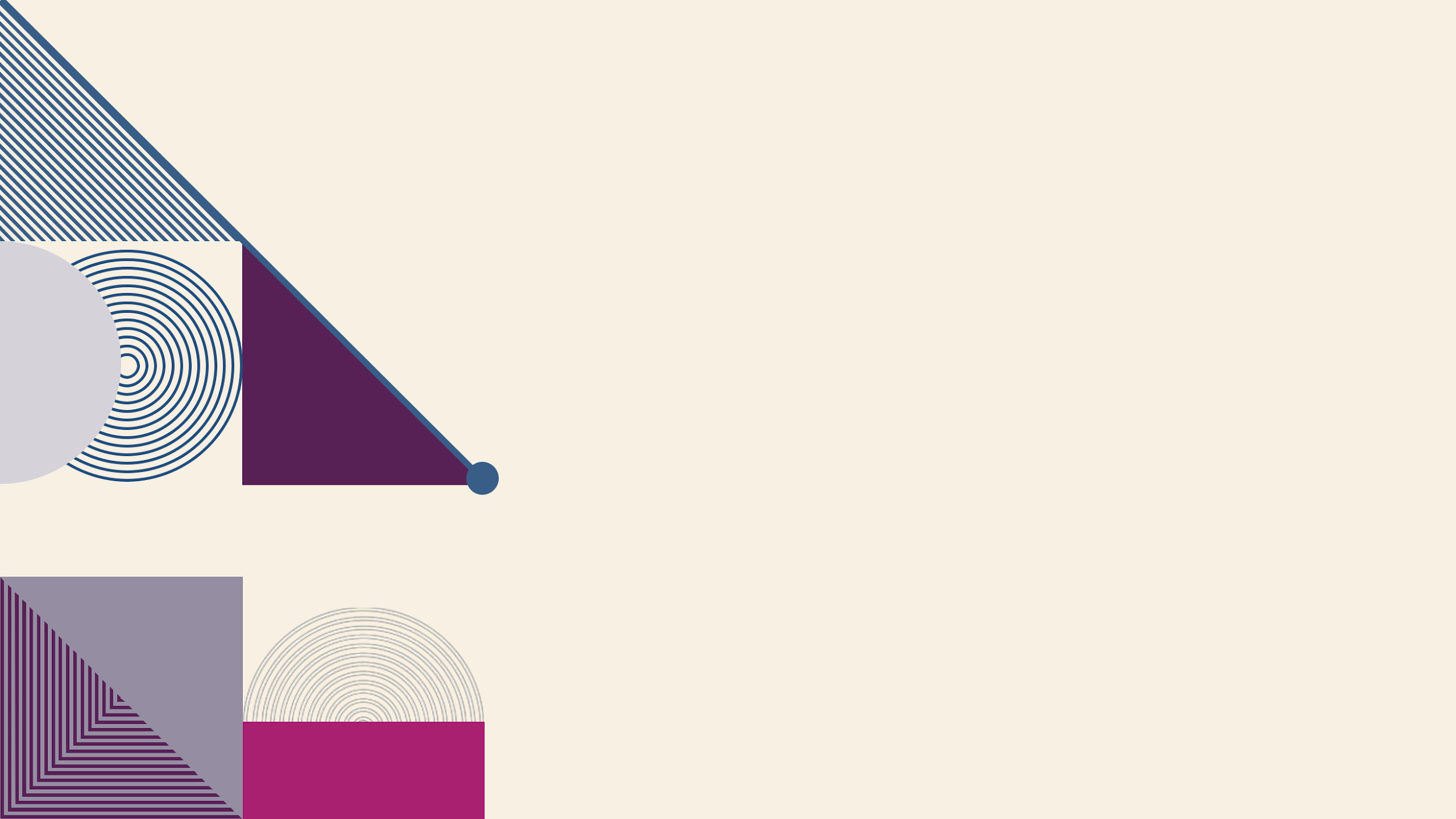


**FRAUD TRANSACTION DETECTION**

DETAILED PROJECT REPORT

Madhvendra Singh and Akanksha

**PROJECT DETAIL**

|  |  |  | **Project Title** | **Fraud Transaction Detection** |
| --- | --- | --- | --- | --- |
|  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  | **Technologies** | **ML | AI | web designing | Flask** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  | **Domain** | **Banking** |
|  |  |  |  |  |
|  |  |  | **Project Difficulty Level** | **Advanced** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

2021 DETAILED PROJECT REPORT 2

**OBJECTIVE**

he purpose of this project is to develop a fraud detection system that can accurately identify fraudulent transactions in a given dataset.

| 2021 | DETAILED PROJECT REPORT | 3 |
| --- | --- | --- |
|  |

1. Introduction

- Purpose of the Project

- Background Information

- Problem Statement

2. Data Collection and Preprocessing

- Data Sources

- Data Description

- Data Cleaning

- Feature Engineering

3. Exploratory Data Analysis (EDA)

- Summary Statistics

- Data Visualization

- Correlation Analysis

- Insights from EDA

4. Machine Learning Algorithms

- Decision Tree

- Model Description

- Hyperparameter Tuning

- Performance Evaluation

- Support Vector Machines (SVM)

- Model Description

- Hyperparameter Tuning

- Performance Evaluation

- Other Algorithms (if applicable)

- Model Descriptions

- Hyperparameter Tuning

- Performance Evaluation

5. Deep Learning Algorithm

- Neural Network Architecture

- Model Training

- Performance Evaluation

6. Results and Evaluation

- Model Comparison

- Evaluation Metrics (Accuracy, Precision, Recall, F1-score)

- Confusion Matrix

- Receiver Operating Characteristic (ROC) Curve

7. Flask Application

- Description of the Application

- User Interface (UI) Design

- Deployment Steps

8. Conclusion

- Summary of the Project

- Achievements

- Limitations and Future Improvements

9. References

- List of Data Sources

- Research Papers

- Online Resources

10. Appendices (if applicable)

- Code snippets

- Additional figures and tables

1. Introduction
   * Purpose of the Project: The purpose of this project is to develop a fraud detection system that can accurately identify fraudulent transactions in a given dataset.
   * Background Information: Provide relevant background information about the significance of fraud detection in financial systems and the challenges associated with it.
   * Problem Statement: Clearly define the problem statement, which is to detect fraudulent transactions accurately using machine learning and deep learning algorithms.
2. Data Collection and Preprocessing
   * Data Sources: Specify the sources from which the transaction data was collected, such as a database or a dataset obtained from a financial institution.
   * Data Description: Provide a brief overview of the transaction data, including the features (columns) available and the target variable indicating fraud or non-fraud.
   * Data Cleaning: Explain the steps taken to clean the data, including handling missing values, outliers, and data formatting issues.
   * Feature Engineering: Describe any additional features created or transformations applied to the data to enhance the performance of the models.
3. Exploratory Data Analysis (EDA)
   * Summary Statistics: Present descriptive statistics of the dataset, such as mean, median, standard deviation, etc., to gain an understanding of the data distribution.
   * Data Visualization: Include visualizations like histograms, box plots, scatter plots, or heatmaps to explore relationships, identify patterns, and detect anomalies.
   * Correlation Analysis: Analyze the correlation between features to identify any strong relationships that could help in fraud detection.
   * Insights from EDA: Summarize the key findings and insights obtained from the exploratory analysis, highlighting any notable trends or patterns.
4. Machine Learning Algorithms
   * Decision Tree: Explain the decision tree algorithm used, including how it splits the data based on certain criteria and handles class imbalance issues.
     + Model Description: Describe the decision tree model architecture and how it makes predictions.
     + Hyperparameter Tuning: Detail the process of tuning the hyperparameters of the decision tree model to optimize its performance.
     + Performance Evaluation: Evaluate the model's performance using appropriate metrics such as accuracy, precision, recall, and F1-score.
   * Support Vector Machines (SVM): Describe the SVM algorithm used for fraud detection and its advantages in handling high-dimensional data.
     + Model Description: Explain the SVM model architecture and how it separates fraudulent and non-fraudulent transactions.
     + Hyperparameter Tuning: Discuss the hyperparameter tuning process for SVM to improve its performance.
     + Performance Evaluation: Assess the model's performance using evaluation metrics and compare it with other models.
   * Other Algorithms (if applicable): Provide descriptions, hyperparameter tuning, and performance evaluation for any additional algorithms you used, such as logistic regression, random forest, or gradient boosting.
5. Deep Learning Algorithm
   * Neural Network Architecture: Describe the architecture of the deep learning model used, including the number of layers, activation functions, and optimization techniques employed.
   * Model Training: Explain the process of training the neural network using the transaction data, including details on loss functions, batch size, and number of epochs.
   * Performance Evaluation: Evaluate the performance of the deep learning model using metrics like accuracy, precision, recall, and F1-score, and compare it with the other models.
6. Results and Evaluation
   * Model Comparison: Compare the performance of all the models used in the project, highlighting their strengths and weaknesses.
   * Evaluation Metrics: Present the evaluation metrics (accuracy, precision, recall, F1-score) for each model, indicating their effectiveness in fraud detection.
   * Confusion Matrix: Display the confusion matrix for each model to show the true positive, true negative, false positive, and false negative predictions.
   * Receiver Operating Characteristic (ROC) Curve: Plot the ROC curve for each model and calculate the area under the curve (AUC) as a measure of model performance.
7. Flask Application
   * Description of the Application: Provide an overview of the Flask application developed for fraud detection, including its functionalities and features.
   * User Interface (UI) Design: Describe the design and layout of the user interface, explaining how users can interact with the application.
   * Deployment Steps: Outline the steps required to deploy the Flask application, including any necessary dependencies or libraries.
8. Conclusion
   * Summary of the Project: Summarize the project's objectives, methodologies, and key findings, emphasizing the success in developing a fraud detection system.
   * Achievements: Highlight the achievements and contributions of the project, such as improved accuracy or efficiency compared to existing methods.
   * Limitations and Future Improvements: Discuss any limitations encountered during the project and suggest potential areas for improvement or future research.
9. References: Provide a list of all the references used in the project, including research papers, online resources, and any external libraries or frameworks utilized.
10. Appendices (if applicable): Include any additional information relevant to the project, such as code snippets, sample outputs, or detailed mathematical formulas.

Health is real wealth in the pandemic time we all realized the brute effects of covid-19 on all irrespective of any status. You are required to analyze this health and medical data for better future preparation.

A dataset is formed taking into consideration information of some individuals. Find key metrics and factors and show the meaningful relationships between attributes of the dataset.

2021 DETAILED PROJECT REPORT 4