2017, 'E-learning, resilience, and change in higher education: Helping a university cope after a natural disaster', *E-Learning and Digital Media* 14(5), 259–274. <https://doi.org/10.1177/2042753017751712>
- Bast, F., 2021, 'Perception of online learning among students from India set against the pandemic', *Frontiers in Education* 6, 1–8. <https://doi.org/10.3389/educ.2021.705013>
- Chandra, S., Verma, S., Kumar, B., Singh, N., Kumari, P. & Verma, A., 2021, 'Knowledge, attitude, and psychological effect on undergraduate/postgraduate students in Lockdown COVID-19 situation', *Journal of Pharmacy and Bioallied Sciences* 13 (suppl 1), S696–S698. https://doi.org/10.4103/jpbs.JPBS_829_20
- Darius, P.S.H., Gundabattini, E. & Solomon, D.G., 2021, 'A survey on the effectiveness of online teaching-learning methods for University and college students', *Journal of The Institution of Engineers (India): Series B* 102, 1325–1334. <https://doi.org/10.1007/s40031-021-00581-x>
- Fatima Afzal & Lynn Crawford, 2022, 'Student's perception of engagement in online project management education and its impact on performance: The mediating role of self-motivation', *Project Leadership and Society* 3, 100057. <https://doi.org/10.1016/j.plas.2022.100057>
- Girik Allo, M.D., 2020, 'Is the online learning good in the midst of Covid-19 Pandemic? The case of EFL learners', *Jurnal Sinestesia* 10(1), 1–10. <https://sinestesia.pustaka.my.id/journal/article/view/24>
- Goudeau, S., Sanrey, C., Stanczak, A., Manstead, A. & Darnon, C., 2021, 'Why lockdown and distance learning during the COVID-19 pandemic are likely to increase the social class achievement gap', *Nature Human Behaviour* 5(10), 1273–1281. <https://doi.org/10.1038/s41562-021-01212-7>
- Hussain, E.T., Daoud, S., Alrabaiah, H. & Owais, A.K., 2020, 'Students' perception of online assessment during the COVID-19 pandemic: The case of undergraduate students in the UAE', in *2020 21st International Arab Conference on Information Technology (ACIT)*, Giza, Egypt, November 28–30, 2020, pp. 1–6.
- Jena, P.K., 2020, 'Online learning during lockdown period for COVID-19 in India', *International Journal of Multidisciplinary Educational Research (IJMER)* 12(7), 82–92. <https://doi.org/10.31235/osf.io/qu38b>
- Kauffman, H., 2015, 'A review of predictive factors of student success in and satisfaction with online learning', *Research in Learning Technology* 23, 1–13. <https://doi.org/10.3402/RLTV23.26507>
- Kitahara, R.T. & Westfall, F., 2007, 'Promoting academic integrity in online distance learning courses', *Journal of Online Learning and Teaching* 3(3), 265–276.
- Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E. & Liu, J., 2020, 'Projecting the potential impact of COVID-19 school closures on academic achievement', *Educational Researcher* 49(8), 549565. <https://doi.org/10.3102/0013189X20965918>
- Luo, S. & Huang, X., 2012, 'A survey research on the online learning adaptation of the college students', in *2012 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet)*, Yichang, China, pp. 107–110.
- Muthuprasad, T., Aiswarya, S., Aditya, K.S. & Jha, G., 2021, 'Students' perception and preference for online education in India during COVID -19 pandemic', *Social Sciences & Humanities Open* 3(1), 100101. <https://doi.org/10.1016/j.ssoho.2020.100101>
- Nayak, J.K., 2018, 'Relationship among smartphone usage, addiction, academic performance and the moderating role of gender: A study of higher education students in India', *Computers & Education* 123, 164–173. <https://doi.org/10.1016/j.compedu.2018.05.007>

- Nguyen, G.H., 2022, 'Non-English majored students' preferences of online learning during the COVID 19 pandemic', *International Journal of TESOL & Education* 2(3), 272–283. <https://doi.org/10.54855/ijte.222319>
- Omar, N.H., Thomas, B., Jusoh, M.Z. & Jalil, S.Z., 2021, 'Students' perception and preference for online learning in Sabah during COVID-19 pandemic', *International Journal of Academic Research in Business and Social Sciences* 11(11), 270–292. <https://doi.org/10.6007/IJARBS/v11-i11/11262>
- Pokhrel, S. & Chhetri, R., 2021, 'A literature review on impact of COVID-19 pandemic on teaching and learning', *Higher Education for the Future* 8(1), 133–141. <https://doi.org/10.1177/2347631120983481>
- Safura, S., 2021, 'E-learning challenges during the COVID-19 pandemic experienced by students', *Ahmad Dahlan Journal of English Studies* 8(2), 135–145. <https://doi.org/10.26555/adjes.v8i2.21332>
- Syauqi, K., Munadi, S. & Triyono, M.B., 2020, 'Students' perceptions toward vocational education on online learning during the COVID-19 pandemic', *International Journal of Evaluation and Research in Education* 9(4), 881–886. <http://doi.org/10.11591/ijere.v9i4.20766>
- Tetyana Rios, Marquita Elliott & Jean Mandernach, B., 2018, 'Efficient instructional strategies for maximizing online student satisfaction', *Journal of Educators Online* 15(3), 1–9. <https://doi.org/10.9743/JEO.2018.15.3.7>
- Valizadeh, M. & Soltanpour, F., 2021, 'Higher education learners' attitudes towards emergency online instruction during COVID-19 pandemic: The context of Turkey', *Advanced Education* 8(19), 123–132. <https://doi.org/10.20535/2410-8286.245672>
- Vinayachandra, Geetha Poornima, K., Rajeshwari, M. & Krishna Prasad, K., 2021, 'Students' perception of the effectiveness of technology assisted online education during COVID-19 pandemic: An empirical study', *International Journal of Management, Technology, and Social Sciences* 6(1), 203–223. <https://doi.org/10.5281/zenodo.4922352>
- Xue, B. & McMunn, A., 2021, 'Gender differences in unpaid care work and psychological distress in the UK COVID-19 lockdown', *PLoS One* 16(3), e0247959. <https://doi.org/10.1371/journal.pone.0247959>
- Zheng, M., Bender, D. & Lyon, C., 2021, 'Online learning during COVID-19 produced equivalent or better student course performance as compared with pre-pandemic: Empirical evidence from a school-wide comparative study', *BMC Medical Education* 21(1), 495. <https://doi.org/10.1186/s12909-021-02909-z>
- Zhou, M. & Kan, M.-Y., 2021, 'The varying impacts of COVID-19 and its related measures in the UK: A year in review', *PLoS One* 16(9), e0257286. <https://doi.org/10.1371/journal.pone.0257286>