A Comprehensive Study of Online vs Offline Modes of Education

Sampurna Patra[1], Trishna Chakrabarty[1], Soumyodeep Manna[1], Satyam Das[1], Subhendu Saha[1]

Suman Das[2]*

[1,2] Department of MCA, Techno College Hooghly, Chinsurah, Hooghly, India

* Corresponding Author Email: srisumandas@gmail.com | ORCID: https://orcid.org/ 0009-0001-0987-8341

Abstract

The proliferation of technology has a notable impact on education, which is why e-learning has become increasingly popular. This study compares and contrasts online and offline learning in terms of learning outcomes, student engagement, instruction quality, and accessibility. In the era of modern technology, e-learning has been extremely useful during COVID-19. As a part of this analysis, we follow a mixed-methods approach, combining quantitative data from surveys and with those qualitative data insights from focus groups. Online education has challenges with student engagement and the digital divide, despite its benefits in terms of affordability, accessibility, and adaptability. Offline education can be limited by time and place, even though it enhanced social engagement and experiential learning. We are investigating the potential benefits of e-learning in the case of a pandemic. A hybrid educational model that maximizes learning outcomes while catering to students' varied needs is presented in the paper's conclusion. Our primary focus is conducting research on all first- through twelfth-grade students who took digital courses during the COVID-19 pandemic.

How to Cite this Article:

Patra, S., Das, S., Saha, S., Das, S., Manna, S. & Chakrabarty, T. (2026). A Comprehensive Study of Online vs Offline Modes of Education. International Journal of Creative and Open Research in Engineering and Management, <i>02</i>(03). https://doi.org/10.55041/ijcope.v2i3.030

https://doi.org/10.55041/ijcope.v2i3.030

Keywords Online Education; Offline Education; Smart Education; Machine Learning.

International Journal of Creative and Open Research in Engineering and Management
ISSN: **3108-1754 (Online)**
Volume 02 Issue 03 March-2026 | Impact Factor: 3.5

I. INTRODUCTION

The foundation of both individual and societal growth is education, and the ways in which it is delivered have changed dramatically over time. With technological progress, e-learning has become one of the most accepted modes of knowledge delivery and offers an alternative to the traditional face-to-face interactions of students and teachers in the traditional classroom arrangement. As we know, in both the modes of education systems, the standard of delivery of knowledge has to be standard in systematic way. We have faced very recent, COVID-19 pandemic phase, due to which we had to synchronize with Online Education. We have found Online Education as very much flexible, without geographical boundaries and without the barrier of specific time scheduling. On the contrary, Offline education, always remains unmatched with its traditional value and always enhances the relationships between students and teachers with easy process of sharing of knowledges. Learning through offline modes, even from ancient times, is always experiential learning because of the face-to-face interactions and strict supervisions over the environment and time. Our research focuses into the comparison of online and offline learning costs involved as course fees, student participation and obviously interactions, environmental availability and conditions, quality of learning, and even the impact on development of skills with moral characters over time. Our study intends to synchronize how the two modes of education can work together and to formulate a more comprehensive hybrid educational system to create the environment related with study, development, even for research , choosing the best parameters from both educational modes.

II. Data Preparation and Research Methodology

2.1Creating the Survey Form

The feedback form was designed to collect information in distance and face-to-face education. While questions in the survey form varied from studies, mental health, availability of gadgets and internet, and other forms of addictions. Survey questions consisting multiple choice questions (MCQ) with grades.

2.2Data Collection

We contacted with the students of different educational institutions to conduct the survey. To ensure high response rate, responses were collected and survey was conducted through online mode using Google Form and offline mode using direct conversations for a prolonged period. Although not enough, but from more than 690 conducted surveys , we can find a dataset which is enough to process the job.

2.3 Data Cleaning

The location of the student has been removed at very first from the dataset i.e. located at corporation, municipality or panchayat.30% of the random has been selected data for testing and 70% for training. 30% data for testing ensures that enough data is suitable for performance evaluation while a main portion of the data is used to train the model.

2.3.1 Why PCA?

Principal Component Analysis (PCA) was used in this study to address the complexity of the dataset and to uncover the underlying factors of Online vs Offline mode of education. PCA is a technique of dimensionality reduction that transforms the data into a set of linearly uncorrelated variables called principal components. The reasons for using PCA in this context are:

Dimensionality Reduction:

The dataset has 27 questions. PCA makes the data easier to analyses and visualize by reducing the number of dimensions while keeping the most important information.

An orthogonal transformation is used in the statistical process called Principal Component Analysis (PCA) to change a set of similar variables into a set of dissimilar variables. PCA is the most popular tool for exploratory data analysis and machine learning for predictive models. Moreover, PCA is an unsupervised learning algorithm that looks into the relationships between variables. Also called general factor analysis where regression finds the line of best fit.

Figure 1. Research Methodology Framework for Comparative Analysis of Online and Offline Education Modes Using Principal Component Analysis (PCA).

Figure 1 illustrates the end-to-end methodological workflow adopted in this study, starting from survey design and data collection, followed by preprocessing, dimensionality reduction using PCA, feature contribution analysis, and final interpretation leading to a hybrid education model recommendation.

Algorithm behind PCA:

Step-by-Step Explanation of PCA (Principal Component Analysis)

Step 1: Standardization

We need to homogenize our dataset to assure that each variable has a mean of 0 and a standard deviation of 1

$$Z=\frac{X-\mu}{\sigma} \tag{1}$$

Here,

μ is the meaning of independent features, $\mu = \{\mu_1, \mu_2, \mu_3 \dots \mu_m\}$,

σ is the standard deviation of independent features, $\sigma = \{\sigma_1, \sigma_2, \sigma_3 \dots \sigma_m\}$

The data is homogenized using Standard Scaler before applying PCA (scaler.fit_transform(selected_train_data)andscaler.transform(selected_test_data)). This ensures that each variable has a mean of 0 and a standard deviation of 1, fulfilling the standardization step.

Step 2: Covariance Matrix Computation

Covariance indicates how much two or more variables change in relation to one another and quantifies the strength of their joint variability. The following formula can be used to determine the covariance.

$$\boldsymbol{COV\,(X_1, X_2) = \frac{\sum_{i=1}^{n}(x_{1i}-x_1)(x_{2i}-x_2)}{n-1}} \tag{2}$$

The value of covariance can be positive, negative, or zero.

- **Positive** covariance states that as the x_1 increases x_2 also increases.
- **Negative** covariance states that as the x_1 increases x_2 also decreases.
- **Zeros**: No direct relation

Though the direct computation of the covariance matrix is not explicitly shown in the code, PCA internally computes the covariance matrix as part of its process. The relationships between variables are captured through the principal components, which are computed using the covariance matrix.

Step 3: Determining Principal Components through Eigenvalues and Eigenvectors

In the third step of PCA, we compute the eigenvalues and eigenvectors of the covariance matrix to identify the principal components. This process involves solving the equation:

$$AX = \lambda X \tag{3}$$

Then λ is a scalar known as the eigenvalue corresponding to eigenvector X. It can also be written as :

$$AX - \lambda X = 0 \tag{4}$$

$$(A - \lambda I)X = 0 \tag{5}$$

Where I represent the identity matrix of the same shape as matrix A. And the above conditions will be true only if (A - λI) will be non-invertible (i.e., singular matrix). This condition can be expressed as:

$$|A - \lambda I| = 0 \tag{6}$$

From the above equation, we can find the eigenvalues λ; after that, the corresponding eigenvector can be found using the equation (4).

In the context of PCA, these eigenvalues indicate the amount of variance captured by each principal component, while the eigenvectors indicate the direction of maximum variance. In practice, PCA uses functions like pca. Fit (scaled_data) [scaled_data = scaler.fit_transform (me1df)]to perform these calculations. The resulting eigenvalues and eigenvectors help identify the principal components, which are the directions in which the data varies the most.

2.3.2 Data Scaling for PCA:

Every feature has an equal impact; the data must be scaled before PCA is performed. This stops the principal components from being dominated by features with wider numerical ranges. Standard Scaler from Scikit-Learn can be used to standardize each feature by setting its mean to

0 and its standard deviation to 1. This step ensures that PCA can accurately identify the basic patterns in the data.

2.3.3 Component Set

For this analysis, 25 principal components were selected based on the explained variance. When combined, these 25 elements explain 90% of the variance in the data. In order to make analysis easier to handle and comprehend, these sections concentrated on capturing the majority of the variability and streamlining the dataset.

2.3.4 Using PCA

- Transforming both the training and test datasets to their principal component scores.

- This transformation allowed us to capture the underlying structure of the data and reduce it to its most informative components.

2.3.5 Analysis

This chapter summarizes the key findings of the research and their implications. It highlights the primary factors identified through Principal Component Analysis (PCA) that influence. The conclusion also reflects on the study's limitations and suggests areas for future research. The two main features we found from this model are - In PCA the most important feature with contribution is 0 with a contribution of 4.825 In testing, the most important feature with contribution is 1 with a contribution of 4.647 The analysis of this research suggests that neither online nor offline education can be deemed universally superior. Also, both modes have some advantages over another which is beneficial to enhance the learning experiences and create the environments to share the knowledges. Rather it is always a good strategy to proceed with a hybrid model consisting of the advantages of both the model i.e. online and offline education system comprising offline interactive classroom environment with online strategical platform. This integration of both the modes always provide the better experience to the educators and also to the students to explore the highest capability of the modes.

Finding the Most Important Contribution from Each Component

Alright, here's how this shakes out in plain English:

Wanna know which features run the show in your PCA setup? Just stare at the loadings matrix—that's your cheat sheet. You'll see, crystal clear, which feature is hogging the spotlight for each of your main components.

Now, "cumulative contributions"—big phrase, simple meaning. That's just you tallying up how much juice each component is giving to the total performance. Then there's "abs_loadings." That's the absolute value of all the loadings, for every feature, for every component. No negatives, just how much each input matters, stripped of the sign.

For the "Total (axis=1)" bit, imagine rolling all those absolute loadings together, for each feature, spread across all your components. You add it up to see which features are flexing the most muscle overall.

And then—because chaos is fun, but order is better—you sort those totals from biggest to smallest. Boom, that's your power ranking. The process spits out a list, showing the heavy hitters at the top, and you'll get the top 46 MVP features for each component contribution, loud and clear. Sit back and let the data brag about what matters most. This procedure helps with feature selection or interpretation by determining which feature has the biggest impact on the PCA.

2.4 Goal

It helps choose the ideal number of components by providing a visual representation of the amount of variance explained by each major component.

2.5 Visualization

A Single bar for each value plotted in the explained variance Itself is a Primary component: Personalization: Label the axes and name, axes more readable for personal productivity Use straws to be able to see the indices of the main components Stick with it; makes the main components nice sticks. Importance : Helps to decide with how many components, the variance recorded by each component needs to be taken so as to get significant dimensionality reduction.

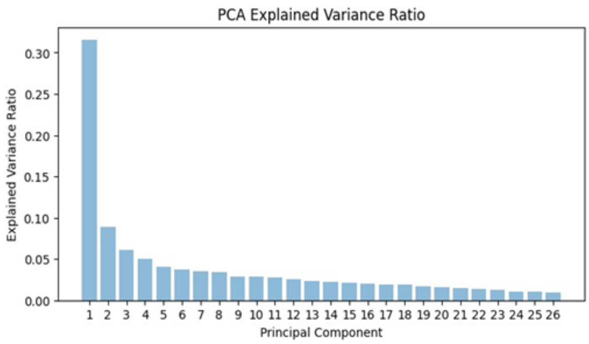

Figure 2. PCA Explained Variance Ratio

2.6 Final Outcome

We ran Principal Component Analysis (PCA)—basically, it is an info filter with a large capacity that filters out everything else and gets to the actual meat of your data mess. If you look at those two bar charts, that is the total "stuff," or variance grabbed by each of these components. And it is SO obvious, those first couple of components own pretty much all the important info and the rest… is like… a background noise? Once maybe third or fifth component, anything is not new, you just fall off the cliff. There will be no water in that stone this leaves us to better appreciate that the patterns in our data are largely a function of only a few features. Following the tracking of these:

Online modes of education are better than offline modes of education.

(অফলাইন শিক্ষার চেয়ে অনলাইন শিক্ষা ব্যবস্থা ভালো।)

Online education is effective for your child.

(অনলাইন শিক্ষা পদ্ধতি আপনার সন্তানের জন্য কার্যকর।)

III. CONCLUSION

This study has provided a comprehensive comparison of online and offline learning environments, especially in light of the COVID-19 pandemic and its effects on students in grades one through twelve. We were able to reduce complex survey data to the essential contributing features that had the biggest impact on the dataset by using Principal Component Analysis (PCA).

According to our analysis, the way parents view the efficacy and quality of online learning significantly influences the general trends. In particular, two questions stood out as the most significant: Is online learning more beneficial than traditional classroom instruction, and is online learning beneficial for their child? These results demonstrate that the success of online education is closely related to how it is delivered, regardless of technical accessibility or logistical issues.

REFERENCES

[1] Means, B., Bakia, M., & Murphy, R. (2014). Learning Online: What Research Tells Us About Whether, When and How. Routledge.

[2] Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). The Difference Between Emergency Remote Teaching and Online Learning. Educause Review.

[3] Allen, I. E., & Seaman, J. (2017). Digital Learning Compass: Distance Education Enrollment Report 2017. Babson Survey Research Group.https://www.sciencedirect.com/science/article/pii/S266637402030011X

[4] The Rise of Hybrid Education Models for Learning Solutions By Praveen Singh. https://cxotoday.com/story/the-rise-of-hybrid-education-models-for learning-solutions

[5] Challenges of online learning amid the COVID-19: College students' perspective Yuefan Xia , Yawen Hu , Chenyi Wu , Ling Yang , Man Lei https://pmc.ncbi.nlm.nih.gov/articles/PMC9815150/