IE 555: Programming for Analytics

Project Proposal Project Title: Data-driven Approach for Surface Roughness Quality Prediction in Manufacturing Processes (Turing Process).

Motivation: The motivation behind this project is to improve the quality of manufactured products by accurately predicting the surface roughness quality in the Turing process. Surface roughness is an important quality factor in manufacturing processes as it affects the functionality, performance, and durability of the products. Traditional methods for predicting surface roughness quality are time-consuming, expensive, and may not be accurate. Therefore, there is a need for a data-driven approach that can accurately predict surface roughness quality in a cost-effective and efficient manner. By developing such an approach, the manufacturing industry can save time and resources while ensuring high-quality products.

Data source: Collecting the data from the actual turning process performed in the workshop at SIT college in India.

Analysis Plan:

Data Collection and Preprocessing:

The first step is to collect the relevant data from various sources such as banks, credit bureaus, and other financial institutions. This data typically includes basic bank details such as income, employment status, age, credit history, loan amount, and other credit-related information. Once the data is collected, it needs to be preprocessed by cleaning, handling missing values, and converting categorical data into numerical data for analysis.

Exploratory Data Analysis:

Exploratory Data Analysis (EDA) is performed to gain insights into the data using data visualization, statistical analysis, and pattern identification using correlation techniques. This step involves visualizing the data, identifying outliers, checking for skewness, and understanding the distribution of the data.

Feature Selection:

Feature selection involves selecting the most relevant features for the model. In this step, various techniques are used to identify the most important features such as classification techniques like Random Forest, Decision tree, or k-nearest neighbors, and recursive feature elimination.

Model Training:

Once the most relevant features are selected, the appropriate machine learning algorithms for classification, such as Random Forest, Decision tree, or k-nearest neighbors, are trained using the preprocessed data. This step involves splitting the data into training and testing sets and training the model on the training set.

Model Evaluation:

After training the model, its performance is evaluated using metrics such as accuracy, precision, recall, and F1 score to assess its effectiveness in predicting creditworthiness. This step involves testing the model on the testing set and comparing its performance to other models.

Comparison of Models:

In this step, the different models are compared with respect to evaluation metrics and the best model is selected for this preprocessed dataset. This step helps to identify the most effective machine learning algorithm for predicting creditworthiness.

Model Deployment:

Once the best model is selected, it is deployed in a production environment for the automatic classification of individuals based on creditworthiness. The system is designed to be scalable and flexible for future updates and changes to the data or model. This step involves integrating the model into an existing system or developing a new system for this purpose.

The **objective** of the proposed project is to develop a data-driven approach for predicting surface roughness quality in manufacturing processes, specifically the Turing process. This will involve collecting and preprocessing data from the actual Turing process performed in the workshop at SIT college in India, performing exploratory data analysis to gain insights into the data, selecting the most relevant features using various techniques, training machine learning algorithms for classification, evaluating model performance using metrics such as accuracy, precision, recall, and F1 score, comparing different models, and deploying the best model in a production environment for automatic classification of surface roughness quality.

Team Members:

1. Avnish Surendra Rai (50488246), avnishsu@buffalo.edu

2. Aditya Vikram Bagaria (50479384), abagaria@buffalo.edu

3. Ayushi Porwal (50442208), ayushipo@buffalo.edu