International Journal of Future Generation Communication and Networking
Vol.14, No. 1, (2021), pp. 3866 – 3877

Use of Data Science during Worldwide Pandemic: an Efficient and Reliable System to Support Decision Making in Different Sectors

Tithi Choudhary[1], Shreya Khedekar[2], Richa Sirwani[3], Muskan Shaikh[4]and Meenakshi A. Thalor[5]

[1] *Information Technology, B.E, AISSMS's Institute of Information Technology, Pune, India*

[1] tithichoudhary@gmail.com, [2] shreya.khedekar17@gmail.com

[3] sirwani.richa@gmail.com, [4] muskanikhat@gmail.com, [5]thalor.meenakshi@gmail.com

Abstract

Any calamities or emergency situations have created drastic and everlasting impacts on mankind since history. Such situations need to be managed in an efficient and effective manner. In this paper using a combination of technologies and tools we will provide a better alternative solution to the intuitive decisions thus making systems more reliable. Data Mining, Data Analytics, Data Visualization and Decision making have key applications demonstrated in this paper.Upon the choice of User Selections, this system will look into the number of deaths and recoveries that have occurred every day from the start of spread of the COVID-19. Based on this information, the goal will be to analyze and formulate a death and recovery analysis, which will in turn help us in understanding the effect of corona virus on different work sectors, while also displaying the GDP of the country. The second phase of the application will be to draw out a prediction of sales in the next upcoming days, based on the current situation in that country. Various business sectors such as apparel chains, grocery shops, etc. will be displayed along with their earnings since past few years until present. This data will be fed as training data for further forecasting process. A recommendation system will be applied to these forecasts, in order to provide best possible solutions for each work sector, so that user can make necessary changes to its business. Different visualization methods will be shown according to the Automated Machine learning system that will generate statistical depiction from the inflow of dynamic data.

***Keywords*- Data Mining, Data Analytics, Data Visualization and Decision Making, Automated Machine Learning, Prediction, UI, UX.**

I. INTRODUCTION

The world has been greatly affected by any disasters, pandemics, epidemics, wars, calamities etc. These situations have created major disturbances in the life of people and have created life changing impacts.

The aim is to develop a web application data intelligently and thus minimize the negative effects of the situation by educating the user of the application with info-graphic details according to the sector the user is a part of. The system provides an insight into the real-time details of the pandemic, in our case COVID -19 and thus shows the tendency of the positive outcomes on making the smartest calculated decision relatively.

Data from the most relevant times of covid-19 spread is analyzed and predictions are used for pattern detection which makes the system learn from the past and ensure the machine learning for the crisis under consideration leads to trusted outputs maximizing the benefits of disaster management. The

ISSN: 2233-7853 IJFGCN

system works with dynamic data fed to the application at runtime supported by Automated Machine Learning.

The sections III, IV, V, and VI i.e., Proposed Work, Implementation, Methodology and Experimental Analysis of this paper provide useful details and focuses on all necessary and relevant information on the system development and execution process.

II. LITERATURE SURVEY

Authors Gunther Schuh [1] describes existing approaches focusing on structuring the Data Mining domain so as to provide decision making to choose suitable Data Mining methods for the user.

Authors Jaya Bajpai [2] explains the problems in Data Cleaning process and presents data cleaning approaches involving different phases such as Data Analysis, Definition of Transformation Work Flow and Mapping Rules, Verification, Transformation, Back Flow Process of Cleaned Data. The tools used for the process of transformation and Data Cleaning have been mentioned.

Authors M. J., Price [4] describes the use of time series models such as ARIMA and Random Forest for prediction. A comparative study is presented which resulted in the fact that Prospective Random Forest model outperformed the ARIMA model over a specified time period.

Authors H. R. Bhapkar [5] proposed a new method for understanding the pattern of recovery and death rates by including the recovery time period.

Authors Mahipal Jadeja [6] present the use of Tree Maps in Data Visualization and transition from traditional representation methods to Tree-Maps. Tree-Maps are widely used to express a variety of nested as well as hierarchical data and data structures. These maps prove to be very useful in finding trends in large data sets. Along with the Tree Maps, other Data Visualization Techniques such as Scatterplot, Sparkline, Pie Chart, Gauge, Heat Map, Histogram, Box Plot etc. are also used to visualize the data by matching the data to right information.

III. PROPOSED WORK

The effect of pandemic has been negative on the economy across the world. During the Spanish influenza in 1918, the consequences on the business sector in the US and Europe were devastating. The overall economic loss incurred to the US business market was huge and many businesses had to shut down. The similar effects can be seen in 2020's covid-19 pandemic. The total loss in the global economy is predicted up to $1Tn. The graph below shows the losses suffered by the share market in various countries between January 10 2020 and March 16 2020. Fig.1.

ISSN: 2233-7853 IJFGCN

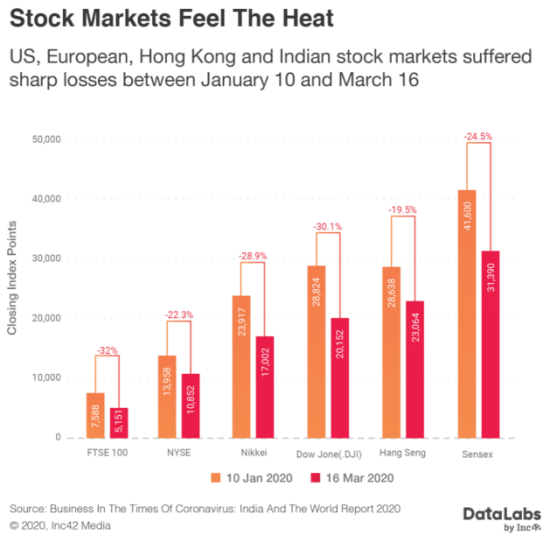

Fig.1.Share Market losses of various countries [3].

This shows how much impact this pandemic has on the economy of countries. The most affected sector is the business sector. According to an article by BBC, there are certain countries that are trying their best to survive this economic crisis (Denmark, Singapore, The US, Rwanda and New Zealand).
The system proposed in this paper, will help various sectors to reduce loss and be resilient to any such situation. The data generated during this pandemic are brought together to get insights and propose better ways for businesses to minimize the effects of macro environmental factors. During the lock-down imposed to control pandemic, many small businesses had to close down, many are still trying to manage and maintain their position in the market. If these businesses are given proper assistance about what, when and where to invest their time and money, it would result in improvised situations.

IV. IMPLEMENTATION

The flow chart depicts the basic workflow of the system architecture. The first step is to attain the relevant datasets for the research and analysis. The datasets aimed are of dynamic dataset of COVID-19, that will help us analyze the recovery and death stats as per the situation; the economic growth and downfall in business sectors (if observed) dataset in that period, etc. The data is achieved from various sources like social media, health departments, economical departments, hence the data is never clean or ready for instantaneous use, it needs to be cleaned using data cleaning techniques. The data cleaning techniques are specified.

Analyzing the past up till present data to get better insights for future development that will eventually help to get a brief overview of trends and patterns where macro environmental factors are affected, composing further decision making easier. The application would start with depiction of recovery and death rates during the pandemic till date by using a calculative formula as proposed by Dr. Bhapkar [5]. The aim is to use this method rather than going with traditional method of statistic calculations to procure a finer and more accurate rate. Meanwhile, per day recovery and death cases would also be represented in info-graphical format. Using automated machine learning algorithms to training and testing of datasets for statistical depiction of dynamic inflow of data regarding sales of every sector in business, such as apparels, automobiles, grocery chains, etc. will also be done. Further, the prediction and analysis procedures to forecast the same will be followed. In addition to this system, another info-graphic

ISSN: 2233-7853 IJFGCN

representation of GDP (Gross Domestic Product) of the country will be depicted. This will give the user an insight as to how the conditions are for business in their sector. Based on these observations, the user can opt to look into the recommendations being provided by our web applications own recommendation system, so that the user can take any necessary steps as required. Once the models are defined and their feasibility gets established, we will be able to develop an apt user interface with the defined methods connected at their backend.

The goal is to let the user review the outcomes for next few days, how the rate of virus spread will be, whilst analyzing the situation for them as to how they can keep their business steady. For ease of operation and understanding, the analysis will be described in visual forms. This will help the clientele to make a thought-through decision for their purposes.

The smooth inflow of data is ensured using Streamlit as middleware which will help note the state change in the application. The ultimate application will assist the client in the emergency situation or calamity that has extreme effects on the world using the past, present data and visualization in least complicated ways, for the measures and better decisions that can be taken to avoid future damage to their business. The macro environment effects on any trade can result in declination of its economic growth, providing assistance in these situations in a user-friendly way is the consequent outcome of this application.

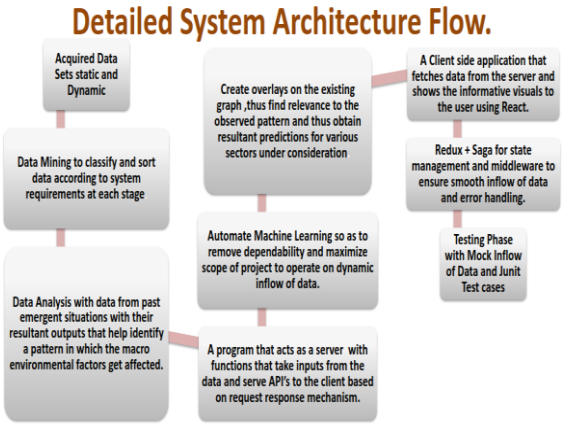

Fig. 2. Implementation Flow Chart [11].

V. METHODOLOGY

Data Cleaning and Mining

Data cleaning is a technique or a process used for eliminating any errors and inconsistencies in the data to upgrade the data quality [8]. Considering the current situation, the data sets regarding any pandemic are huge and changes in the trends take place gradually. This bulk source of data contains patterns of uncertainties that may hinder the further process of prediction [2]. In order to avoid such situations, we need various methods of cleaning Depending upon the need of the data sets, there are few tools (ETL tools) for data cleaning among which the one that fits well must be used: COPYMANAGER (Information Builders), DATASTAGE(Informix/Ardent), EXTRICATE (ETI), POWERMART, DECISIONBASE (CA/Platinum).

These are more advanced methods of data processing; and instead, we can make use of some simpler means of cleaning of data, like clearing null values of certain columns by either discarding the data, which may lead to major loss. Hence another option of normalizing the data by means of statis-

ISSN: 2233-7853 IJFGCN

tics is also available. We can substitute the null values with the mean, median or mode of the column. This will not only give us a precise value of calculation, but will also not lead to data loss overall.

Data Analytics (Favorable algorithm prediction)

The application requires two major analytical approach to represent the per day recovery to death ratio as well as prediction of sales and earnings by each business sector of the country. Hence for these actions, certain research on types of algorithm and methods was required. In the current case-scenario, a very less amount of data is available. Due to this severely fast developing epidemic conditions, our model needs to be working among high uncertainties. Thus, in order to select the most appropriate and accurate algorithm, we will be making use of an extensive selection method, named as Group of Optimized and Multisource Selection, abbreviated as GROOMS [9]. This method is specifically designed in order to achieve highest rates of prediction with limited resource. This design is being used for group forecasting, by making use of algorithms, some of which are capable of taking multiple sources of input.

After comparison among various optimized methods, the one that produces highest prediction accuracy will be implemented for defining the forecast on rate of growth/decline in spread of the virus.

Group of Optimized and Multi-source Selection. In this method, major three step process is implemented. Step one starts with the data set, no matter how small or large in scale, is being fed to the collection of models defined by us. In the next step, each candidate model will present its best performance, either by optimizing the parameters, or by including other data sources. The characteristics of algorithms being used in this method are roughly categorized as: non-parametric (e.g., linear/logistic regression, simple statistics) that may not require fixture of parameters: parametric model (e.g. activation functions for neural networks, splitting criteria for decision tree and random forest classifiers); and dual models whose performances are sensitive to both types of input sources as well as selection of parameters. Step wise allocations for GROOMS are shown in Fig. 3.

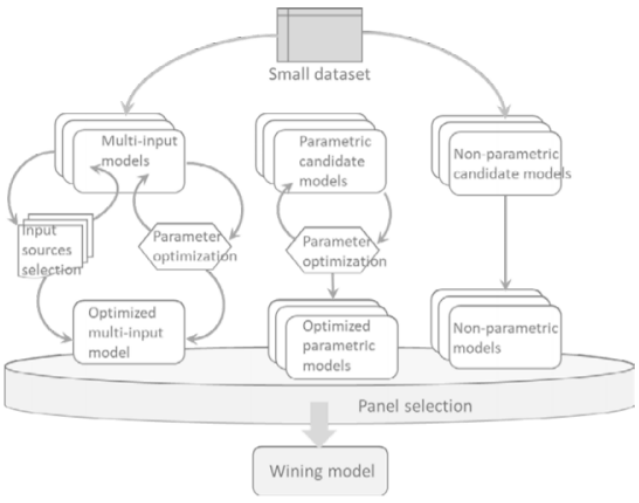

Fig. 3. GROOMS top to bottom approach [7].

All though GROOMS methodology analyses algorithms of various segregation, we chose this method to analyse various linear models of regression itself. The linear regression models that were experi-

ISSN: 2233-7853 IJFGCN

mented were as follows: ElasticNet CV, Linear Regression, SGD Regressor, Bayesian Ridge, Lasso Lars, ARD Regression, Passive Aggressive Regressor and TheilSen Regressor.

Data Visualization (Infographic Visuals of analysis and prediction)

Visualizing information in the form of graphs, charts, diagrams that help getting a better overview of the large data is called Data Visualization. We need data visualization because the visual summary of any information makes it easier to understand the trends and patterns of data.[10] Analysis of data is done to get insights from data, visualization adds to the betterment of data analysis. Data visualization of unstructured data does not give good insights hence it is necessary to have clean data, which helps in visualization.[12].

Types of data visualization techniques used in this system are as follows:

a) Line Chart
b) Tree Map
c) Polar Chart

Line Chart.

Line chart is widely used for visualizing patterns, trends and fluctuations in data that changes continuously. We can easily compare two or more datasets that are different yet have similar trends [13]. Through this representation, we can easily display the per day recoveries and deaths happening. From python programming perspective, there are many ways of plotting the graphs, some statically, and some animated in nature. One such library of graph, known as plotly is being used in the system.

Tree Map.

Tree is used to display hierarchical data while Tree map is used to display hierarchy and compare values between categories and subcategories [6]. The concept of tree maps is simple: allocate one box for the parent node and children of the node are drawn as boxes within it. This type of representation would be effective for showing business economy state-wise in the applications case.

*Polar Chart***.**

Polar charts are similar to pie charts but the angles measure the same and there is a change in radius depending on the value [14]. Polar charts here will show whether the certain sector should open for business and will it earn profits amid lockdown or is it better to keep low and not add-up extra costs. This is based on data received in the past and government guidelines at that time [16]. The tree map below shows the after ease of lockdown some sectors were able to earn revenue. Fig.4.

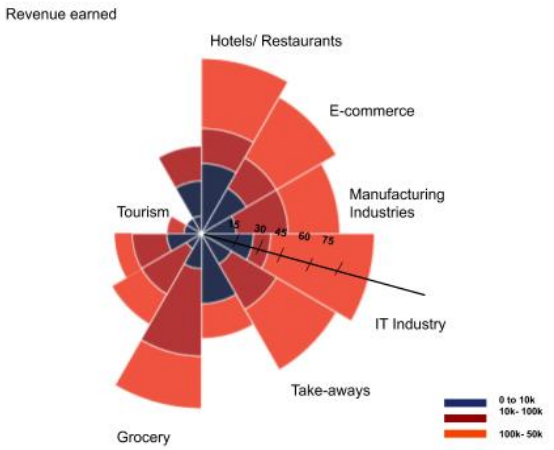

Fig. 4. Polar Chart showing the revenue earned by various sectors during Lockdown [15].

ISSN: 2233-7853 IJFGCN

VI. EXPERIMENTAL ANALYSIS

The first step in the application is to have a calculative source, so as to display the recovery and death rate analysis during this pandemic. For that to happen, the main initial process is to find a relative dataset, and looking into the cleaning process of it. The dataset being used for recoveries vs deaths analysis is from World Bank Group. The dataset for sales and revenue in various sectors are being taken from FRED Economic Research website that indulge in providing historical and present data for various aspects. In both the cases, the cleaning and processing of data is being done by simpler methods of statistics. The null values in a column are substituted with mean values of the entire column, giving an even palette to work on.

The next step is calculating the rate of recoveries and deaths occurred during the pandemic. After cleaning and normalization, the data is useful for the inspection. For this process, we make use of a formula as defined by Dr. Bhapkar in his searches. According to the formula, since mortality rate changes over time during COVID-19 pandemic, we call it Progressive Mortality Rate (PMR). In general, for any case/pandemic, C with an average recovery time period of P_{avg}, PMR can be computed as,

$$PMR = \frac{D_C}{N_{Pavg}} \times 100 \ ;$$

where D_C denotes the total number of death cases due to C disease and N_{Pavg} refers to a total number of infections before average recovery time period P_{avg} [5].

Similarly, Recovery Rate (RR) used to follow classical mathematical equation, $RR = \frac{R}{N} \times 100$, where R refers to total number of recovered cases.

As before, when recovery time period is taken into account, Progressive Recovery Rate (PRR) is,

$$PRR = \frac{R_C}{N_{Pavg}} \times 100 \ ;$$

where R_C refers to total number of recovered cases till date [5].

According to the analysis, the general recovery time is 14 days. But in our case, we make the P_{avg} modification to 15 days for calculative simplicity. Based on the formula described above, an algorithm for the rate calculation is formulated.

In the initial stages, we make use of simple method of rate analysis just by division of deaths caused and total number of infections recorded. Alongside, we also implement Dr. Bhapkar's method of analysis.

On establishing a visualization in the python programming, itself, we implement a representation of recoveries and death per day, in comparison to the total confirmed cases of any given day.

The main purpose of the deceased and recovered rates are to understand the impact of COVID-19 on business sector, by looking into the mortality against the disease. The aim is that while the user can view the per day numbers of confirmed cases, along with recoveries and deaths, they would also be able to see the rate with which the recoveries are occurring in the UI. It will be precise as we are taking in consideration the 15 days recovery period.

The second functionality of the application is analyzing the sales and revenues that will be gained by each of the sectors, while the situation of crisis still exists. Our goal is to work through the past and present data of sales acquired by each of the working sector, and based on this, we will forecast the sales for upcoming days. For this process, as mentioned in V section of the paper, we make use of a methodology that will help us decide the best fit model for the prediction, i.e., GROOMS (Group of Optimized and Multi-source Selection). In this, we make use of linear regression algorithms for

ISSN: 2233-7853 IJFGCN

smooth inflow of data during the visual representation of the forecasting. Not to mention, instead of using complex algorithms of ML and Deep Learning like SVM, K-Means or ANN that may lead to overfitting of data which is already not large in size, it's easier to use simpler methods for analysis.

For GROOMS method of selection, we chose algorithms such as ElasticNet CV, Linear Regression, SGD Regressor, Bayesian Ridge, Lasso Lars, ARD Regression, Passive Aggressive Regressor and TheilSen Regressor. We implement training and testing on the same data for all the algorithms. As mentioned earlier, the data being used is from FRED Economic Research website, and in the sector of apparel chains and their retail sales [12]. After testing for all the methods, we look into the Mean Squared Errors, Root Mean Squared Errors and R^2 scores of each of them, which lets us finally choose the desired method of implementation. The lowest the MSE & RMSE value and the highest the R^2 values is, the better the algorithm. A tabulation of the algorithms implemented along with their MSE, RMSE and R^2 scores are represented in table 1.

Table 1. Algorithm Comparison on the basis of GROOMS analysis.

S.No.	Algorithm	MSE	R2	RMSE
1.	Random Forest Regressor	1393970.9471	0.977735	1180.6655
2.	SGD Regressor	2187731.7462	0.872501	1479.0983
3.	ARD Regression	2167423.6479	0.873684	1472.2173
4.	Lasso Lars	2212334.4224	0.871182	1487.3918
5.	Linear Regression	2201089.6878	0.871658	1483.6070
6.	Bayesian Ridge	2202189.1251	0.871722	1483.9775
7.	ElasticNet CV	2190584.1296	0.872335	1479.0983
8.	TheilSen Regressor	2095565.7386	0.877872	1447.6069
9.	Passive Aggressive Regressor	2210354.6429	0.871067	1486.7261

From the above observation, we can see that Random Forest Regressor produces minimum MSE and maximum R^2 score. Thus, the best fit model for the prediction is considered to be Random Forest Regressor. Our aim is that for every different set of data, we implement an ensemble method, via which the algorithm having highest R^2 and lowest MSE and RMSE will be selected for forecasting. The approach is to introduce an automation which will consider an algorithm based on the data requirements.

The other method of implementation for analyzing and forecasting data of grocery stores as well as ecommerce, that we will be using is Prophet algorithm. It is a procedure for forecasting time series data based on an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects. It works best with time series that have strong seasonal effects and several

ISSN: 2233-7853 IJFGCN

seasons of historical data. Prophet is robust to missing data and shifts in the trend, and typically handles outliers well.

As mentioned above, there are three main components to this model: trend, seasonality and holidays. They are combined in the following way:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t$$

where,

g(t): piecewise linear or logistic growth curve for modelling non-periodic changes in time series

s(t): periodic changes (e.g., weekly/yearly seasonality)

h(t): effects of holidays (user provided) with irregular schedules

ε_t: error term accounts for any unusual changes not accommodated by the model [17].

We apply the model prophet on training data, which consist of sales made by Walmart from 2013 up until October of 2020. In return, we consider on finding whether the model predicts correct or accurate values of sales based on the training data, meanwhile also forecasting the revenue that will be collected for the next 12 months.

The result of the predict() function is a DataFrame that contains many columns. Perhaps the most important columns are the forecast date time (‘ds‘), the forecasted value (‘yhat‘), and the lower and upper bounds on the predicted value (‘yhat_lower‘ and ‘yhat_upper‘) that provide uncertainty of the forecast [18], which can be seen as in table.

Table 2. Predicted value depiction.

ds	yhat	yhat_lower	yhat_upper
2021-05-31	1506028.89	1475096.77	1539521.06
2021-06-30	1522551.63	1489504.50	1565302.87
2021-07-31	1545800.69	1511274.67	1578982.63
2021-08-31	1490185.37	1456465.21	1524025.88
2021-09-30	1527254.71	1487929.07	1567131.07

The above predicted data, in relation to the upper and lower bounds as well as forecasted data can be displayed by training data are represented as black dots and the forecast is a blue line with upper and lower bounds in a blue shaded area. Thus, this makes it clear that there is a possible gradual increase in the sales in the next 12 months (12 months, since the period specified is equivalent to 12) [20]. The above observations can be viewed in fig. 5.

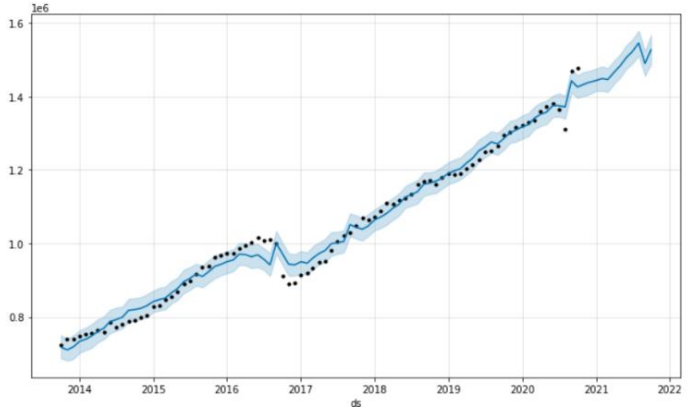

Fig. 5. Plotting Original as well as predicted value using upper and lower bounds of Prophet [19].

ISSN: 2233-7853 IJFGCN

In comparison with the new predicted and forecasted data, the original data contained revenue sales up till October 2020; whereas now, after a steady trend analysis and then predicting based on the period defined, now the data results upto sales being displayed till September of 2021, as shown in fig. through this implementation, a user working in retail sales (specifically Walmart) can definitely have a vision as to how to sustain them selves to maintain their revenues according to the forecasted details (fig. 6). As mentioned earlier, based on this result, the system will also be able to provide a definitive strategy or in layman's language, certain recommendations for their work sectors.

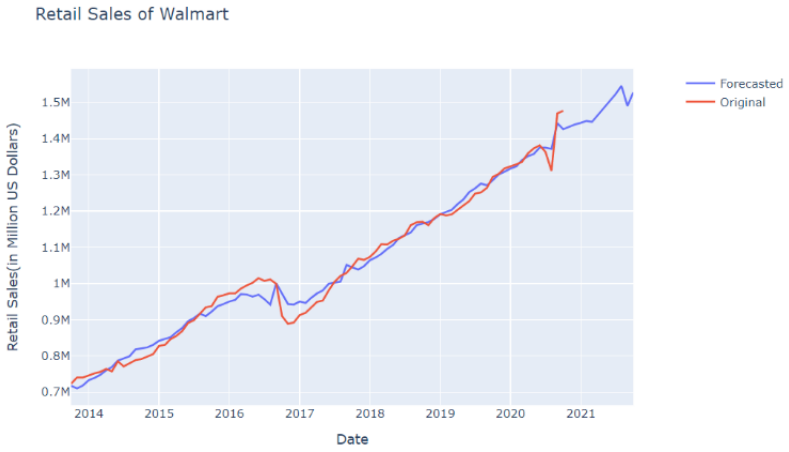

Fig. 6. Comparison between Original and predicted sales, as well as forecasted value representation [19].

Moving towards the forecasting section, our aim is to predict the sales analysis for 2 months from the current date of data present. For this, we first look into the trends of the past data; whether at certain conditions, how the revenues or sales were getting affected. This becomes the basis of our analysis. When we look into such patterns, we can also see that what steps were taken to sustain the business or given sector, as per user's discretion. By this theory we can easily implement a support system, which will be based on switch case setup. The logic is to produce certain recommendations as per the sector, when those conditions arise. For instance, if we take the hospitality management sector during the corona virus pandemic, lot of people were hesitant to use the establishments such as hotels and restaurants. For the customers to feel safe, the management came with the solution of sanitizing and creating as isolated environment as possible. Thus, such relative recommendations can be provided by the system according to the faction of work, and it will produce fruitful results.

VII. CONCLUSION

The paper thus, provides an informative insight into the real-world application of data science and analytics. The demonstrations of problem solving using a combination of different methods and tools and how they will be beneficial for the business sector. Analyzing the effect of covid-19 by looking into the aspects of recovery and death rates, how the revenues and sales have been in the past, present, and how will the situation be affected in future. Whilst predicting the future sales, also providing a solution to maintain their stability in the field of their work is an important aspect of the system, for the user.

These insights will not only provide information to its business users, but will also be easy to understand due to its highly info-graphic visual representation, along with a recommendation system to provide the necessary acknowledgement.

ISSN: 2233-7853 IJFGCN

VIII. SCOPE OF IMPROVEMENT

To make the system completely self-reliant and accurate, there may be additions to the set of tools and technologies used. The advanced versions, updated algorithms that are successful on boundary conditions and exceptions may be used for newer versions of the system with time. A major breakthrough could be applying the system for various other job sectors such as education, banking and finance as well as other tertiary sectors. These solutions can be implemented for any kind of calamities, as well as pandemic situations. Business Intelligence powered with AI will serve as a decision support system, that will help user make certain decisions for their business according to the type and give them fruitful results.

Another important addition could be creation of admin panel, that will enable them to add data as per requirement, as well selecting of apt algorithm as per the data. Because every data has different variations, thus algorithms per data can also change significantly.

REFERENCES

[1] Schuh, Günther & Prote, Jan-Philipp & Hünnekes, Philipp. (2020). Data mining methods for macro level process planning. Procedia CIRP. 88. 48-53. 10.1016/j.procir.2020.05.009.
[2] Jaya Bajpai Pravin S. Metkewar(2016), Data Quality Issues and Current Approaches to Data Cleaning Process in Data Warehousing, published in GRD Journals- Global Research and Development Journal for Engineering.
[3] Inc42, 2020. Stock Markets Feel the Heat [Online]. Delhi. Available from: https://inc42.com/datalab/business-in-a-pandemic-how-coronavirus-will-change-india-the-world/ [Accessed 7 November 2020].
[4] Kane, M. J., Price, N., Scotch, M., & Rabinowitz, P. (2014). Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. BMC Bioinformatics, 15(1), 276. doi:10.1186/1471-2105-15-276.
[5] Bhapkar, H.R., Mahalle, P.N., Dey, N. et al. Revisited COVID-19 Mortality and Recovery Rates: Are we Missing Recovery Time Period?. J Med Syst 44, 202 (2020). https://doi.org/10.1007/s10916-020-01668-6.
[6] Jadeja, Mahipal & Shah, K.. (2015). Tree-map: A visualization tool for large data. CEUR Workshop Proceedings. 1393. 9-13..
[7] Ruben Gonzalez Crespo, 2020. GROOMS with top-down optimization process [Online]. Available from: https://www.researchgate.net/figure/GROOMS-with-top-down-optimization-processes_fig1_339107651 [Accessed 9 November 2020].
[8] X. Xu, Y. Lei and Z. Li, "An Incorrect Data Detection Method for Big Data Cleaning of Machinery Condition Monitoring," in IEEE Transactions on Industrial Electronics, vol. 67, no. 3, pp. 2326-2336, March 2020, doi: 10.1109/TIE.2019.2903774..
[9] Shinde, G. R., Kalamkar, A. B., Mahalle, P. N., Dey, N., Chaki, J., & Hassanien, A. E. (2020). Forecasting Models for Coronavirus Disease (COVID-19): A Survey of the State-of-the-Art. SN Computer Science, 1(4). doi:10.1007/s42979-020-00209-9.
[10] "What is data visualization" https://www.klipfolio.com/resources/articles/what-is-data-visualization#DataViz2
[11] Muskan Shaikh, 2020. Detailed System Architecture Flow. Pune [Accessed 15 August 2020].
[12] "What is Data Visualization and Why Is It Important?" https://www.import.io/post/what-is-data-visualization/
[13] Y. Wang, F. Han, L. Zhu, O. Deussen and B. Chen, "Line Graph or Scatter Plot? Automatic Selection of Methods for Visualizing Trends in Time Series," in IEEE Transactions on Visualization

ISSN: 2233-7853 IJFGCN

and Computer Graphics, vol. 24, no. 2, pp. 1141-1154, 1 Feb. 2018, doi: 10.1109/TVCG.2017.2653106.
[14] "Polar chart"- https://doc-ar-chives.microstrategy.com/producthelp/10.6/AdvancedReportingGuide/WebHelp/Lang_1033/Content/AdvancedReporting/Polar_chart.htm
[15] Richa Sirwani, 2020. Polar Chart: Revenue Earned. Pune [Accessed 21 september 2020].
[16] N. Ari and M. Ustazhanov, "Matplotlib in python," 2014 11th International Conference on Electronics, Computer and Computation (ICECCO), Abuja, Nigeria, 2014, pp. 1-6, doi: 10.1109/ICECCO.2014.6997585.
[17] A Quick Start of Time Series Forecasting with a Practical Example using FB Prophet - https://towardsdatascience.com/a-quick-start-of-time-series-forecasting-with-a-practical-example-using-fb-prophet-31c4447a2274
[18] "Time Series Forecasting With Prophet in Python" - https://machinelearningmastery.com/time-series-forecasting-with-prophet-in-python/
[19] Tithi Choudhary, 2020. Retail Sales of Walmart. Pune. Available from: https://colab.research.google.com/drive/1PVRgcQL7SLzn9o7C_xLO2A-YR1XOrTzG#scrollTo=eU2UzlY-SU8K [Accessed 13 January 2021].
[20] Generate Quick and Accurate Time Series Forecasts using Facebook's Prophet (with Python & R codes) - https://www.analyticsvidhya.com/blog/2018/05/generate-accurate-forecasts-facebook-prophet-python-r/

ISSN: 2233-7853 IJFGCN