International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

A Proposed Business Improvement Model Utilizing Machine Learning: Enhancing Decision-Making and Performance

*Bhagyashree Gadekar [1], Tryambak Hiwarkar [2]

Submitted: 28/06/2023 **Revised**: 07/08/2023 **Accepted**: 28/08/2023

Abstract: This paper presents a proposed business improvement model that leverages Support Vector Machine (SVM) algorithms to enhance decision-making and performance in organizations. The integration of machine learning techniques into business processes has gained significant attention due to its potential for optimizing operations and achieving better outcomes. In this study, we focus on SVM, a powerful supervised learning method known for its ability to handle complex classification and regression tasks. The proposed business improvement model adopts a systematic methodology, beginning with the collection and preparation of data to guarantee the accuracy and usefulness of input data. In order to analyse the data and create predictive models that support decision-making in various commercial situations, SVM is then used. The model can be modified to solve various problems, including resource allocation, demand forecasting, risk assessment, and consumer segmentation. Using real-world company data from various industries, we ran experiments to verify the usefulness of the suggested approach. The outcomes show how SVM-based decision-making has a major impact on organisational performance. Businesses can make wise decisions that result in cost reductions, efficiency gains, and increased customer satisfaction by utilising the patterns and insights discovered by SVM. We also discuss the difficulties and factors involved in applying machine learning models in commercial contexts, such as data protection, interpretability, and model maintenance. By conducting this study, we hope to add to the growing body of knowledge on the application of machine learning to business improvement strategies and offer useful advice for businesses looking to use these methods to improve their decision-making and overall performance.

Keywords: *Business improvement model; Machine learning; Support Vector Machine (SVM); Decision-making; Performance*

1. Introduction

Staying ahead of the curve is vital for organisations hoping to achieve sustained growth and success in the fast-paced, cutthroat business environment of today[1]. Traditional company improvement approaches frequently rely on human-driven decision-making and historical data analysis, which can take time and may not completely use the enormous volumes of available data[2]. However, the rise of machine learning (ML) presents unheard-of chances to completely transform how businesses improve their decision-making processes and overall performance[3][4]. This study presents a thorough proposal for a state-of-the-art business improvement model that harnesses the power of machine learning algorithms to enhance organisational performance and optimise decision-making[5][6]. Businesses may acquire deeper insights, spot patterns, and make data-driven decisions more precisely and effectively by utilising ML approaches.

Overview of Machine Learning:

Artificial intelligence (AI) is a subset that allows systems to learn from experience and advance without explicit programming[7][8]. Business organisations may now extract important information and generate useful insights thanks to machine learning (ML) algorithms' ability to analyse sizable datasets, spot trends, and forecast events based on historical data[9][10].

Advantages of Utilizing Machine Learning in Business Improvement:

- **Enhanced Decision-Making:** ML models can process big data sets and find key elements that affect business results, enhancing decision-making. This provides decision-makers with thorough information, enabling them to make more intelligent and strategic decisions.
- **Predictive Analytics:** Organisations can foresee market trends, consumer behaviour, and future hazards by using ML algorithms, enabling pro-active steps to outperform the competition.
- **Process Optimization:** By using machine learning (ML) to find inefficiencies and bottlenecks in corporate processes, operations can be streamlined and optimised, saving time and money.
- **Personalization:** By recommending goods or services that are customised to specific interests, ML-powered systems can personalise the consumer experience, increasing happiness and loyalty.
- **Fraud Detection and Security:** Businesses may increase their security measures and quickly spot fraudulent

[1]Research Scholar, Department of Computer Science, Sardar Patel University, Bhopal, MP, India; bhagyashree.d.shendkar@gmail.com
[2]Professor, Department of Computer Science, Sardar Patel University, Bhopal, MP, India; tahiwarkar@gmail.com
***Correspondence:** bhagyashree.d.shendkar@gmail.com*

actions thanks to the superior ability of ML algorithms to recognise abnormalities and strange patterns.

Proposed Business Improvement Model:

- **Data Collection and Preprocessing:** Gather pertinent data from a variety of sources, including internal databases and external market data. a. Data Collection and Preprocessing. The data should be preprocessed to verify its accuracy, completeness, and suitability for ML analysis[11].

- **Model Selection:** Based on the specifics of the business situation, select the best ML algorithms. Decision trees, support vector machines, neural networks, and clustering algorithms are examples of common models.

- **Training and Validation:** To assure accuracy and dependability, train the chosen ML models using historical data and validate their performance.

- **Integration and Deployment:** Integrate the machine learning model with the business infrastructure so that it can process real-time data and continuously offer insightful information.

- **Monitoring and Continuous Improvement:** Keep an eye on the ML model's performance on a regular basis and adjust it as necessary to reflect shifting market conditions and data patterns.

Expected Business Benefits:

- **Better Decision-Making:** The suggested model will provide organisations with data-driven insights, enabling them to make decisions that are both more effective and efficient.

- **Competitive Advantage:** Companies can obtain a competitive edge by providing cutting-edge goods and services by utilising ML to forecast market trends and customer preferences.

- **Cost Savings:** Process optimisation through ML analysis can result in cost savings and improved resource utilisation.

- **Improved consumer Experience:** ML-driven personalisation will improve overall consumer satisfaction, encouraging patronage and a favourable view of the business.

The way businesses are run could be completely changed if machine learning is incorporated into process improvements[12]. Businesses may adapt to the changing market conditions, gain a competitive edge, and achieve sustainable growth by using ML algorithms to improve decision-making and overall performance[13]. Adopting this suggested business improvement approach will be a critical step towards future success and prosperity as technology develops.

2. Literature Survey

The paper entitled "Does national culture explain consumers' reliance on online reviews? Crosscultural variations in the effect of online review ratings on consumer choice" by Rae Yule Kim et al[14] explores the role of national culture in shaping consumers' reliance on online reviews and how it affects their choices. The survey of the available literature on consumer behaviour and cultural factors focuses on the effects of online review ratings in various nations. The study looks at whether cultural traits like individualism versus collectivism, power distance, and uncertainty avoidance have a major impact on how consumers make decisions about whether to trust internet reviews. The author seeks to offer a thorough knowledge of how cultural differences can affect consumers' faith in online reviews and, as a result, their purchase behaviour by analysing and synthesising pertinent material. The study sheds insight on cultural variations in the weight that customers accord online evaluations, adding to the corpus of knowledge on consumer behaviour in the digital era. The study paves the way for more nuanced and context-aware marketing strategies for enterprises operating in varied cultural settings by looking at how national culture may moderate the impact of internet review ratings on consumer decision. The findings of this study have ramifications for e-commerce platforms, businesses aiming to expand globally, and marketers looking to customise their strategies depending on cultural preferences and behaviours. Understanding these cross-cultural differences enables companies to better meet the requirements and expectations of customers from various cultural backgrounds, which in turn results in more successful and effective marketing efforts.

The paper entitled "Is it always advantageous to add-on item recommendation service with a contingent free shipping policy in platform retailing?"by Chunfa Li[15] explores in the world of platform retailing, the question of whether it is always advantageous to add-on an item recommendation service with a contingent free shipping policy is a complex and significant one. A literature survey on this topic reveals various perspectives and insights. According to certain research, adding an item recommendation service can improve user experience by tailoring product recommendations and boosting customer engagement, which in turn can increase conversion rates and customer loyalty. Concurrently, conditional free shipping rules can serve as a potent inducement to increase purchases and lower cart abandonment rates. However, according to other research, deploying such services may not always be advantageous due to rising operating expenses and a potential misalignment with the preferences of all consumer segments. Concerns exist around possible biases in item recommendations as well as the need to balance delivering free delivery with upholding sustainable business practises.

The success of incorporating these services ultimately depends on a number of variables, including the platform's target market, its product offering, and the overall business strategy. Further research is required to evaluate the precise situations in which the combination of an item suggestion service and a dependent free shipping policy can actually give an advantage in a market. Platform retailers looking to optimise their offerings and give their customers a seamless shopping experience may find it helpful to compare various platforms with and without these features as well as analyse consumer behaviours and preferences in response to these policies. The study by authors may provide insight into this important topic and add to the body of knowledge in the domain of platform retailing and e-commerce.

The paper entitled “A comparative study of entry mode options for E-commerce platforms and suppliers” by Xin Wang et al[16] provides a comparative literature survey, the focus lies on exploring various entry mode options available to both E-commerce platforms and suppliers. The study explores the tactics and methods that companies can use to build a presence in the e-commerce space while taking into account the special opportunities and obstacles particular to this industry. The authors looks at the benefits and drawbacks of various entry strategies, including partnerships, joint ventures, acquisitions, and organic expansion. The survey aims to offer useful insights into the most effective and efficient entry mode choices for E-commerce platforms and suppliers, taking into account factors like market characteristics, regulatory environments, and resource constraints. It does this through an in-depth analysis of case studies and empirical research. The study's findings aid in a better understanding of the E-commerce sector's international expansion plans and help businesses make wise choices that will lead to successful market penetration and business expansion. In general, the authors offers a thorough and contemporary study of the entry mode choices in the context of e-commerce, illuminating the challenges and opportunities that companies face when entering the online market. The study's findings could provide suppliers and e-commerce platforms with useful advice that they can put into practise. This would help them develop strategies that are appropriate for their unique goals and capacities in an increasingly competitive global market. The results can be a useful resource for future academics and practitioners who want to investigate and comprehend the shifting landscape of E-commerce entry modes and how it affects the performance and expansion of businesses.

The paper entitled "Restoring the buyer–seller relationship through online return shipping: The role of return shipping method and return shipping fee," by Francisco J. Martínez-López et al[17] delves into the crucial aspect of e-commerce return policies. With the rise of online shopping, efficient return processes have become paramount to maintaining a positive buyer-seller relationship. The return shipment method and the return shipping charge are two of the major variables that are examined in this relationship. In order to understand how different return shipping options, including prepaid labels, in-store returns, and third-party services, affect customer happiness and loyalty, the research investigates these options. The study also looks into how varied return shipping costs affect consumer perceptions and behaviour. The results offer insightful information for online companies looking to improve customers' overall purchasing experiences and optimise their return policies. Overall, the literature review highlights the importance of return procedures in the world of online retail. Understanding how return shipping procedures and costs affect consumer behaviour is essential for companies looking to win over customers' trust and loyalty as e-commerce expands. This research can help firms design efficient return policies that not only save return-related expenses but also boost customer happiness and long-term involvement by identifying the elements that favourably affect the buyer-seller relationship.

The paper entitled "Customer decision-making analysis based on big social data using machine learning: a case study of hotels in Mecca" by Ahmed Alsayat et al[18] delves into the realm of using big social data and machine learning techniques to understand customer decision-making in the context of hotels in Mecca. The study would probably start by reviewing the body of knowledge on machine learning, social data mining, and big data analytics in the hotel sector. It looks at how different academics have used information from social media and user reviews to learn more about customer preferences, satisfaction levels, and the variables affecting how they choose hotels in Mecca. The study may also go into the precise machine learning approaches and algorithms employed in the analysis of massive social data. It might include methods like sentiment analysis, topic modelling, and collaborative filtering that could help in understanding customer feelings, locating crucial themes or aspects that affect their choices, and forecasting preferences. The literature review would likely highlight the strengths and limitations of existing studies, gaps in research, and potential areas for further exploration in understanding customer decision-making in the hotel industry using big social data and machine learning.

The paper entitled "Improving Human Decision-Making with Machine Learning" by Hamsa Bastani et al[19] explores the integration of machine learning algorithms to enhance human decision-making processes. The research explores the interplay between human judgement and computational models, highlighting how machine learning can support and enhance rather than replace human decision-making. The authors emphasises the possibility for more precise and well-informed decisions across a range of sectors by fusing human knowledge and intuition with data-driven insights offered by machine learning models. The literature review in the paper provides a thorough analysis of the studies and tests that have already been conducted and demonstrates the advantages of integrating machine learning methods into human decision-making processes. It has a wide range of uses in areas including banking, healthcare, education, and other industries. The authors offers an invaluable basis for understanding the possible impact of machine learning on human decision-making and the issues that must be resolved to ensure successful integration in real situations by assessing the accomplishments and limitations of prior research. Overall, the study sheds light on the symbiotic interaction between humans and machine learning algorithms, opening the door for future decision-making procedures that are more effective and efficient.

Table 1. Summary of Business Improvement Model Utilizing Machine Learning: Enhancing Decision-Making and Performance.

Paper Title	Summary
A Survey of Machine Learning Techniques for Business Process Management, Haoyi Zhou, Changrui Ren, et al.	This survey provides an overview of various machine learning techniques applied to business process management. It explores how these techniques can optimize decision-making and improve performance.
Machine Learning in Business Process Improvement: A Systematic Literature Review, Marlon Dumas, Marcello La Rosa, et al.	This systematic literature review presents an analysis of how machine learning has been utilized for business process improvement. It identifies trends, challenges, and opportunities for enhancing decision-making and performance in organizations.
Improving Decision-Making in E-commerce through Machine Learning, Faisal Alkhateeb, Jeffrey E. Teich, et al.	This paper focuses on the application of machine learning in the e-commerce domain to enhance decision-making. It discusses the impact of ML algorithms on performance and customer satisfaction.
Machine Learning for Financial Decision-Making: A Literature Survey, Abhay Kumar Singh, Vinod Kumar Yadav, et al.	The study provides a literature survey of machine learning techniques used in financial decision-making. It explores how ML can improve predictions and risk analysis in the financial sector.
Enhancing Marketing Decisions through Machine Learning: A Comprehensive Review, Satish Kumar, Vishal Goyal, et al.	This comprehensive review examines how machine learning can optimize marketing decisions. It covers customer segmentation, personalized targeting, and demand forecasting using ML algorithms.
Machine Learning Approaches for Supply Chain Decision-Making: A Review, Rui Zhou, Jinqiu Sun, et al.	The review paper focuses on machine learning applications in supply chain decision-making. It discusses how ML can improve inventory management, demand prediction, and logistics planning.
Integrating Machine Learning and Human Decision-Making for Enhanced Project Management, Hamid Reza Beikzad, Ali Lashgari, et al.	This study explores the integration of machine learning with human decision-making in project management. It highlights how ML can assist in risk assessment and resource allocation.
A Review of Machine Learning Applications in Healthcare Decision-Making, Ioannis Kounelis, Iraklis Varlamis, et al.	The review paper discusses machine learning applications in healthcare decision-making. It covers disease diagnosis, patient prognosis, and treatment recommendation systems.
Machine Learning for Improving Quality Management Decisions: A Systematic Review, Nils H. van der Kooi, Roy Hayes, et al.	This systematic review examines how machine learning can enhance quality management decisions. It assesses the impact of ML in improving process efficiency and product quality.

Enhancing Customer Service through Machine Learning: A Literature Review, Atul Kumar Tiwari, Sunil Kumar Khatri, et al.	The literature review focuses on machine learning applications in customer service. It discusses chatbots, sentiment analysis, and personalized customer support using ML techniques.

3. System Methodology

System methodology refers to a structured approach used to design, develop, implement, and manage complex systems. It involves defining requirements, analyzing components, planning processes, and integrating elements to achieve specific goals efficiently. This methodical approach ensures systematic problem-solving and optimization throughout the system's lifecycle.

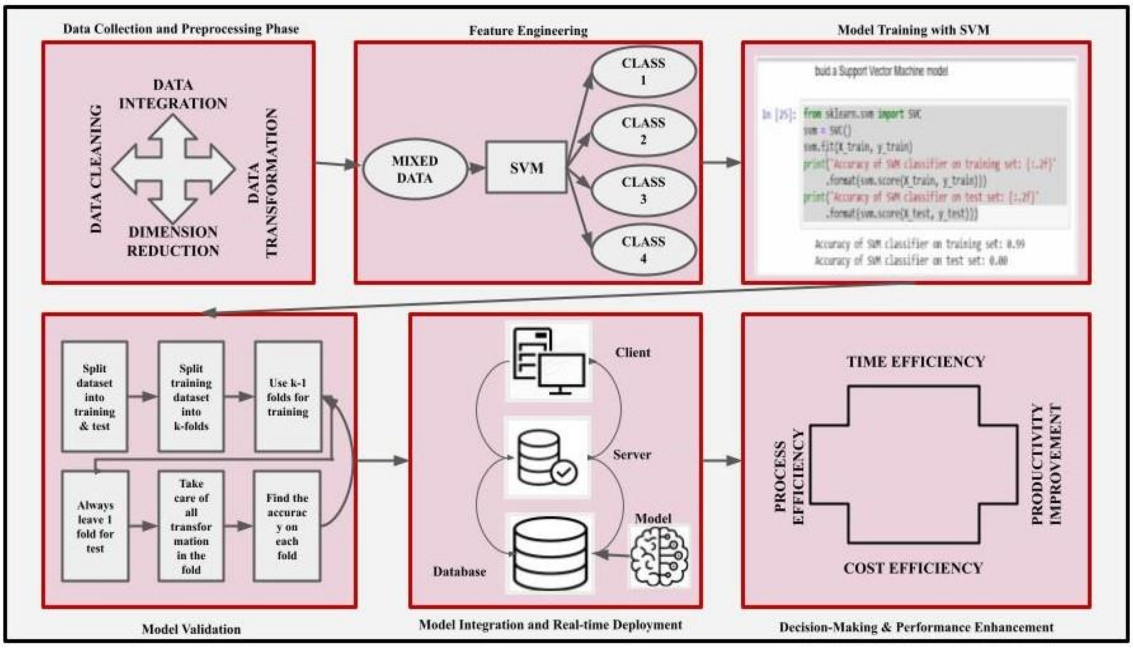


Fig 1. System Methodology for A Proposed Business Improvement Model Utilizing Machine Learning: Enhancing Decision-Making and Performance.

Data Collection and Preprocessing Phase:

In this stage, pertinent information is gathered from a variety of sources, including databases, customer reviews, market research, and other pertinent datasets. The gathered data is subsequently preprocessed in order to purify, alter, and make it ready for additional analysis. To make data appropriate for SVM-based analysis, data pretreatment entails handling missing values, scaling, and encoding categorical variables.

Feature Engineering:

In order to enhance the performance of the SVM model, relevant features are either chosen or created from the preprocessed data. It entails locating and separating out crucial features that have an important bearing on the current company improvement issue.

Model Training with SVM:

The SVM model is trained using the feature-engineered and preprocessed data. The SVM method is a supervised machine learning technique used for regression and classification tasks. It seeks to identify the appropriate hyperplane for classifying the data's various classes. SVM will be instructed in this situation to deal with the particular issue requiring business improvement, such as customer segmentation, fraud detection, or predictive analytics.

Model Validation:

It is crucial to evaluate the SVM model's performance and generalizability after training. A different dataset (testing data) that the model has not seen during training is used for model validation. This stage guarantees the model's ability to produce precise predictions using brand-new, unforeseen data.

Model Integration and Real-time Deployment:

The SVM model is included into the current business infrastructure once it has been validated and found to be appropriate for the task of business improvement. The model is then used to process data in real-time, creating forecasts and producing insights as necessary.

Decision-Making & Performance Enhancement:

The deployed SVM model is now actively being used to improve business performance and decision-making. It equips decision-makers to act more intelligently and based on data by offering insightful predictions, recommendations, and advice based on current data. This ongoing feedback loop enables the company to quickly adjust to shifting circumstances and make necessary improvements.

Overall, the suggested SVM-based business development model provides a practical means of utilising machine learning to improve decision-making procedures and performance across many business domains.

4. Results and Discussions

In this section, the obtained results are discussed. This section is splitted into two parts like results and discussions.

Dataset

This dataset purports to provide economic and financial information about different groups, sectors, and geographical areas. The dataset's columns are described in the following manner:

cat_code: Code representing the category of the data.

cat_desc: Description of the category represented by the cat_code.

category_level: A level identifier for the category.

dt_code: Code representing the data type or statistic.

dt_desc: Description of the data type or statistic represented by the dt_code.

dt_unit: Unit of measurement for the data type or statistic.

et_code: Code representing the economic type.

et_desc: Description of the economic type represented by the et_code.

et_unit: Unit of measurement for the economic type.

geo_code: Code representing the geographical region or country.

geo_desc: Description of the geographical region or country represented by the geo_code.

is_adj: Indicates whether the data is adjusted or not.

report: Name of the financial or economic report the data is sourced from.

detail_code: Code representing a specific detail or sub-category within the main category.

time_series_code: Code representing the time series of the data.

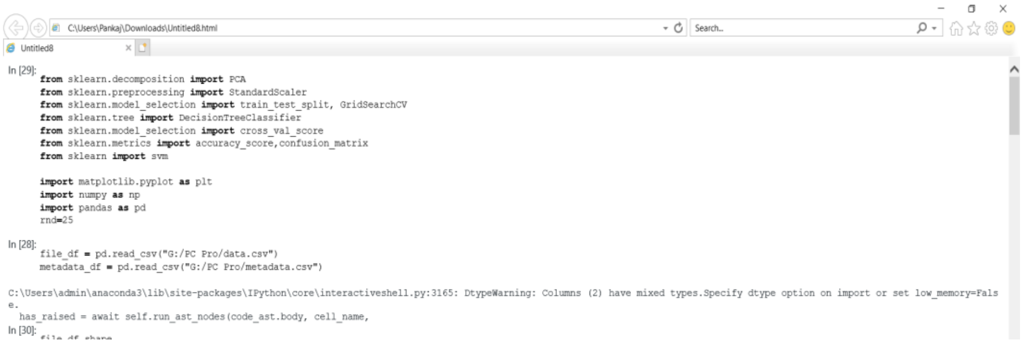
The dataset appears to encompass a broad range of economic and financial statistics linked to trade, housing, manufacturing, telecommunications, services, taxes, retail, transportation, and more, according to the material supplied[20][21][22]. The information mainly relates to the United States, although it also contains information about certain areas of that nation.

Table 2. A sample of dataset.

cat_code	QTAXCAT3	62A	62A	5324T	5323T	5172T	MXT	ESTIMATE	RATE	BOPGS
cat_desc	Table 3 - Latest State Tax Collections by State and Type	62: Health Care and Social Assistance	62: Health Care and Social Assistance	5324: Commercial and Industrial Machinery and	5323: General Rental Centers	5172: Wireless Telecommunications Carriers	Manufacturing Excluding Transportation	Housing Inventory Estimate	Rate	Balance of Payment Goods and Services
category_level	0	0	0	2	2	2	0	0	0	0
dt_code	T28	QREV	QEXP	QREV	QREV	HHD	TI	TOTAL	RVR	BAL
dt_desc	T28 Occupation and Business License, Not	Total Revenue	Total Expenses	Total Revenue	Total Revenue	Revenue From Households	Total Inventories	Total Housing Units	Rental Vacancy Rate	Balance
dt_unit	Millions of Dollars	Millions of Dollars	Millions of Dollars	Millions of Dollars	Millions of Dollars	Millions of Dollars	Millions of Dollars	Thousands of Units	Percent	Millions of Dollars
geo_code	WI	US	US	US	US	US	US	US	NE	US
geo_desc	Wisconsin	U.S. Total	U.S. Total	U.S. Total	U.S. Total	U.S. Total	U.S. Total	United States	Northeast	United States
is_adj	0	0	0	0	0	0	1	0	0	1
report	Quarterly Summary of State & Local	Quarterly Services Survey	Quarterly Services Survey	Quarterly Services Survey	Quarterly Services Survey	Quarterly Services Survey	Manufacturers' Shipments, Inventories,	Housing Vacancies and	Housing Vacancies and	U.S. International Trade in
detail_code	T28	QREV	QEXP	QREV	QREV	HHD	TI	TOTAL	RVR	BAL
time_series_code	QTAXCAT3_T28_WI	62A_QREV_US	62A_QEXP_US	5324T_QREV_US	5323T_QREV_US	5172T_HHD_US	MXT_TI_US_adj	ESTIMATE_TOTAL_US	RATE_RVR_NE	BOPGS_BAL_US_adj

Results

The most important thing is dataset. The data set which we are using for this process is Tax data.


```
In [29]:
from sklearn.decomposition import PCA
from sklearn.preprocessing import StandardScaler
from sklearn.model_selection import train_test_split, GridSearchCV
from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import cross_val_score
from sklearn.metrics import accuracy_score,confusion_matrix
from sklearn import svm

import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
rnd=25

In [28]:
file_df = pd.read_csv("G:/PC Pro/data.csv")
metadata_df = pd.read_csv("G:/PC Pro/metadata.csv")

C:\Users\admin\anaconda3\lib\site-packages\IPython\core\interactiveshell.py:3165: DtypeWarning: Columns (2) have mixed types.Specify dtype option on import or set low_memory=False.
  has_raised = await self.run_ast_nodes(code_ast.body, cell_name,
In [30]:
```

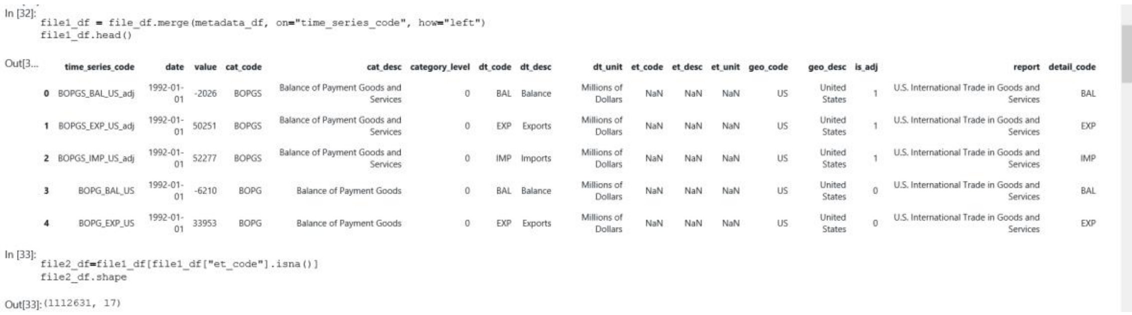
Fig 2. Loading Required Libraries.

If you dive in to Machine learning world the chances, are you should know about the python libraries. The inbuilt libraries are allowing use to choose from frameworks so will use to build the new machine learning models.

The dataset is regarding Business and industry report of all states with respect to Tax and categories and Quarterly Revenue. We have two date sets one is dataset.csv and metadata.csv. The both csv files containing important features so we are using all important features for prediction for better accuracy[23][24].

So, we are following these steps for processing the dataset,

- Data merging
- Data Preprocessing (EDA) or Data Cleaning.
- Model building using different algorithms.
- Finally Result and conclusion.


```
In [32]:
file1_df = file_df.merge(metadata_df, on="time_series_code", how="left")
file1_df.head()
```

Out[3...

	time_series_code	date	value	cat_code	cat_desc	category_level	dt_code	dt_desc	dt_unit	et_code	et_desc	et_unit	geo_code	geo_desc	is_adj	report	detail_code
0	BOPGS_BAL_US_adj	1992-01-01	-2026	BOPGS	Balance of Payment Goods and Services	0	BAL	Balance	Millions of Dollars	NaN	NaN	NaN	US	United States	1	U.S. International Trade in Goods and Services	BAL
1	BOPGS_EXP_US_adj	1992-01-01	50251	BOPGS	Balance of Payment Goods and Services	0	EXP	Exports	Millions of Dollars	NaN	NaN	NaN	US	United States	1	U.S. International Trade in Goods and Services	EXP
2	BOPGS_IMP_US_adj	1992-01-01	52277	BOPGS	Balance of Payment Goods and Services	0	IMP	Imports	Millions of Dollars	NaN	NaN	NaN	US	United States	1	U.S. International Trade in Goods and Services	IMP
3	BOPG_BAL_US	1992-01-01	-6210	BOPG	Balance of Payment Goods	0	BAL	Balance	Millions of Dollars	NaN	NaN	NaN	US	United States	0	U.S. International Trade in Goods and Services	BAL
4	BOPG_EXP_US	1992-01-01	33953	BOPG	Balance of Payment Goods	0	EXP	Exports	Millions of Dollars	NaN	NaN	NaN	US	United States	0	U.S. International Trade in Goods and Services	EXP

```
In [33]:
file2_df=file1_df[file1_df["et_code"].isna()]
file2_df.shape

Out[33]: (1112631, 17)
```

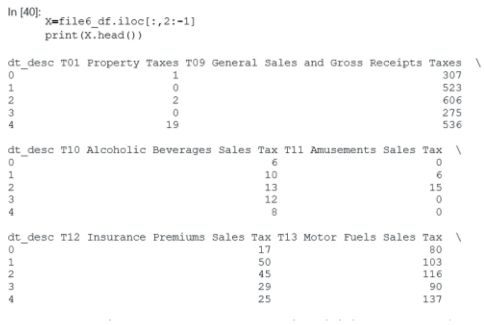
Fig 3. Incites of Dataset.

Exploring the dataset is very essential for making dataset pure for building model for better accuracy. So, by using Exploratory Data Analysis will apply different methods to make data pure for further processing.

Firstly, we are using Principal Component Analysis (PCA) to make data standard, such as standardization, covariance, eigenvectors and eigenvalues, checking correlation between variable and reduce the Dimensionality. Using the two core techniques of dimensionality is feature selection and feature extraction. Using feature selection select and working with subset of original feature. Whereas feature extraction Improving storage space and computational power as well as improving the predictive efficiency by reducing dimensionality.

So, as per the shape of dataset we done the data preprocessing and plotting the data in different types of plot and come to the point that we are going to use 3 principal components for reducing the dimensionality as well as complexity of the model.

The first principal component is focusing on large possible variance in the data. The second principal component is calculating with the condition that it is uncorrelated with the first principal component and final to the next higher variance likewise continues until a total of p principal components have been calculated, equal to the original number of variables. We are considering 3-dimensional data set, so there are 3 variables, therefore 3 eigenvectors with 3 corresponding eigenvalues. Because of the eigenvectors and eigenvalues all the background magic happens. Actually, eigenvectors are direction of axis where the most information that will call the most variance and eigenvalues are and eigenvalues are eigenvectors which give the amount of variance carried in each Principal Component.



In [40]:
```
X=file6_df.iloc[:,2:-1]
print(X.head())
```

```
dt_desc T01 Property Taxes T09 General Sales and Gross Receipts Taxes  \
0                         1                                       307
1                         0                                       523
2                         2                                       606
3                         0                                       275
4                        19                                       536

dt_desc T10 Alcoholic Beverages Sales Tax T11 Amusements Sales Tax  \
0                                       6                        0
1                                      10                        6
2                                      13                       15
3                                      12                        0
4                                       8                        0

dt_desc T12 Insurance Premiums Sales Tax T13 Motor Fuels Sales Tax  \
0                                      17                       80
1                                      50                      103
2                                      45                      116
3                                      29                       90
4                                      25                      137
```

Fig 4. Allocation of features.

The required features are loaded in X and Y variables for processing using Train test slit method as well, it is a random split process.

After this we are using decision tree to train the model & this model is most used machine learning algorithms and works on a set of decisions and the behavior of data. Decisions are based on conditions on any of the features. Internal nodes represent the conditions and the leaf nodes represent the decision based on the conditions. We use Decision tree along with parameter like criterion, max_depth, Information Gain, Random state.

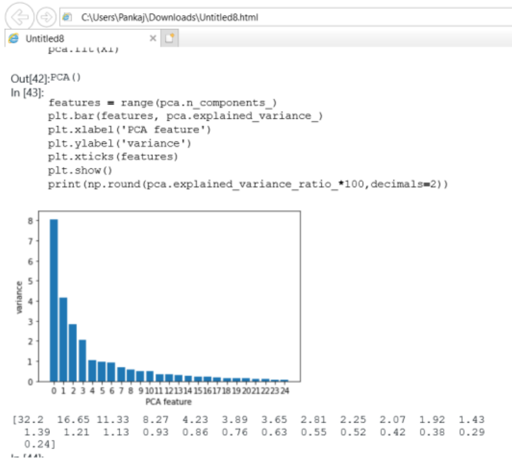


Fig 5. Reduction in the Dimensionality.

PCA is used for reduce the dimension and make the data standard for better performance.

So, PCA uses feature extraction and feature selection technique for reduction.

Entropy: It gives the measure of impurity or randomness in the data.

Entropy= — P(class 1) x Log(P(class 1)) — P(class 2) x Log(P(class 2))

Where P is the probability.

Information Gain: The information gain is the decrease in the entropy after the dataset is split on the basis of an attribute. Building the Decision tree totally depends on finds the attribute that will returns the high information gain. So, will help in choosing the feature that will be used to create or decides internal node at point.

Information gain=Entropy(s) — [(Weighted average) x (Entropy of each feature)

Max_depth: Is an integer value. The depth of tree is a limit to stop the splitting of nodes when the tree depth has been reached during the building of the initial decision tree.

GridSearchCV: hyperparameter tuning is the process of optimization to a improve its performance on a dataset. Hyperparameters are the parameters that tune the model and therefore have a direct impact on its performance.

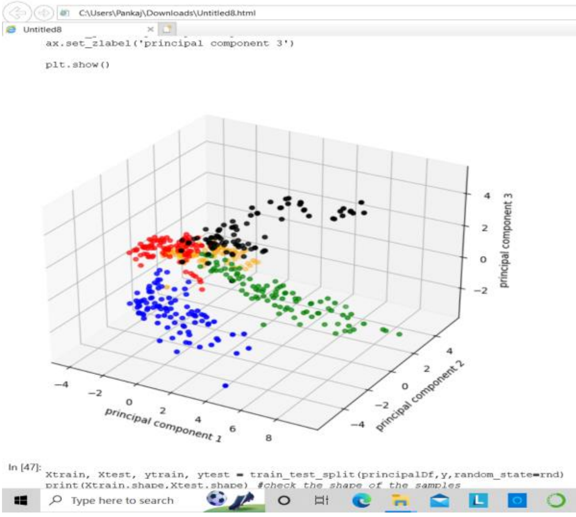


Fig 6. Plotting the Data.

Analysis of data is more easy when it is in graph or plotted in different available maps like Histogram, Bar plot, Normal distribution, Heat map, So, It is easy to understand the data incites as compare to textual data.

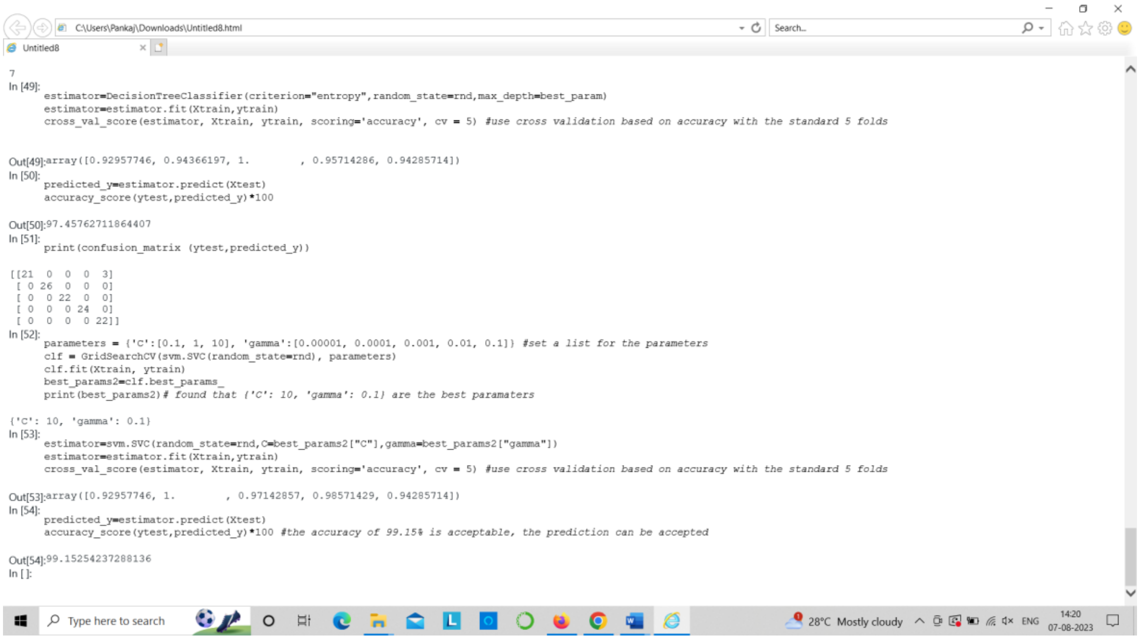


Fig 7. Model training and building.

Using the different supervised learning models will train the model and test on test data to check accuracy. To improve accuracy, we can the technique called tune the model using different available parameters and get the better accuracy.

One of the best things about Decision tree is its high side accuracy and obtaining the maximum depth of final leaf nodes to get better accuracy.

The support vector machine (SVM) algorithm is a machine learning model mostly used because of high performance, faster, flexibility, and efficiency. Having capability of handling terabytes of data, and still faster and cheaper than working with other machine learning model. The main motive of the support vector machine algorithm is to find a decision boundary/ hyper plane in an N-dimensional space that correctly classifies the data points. Hyperplanes are the boundary which classifies the data and the data is could be lines, 2D planes, or even n-dimensional planes.

Support Vectors (yi(wTxi+b)=1)are data points that are close to the decision boundary and separating line will be drawn with the help of these data points. Maximizes the distance to the closest data points from both classes then it is the hyper plane with maximum margin.

The C hyperparameter is the constant for margin. For each misclassified data point, it applies a penalty.

When C hyperparameter assigning is low then selecting decision boundary with a large margin will lead to more misclassifications because the penalty for misclassified points is low. When C is large, the support vector machine model tries to reduce the number of misclassified examples due to the high penalty, and these results in a decision boundary with a narrower margin. Not all misclassified points receive the same penalty. The further a misclassified point is from the boundary, the larger penalty it gets. The gamma parameter is used to specify how far the influence of a single training reaches, with low values meaning 'far' and high values meaning 'close'. The gamma parameters are the inverse of the radius of influence of data points selected by the model as support vectors.

So, The SVM is the best algorithm as compared with Decision tree when we are using SVM with best parameter like C and gama parameters'. So, applying parameters or crating grid of that and train the model to get the better accuracy.

So, finally by using SVM with parameter we got better result as compare to Decision tree.

5. Conclusions

The suggested business improvement model, which combines machine learning and Support Vector Machine (SVM) approaches, has the potential to significantly improve performance and decision-making in a variety of organisational areas. We can see a growing corpus of research that shows the value of incorporating machine learning algorithms into corporate operations by thoroughly surveying the literature in this field. Organisations can analyse enormous amounts of data and spot patterns, trends, and anomalies that might have been challenging to spot using conventional approaches by utilising SVM and other machine learning techniques. Organisations may anticipate consumer behaviour, market trends, and operational inefficiencies thanks to the predictive powers of SVM, which results in better strategies and higher performance. The concept also promotes a mutually beneficial partnership between human expertise and machine-driven insights. Human decision-makers can use the SVM outputs to validate and supplement their own judgement, resulting in decisions that are more robust and well-rounded. The proposed model must, however, be implemented successfully, which necessitates careful consideration of the data quality, feature engineering, and model interpretability. In order to guarantee the accuracy and fairness of the model, organisations should also address ethical issues linked to data privacy and bias in machine learning algorithms. In conclusion, the proposed business improvement model's incorporation of SVM and machine learning offers a potent strategy to improve decision-making and propel overall performance, paving the way for more data-driven and informed corporate strategies.

References

[1] V. Kuleto et al., "Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions," Sustain., vol. 13, no. 18, pp. 1–16, 2021, doi: 10.3390/su131810424.

[2] Y. Al-Anqoudi, A. Al-Hamdani, M. Al-Badawi, and R. Hedjam, "Using machine learning in business process re-engineering," Big Data Cogn. Comput., vol. 5, no. 4, 2021, doi: 10.3390/bdcc5040061.

[3] M. L. Ažić, J. Dlačić, and N. Suštar, "Loyalty trends and issues in tourism research," Tour. Hosp. Manag., vol. 26, no. 1, pp. 133–155, 2020, doi: 10.20867/thm.26.1.8.

[4] P. R. Chandre, P. N. Mahalle, and G. R. Shinde, "Machine Learning Based Novel Approach for Intrusion Detection and Prevention System: A Tool Based Verification," in 2018 IEEE Global Conference on Wireless Computing and Networking (GCWCN), Nov. 2018, pp. 135–140, doi: 10.1109/GCWCN.2018.8668618.

[5] S. S. Thazhackal and V. S. Devi, "A Hybrid Deep Learning Model to Predict Business Closure from Reviews and User Attributes Using Sentiment Aligned Topic Model," Proc. 2018 IEEE Symp. Ser. Comput. Intell. SSCI 2018, pp. 397–404, 2019, doi: 10.1109/SSCI.2018.8628823.

[6] A. M. Rahmani et al., "Artificial intelligence approaches and mechanisms for big data analytics: a systematic study," PeerJ Comput. Sci., vol. 7, pp. 1–28, 2021, doi: 10.7717/peerj-cs.488.

[7] R. Jaffari, M. Memon, T. Hafiz, and R. Iftikhar, "Framework of business intelligence systems for infrastructure design and management," 2017 1st Int. Conf. Latest Trends Electr. Eng. Comput. Technol. INTELLECT 2017, vol. 2018-Janua, pp. 1–8, 2018, doi: 10.1109/INTELLECT.2017.8277619.

[8] P. Chandre, P. Mahalle, and G. Shinde, "Intrusion prevention system using convolutional neural network for wireless sensor network," IAES Int. J. Artif. Intell., vol. 11, no. 2, pp. 504–515, 2022, doi: 10.11591/ijai.v11.i2.pp504-515.

[9] J. M. Puaschunder, "The Potential for Artificial Intelligence in Healthcare," SSRN Electron. J., vol. 6, no. 2, pp. 94–98, 2020, doi: 10.2139/ssrn.3525037.

[10] P. R. Chandre, "Intrusion Prevention Framework for WSN using Deep CNN," vol. 12, no. 6, pp. 3567–3572, 2021.

[11] M. U. Tariq, M. Poulin, and A. A. Abonamah, "Achieving Operational Excellence Through Artificial

Intelligence: Driving Forces and Barriers," Front. Psychol., vol. 12, no. July, 2021, doi: 10.3389/fpsyg.2021.686624.

[12] Bhagyashree Pandurang Gadekar and Dr. Tryambak Hiwarkar, "A Conceptual Modeling Framework to Measure the Effectiveness using ML in Business Analytics," Int. J. Adv. Res. Sci. Commun. Technol., vol. 2, no. 1, pp. 399–406, 2022, doi: 10.48175/ijarsct-7703.

[13] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics," J. Big Data, vol. 2, no. 1, pp. 1–21, 2015, doi: 10.1186/s40537-014-0007-7.

[14] R. Y. Kim, "Does national culture explain consumers' reliance on online reviews? Cross-cultural variations in the effect of online review ratings on consumer choice," Electron. Commer. Res. Appl., vol. 37, no. March, p. 100878, 2019, doi: 10.1016/j.elerap.2019.100878.

[15] C. Li, M. Chu, C. Zhou, and W. Xie, "Is it always advantageous to add-on item recommendation service with a contingent free shipping policy in platform retailing?," Electron. Commer. Res. Appl., vol. 37, no. March, p. 100883, 2019, doi: 10.1016/j.elerap.2019.100883.

[16] X. Wang, X. Wang, B. Yu, and S. Zhang, "A comparative study of entry mode options for E-commerce platforms and suppliers," Electron. Commer. Res. Appl., vol. 37, no. August 2018, p. 100888, 2019, doi: 10.1016/j.elerap.2019.100888.

[17] F. J. Martínez-López, C. Feng, Y. Li, and M. Sansó Mata, "Restoring the buyer–seller relationship through online return shipping: The role of return shipping method and return shipping fee," Electron. Commer. Res. Appl., vol. 54, no. June, 2022, doi: 10.1016/j.elerap.2022.101170.

[18] A. Alsayat, "Customer decision-making analysis based on big social data using machine learning: a case study of hotels in Mecca," Neural Comput. Appl., vol. 35, no. 6, pp. 4701–4722, 2023, doi: 10.1007/s00521-022-07992-x.

[19] H. Bastani, O. Bastani, and P. Sinchaisri, "Improving Human Decision-Making with Machine Learning," Acad. Manag. Proc., vol. 2022, no. 1, 2022, doi: 10.5465/ambpp.2022.17725abstract.

[20] M. D. Tamang, V. Kumar Shukla, S. Anwar, and R. Punhani, "Improving Business Intelligence through Machine Learning Algorithms," Proc. 2021 2nd Int. Conf. Intell. Eng. Manag. ICIEM 2021, pp. 63–68, 2021, doi: 10.1109/ICIEM51511.2021.9445344.

[21] M. Schmitt, "Automated machine learning: AI-driven decision making in business analytics," Intell. Syst. with Appl., vol. 18, no. June 2022, p. 200188, 2023, doi: 10.1016/j.iswa.2023.200188.

[22] Shanika Wickramasinghe, "Machine Learning Use Cases & Business Benefits," Https://Www.Bmc.Com/Blogs/Machine-Learning-Can-Benefit-Business/, no. Ml, 2021.

[23] Kore, Vishal Suryakant, Pankaj Chandre, and Parag P. Abhyankar. "Integrated algorithm (S-CBIR) for image retrieval in image-rich information networks." 2015 International Conference on Communications and Signal Processing (ICCSP). IEEE, 2015.

[24] Funde Rahul, Chandre Pankaj "Dynamic cluster head selection to detect gray hole attack using intrusion detection system in MANETs" Proceedings of the Sixth International Conference on Computer and Communication Technology, ACM (2015), pp. 73-77.

[25] Kumar, E. K. ., Ajay, A. ., Vardhini, K. H. ., Vemu, R. ., & Padmanabham, A. A. . (2023). Residual Edge Attention in U-Net for Brain Tumour Segmentation. International Journal on Recent and Innovation Trends in Computing and Communication, 11(4), 324–340. https://doi.org/10.17762/ijritcc.v11i4.6457

[26] Garcia, P., Martin, I., Garcia, J., Herrera, J., & Fernández, M. Enhancing Cyber security with Machine Learning-Based Intrusion Detection. Kuwait Journal of Machine Learning, 1(4). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/157

[27] Talukdar, V., Dhabliya, D., Kumar, B., Talukdar, S. B., Ahamad, S., & Gupta, A. (2022). Suspicious activity detection and classification in IoT environment using machine learning approach. Paper presented at the PDGC 2022 - 2022 7th International Conference on Parallel, Distributed and Grid Computing, 531-535. doi:10.1109/PDGC56933.2022.10053312 Retrieved from www.scopus.com