**NITTE MEENAKSHI INSTITUTE OF TECHNOLOGY**

(AN AUTONOMOUS INSTITUTION, AFFILIATED TO VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELGAUM,

APPROVED BY AICTE & GOVT.OF KARNATAKA

Logo

Description automatically generated

**Learning Activity Project Proposal**

on

**Title: BANKING and MARKETING ANALYSIS DATASET**

*Submitted in partial fulfilment of the requirement for the award of Degree of*

*Bachelor of Engineering*

*in*

*Computer Science and Engineering*

**Faculty In-charge: Dr. VANI VASUDEVAN**

**Submitted by:**

N. NAGA JAYANTH 1NT19CS118

C.VINEETH 1NT19CS061

T.EKESWAR REDDY 1NT19CS198

V.LAXMI DEEPAK 1NT19CS214

**NITTE MEENAKSHI INSTITUTE OF TECHNOLOGY**

(AN AUTONOMOUS INSTITUTION, AFFILIATED TO VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELGAUM

, APPROVED BY AICTE & GOVT.OF KARNATAKA)

Department of Computer Science and Engineering

**(Accredited by NBA Tier-1)**

**Logo

Description automatically generated**

**CERTIFICATE**

This is to certify that the **BANKING AND MARKETING ANALYSIS**

is an authentic work carried out by **N. NAGA JAYANTH(1NT19CS118), C. VINEETH(1NT19CS061), V. LAXMI DEEPAK(1NT19CS214) and T. EKESWAR REDDY(1NT19CS198)** Bonafede students of **Nitte Meenakshi Institute of Technology**, Bangalore in partial fulfilment for the award of the degree of ***Bachelor of Engineering*** in COMPUTER SCIENCE AND ENGINEERING of Visvesvaraya Technological University, Belagavi during the academic year ***2021-2022.*** It is certified that all corrections and suggestions indicated during the internal assessment has been incorporated in the report. This project has been approved as it satisfies the academic requirement in respect of project work presented for the said degree.

|  |  |  |
| --- | --- | --- |
| **Internal Guide** | **Signature of the HOD** | **Signature of Principal** |
|  |  |  |
| Dr. Vani Vasudevan  Designation: IEEE SM|MACS  CP|ABET PEV|Professor-CSE | Dr. Sarojadevi H.  Professor, Head, Dept. CSE, NMIT Bangalore | Dr. H. C. Nagaraj  Principal,  NMIT,  Bangalore |

**DECLARATION**

We hereby declare that

(i) The project work is our original work

(ii) This Project work has not been submitted for the award of any degree or examination at any other university/College/Institute.

(iii) This Project Work does not contain other persons’ data, pictures, graphs, or other information, unless specifically acknowledged as being sourced from other persons.

(iv) This Project Work does not contain other persons’ writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

a) their words have been re-written, but the general information attributed to them has been referenced.

b) where their exact words have been used, their writing has been placed inside quotation marks, and referenced.

(v) This Project Work does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

|  |  |  |
| --- | --- | --- |
| **NAME** | **USN** | **Signature** |
| N. NAGA JAYANTH | 1NT19CS118 |  |
| C.VINEETH | 1NT19CS061 |  |
| T. EKESWAR REDDY | 1NT19CS198 |  |
| V.LAXMI DEEPAK | 1NT19CS214 |  |

Date: 17/01/2022

**ACKNOWLEDGEMENT**

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of the people who made it possible, whose constant guidance and encouragement crowned our effort with success. I express my sincere gratitude to our Principal **Dr. H. C. Nagaraj**, Nitte Meenakshi Institute of Technology for providing facilities.

We wish to thank our HOD**, Dr. Sarojadevi H.** for the excellent environment created to further educational growth in our college. We also thank him for the invaluable guidance provided which has helped in the creation of a better project.

I hereby like to thank our Dr. Vani V, Designation: IEEE SM|MACSCP|ABET PEV |Professor-CSE Department of Computer Science & Engineering on **her** periodic inspection, time to time evaluation of the project and help to bring the project to the present form.

Thanks to our Departmental Project coordinators. We also thank all our friends, teaching and non-teaching staff at NMIT, Bangalore, for all the direct and indirect help provided in the completion of the project.

|  |  |  |
| --- | --- | --- |
| **NAME** | **USN** | **Signature** |
| N. NAGA JAYANTH | 1NT19CS118 |  |
| C.VINEETH | 1NT19CS061 |  |
| T. EKESWAR REDDY | 1NT19CS198 |  |
| V.LAXMI DEEPAK | 1NT19CS214 |  |

**CONTENT**

**DECLARATION**

**ACKNOWLEDGEMENT**

**ABSTRACT**

**TABLE OF CONTENTS**

**CHAPTER 1: INTRODUCTION**

* 1. Motivation
  2. Problem Domain
  3. Aim and Objectives

**CHAPTER 2: DATA SOURCE AND DATA QUALITY**

* 1. Dataset Used
  2. Data pre-processing

**CHAPTER 3: METHODS & MODELS**

3.1Data mining questions

3.2 Data mining algorithms

3.3 Data Mining Models

**CHAPTER 4: MODEL EVALUATION & DISCUSSION (WITH VISUALIZATION)**

**CHAPTER 5: CONCLUSION &FUTURE DIRECTION**

**CHAPTER6: REFLECTION PORTFOLIO**

**REFERENCES**

**APPENDICES**

Abstract:

Term deposits are a major source of income for a bank. A term deposit is a cash investment held at a financial institution. Your money is invested for an agreed rate of interest over a fixed amount of time, or term. The bank has various outreach plans to sell term deposits to their customers such as email marketing, advertisements, telephonic marketing, and digital marketing.

Telephonic marketing campaigns remain one of the most effective ways to reach out to people. However, they require huge investment as large call centers are hired to execute these campaigns. Hence, it is crucial to identify the customers most likely to convert beforehand so that they can be specifically targeted via call. So here we will detect the people who will be more likely to be taking the term deposit from a bank.

# 1. Introduction

## 1.1 Motivation:

* The data mining process analyses data quantities, extracts and interprets the information obtained using mathematical and statistical models, thus helping banks benefit from a competitive advantage. Data mining techniques have a wide applicability in retail banking and are used in customer relationship and marketing management, risk management, fraud detection, money laundering prevention.

## 1.2 Problem domain:

* The invention of modern computers and application of data mining techniques meant businesses could finally analyse exponential amounts of data and extract valuable insights, forecasting likely business outcomes, mitigating risks, and taking advantage of newly identified opportunities.
* Due to its usefulness across many industries, and its critical role in business success, data mining is a promising career path. Companies need data scientists skilled in mining techniques who can present their findings in understandable ways.

## 1.3 Aim and Objectives:

* The main aim is to determine whether the customer is willing to take the term deposit or not.
* We determine it by using data mining technique like classification, in which we use various types of classification algorithms like logistic regression, KNN algorithm, Decision Tree algorithm to predict the output of categorical variables.

# 2. Data source and Data Quality

## 2.1 Dataset Used

* The data is related with direct marketing campaigns of a Portuguese banking institution, here we are using the raw data in training dataset i.e., train.csv to train the model and test.csv to test the model.
* The used attributes are
  + age
  + job
  + marital
  + education
  + balance
  + housing
  + loan
  + month
  + day
  + campaign
  + pdays
  + previous
  + y
* We collected the data form the link mentioned below.

[Banking Dataset - Marketing Targets](https://www.kaggle.com/prakharrathi25/banking-dataset-marketing-targets)

## 2.2 Data Pre-processing:

* Data pre-processing is a process of preparing the raw data and making it suitable for a machine learning model.
* The steps involved in this are
* Data pre-processing is a data mining technique used for transforming of raw data to a useful information.
* Using pre-processing algorithms in the data mining,
* Number of the missing values in each row is calculated (the are no missing value in our dataset)
* Removal of rows containing missing values is done.
* Duplicate values are also calculated and after applying algorithm we came to know that there are zero duplicate values.

# 3. Methods and Models

## 3.1 Data Mining Questions

* Find the unwanted columns in the data sets.
* What are the missing values?
* Find features with one value.
* Unique values in each column
* Checking the dataset is balanced or not.

## 3.2 Data Mining Algorithms

* The data mining Algorithms that are used in this project is
* **Decision Tree algorithms**: - It is a supervised learning Algorithm, this algorithm builds decision trees using a top-down, greedy approach.
* **KNN algorithm**: - Used to solve both classification and regression problems. It's easy to implement and understand.
* **Random Forest Algorithm**: - It is a supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of **ensemble learning.**
* Plotted the graph between the categorical features and the target values

## 3.3 Data Mining Models.

* Analysing the data based on the data mining algorithms. In the banking analysis decision tree classifier, random forest classifier and K-NN classifier accuracy is tested. We test the results of the accuracy for both test and training set which is chosen randomly from available dataset.

# CHAPTER4. Model Evaluation & Discussion (with necessary visualization)

* Model Evaluation is the process through which we quantify the quality of a system's prediction.
* To predict or evaluate a model, we use some models as a Confusion matrix, Accuracy, Precision, and Recall

**Data Visualization: -**

### **Categorical feature distribution:**

Chart, bar chart

Description automatically generatedChart, bar chart

Description automatically generatedChart, bar chart

Description automatically generatedChart, bar chart

Description automatically generatedChart, bar chart

Description automatically generatedChart, bar chart

Description automatically generatedChart, bar chart

Description automatically generatedChart, bar chart, treemap chart

Description automatically generated

Chart, bar chart

Description automatically generated

**5. Conclusion & Future Direction:**

* We have successfully implemented all the necessary algorithms and models to determine whether the customer will take the term deposit or not in the direct marketing campaigns of a Portuguese banking institute.
* These methods can also be used in marketing and in supermarkets to increase their sales, etc.

# 6. Reflection Portfolio:

We have learnt a lot from the project. We came to know how to use the algorithms and implement the algorithms to a dataset practically. We analysed the relationship between the attributes and plotted its corresponding graph and could extract useful information from the dataset.

**REFERENCES**

1. https://www.kaggle.com/prakharrathi25/banking-dataset-marketing-targets
2. [**https://www.tutorialspoint.com/data\_mining/dm\_quick\_guide.htm**](https://www.tutorialspoint.com/data_mining/dm_quick_guide.htm)
3. [**https://www.javatpoint.com/data-mining**](https://www.javatpoint.com/data-mining)

**APPENDICES**

APPENDIX A: - <https://www.kaggle.com/prakharrathi25/banking-dataset-marketing-targets>

APPENDIX B: jupyter notebook