SYNOPSIS

**LOAN APPROVAL PREDICTION USING MACHINE LEARNING**

**Abstract:** Loan approval is a critical process in the financial sector, which influences both banking institutions and borrowers. Traditional loan approval approach is dependent on manual assessment and rule based systems, which is very much time consuming and are prone to errors. This dissertation makes promising use of Machine Learning (ML) techniques which helps to automate faster and enhances the loan approval process.

**Objective:** The primary objective of this dissertation is to develop a machine learning model to predict whether an applicant’s loan will be approved or not based on various features or variables of that applicant. The model will be trained on historical data and its accuracy will be calculated on different machine learning algorithms and to check which algorithm performs best in approval of loan. This project helps to identify key factors which influences loan approval decisions like Credit history, Applicant income, Co-applicant income, etc. and to provide insights into the loan approval process and help financial institutions make more informed decisions about lending.

**Data Description:** The project makes use of historical Loan Approval dataset which has been downloaded from Kaggle which involves 598 observations and 12 variables. The variables include:

* **Applicant Information:** gender, marital status, number of dependents, education, property area, applicant income and credit history.
* **Loan Information:** loan amount, loan tenure, co-applicant income and employment status.
* **Approval Status**: whether the loan was approved or rejected.

**Methodology:** The project makes use of structured Machine learning pipeline which includes the following steps:

* **Data Preprocessing:** It includes handling missing values, outlier detection and encoding categorical features into numeric format.
* **Exploratory Data Analysis (EDA)**: Identifying correlations and visualizing the key patterns in the data.
* **Model Selection**: Implementing various Machine Learning algorithms like Logistic Regression, K-Nearest Neighbour classifier (KNN), Decision Tree classifier, Random Forest, Support Vector Classifier (SVC) and Naïve Bayes classifier.
* **Model Evaluation**: Assessing model performance like accuracy and AUC-ROC score.