LOG BOOK

Athulya Shanty (C00313623)

I started the notebook from scratch.		
Filename	Dataset	Steps
		Load the CSV file into a Data Frame
z.Data,Chandrig, igyrib	Alzheimers Data	Check the number of rows and columns in the DataFrame
		Display basic information about the dataset
		Display summary statistics of the dataset
		Chack for missing values Chack if the data is imbalanced
		Check for dualizate must in the dataset
		Sive the cleaned data to a CSV file
	Student Performance Data	Load the CSV file into a Data Frame
		Check the number of rows and columns in the DataFrame
		Display basic information about the dataset
		Display summary statistics of the dataset
		Check for missing values
		Check for duplicate rows in the dataset
		Save the cleaned data to a CSV file
2.Equipatory, Data, Analysis, jayro		Load the CSV file into a Data Frame
	Alzheimers Data	Plot the distribution of the Diagnosis column
		Plot the top 10 countries with the highest Alzheimer's cases
		Plot the gender distribution for diagnosed cases Plot the distribution of age groups by diagnosis
		Plot the distribution of the Family History of Alzheimer's
		Categoriae BMI and glot the distribution by diagnosis
		Identify numerical & Categorical columns
		Outlier Detection
	Student Performance Data	Load the CSV file into a DataFrame
		Plotted Relation between hours studied and performance index
		Plot the relation between sleep hours and performance index
		correlation heatmap
		Outlier Detection
3. Data, Proprocessing Synto	Alzheimers Data	Load the CSV file into a DataFrame
		Identify numerical & Categorical columns
		Normalize the numerical columns using StandardScaler()
		Label Encoding
		Save the processed data to a CSV file Spit the data into training (70%) and remaining (30%)
		Spit the cartainto training (10%) and remaining (30%) Spit the remaining data into test (50% of remaining) and validation (50% of remaining)
		Spit the remaining data into test (some or remaining) and valuation (some or remaining) Define the features and target variable for training, testing & validation
		Save the features and target for training, testing & validation as pickle file
	Student Performance Data	Load the CSV file into a Data Frame
		Identify numerical & Categorical columns
		Normalize the numerical columns using StandardScaler()
		Label Encoding
		Split the data into training (70%) and remaining (30%)
		Spit the remaining data into test (50% of remaining) and validation (50% of remaining)
		Define the features and target variable for training, testing & validation
		Save the features and target for training, testing & validation as pickle file
4_Linear_Regression.jpynb	Student Performance Data	Load the features and target for training, testing & validation
		Create a linear regression model, Train the model using the training data, Predict & Evoluate RMSE and R*2 score using the test and validation data
		Plot actual vs predicted for test data
		Plot actual vs predicted for validation data
5_Logistic_Riagression.ipyeb	Alzheimers Data	Load the features and target for training, testing & validation Created a function calculate rinse and predict) to calculate the innse value and predict the tarest values
		Created a function calculate _mse_and _predict) to calculate the mse value and predict the target values Created a function train and evaluate (osistic repression) for loaistic repression model
		Created a function trans_and_evaluate_registarc_regression() for logistic regression model Calculate the Test Accuracy
		Calculate RMSE for test & validation data
		Save the accuracy, RMSE on test & validation data as a pickle file
		Load the features and target for training, testing & validation
R_Decision_Trees.lpynb	Alzheimers Data	Created a function calculate_rmse_and_predict() to calculate the mise value and predict the target values
		Defined a range (1,15) to find optimal depth of decision tree by Plotting the mean cross-validated accuracy for each depth. Optimal depth was calculated to be 5
		Train the Decision Tree Classifier with the optimal depth, Predict on the test data, Calculate the accuracy
		Calculate RMSE for test & validation data
		Plot the decision tree with a depth of 2 levels (opted 2 for better visualization)
		Update the accuracy, RMSE on test & validation data to the pickle file
		Load the features and target for training, testing & validation
7_Random_Forest.joynb	Alzheimers Data	Created a function calculate_rmse_and_predict() to calculate the rmse value and predict the target values
		Initialities the Random Forest Classifier, Train the classifier, Predict on the test data, Calculate the accuracy Calculate PMSE for test & validation data
		Calculate RMSE for test & validation data Update the accuracy, RMSE on test & validation data to the pickle file
		Update the accuracy, MMSE on test a valuation data to the prise tile Load the features and target for training, testing & validation
0_INNLigato	Alzheimes Data	Created a function calculate _mse_and_predict() to calculate the mse value and predict the target values
		Performed K Fold Cross Validation to find the optimal kvalue and optimal in- 17
		Created a function train_knn_model() for KNN model
		Implemented KNIN with Euclidean distance & Calculated RMSE for test & validation data
		Implemented KNN with Manhattan distance & Calculated RMSE for test & validation data
		Implemented KNN with Minkowski distance & Calculated RMSE for test & validation data
		Update the accuracy, RMSE on test & validation data to the pickle file
		Load the features and target for training, testing & validation
		Created a function calculate_rmse_and_predict() to calculate the rmse value and predict the target values
9_SVM.ipyrib	Alzheimers Data	Created a function train_and_evaluate_swm() for SVM model
9_SVM.lipynib	Alzheimers Data	Created a faccifice train, and, evaluate, sum() for SMM model: Implemented SMM with binear learned & Calculated RMSE for sext & available on data
2_SVALignto	Alzheimers Data	Created a function train_and_evaluate_swm() for SVM model
9,3/ALigyrib	Alzheimers Data	Cleated a function trade, and evaluate, sumplier 95M model traplamented 95M with Transact recent & Calculated RMSE for tree & validation data traplamented 95M with Transact A Calculated RMSE for tree & validation data
9,3/M.ligynb	Alzheimers Data	Cleated a function train, and evaluate, sunty for SMM model Implemented SMM with linear learned & Calculated RMSE for text & validation data Implemented SMM with Total rearned & Calculated RMSE for text & validation data Updates the accouncy, RMSE on text & validation data to the profile file Updates the accouncy, RMSE on text & validation data to the profile file
		Creates a function train, and creations, send for 150 M model Integenment at 50 M with Train a train of Contract of 150 M to 150 M available on that Implement at 50 M with Train a train of Contract of 150 M to 150 M available on data Update the excessive (Find on the Available on the set of 150 M available on the Lated the Mulanes and Engels for training, bening & vestilation
9.30/4.ipyro 10.30/4.ipyro	Alzheimers Data Alzheimers Data	Cleated a function state, and evaluate, samp for 50M model Insperiment 450M with linear invent & Calculated PMEE for text & validation data Insperiment 450M with 150M at 150M
		Created a function study, and evaluate, semily for SMM model implamented SMM with linear isomal & Casinulated RME for treat & validation data implamented SMM with RME Semented, Colonialed RME for treat & validation data Update the accuracy, RME or the & validation data to the parks file Update the accuracy, RME or the & validation data to the parks file Accuracy RME or the development of the students for the s
		Cleases a function to app, and c, evaluate, semily for 1994 model Integrenment 2994 with Time are used, E.C. calculate 6996 for the sit is validation data Update the accounty, PMC of the sit validation 6995 for the sit is validation data Update the accounty, PMC of the sit validation data to the probabilities Load the function account, manage, and product (is calculated for time validation data) Cleases as function accounts, manage, and product (is calculated for time validation data) Calculated for the control of the site
		Creates a function train, and creature, serily for SMM model Integrences SMM with linear branch & Calculated PMEE for test & violation data Implainment SMM with PMEE for test Management of SMEE for test & violation data Update the accuracy, RMM or test A violation data to the problet for Lead the function of test and violation data to the problet for Lead the function circulation, remaining, testing & violation Constant an internal conduction, remaining, testing & violation Constant an internal conduction, remaining, testing & violation Constant and conduction circulation, remaining, testing & violation Applied the General Market SMEE or conduction data conduction the service of testing or violation data Applied the General Market SMEE or conduct & Calculated Accuracy
		Creates a function to are, and, evaluate, semily for 1994 model Integenence 2994 with 1987 are travel. A Conducted PREE for text 4 violation data Unglamented 2994 with 1987 are and, Conducted PREE for text 4 violation data Unglamented 2994 with 1987 are and, Conducted PREE for text 4 violation data Unglamented 2994 with 1987 are a violation data to the probability of the conducted and the southern and travel to the violation of the conducted and the southern and travel to the violation and travel to the violation data to the probability of the conducted and the southern and travel to the southern and travel to the southern and product the target values Applied the Southern and Super model & Conducted Accuracy Conducted PREE for text it wild-time data to the prior the southern and super model & Conducted Accuracy Underst the accuracy and Super model & Conducted Accuracy Underst the accuracy and Super model & Conducted Accuracy Underst the accuracy and Super model & Conducted Accuracy Applied Edow method of find the optimism and are of by place Super method of find the optimism and are of 2
10 Naiws, Bayesian lyynb	Alphoimers Duta	Creates a function trans, and, evaluate, semily for 99M model Insperiment 99M with Time a trans. A Conducted 99EE for the sit a validation data superiment 99M with 19EE and A Conducted 99EE for the sit a validation data Update the accounty, 99EE of the sit a validation data Update the accounty, 99EE of the sit a validation data to the probe 19EE Lacidath the summa and target for training reporting 4 validation Counted a function calculation, man, and predictly 1 to calculation the mans value and product the target values Applies the Source shall believe growth A. Colladate Accounty Counted to Model for treat A validation data Applies the Bounced States Stay on coulds. Excludated Accounty Lacidate 19EE on the Validation data Applies the Bounced States Stay on coulds. Excludated Accounty Lacidate the account, 96EE on that a validation data is to be published to Import the dataset and review Displaces cultume for countering