A STUDY ON THE IMPACT OF COVID-19 ON STUDENT BEHAVIOR AND LEARNING IN INDIA

BHAGYA K.S.

DATE:

EMAIL: bhagyakarthika25@gmail.com

CONTENT

ABSTRACT

2020 has witnesses a different scenario in the world which has never been recorded in the history of any country in the world. Economy all over the world including India also been very badly affected. Production sector financial system at one side and education and other important services at another end. Due to coronavirus pandemic the state governments across the country temporarily started shutting down schools and colleges. As per the present situation, there is an uncertainty when schools and colleges will reopen. No doubt, this is the crucial time for education sector because entrance tests of several universities and competitive examinations are held during this period. Along with them how can we forget about board examinations and admissions into new courses and foreign education trips etc. The immediate solution of coronavirus is necessary or if like these days pass then closure of schools and colleges does not even have short term impact in India but can even cause far-reaching economic and societal consequences. Due to the close down of educational institutes it is estimated to affect around 600 million learners across the world. Considering the statistical facts the universities, colleges and the complete education system has adopted the online impartment of education which has been adopted by learners as well and as a result many such online platforms has been initiated. This article is an attempt to throw some light on to the problems being faced by the online teaching and learning and also the changes in the students behavior due to the pandemic in India.

OBJECTIVE

The objective of the project is to create a machine learning model to study the impact of COVID-19 on students behavior and learning. COVID-19 pandemic has drastically affected the life of students. All the way from classroom learning to e online classes in the lock down, their lifestyle has changed dramatically. Here we are trying to understand the impact on a small data with few samples.

TECHNOLOGY USED

PYTHON

Python is a high-level, general-purpose programming language. Its design philosophy emphasizes code readability with the use of significant indentation. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small- and large-scale projects. Python is open-source, so it is free to use, modify and distribute the python source code. Python is a high-level programming language that has English-like syntax making it easier to read and understand the code also the standard library of python is very big, that any and all function needed for a project can be found minimizing the use of external libraries. Different python packages like pandas, numpy, matplotlib, seaborn will used in the project.

MACHINE LEARNING

Machine learning is the study of computer algorithms that can improve automatically through experience and by the use of data. It is seen as a part of artificial intelligence. Machine learning algorithms build a model based on sample data, known as training data, in order to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of applications where it is difficult to develop conventional algorithms to perform the needed tasks.

LIBRARIES USED

NumPy

numpy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high -level mathematical functions to operate on these arrays. it is a open-source software and has many contributors.

Pandas

pandas is a software library written for the python programming language for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series. It is a free software released under the three-clause BSD license. The name derived from the term *"panel data"* , an econometrics term for data sets that include observations over multiple time periods for the same individuals.

Matplotlib

matplotlib is an amazing visualization library in Python for 2D plots of arrays. It is a multi-platform data visualization library built on NumPy arrays and designed to work with the broader scipy stack. It was introduced by John Hunter in the year 2002. One of the greatest benefits of visualization is that it allows us visual access to huge amounts of data in easily digestible visuals. It consists of several plots like line, bar, scatter, histogram etc

Seaborn

Seaborn is a library mostly used for statistical plotting in Python. It is built on top of Matplotlib and integrates closely with pandas data structures. It provides beautiful default styles and colour palettes to make statistical plots more attractive. Seaborn helps you explore and understand your data. Its plotting functions operate on data frame and arrays containing whole datasets and internally perform the necessary semantic mapping and statistical aggregation to produce informative plots. Its dataset-oriented, declarative API lets you focus on what the different elements of your plots mean, rather than on the details of how to draw them.

Sklearn

Scikit-Learn is a free machine learning library for Python. It supports both supervised and unsupervised machine learning, providing diverse algorithms for classification, regression, clustering, and dimensionality reduction. The library is built using many libraries such as NumPy and SciPy. It also plays well with other libraries, such as Pandas and Seaborn.

ALGORITHM

Explanatory Data Analysis (EDA) in statistics is an approach to analyzing data sets to summarize their main characteristics, often with visual methods. A statistical model can be used or not, but primarily EDA is for seeing what the data can tell us beyond the formal modeling or hypothesis testing task. Exploratory data analysis was promoted to encourage statisticians to explore the data, and possibly formulate hypotheses that could lead to new data collection and experiments. EDA is different from initial data analysis (IDA), which focuses more narrowly on checking assumptions required for model fitting and hypothesis testing, and handling missing values and making transformations of variables as needed. EDA encompasses IDA. Exploratory data analysis, robust statistics, non parametric statistics, and the development of statistical programming languages facilitated statisticians' work on scientific and engineering problems. There are a number of tools that are useful for EDA, but EDA is characterized more by the attitude than by particular techniques. Typical graphical techniques used in EDA are: Box plot, Histogram, Multi-vari chart, Run chart, Pareto chart, Scatter plot, Stem-and-leaf plot, Parallel coordinates, etc.

1. Observe dataset

The first step to conducting exploratory data analysis is to observe the dataset at a high level. Start by determining the size of dataset, including how many rows and columns it has. This can help you predict any future issues that might have with the data.

2. check for missing values

3. Visualizing and replacing the missing values

With the help of heat map, we can see the amount of data that is missing from the attribute. With this, we can make decisions whether to drop these missing values or to replace them. Usually dropping the missing values is not advisable but sometimes it may be helpful too. I replaced these missing values with mean because the number of missing values is less(we can use median too).

4. Asking Analytical Questions and Visualizations

I asked questions like, What devices that the students use for online learning?,How students spent their study time?, What is student's favorite stress buster, How their health has been impacted by various factors?, Does number of meals affect student's weight?,How they utilized their free time? Etc and visualized them.

DATA COLLECTION

Data is mainly collected from internet. Relevant data can be obtained from websites like kaggle.
Data collection is also possible by survey using google forms.

RESULT

CONCLUSION

The aim of the project is to conduct a study on the changes that brought in the behavior and learning of the students in India due to COVID-19 pandemic. This project will help to analyze the mental as well as physical health of the students, their free time activities, average class score before and after COVID-19 and helps in finding the impact of COVID-19 on student behavior and learning.