

# ***Project Report***

## ***Project Overview***

*The project encompasses four distinct projects related to insurance , credit card data analysis, image classification and Activity Detection. Each project involves various tasks, including data preprocessing, visualization, and the implementation of machine learning models. Here is a comprehensive report detailing the key aspects of each project. **Project 1: Insurance Data Analysis***

### ***Data Preprocessing***

**Duplicate Rows:** Detected and removed duplicate rows from the insurance data.

**Outliers:** Conducted outlier detection, but no outliers were identified.

#### **Descriptive Statistics:**

Minimum, Maximum, Mean, Standard Deviation for each numerical feature.

#### **Visualization:**

Utilized various visualizations, such as histograms, box plots, and scatter plots, to explore and understand the data distribution.

#### **Categorical to Numerical Conversion:**

Converted categorical variables to numerical using appropriate encoding methods.

## ***Modeling***

### ***Linear Regression:***

Achieved an accuracy of 80%.

### ***Random Forest:***

Applied hyperparameter optimization, achieving an accuracy of 90.3%.

### ***SVM (Support Vector Machine):***

Obtained an accuracy of 71.2%.

### ***KNN (K-Nearest Neighbors):***

Conducted hyperparameter tuning, resulting in an accuracy of 86.56%.

### ***Decision Tree:***

Implemented a decision tree model.

## ***Project 2: Activity Detection***

### ***Data Overview***

No Nulls or Duplicates: Ensured the absence of null values and duplicates in both train and test datasets.

### ***Modeling***

1-KNN (K-Nearest Neighbors)

2-Logistic Regression

3-Decision Tree

4-SVM (Support Vector Machine)

5-Random Forest

## ***Project 3: Credit Card Data Analysis***

### ***Data Preprocessing***

#### ***Handling Null Values:***

Checked for null values and addressed accordingly.

#### ***Data Description:***

Computed mean, median, and standard deviation.

Converted class values to 0 and 1.

#### ***Time Column:***

Divided the time column into minutes and hours.

### ***Correlation Analysis:***

Conducted correlation analysis among features.

### ***Boxplot:***

Utilized boxplots for visualizing feature distributions.

## ***Modeling and Results***

### ***1-KMeans:***

Accuracy: 0.074

## ***Project 4: Image Classification***

### ***Data Overview***

Training Data: Located at 'E:\level 4\machine learning\projects\seg\_train'

Testing Data: Located at 'E:\level 4\machine learning\projects\seg\_test'

Prediction Data: Located at 'E:\level 4\machine learning\projects\seg\_pred'

### ***Data Exploration and Preprocessing***

#### ***Class Names:***

Extracted class names from the training and testing data directories.

#### ***Image Size Analysis:***

Examined the size distribution of images in the training and testing datasets.

#### ***Class Labeling:***

Mapped class names to numerical labels (0 to 5). ***Data***

### ***Split and Visualization***

#### ***Train-Test Split:***

Divided the data into training and testing sets.

#### ***Data Visualization:***

Displayed random images from the training and testing sets along with their corresponding class labels.

Model Architecture

#### ***Convolutional Neural Network (CNN):***

Designed a CNN model with multiple convolutional and dense layers.

Utilized dropout for regularization.

### ***Model Evaluation***

Test Accuracy:

Evaluated the model on the test set.

## ***Model Saving and Loading***

### ***Save Model:***

Saved the trained model to 'E:\level 4\machine learning\projects/my\_model.hdf5'. ***Load***

### ***Model:***

Loaded the saved model for deployment.

## ***Prediction***

Prediction Visualization:

Visualized predictions on a subset of the prediction data.

## ***Model Performance Analysis***

### ***Confusion Matrix:***

Computed and visualized the confusion matrix.