

VISVESVARAYA TECHNOLOGICAL UNIVERSITY
“JNANA SANGAMA” BELAGAVI-590018.



An Internship Report On

“BANK LOAN STATUS PREDICTION”

**SUBMITTED IN PARTIAL FULFILMENT FOR THE AWARD OF THE DEGREE OF BACHELOR
OF ENGINEERING**

IN

COMPUTER SCIENCE AND ENGINEERING

By

INCHARA S GOWDA[1JB20CS039]

UNDER THE GUIDANCE OF

Internal Guide

Dr. Krishna A N
Prof and Head
Dept. of CSE, SJBIT
Bengaluru

External Guide

Dr. Ganga Holi
Supervisor
Vaazu India Pvt Ltd.
Bengaluru



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SJB INSTITUTE OF TECHNOLOGY

#67, BGS HEALTH AND EDUCATION CITY, Dr Vishnuvardhan Road, Kengeri, Bengaluru-560060.

2023-2024

||Jai Sri Gurudev||
Sri Adichunchanagiri Shikshana Trust ®

SJB INSTITUTE OF TECHNOLOGY

#67, BGS Health & Education City, Dr Vishnuvardhan Road, Kengeri, Bangalore-560060.



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

Certified that the Internship work entitled “**BANK LOAN STATUS PREDICTION**” carried out by **Inchara S Gowda [1JB20CS039]** is a bonafide student of SJB Institute of Technology in partial fulfillment for the award of “**BACHELOR OF ENGINEERING**” in **COMPUTER SCIENCE AND ENGINEERING** as prescribed by **VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI** during the academic year 2023-2024. It is certified that all corrections /suggestions indicated for internal assessment have been incorporated in the report deposited in the departmental library. The internship report has been approved as it satisfies the academic requirements in respect of Internship work prescribed for the said degree.

Signature of Guide

Dr. Krishna A N
Prof and Head
Dept. of CSE, SJBIT

Signature of HOD

Dr. Krishna A N
Prof and Head
Dept. of CSE, SJBIT

Signature of Principal

Dr. K V Mahendra Prashanth
Principal
SJBIT, Bangalore

EXTERNAL VIVA

Name of the Examiners

Signature with Date

1. _____

2. _____



ACKNOWLEDGEMENT

I would like to express my profound gratefulness to His Divine Soul Jagadguru Padmabhushana **Sri Sri Sri Dr. Balagangadharanatha Maha Swamiji** and His Holiness Jagadguru **Sri Sri Sri Dr. Nirmalanandanatha Maha Swamiji** for providing me an opportunity to complete my academics in this esteemed institution.

I would like to express my profound thanks to **Revered Sri Sri Dr. Prakashnath Swamiji**, Managing Director, SJB Institute of Technology, for his continuous support in providing amenities to carry out this project in this admired institution.

I express my gratitude to **Dr. K. V. Mahendra Prashanth**, Principal, SJB Institute of Technology, for providing me an excellent facilities and academic ambience, which has helped me in the satisfactory completion of the Internship.

I extend my sincere thanks to **Dr. Krishna A N**, Professor and Head, Department of Computer Science and Engineering, SJB Institute of Technology, for providing me an invaluable support throughout the period of Internship.

I express my deepest gratitude and sincere thanks to my guide **Dr. Krishna A N**, Professor and Head, Department of Computer Science and Engineering, SJB Institute of Technology, I also would like to thank **Dr. Ajay B N**, Associate Professor, Department of Computer Science and Engineering, SJB Institute of Technology and Internship Coordinator, who has extended his support throughout the internship duration and **Dr. Ganga Holi**, Supervisor, Vaazu India Pvt Ltd. for their valuable guidance and support.

Finally, I take this opportunity to extend my earnest gratitude and respect to my parents, the teaching & non-teaching staff of the department, the library staff and all my friends, who have directly or indirectly supported me during the period of my internship.

Regards,
INCHARA S GOWDA [1JB20CS039]

COMPANY CERTIFICATE

1084 6th Main 3rd Stage
BEML Layout
RajaRajeshwari Nagar
Bangalore 560098
info@vaazuinc.com
CIN: U72200KA2012PTC066439



Date : 26/09/2023

CERTIFICATE OF INTERNSHIP

Congratulations!

This certificate is provided to **INCHARA S GOWDA** of **SJB Institute Of Technology**, for successful completion of internship at **Vaazu India Private Limited, Bangalore** under the guidance of **Dr Ganga Holi** for 4 weeks i.e., from 16th Aug 2023 to 15th Sept 2023.

During the period of internship with us she had been exposed to the practicality and implications of the concepts of "**Machine learning with Python**" which included training and hands on experience.

We wish you all the success in your future endeavors.

Yours,

For Vaazu India Pvt Ltd,

A handwritten signature in black ink, appearing to read "Sudheer Akkisetty", with a horizontal line extending to the right.

Sudheer Akkisetty

Chief Executive Officer

Confidential.

DECLARATION BY THE STUDENT

I, **INCHARA S GOWDA [1JB20CS039]**, student of 7th semester Computer Science and Engineering, **SJB INSTITUTE OF TECHNOLOGY**, Bangalore, hereby declare that the internship entitled “**BANK LOAN STATUS PREDICTION**” submitted to the **Visvesvaraya Technological University, Belagavi** during the academic year **2023-24**, is a record of an original work done by me under the guidance of my Internal guide **Dr. Krishna A N**, Prof and Head, Department of Computer Science and Engineering, SJB Institute of Technology, Bangalore and my external guide, **Dr. Ganga Holi**, Supervisor, Vaazu India Pvt Ltd. This internship dissertation report is submitted in partial fulfillment for the award of Computer Science and Engineering. The results embodied in this report have not been submitted to any other University or Institute for the award of any degree.

Date: 23/12/2023

Place: Bangalore

INCHARA S GOWDA[1JB20CS039]

EXECUTIVE SUMMARY

The internship at Vaazu India Private Limited, provided a comprehensive overview of machine learning focused on predicting bank loan status using various models implemented in Python with the assistance of the sklearn library. During my one-month internship, I engaged in a comprehensive exploration of machine learning applications using Python, specifically focusing on the domain of banking for loan status prediction. The primary objective of this endeavor is to enhance the accuracy and efficiency of loan approval processes within financial institutions.

The internship involved not only the implementation of machine learning models but also a meticulous process of data-preprocessing, employing algorithms such as decision trees, support vector machines and others and evaluating each model's performance to ascertain their effectiveness in predicting loan outcomes. The most suitable model was determined based on the ability to accurately predict loan statuses.

TABLE OF CONTENTS

Acknowledgement	i
Company certificate	ii
Declaration by the student	iii
Executive summary	iv
Table of contents	v
List of figures	vi

CHAPTER NO	PARTICULARS	PAGE NO
1	ABOUT THE COMPANY	1-3
	1.1 Technology services	1-3
	1.2 Recruitment services	3-3
2	ABOUT THE DEPARTMENT	4-5
	2.1 Key benefits	4-5
3	TASKS PERFORMED	6-12
	3.1 Introduction	6
	3.2 Problem statement	6
	3.3 System specifications	6-7
	3.4 Design	7-9
	3.5 Implementation	9-12
4	REFLECTION NOTES	13-15
	4.1 Specific outcomes	13-15
	CONCLUSION	16
	REFERENCES	17

LIST OF FIGURES

FIGURE NO	FIGURE DESCRIPTION	PAGE NO
Fig 1.1	Logo of Vaazu India Private Limited	1
Fig 3.1	Flow chart of Bank loan status Predictor	9
Fig 4.1	Testing set accuracy for various models	13
Fig 4.2	Comparison of accuracy values	14
Fig 4.3	Loan approved status for logistics regression prediction	14
Fig 4.4	Loan rejected status for logistics regression prediction	14
Fig 4.5	Loan approved status for decision tree prediction	15
Fig 4.6	Loan rejected status for decision tree prediction	15

CHAPTER 1

ABOUT THE COMPANY

Vaazu India Private Limited is a private company incorporated on 25 October 2012. It is classified as Non-govt company and is registered at Registrar of Companies, Bangalore. It is involved in Software publishing, consultancy and supply. Software publishing includes production, supply and documentation of ready-made (non-customized) software, operating systems software, business & other applications software, computer games software for all platforms. Consultancy includes providing the best solution in the form of custom software after analysing the user's needs and problems. Custom software also includes made-to-order software based on orders from specific users. Also, included are writing of software of any kind following directives of the users; software maintenance, web-page design. The company provides high- value outsourced global workforce solutions and consulting services.



Fig 1.1: Logo of Vaazu India Private Limited

1.1 TECHNOLOGY SERVICES

➤ Software Development

Vaazu provides a wide range of experienced developers, fully licensed in key areas like desktop, mobile, reporting, and web application development. Delivers custom, scalable, cost-effective cutting-edge solutions for global organizations and technology startups. Vaazu's expertise spans all major technologies and platforms, and advances to innovative technology trends. The solutions help organizations gain better customer understanding to foster innovation for their brand, products, and consumer experiences while solving both strategic and tactical business challenges.

- Enterprise Applications
 - Enterprise Data Management
 - Collaboration and Document Management solutions
 - Business Intelligence
 - Software Product Development
-

➤ **Blockchain Consulting**

Rewrite the rules, unlock the value and enable real-time, frictionless data exchange with our extremely efficient blockchain applications. Giving solution, though comprehensive, is an MVP—minimum viable product—version that is intended for rapid deployment, testing, and adjustment.

- Consulting & POC
- DApp & Web3 Services
- NFT And Marketplace
- New Cryptocurrency/Token
- DeFi (Decentralized Finance)

➤ **Managed IT Services**

Ensure every component of your office network runs smoothly with Managed IT Services. Focus on your core business and not waste time on tools and technology, provide managed IT solutions to the businesses, which help them focus more on their core business instead of encountering day-to-day technology operations and management issues. Microsoft Azure, AWS, Google Cloud are few of the renowned technology partners, which aids us in providing effective infrastructure services on cloud to our associated enterprises.

- Web Filtering & Protection
- Anti-Virus and Anti-Malware Protection
- Cloud Management
- Remote Monitoring & Maintenance
- HelpDesk Support Service

➤ **Microsoft Workspace Solutions**

Vaazu provides cloud applications development, cloud architecture design, cloud migration, DevOps, ML & AI powered applications such as SaaS, PaaS, and IaaS. Optimizing Microsoft technology investments with Microsoft modern workplace solutions and provide a Cloud-focused assessment, strategy, and roadmap for a successful Cloud project.

- Application Modernization
- Cloud Migration Services
- Software Support and Maintenance
- Power Platform Services
- Share Point Mitigation

- Modern Collaboration Tools
- Support Services

1.2 RECRUITMENT SERVICES

➤ **IT Staffing**

The Vaazu IT Staffing team deploys highly trained, certified IT specialists backed by experienced team of professionals. As a technology company, the company understands client needs from a technical perspective, ensure that every engagement is successful.

➤ **Recruitment Processing Outsourcing**

Our Enterprise RPO solution looks after the complete recruitment process taking care of every piece of the talent puzzle – attracting candidates, screening, and assessing them, scheduling interviews, offer negotiations and onboarding colleagues, monitoring compliance, and conducting exit interviews.

CHAPTER 2

ABOUT THE DEPARTMENT

Vaazu India Private Limited offers a valuable internship opportunity focused on Recruitment Process Outsourcing (RPO) in the field of Machine learning using Python. Gaining practical experience, enhancing skills in a supportive learning environment and exploring the machine learning technologies with hands-on projects elevates the career prospects with this insightful internship. Internship was provides under Recruitment Services department.

- **Enterprise RPO:**

Outsource your talent acquisition process and we create tailored solutions designed to meet your organization's specific challenges.

- **Project RPO:**

Get flexible recruitment support during peak and slow times for short- term, seasonal, or specialty hiring needs. We will match the right talent for your special circumstances.

- **On-demand RPO:**

When speed is critical, leverage our Rapid Response on-demand team of accomplished recruiters to fill your short-term need.

- **Hybrid RPO:**

Maintain your desired degree of control and oversight by outsourcing some of the more challenging or time-consuming aspects of recruitment.

- **Bench Marketing:**

Speed up the placements with our bench marketing expertise! We map your bench candidates with their potential employers easily and quickly.

- **Recruiter On Demand:**

Access a global network of sources' and recruiters to help meet your recruitment and business goals. Our team uses your branding and processes for fast yet seamless talent acquisition.

2.1 KEY BENEFITS

- **Cost savings:**

Reduce recruitment costs, with a significant reduction in the use of third-party agencies. Cost-saving in RPO is achieved through streamlined processes, efficient candidate sourcing, and optimized recruitment strategies.

- **Global reach:**

Get immediate access to the world's best talents anywhere and anytime you need them with the help of our stunning recruiting capabilities.

- **Brand awareness:**

Increase your employer brand value and position your organization as an employer of choice.

- **Faster hiring:**

Build a range of talent pools of skilled, 'at-the-ready' candidates to meet your future talent needs.

- **Data & insights:**

Leverage reporting, data and insights to create operational excellence and superior risk management.

- **Easy to implement:**

Some projects can be up and running within a matter of weeks. Test drive RPO by starting with a project

CHAPTER 3

TASKS PERFORMED

The project Bank loan status prediction focuses on exploring and implementing different machine learning models to predict whether a loan application is likely to be approved or rejected based on historical data. The purpose of this project is to develop and implement machine learning models using Python for predicting bank loan statuses. By leveraging various machine learning algorithms, we aim to enhance the efficiency and accuracy of loan approval processes, thereby reducing the risk associated with lending decisions.

3.1 INTRODUCTION

Bank loan status prediction is a crucial application of machine learning in the financial sector. Accurate prediction of whether a loan will be approved or not is essential for both banks and customers. We have designed, developed and implemented a Python based Classification prediction Model which will learn from the given training data and then predict the Loan Approval Status classification values for the provided Customer data.

3.2 PROBLEM STATEMENT

To design and develop “Bank Loan Status Predictor” having the following features:

- a) To read, analyse, preprocess the given training dataset producing the Pre-processed data.
- b) To execute and verify all the applicable Classifier algorithms on the pre-processed dataset.
- c) Identify the most suitable classifier out of these classifiers based on the testing data accuracy values.
- d) To predict the Bank Loan Approval status classification values for the given Customer data using the finalized classifier
- e) To get the predicted Bank Loan Approval Status values based on the inputs provided

3.3 SYSTEM SPECIFICATIONS

➤ Software requirements

- **Operating System:** Windows 10 / Higher Version
 - **Python:** Install the latest version of Python (3.x) as the primary programming language.
 - **Integrated Development Environment (IDE):** Choose a suitable IDE for Python development. Popular choices include Google Colab, Jupyter Notebooks, or VSCode.
-

- **Machine Learning Libraries:** Install essential libraries like NumPy, Pandas and Matplotlib for data manipulation and visualization. Install sklearn for machine learning model implementation.
- Hardware requirements
 - **Processor (CPU):** A multi-core processor (quad-core or higher) is beneficial for parallel processing during model training.
 - **RAM:** Minimum 4GB of RAM is recommended, and larger datasets may require more.
 - **Storage:** Adequate storage space for datasets and model files. An SSD is preferable for faster data access.
 - **Internet Connection:** A stable internet connection is required for downloading datasets, libraries, and updates.

3.4 DESIGN

Machine Learning (ML)

Machine learning is programming computers to optimize a performance criterion using example data or past experience. We have a model defined up to some parameters, and learning is the execution of a computer program to optimize the parameters of the model using the training data or experience.

Machine Learning Algorithms used in Bank loan status predictor:

There are a number of algorithms used to command and train machines to perform sentiment analysis. Below are some of the algorithms used:

- **Logistic regression:** It is a statistical method used for analyzing a dataset in which there are one or more independent variables that can be used to predict the outcome of a categorical dependent variable. It's important to note that despite its name, logistic regression is actually a classification algorithm, not a regression algorithm.
- **Naive Bayes:** Naive Bayes classifiers are a collection of classification algorithms based on Bayes' Theorem. It is not a single algorithm but a family of algorithms where all of them share a common principle, i.e. every pair of features being classified is independent of each other.
- **Decision Tree:** It is a supervised machine learning algorithm used for both classification and regression tasks. It creates a model that makes decisions based on the features of the data.

Decision trees are easily interpretable and can handle both categorical and numerical data which makes it unique. They also serve as building blocks for ensemble methods like Random Forest and Gradient Boosting.

- **Support Vector Classifier:** It is sometimes used interchangeably with "Support Vector Machine" (SVM) in certain contexts. Both refer to a type of supervised learning algorithm that can be used for both classification and regression tasks. SVCs are widely used in various fields such as computer vision, natural language processing, and bioinformatics. They are particularly effective in cases where the decision boundary is non-linear or when a clear margin of separation exists between classes.

- **Random forest:** A random forest classifier is a machine learning algorithm that uses decision trees for classification, regression, and other tasks. It's an ensemble classifier that combines the predictions of multiple decision trees to estimate a result.

Bank loan status prediction typically involves a combination of data preparation, model selection, and evaluation. Here's a high-level methodology for predicting loan status[2]:

➤ **Data Collection:**

- Gather historical loan application data, including approved and denied loans, along with relevant borrower information.

➤ **Data Preprocessing:**

- Clean and preprocess the data, addressing missing values, outliers, and inconsistencies.
- Encode categorical variables (e.g., one-hot encoding) and scale numerical features within the range 0 to 1.

➤ **Model Selection:**

- Choose appropriate machine learning models. Common choices include logistic regression, decision trees, random forests, KNeighbors and SVM.
- Experiment with different models and hyperparameters to find the best-performing model.

➤ **Model Training:**

- Train the selected models on the training data using the chosen features.

➤ **Model Evaluation:**

- Evaluate the models on the validation set using relevant metrics like accuracy.

➤ **Monitoring and Maintenance:**

- Continuously monitor the model's performance in a real-world environment and retrain it periodically to account for changing conditions.

This methodology provides a structured approach to building and deploying a bank loan status prediction system that can help banks make informed lending decisions while adhering to regulatory and ethical standards.

3.5 IMPLEMENTATION

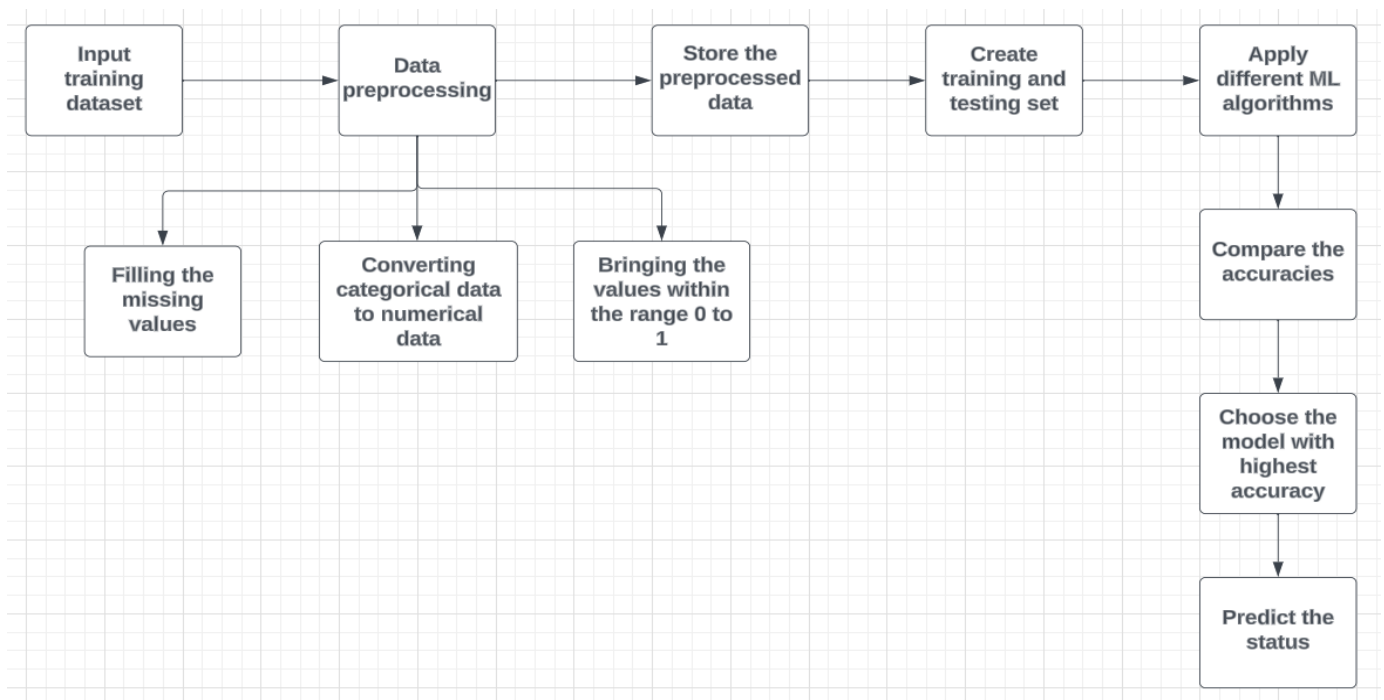


Fig 3.1 Flow chart of Bank loan status Predictor

The project involves the following key steps[2]:

➤ **Input Training Dataset:**

Begin by importing the historical loan dataset, containing information about applicants and their loan outcomes. This forms the foundation for building and training machine learning models.

```
data = pd.read_csv('loan_data.csv')
```

➤ **Data Preprocessing:**

Clean and preprocess the dataset to handle missing values and inconsistencies. This step ensures the dataset's quality and prepares it for effective training and testing. It involves

- **Filling the Missing Values:-** Employ strategies, such as mean or median

imputation, to fill in missing values within the dataset. This ensures completeness in the training data, preventing potential bias in the machine learning models.[5]

```
data['Gender'].fillna(data['Gender'].mode()[0], inplace=True)
data['Married'].fillna(data['Married'].mode()[0], inplace=True)
data['Dependents'].fillna(data['Dependents'].mode()[0], inplace=True)
data['Self_Employed'].fillna(data['Self_Employed'].mode()[0], inplace=True)
data['LoanAmount'].fillna(data['LoanAmount'].mode()[0], inplace=True)
data['Loan_Amount_Term'].fillna(data['Loan_Amount_Term'].mode()[0],
inplace=True)
data['Credit_History'].fillna(data['Credit_History'].mode()[0], inplace=True)
```

- **Converting Categorical Data to Numerical Data:-** Convert categorical variables into numerical representations, facilitating the compatibility of these features with machine learning algorithms. Techniques like one-hot encoding are commonly used for this purpose.[5]

```
data['Gender']=data['Gender'].map({'Male':0,'Female': 1})
data['Married']=data['Married'].map({'No':0, 'Yes': 1})
data['Dependents']=data['Dependents'].map({'0':0,'1':1, '2': 2, '3+': 3})
data['Education']=data['Education'].map({'Graduate': 1, 'Not Graduate': 0})
data['Self_Employed']=data['Self_Employed'].map({'No': 0, 'Yes': 1})
data['Property_Area']=data['Property_Area'].map({'Rural': 0, 'Semiurban': 1,
'Urban': 2})
data['Loan_Status']=data['Loan_Status'].map({'N': 0, 'Y': 1})
```

- **Bringing Values Within the Range 0 to 1:-** Normalize or scale numerical features to a standardized range (e.g., 0 to 1) using the standard normalisation formula. This step ensures that features with different scales contribute equally to the machine learning models, preventing dominance by certain variables in their column.

```
for i in data.columns[1:]:
```

```
    data[i]=(data[i]-data[i].min())/(data[i].max()- data[i].min())
```

➤ **Store the Preprocessed Data:**

Save the cleaned and preprocessed dataset to ensure consistency in subsequent analyses. Storing this processed data allows for efficient reuse during model training, testing, and future predictions. The preprocessed data is stored in another .csv file.

```
data.to_csv('preprocess_loan_data.csv')
```

➤ **Create Training and Testing Set:**

Split the preprocessed dataset into training and testing sets. The training set is used to train the machine learning models, while the testing set evaluates their performance on unseen data, gauging their generalization capabilities.

```
X_train, X_test, y_train, y_test = train_test_split(X, y, stratify = data['Loan_Status'],
test_size = 0.2)
```

➤ **Apply Different ML Algorithms:**

Implement various machine learning algorithms, such as Decision Trees, Support Vector Machines, Logistic regression, KNeighbors and Random Forests, on the training dataset. This step explores diverse approaches to predict loan status and identifies the most effective models.[3]

```
from sklearn.linear_model import LogisticRegression
model1=LogisticRegression()
model1.fit(X_train.values,y_train)
accLR=model1.score(X_test,y_test)*100
print("Accuracy:",accLR)

from sklearn.neighbors import KNeighborsClassifier
model2=KNeighborsClassifier()
model2.fit(X_train.values,y_train)
accKN=model2.score(X_test,y_test)*100
print("Accuracy:",accKN)
```

➤ **Compare the Accuracies:**

Evaluate the performance of each machine learning model on the testing set by comparing accuracy metrics. This step provides insights into how well each algorithm predicts loan statuses in real-world scenarios.[3]

```
algo = ['LogisticRegression','KNeighbors','RandomForest','SVC','DecisionTree']
accuracy = [accLR,accKN,accRF,accSVC,accDT]
compare = pd.DataFrame({'Model' : algo, 'Accuracy' : accuracy}, index = [i for i in
range(1,6) ])
compare.T
```

➤ **Choose the Model with Highest Accuracy:**

Select the machine learning model that exhibits the highest accuracy on the testing set. This model is deemed the most effective for predicting loan statuses and becomes the primary candidate for deployment. Plotting the graph of Model VS Accuracy will make the job easier.

➤ **Predict the Status:**

Utilize the chosen machine learning model to predict loan statuses for new applicants. The model, trained on historical data, can now provide valuable insights into whether an applicant is likely to be approved or rejected for a loan.[3]

```
a=model1.predict([[1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,0.0]])
```

```
if a!=0:
```

```
    print("loan approved")
```

```
else:
```

```
    print("loan rejected")
```

CHAPTER 4

REFLECTION NOTES

During the internship with Vaazu India Pvt Ltd, I gained valuable insights into the practical application and the concepts of "Machine Learning with Python". The internship curriculum focused on equipping me with a solid foundation in the basics of machine learning, emphasizing key components such as data preprocessing, model training, and evaluation. One of the pivotal aspects of the internship involved mastering the art of importing various machine learning models. The hands-on approach not only deepened my understanding of machine learning but also enhanced my ability to implement these concepts in real-world scenarios. As I move forward, I am confident that the skills and insights gained during this internship will serve as a strong foundation for my continued journey in the field of machine learning.

Overall, this internship significantly enriched my skill set, equipping me with the expertise required for tackling complex challenges in the dynamic field of machine learning. This experience not only sharpened my technical proficiency in Python and machine learning but also provided a deep understanding of the practical nuances associated with real-world applications in the financial sector. The hands-on approach not only deepened my understanding of machine learning but also enhanced my ability to implement these concepts in real-world scenarios. This exposure significantly enables to apply theoretical knowledge to practical challenges.

4.1 SPECIFIC OUTCOMES

The outputs of various stages involved in the development of Bank Loan Status Prediction is as follows:

Stage 1: Pre-processed Training Data.

Stage 2: Most suitable classifier is identified.

Stage 3: Customer data with predicted Bank Loan Approval Status values

On executing the finalized algorithm, we can predict the Bank Loan Approval Status values for the given Customer data.

	1	2	3	4	5
Model	LogisticRegression	KNeighbors	RandomForest	SVC	DecisionTree
Accuracy	79.674797	78.861789	76.422764	76.422764	79.674797

Fig 4.1: Testing set accuracy for various models

The above Fig 4.1 contains accuracy values of the different machine learning models for our pre-processed bank loan dataset. These accuracy values are for the testing data which was obtained on applying train_test_split() function.

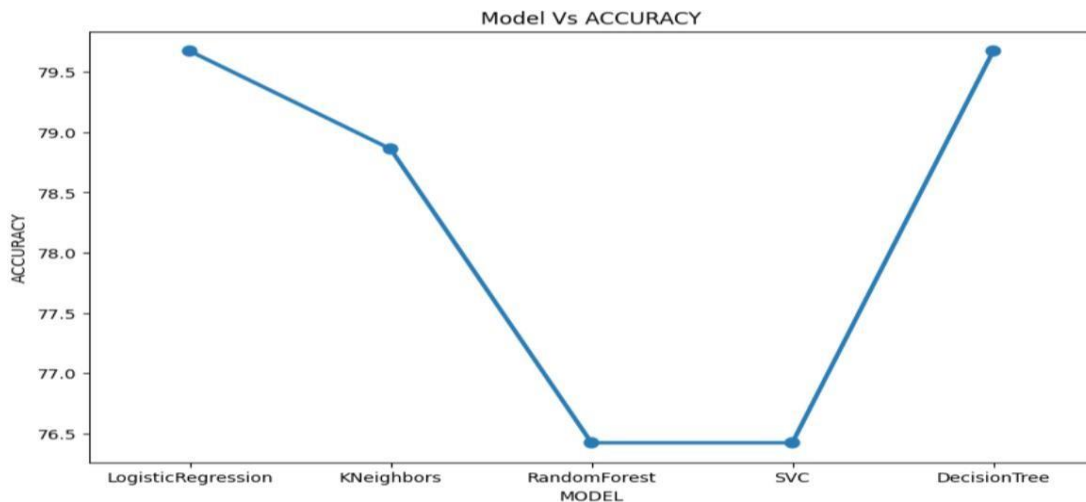


Fig 4.2: Comparison of accuracy values

Fig 4.2 contains the testing data accuracy results which is further used to plot the graph and compare the best model suitable for our future prediction. From the above comparison graph Fig 4.2, we can see that Logistic Regression and Decision Tree models have more accuracy which gives us the best performance. Hence they are the best possible models to predict.

```
a=model1.predict([[1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,0.0]])
if a!=0:
    print("loan approved")
else:
    print("loan rejected")
```

loan approved

Fig 4.3: Loan approved status for logistics regression prediction

Utilizing logistic regression for loan status prediction yielded an accuracy of 79.67%. The model, trained on historical data, predicts loan approval status in Fig 4.3, providing a valuable tool for assessing applicant eligibility in taking loan.

```
a=model1.predict([[0.0,1.0,0.0,1.0,0.0,0.0,0.0,1.0,0.0,0.0,1.0]])
if a!=0:
    print("loan approved")
else:
    print("loan rejected")
```

loan rejected

Fig 4.4: Loan rejected status for logistics regression prediction

Using logistic regression, the loan rejection status was predicted with an accuracy of 79.67% in Fig 4.4. This model successfully identifies applicants likely to face loan rejection based on the historical training dataset.

```
a=model5.predict([[1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,0.0]])
if a!=0:
    print("loan approved")
else:
    print("loan rejected")
```

loan approved

Fig 4.5: Loan approved status for decision tree prediction

Using the Decision Tree algorithm for loan status prediction achieved an accuracy of 79.67%. The model successfully identified whether a loan application would be approved in Fig 4.5. This accuracy underscores the effectiveness of the Decision Tree model in making accurate predictions regarding loan approval status.

```
a=model5.predict([[0.0,1.0,0.0,1.0,0.0,0.0,0.0,1.0,0.0,0.0,1.0]])
if a!=0:
    print("loan approved")
else:
    print("loan rejected")
```

loan rejected

Fig 4.6: Loan rejected status for decision tree prediction

In the Fig 4.6, the Decision Tree model predicted loan rejection status with an accuracy of 79.67%, showcasing its effectiveness in discerning factors contributing to loan denial based on historical data.

CONCLUSION

In the domain of Bank Loan Status Prediction, our machine learning approach has culminated in a comprehensive analysis of various models, with logistic regression and decision tree models emerging as the most accurate, boasting a commendable 79.66% accuracy. This outcome signifies the efficiency of these models in discerning patterns within historical loan data. Notably, when compared to alternative models such as Naive Bayes, Random Forest, and Decision, the selected models outperformed, showcasing their robustness in predicting loan outcomes. The success of logistic regression emphasizes its suitability for binary classification tasks, while decision trees reveal their interpretability and capacity to capture complex decision boundaries.

This product conclusion underscores the practical applicability of machine learning in enhancing the efficiency of loan approval processes, providing financial institutions with valuable insights for informed and accurate decision-making. The chosen model demonstrated high accuracy and reliability, showcasing its potential for application in real-world lending scenarios. The integration of such predictive models into banking systems has the potential to streamline loan approval processes, reduce risks, and enhance overall operational efficiency. The implementation of the chosen machine learning model into the bank's existing loan approval system is recommended. Continuous monitoring and periodic model retraining should be conducted to adapt to changing trends and ensure sustained performance.

REFERENCES

- [1] Tom M Mitchell, "Machine Learning", 1st Edition, McGraw Hill Education, 2017
- [2] A. Gupta, V. Pant, S. Kumar and P. K. Bansal, "Bank Loan Prediction System using Machine Learning," 2020 9th International Conference System Modeling and Advancement in Research Trends (SMART), Moradabad, India, 2020.
{<https://ieeexplore.ieee.org/document/9336801>}
- [3] <https://www.geeksforgeeks.org/machine-learning/>
- [4] <https://www.coursera.org/learn/launching-machine-learning>
- [5] https://www.w3schools.com/python/python_ml_getting_started.asp