|  |  |
| --- | --- |
| SR No. 01 | Methodology:   * + used Exploratory data Analysis to Analyse data.   Technique:   * + To begin this exploratory analysis, first use matplotlib to import libraries and define functions for plotting the data.   + Sklearn.preprocessing (For pre-processing)   + Matplotlib.pyplot (For plotting graph)   + Numpy (For linear algebra)   + Pandas (for data pre-processing on csv file)   Findings:   * + Correlation matrix for data   + Scatter and density plot   Link   * + <https://www.kaggle.com/vanwest/starter-students-academic-performance-c514bf95-6> |
| 02 | Methodology   * + In this they didn’t used full dataset column thay used only [raisedHand, VisITedResources, AnnouncementsView, Discussion] column based on that thay are predicting the performance of students. And thay are predicting that only for top 4 nationality for both male and female.   Technique   * + Numpy (For linear algebra)   + Pandas (for data pre-processing on csv file)   + Matplotlib.pyplot (For plotting graph)   Findings   * + Hand raised by gender   + Hand raised by gender with top 4 country   + Hand raised by nationality   + VisITedResources by nationality   + AnnouncementsView by nationality   + Discussion by nationality   + Hand raised by nationality on each subjects   Link   * + <https://www.kaggle.com/d50stuck/kalboard-360-use-case> |
| 03 | Methodology   * + *LogsticRegression*   + *XGBClassifier*   Technique   * + Numpy (For linear algebra)   + Pandas (for data pre-processing on csv file)   + Matplotlib.pyplot (For plotting graph)   + Seaborn (for Data Analytics and plotting graph)   + sklearn.linear\_model (for linear regression)   + Xgboost (for XGBClassifier)   + sklearn.metrics (for classification report)   Findings   * + Plotting on linear regression   + Plotting on XGBClassifier   Link   * + <https://www.kaggle.com/hyojungshin/py300-part1-2-xapi-edu-data> |
| 04 | Methodology   * + Exploratory Data Analysis   + *LogsticRegression*   + *XGBClassifier*   Technique   * + Numpy (For linear algebra)   + Pandas (for data pre-processing on csv file)   + Matplotlib.pyplot (For plotting graph)   + Seaborn (for Data Analytics and plotting graph)   + sklearn.linear\_model (for linear regression)   + Xgboost (for XGBClassifier)   + sklearn.metrics (for classification report)   Findings   * + Plotting on linear regression   + Plotting on XGBClassifier   Link   * + <https://www.kaggle.com/songhunhan/classification-studygrade> |
| 05 | Methodology   * + Exploratory Data Analysis   Technique   * + Numpy (For linear algebra)   + Pandas (for data pre-processing on csv file)   + Matplotlib.pyplot (For plotting graph)   + Seaborn (for Data Analytics and plotting graph)   Findings   * + Boxplot of behavioral attributes of frequency of [raisedhands, VisITedResources, AnnouncementsView, Discussion]   + Plotting on gender   + Plotting on nationality   + Plotting on place of birth   + Plotting on stageID   + Plotting on graceID   + Plotting on sectionID   + Plotting on Topic   + Plotting on semester   + Plotting on relation   + Plotting on ParantsansweringSurway   + Plotting on ParantsSchoolSatisfation   + Plotting on StudentsAbsenceDays   + Plotting on class   + Pairplot of behavioral numerical attributes of frequency of [raisedhands, VisITedResources, AnnouncementsView, Discussion]   + Heatmap of the correlation between numerical behavioral attributes of [raisedhands, VisITedResources, AnnouncementsView, Discussion]   Link   * + <https://www.kaggle.com/songhunhan/classification-studygrade> |
| 06 | Methodology   * + RandomForestClassifier   Technique   * + Numpy (For linear algebra)   + Pandas (for data pre-processing on csv file)   + Sklearn (for classification)   Findings   * + Model accuracy score(0.6916666666666667)   + Hyperparameter based on random classifiers   Link   * + <https://www.kaggle.com/aristotle609/academic-education-model> |
| 07 | Methodology   * + Exploratory Data Analysis   Technique   * + Numpy (For linear algebra)   + Pandas (for data pre-processing on csv file)   + Matplotlib.pyplot (For plotting graph)   Findings   * + Analysis on Categorical Data   + ananlysis on Numerical Data   + Plots onDiscussions in Science and Liberal fields   + Plots on Class Activity of [Raise Hands, discussion, VisITedResources] (for Absenses Under – 7 & Over – 7)   Link   * + <https://www.kaggle.com/aristotle609/academic-performance-eda> |
| 08 | Methodology   * + MLPClassifier   + RandomForestClassifier   Technique   * + Numpy (For linear algebra)   + Pandas (for data pre-processing on csv file)   + Matplotlib.pyplot (For plotting graph)   + Sklearn (for classification)   + FeatureImportances (for feature mapping)   Findings   * + Plotting on [class, gender, Nationality, StageID, GradeID, SectionID, ParentAnsweringSurvey, ParentSchoolSatisfaction]   + MLPClassifier Model   + Confusion matrix   + Features importents of 58 features using RandomForestClassifier   Link   * + <https://www.kaggle.com/ashup1501/notebook4776b69f36> |
| 09 | Methodology   * + Managing Overfitting   + Regression   + Classification   Technique   * + Numpy (For linear algebra)   + Pandas (for data pre-processing on csv file)   + Matplotlib.pyplot (For plotting graph)   + Seaborn (for Data Analytics and plotting graph)   + Sklearn (for classification & regression)   Findings   * + Pair Plotts of [raisedhands, VisITedResources, AnnouncementsView, Discussion]   + Plotting of [Distribution of Student Participating, Distribution of Student Visiting Resources]   + Andrew Curves for Parents Satisfaction   + Plotting on Parent Satisfaction on Grades   + Plotting on K-Nearest Neighbour   + Plotting on KMeans   + Plotting on Hierachical Clustering   + AgglomerativeClustering   + Model   + plotting of results   + Confusion matrix   + Solution of overfitting problem   + Top10 Models and Dataset   + feature\_scores   + Plotting on feature scores of the features   Link   * + <https://www.kaggle.com/afrologicinsect/manage-overfitting-in-predicting-acad-performance> |
| 10 | Methodology   * Used seaborn to Data Visualization and Exploration * the perceptron and SVC will be trained on the continuous data (Classification)   Technique   * Seaborn (To visualization of the data) * Pandas (For the creating of different data frames and cleaning of data) * NumPy (Used for mathematical purpose) * matplotlib (To plot the graphs based on the data) * Sklearn (To train the models based on the dataset)   Findings   * Misclassified Samples * Accuracy score of the model (Perceptron , MLPClassifier , SVC)   Link   * <https://www.kaggle.com/dan195/classification-of-student-marks> |
| 11 | Methodology   * Exploratory data Analysis   Technique   * NumPy ( For linear Algebra ) * Pandas ( For data Processing ) * Matplotlib ( For plotting the graph )   Findings   * The relation between the raise hand and absence in class. The outcome is that the student who Is absence for more than 7 days he/she has the very low raise hand and the student who is absence for less than 7 days has more.   Link   * <https://www.kaggle.com/sohailkhan/exploratory-analysis-1> |
| 12 | Methodology   * A benefit of EDA is that is allows you to examine the data as it is without making any assumptions * Just Analyzing the data in the dataset to understand the relation between the data to each other and many more.   Technique   * NumPy ( Used for the Linear Algebra ) * Pandas ( Used for the data processing )   Findings   * By applying such methods they have plotted the graphs and based on that we have to understand the insight of the data. * They have compared many fields with each others to identify the correlation between them like Raise hand → Discussion , Announcement → raise hand , Discussion → Visit Resources etc.…   Links   * <https://www.kaggle.com/wittykitty/exploratory-data-analysis-for-beginners> |
| 13 | Methodology   * EDA ( Exploratory data Analysis ) * Sklearn to predict the model values on the dataset   Technique   * NumPy (Used for the Linear Algebra ) * Pandas ( Used for data preprocessing ) * ggplot( To plot the data ) * Sklearn ( For model execution and finding the accuracy )   Findings   * By applying the EDA he has find all the relationship of the data each and every possible combination of data with each other. * By applying this in the histogram , tilegraph , boxplot (With every possible combination) and many other kind of the graph to visualize the data in the appropriate way. * He has even visualized the data in the Decision tree and train the models like Decision tree , Random Forest , Support Vector Machine and at the end he has used Ensemble Learning to combine all the three models output and learn from that. * At the end he has find the accuracy for that. Which is 88.43 %   Link   * <https://www.kaggle.com/arathee2/predicting-classes/report> |
| 14 | Methodology   * Use Machine Learning to predict the student Performance on the various data like raise hand etc..   Technique   * NumPy (Used for the Linear Algebra ) * Pandas ( Used for data preprocessing ) * Matplotlib ( Used to plot the graphs ) * Sklearn ( For Machine Learning models executions )   Findings   * Plot the graph for the pivot table * Plot the graphs for correlation of data with each other * Plot other Histograms for better data visualization of data based , To identify the which data column is important and which is not * Used Machine Learning Algorithms like Random Forest , Logistic Regression , SVM , stratified KFold for KFold validation to get the better accuracy and many more algorithms. * Then plot the graph for all the algorithms accuracy score to identify easily. * Again after some modification in the dataset using transformation he has applied All the algorithms again and find the accuracy score for that all. * And again plot the graph for that algorithm again and get the better result.   Link   * <https://www.kaggle.com/rmalshe/student-performance-prediction> |
| 15 | Methodology   * Find the prediction rate of the Logistic Regression and plot that accuracy in the Confusion metrics.   Technique   * NumPy ( Used for the Linear Algebra ) * Pandas ( Used for the Processing of data ) * Sklearn ( To implement the Logistic Regression & Plot the confusion metrics )   Findings   * Accuracy score of the Linear Regression * Plot the confusion metrics   Links   * <https://www.kaggle.com/jpotts18/logistic-regression-1> |
| 16 | Methodology   * Correlation of data using Machine Learning Algorithm * PCA ( Principle Component Analysis )   Technique   * Numpy ( For Linear Algebra ) * Pandas ( For Processing of data ) * Seaborn ( To visualize the EDA ) * Matplotlib ( To plot the graph of different type )   Findings   * The correlation graph between different student attributes * PCA graph to visualize the data ( Individual explained Variance )   Links   * <https://www.kaggle.com/isaacinyang/notebookf9a5a17c71> |
| 17 | Methodology   * The bar chart to display the nationality of student from the dataset   Techniques   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization into the graph )   Findings   * The graph that represent how many students are from which country in the histogram graph and Pie chart.   Link   * <https://www.kaggle.com/shashanksingh/charting-nationality> |
| 18 | Methodology   * Visualize the data in the histogram to with each other to understand the data * Plot the Heatmap of correlation with the help of seaborn.   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For visualizing the data in graph ) * Seaborn ( For visualizing the heap map of correlation )   Finding   * A histogram of Importance of feature to identify the feature which is impacting more on the training dataset. * A heat map of correlation which is identify the correlation of the features with each other. The highly correlated the high impact on the training dataset.   Link:   * <https://www.kaggle.com/jeffreynghm/academic-performances> |
| 19 | Methodology   * Data Exploration   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing )   Findings   * All the classes available in the dataset * All the classes with the number of occurrence in the dataset   Link :   * <https://www.kaggle.com/jjconde/quick-data-exploration> |
| 20 | Methodology   * Random Forest and Ensemble Learning Algorithm to find the Accuracy. * Used the Hyperparameter to find the appropriate parameter for that dataset according to the model. * Feature selection to find the important feature from the dataset.   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Sklearn ( To implement the Algorithm ) * Used L1-based Feature selection   Findings   * The accuracy score of the Random Forest classifier and Bagging classifier with hyperparameter tuning and without hyperparameter tuning   Link:   * <https://www.kaggle.com/chandiniunnikrishnan/student-grades-classification> |
| 21 | Methodology   * Data Exploration * Machine Learning Algorithm * Cross-Validation for hyper parameter tuning   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data Processing ) * Matplotlib ( For Visualizing the data in the graph ) * Sklearn ( For Machine Learning )   Findings   * Best Hyper-parameter for Random Forest Classifier * Feature important table represent the feature importance in the dataset. * The graph of importance of the feature in a form of histogram   Link   * <https://www.kaggle.com/gemunu/student-perfomance> |
| 22 | Methodology   * Data Analysis * Data Visualization   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization in graphs )   Findings   * The heat map of correlation between raise hand , Visit Resource , Announcement , Discussion. * Seaborn to represent the graphs in different ways. * The graph of Discussion -> raise hand , Visit Resource , Announcement , Discussion   Link:   * <https://www.kaggle.com/liaopan/data-analysis-in-students-academy> |
| 23 | Methodology   * Regression used to predict the scores of the students * GridSearchCV to find the hyperparameter of the respective model. * kFold validation to find the Folds the dataset   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Sklearn ( For Regression and Hyperparameter tuning ) * Matplotlib ( For visualizing the data )   Findings   * The best hyper-parameter for the Regression algorithms * And the prediction score of all the model based on the dataset   Link:   * <https://www.kaggle.com/raghavrastogi75/student-scores-analysis-and-prediction/script> |
| 24 | Methodology   * Decision Tree and Random Forest * EDA   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Seaborn ( For EDA ) * Matplotlib ( For Visualizing graphs ) * Sklearn ( For Decision Tree and Random Forest )   Findings   * Confusion matrix for the Decision Tree and Random Forest * Heat map of the Classification Report * Accuracy score for the Classification   Link   * <https://www.kaggle.com/miya0930/decisiontree-and-randomforest> |
| 25 | Methodology   * Data Analysis * Machine Learning ( Random Forest Classifier , SVC and Decision Tree )   Technology   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Sklearn ( For Random Forest )   Findings   * Graph of raise hand -> Class , class -> Parent , class -> semester , class -> Student Absence , class -> Gender , class -> Parent school satisfaction and many more… * Accuracy Score for the Decision Tree , Random Forest and SVC.   Link   * <https://www.kaggle.com/ritgoe/student-data-analytics> |
| 26 | Methodology   * EDA ( Exploratory Data Analysis ) * Logistic and XGBoost   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Sklearn ( For Linear and XgBoost ) * Seaborn ( For EDA )   Findings   * Get the graph of gender , Place of Birth , student Absence Days etc… * Logistic Regression results , score and classification Report and same for XGBoost. * The graph of important feature   Link   * <https://www.kaggle.com/rahulvks/exploratory-data-analysis-logistic-and-xgboost> |
| 27 | Methodology   * Data visualization * Seaborn   Technique   * Pandas ( For data processing ) * Seaborn ( For data visualization )   Findings   * Count Plots of every feature available in the dataset with horizontal and vertical graph   Link   * <https://www.kaggle.com/slamnz/seaborn-tutorial-count-plots> |
| 28 | Methodology   * EDA * Machine Learning ( Decision Tree and Random Forest Classifier )   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Sklearn ( For Machine Learning ) * Seaborn ( For EDA )   Findings   * The graphs which represents the data in the very good manner * The Accuracy score for the Decision Tree Classifier and Random Forest Classifier   Link   * <https://www.kaggle.com/ritgoe/notebookc795091ab7> |
| 29 | Methodology   * Data visualization * Tree Bases classification   Technique   * NumPy ( For Linear Algebra ) * Panada ( For data processing ) * Matplotlib ( For data visualization ) * Sklearn ( For Logistic Regression and Decision Tree )   Findings   * The graphs for the different types of comparison of the data with each other * Done the data exploration * Get the accuracy score , Classification Report , Confusion metric for the Random Forest and Logistic Regression * Done the Gradient Boosting for the getting the better result. * By adding some more hyper-parameter find the better accuracy model for the Gradient Boosting   Link   * <https://www.kaggle.com/rruizendaal/visualizations-tree-based-classification-methods> |
| 30 | Methodology   * Machine Learning ( Find best classification algorithm for the dataset)   Technique   * Pandas ( For data processing ) * Sklearn ( For Classification Algorithms )   Findings   * All the classification algorithms Average score and Standard Deviation to find which is best of all them   Link   * <https://www.kaggle.com/slamnz/what-is-the-best-classifier-to-start-with> |
| 31 | Methodology   * Performance analysis of the student * Cor-relation of the data with each other * Machine Learning ( perceptron , SVC , MLP Classifier , Random Forest and XGBoost )   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Sklearn ( For Machine Learning ) * Seaborn ( For EDA )   Findings   * Histogram graph for the different data available in the dataset * Heat Map of co-relation of data available in the dataset * Regression plot of the dataset * Classification Report , Accuracy score for the Machine Learning Algorithm * The Tile graph for the feature importance according to XGB ( XGBoost ) algorithm.   Link   * <https://www.kaggle.com/satadru5/student-peroformance-analysis> |
| 32 | Methodology   * Predicting Student Academic Performance using artificial neural networks   Technique   * Same as no.8   Link   * <https://www.kaggle.com/akinoladejodamilare/predicting-student-performance-v2> |
| 33 | Methodology   * Apply Support Vector Machines and Random Forest seperately to the student academic performance dataset, predict the class of the student. The classes represent their academic level, (L = low, M = medium, H = hi). Then, use them in an ensemble and find the accuracy for each. Finally, use Random Forest feature importances and plot the most important features for the prediction.   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Sklearn ( For Machine Learning ) * Seaborn ( For EDA )   Findings   * Model for Support vector machines with Accuracy: 0.6875 * Model for Random forest with Accuracy: 0.7361111111111112 * Plotting of most important features for the prediction * Predictions result * Computing error metrics   Link   * <https://www.kaggle.com/fatemahkhan/practice-21-1-2021> |
| 34 | Methodology   * RandomForestClassifier * LogisticRegression * SVM   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Sklearn ( For Machine Learning )   Findings   * Models of [LogisticRegression, Randomforest, SVM]   Link   * <https://www.kaggle.com/askal100/notebookb7e6e1de64> |
| 35 | Methodology   * Data Visualization   Technique   * Libraries [readr, ggplot2, dplyr, viridis, gridExtra, rpart, tree, rpart.plot, randomForest, corrplot, ggthemes, stringr, scales, plotly, funModeling, corrplot, fBasics, kableExtra]   Findings   * Plotting of [raisedhands, visitedresources, , announcementsview, discussion] * Plotting on count of gender * Plotting on raisedhands by nationality on relation * Plotting on count of [raisedhands, visitedresources, , announcementsview, discussion] based on L,M,H classes.   Link   * <https://www.kaggle.com/duyguatasever/students-academic-performance-visualization> |
| 36 | Methodology   * Cluster and Random Forest Analysis of Student Performance * PCA ( Principal Component Analysis ) * Decision Tree Analysis   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * SKlearn ( For Decision Tree ) * Matplotlib ( For plotting graph ) * Seaborn ( For PCA )   Findings   * The graph of class -> Decision and class -> Raise Hand * Cluster plotting using KMean Algorithm * Plotting of data based on the Principal Component Analysis * The Mean Square error for the Random Forest Classifier and KMean Algorithm   Link   * <https://www.kaggle.com/znielsen/analysis-and-prediction-of-student-performance> |
| 37 | Methodology   * Feature Classification * Data Preparation   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing )   Findings   * Numerical Features from the dataset * Categorical Features from the dataset   Link   * <https://www.kaggle.com/slamnz/correlation-table-for-student-academ> |
| 38 | Methodology   * Linear Classifier * DNN ( Deep Neural Network )   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * TensorFlow ( For Linear Classifier and DNN ) * Seaborn ( For EDA )   Finding   * The graph for the Class -> Discussion and Class -> Visit Resources. * The Accuracy score for the Linear Classifier and Deep Neural Networks.   Link   * <https://www.kaggle.com/biscuitlickz/student-grade-prediction> |
| 39 | Methodology   * Random Forest Classifier for predicting the student performance * GridSeachCV   Technology   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Sklearn ( For Random Forest and GridsearchCV )   Findings   * The confusion matrix and classification report for the Random Forest Classifier * Hyper-Parameter which will suite best to the current dataset   Link   * <https://www.kaggle.com/davified/random-forest-classifer-0-94-recall-rate> |
| 40 | Methodology   * Data Analysis * Feature Importance   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Seaborn ( For EDA )   Findings   * Count plot on the bases of [ Gender , class ] * The accuracy score for the Decision Tree Classifier * Apply the KFold Validation * List of all features with importance percentage while training model   Link   * <https://www.kaggle.com/pizero/general-dataset-analysis> |
| 41 | Methodology   * EDA ( Exploratory Data Analysis )   Technique  Findings   * The count graph for the Integer Variable. * The count graph for the factor Variable. * Relationship with class variable and Integer Variable. * Relationship with class variable and factor Variable. * The co-relation graph for all the features. * Visualizing the Decision tree based on the dataset * Importance of each columns by RandomForest   Link   * <https://www.kaggle.com/c34klh123/eda-for-student-data/report> |
| 42 | Methodology   * Data Visualization   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data Visualization ) * Seaborn ( For EDA )   Findings   * Sworn Plot [ High ,middle , low level student -> Raise hand , Visit Resource Variable , Announcement View ] * The average score of the Jordan grades * The average success score of the Chemistry Topic * Relation with mother or father affect success of student * Sworn Plot Discussion group participation * Graph of Absence effect on success * Positive and Negative Effect of these two variables [ Being Male , Live in Jordan , Work on chemistry , Relation with Mother , Participation Discussion , Absence Days (above 7) ]   Link   * <https://www.kaggle.com/kanncaa1/factors-affecting-success-in-school/notebook> |
| 43 | Methodology   * Hypothesis testing and Visualization   Technique   * Ggplot2 ( For plotting the graph )   Findings   * The graph based on the Number of Raise Hand and Gender   Link   * <https://www.kaggle.com/syuzanna/hypothesis-testing-and-visualization> |
| 44 | Methodology   * Random Forest Classification Algorithm   Technique    Findings   * Mean Decrease Accuracy of all the features * Mean Decrease Gini of all the features   Link   * <https://www.kaggle.com/syuzanna/randomforest> |
| 45 | Methodology   * Decision Tree Algorithm   Technique   * Caret   Findings   * Visualization of Decision Tree based on the features in the dataset   Link   * <https://www.kaggle.com/syuzanna/decision-tree/notebook> |
| 46 | Methodology   * Naive Bayes Algorithm   Technique   * Caret   Findings   * The graphs of the naïve bayes algorithm with True Positive and False Positive   Link   * <https://www.kaggle.com/syuzanna/naive-bayes> |
| 47 | Methodology   * Data Analysis * Machine Learning Algorithm * Dimensionality Reduction   Techniques   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Seaborn ( For EDA ) * Matplotlib ( For data visualization ) * Sklearn ( For Machine Learning Algorithm )   Findings   * Count Plot for visualization of data [ Gender , Relation , Topic , Section ID , Grade ID , Nationality , Class , Stage ID , Semester , Parent Answering Survey , Parent School satisfaction , Student Absence Day ] * The graphs of Dimensionality Reduction of all the feature available * The graphs of Scaled Algorithm Comparison with all the trained model [ LR , Lasso , EN , KNN , DTR , SVR ] * Tuning the Lasso algorithm * Scaled Ensemble Algorithm Comparison with Standard Scaler and [ AB , GBR , RF , ET ] and KFold validation. * Men Square Error for the AdaBoost and Lasso   Link   * <https://www.kaggle.com/rocki37/students-scores-analysis> |
| 48 | Methodology   * Data Analysis * Predicting the Values with the help of Machine Learning   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Seaborn ( For EDA ) * Sklearn ( For Machine Learning Algorithm )   Findings   * Get the Accuracy for the Logistic Regression , SVC , Gaussian NB , Decision Tree Classifier , Decision Tree Classifier , Random Forest Classifier and get the confusion matrix of all the algorithms.   Link   * <https://www.kaggle.com/kartikjoshi09/academic-performance-student-xapi> |
| 49 | Methodology   * Data Visualization * Classification Algorithm   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Sklearn ( For Classification Algorithm )   Findings   * Count Plot of the [ Number of Student by gender , Most Popular Topic , Most Present Nationality , School Level , Number of Discussion ] * The accuracy score , Precision Score , Confusion matrix and f1 score for the Logistic Regression , Naïve Bayes , Linear SVM , KNN , Random Forest.   Link   * <https://www.kaggle.com/pedroharagao/classification-algorithms-applied-to-educational> |
| 50 | Methodology   * Students' Performance Prediction Using Deep Neural   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * plotly.express ( For Plotting graph ) * Sklearn ( For Classification Algorithm ) * Tensorflow (For Machine Learning)   Findings   * Model with Accuracy : 0.9902 * Plotting of Training and Validation Loss   Link   * <https://www.kaggle.com/parth899/students-performance-prediction-using-deep-neural> |
| 51 | Methodology   * Prediction Using XgBoost   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Sklearn ( For Classification Algorithm ) * Seaborn ( For EDA ) * XgBoost (For Machine Learning)   Findings   * SVM Model with Accuracy : 0.7083333333333334   Link   * <https://www.kaggle.com/asheshmjoshi/student-academic-prediction-xgboost> |
| 52 | Methodology   * Predicting gender based on educational performance with [DecisionTreeClassifier, RandomForestClassifier, KNeighborsClassifier, MLPClassifier]   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Sklearn ( For Machine Learning Algorithm )   Findings   * Plotting on scatter matrix * Failed Model we can’t predict with high precision student's gender by educational performance.   Link   * <https://www.kaggle.com/michakapaa/gender-competition> |
| 53 | Methodology   * Predicting Academic Performance with TensorFlow ANN   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * plotly ( For Plotting Graph ) * Sklearn ( For Machine Learning Algorithm ) * TensorFlow (For Machine Learning )   Findings   * Plotting on training and validation loss * ANN model with accuracy 0.6979   Link   * <https://www.kaggle.com/gcdatkin/predicting-academic-performance-with-tensorflow> |
| 54 | Methodology   * In this, I am interested with the effect of the raising hands, visiting resources and viewing announcements. * Then found two of students have low level grade although they have higher values of raising hands, visiting resources and viewing announcements and start researching it **why !!!.**   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Seaborn ( For EDA )   Findings   * Plotting on High, middle and low level students of [raising hands, visiting resources, viewing announcements] * Plotting on Students Success on different topics * Plotting on Relation with father or mother affects success of students * Plotting on Discussion group participation * Plotting on Absence effect on success * Positive and negative effects on success of these two students can be seen.as   Link   * <https://www.kaggle.com/habibmrad1983/factors-affecting-success-in-school-habib-mrad> |
| 55 | Methodology   * Data Visualization * EDA * Feature Engineering   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * plotly ( For data visualization ) * Seaborn ( For EDA )   Findings   * Plotting on [Student Success for Topic, Student Success for NationalITy, Student Success for Student Absence Day, ParentschoolSatisfaction, Relation, StageID] * Plotting on student Success by nationality * Box plot of [raisedhands, VisITedResources, AnnouncementsView, Discussion, studentsuccess] * Relationships Between Numeric Variables [raisedhands, VisITedResources, AnnouncementsView, Discussion] * Correlation matrix * Testing results * Plotting of PCA   Link   * <https://www.kaggle.com/afranur/data-visualization-eda-feature-engineering> |
| 56 | Methodology   * RandomForestClassifier * CalibratedClassifierCV * LogisticRegression   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Sklearn ( For Machine Learning Algorithm )   Findings   * Random Forest Model   Link   * <https://www.kaggle.com/khamaralmaskhan/studentperformancedataset> |
| 57 | Methodology   * DecisionTreeClassification * RandomForestClassifier   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Sklearn ( For Machine Learning Algorithm )   Findings   * Plotting on Number of times vs number of students raised their hands in a certain time. * Plotting on Number of times vs number of student visted resource in a certain time. * Plotting of categorical features with numerical features * Model of Decision Tree with Accuracy : 0.6875 * Model of Naive Bayes with Accuracy: 0.6666 * Model of Random Forest Classifier with Accuracy: 0.7708333333333334 * Plotting of Feature Importance of Randomn Forest Classifier   Link   * <https://www.kaggle.com/zeeshankhand/machine-learning-course-project> |
| 58 | Methodology   * RandomForestClassifier * EDA * KNeighborsClassifier * LogisticRegression * Seaborn ( For EDA )   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Sklearn ( For Machine Learning Algorithm )   Findings   * Plotting on How students perform in different topics * Plotting on how satisfied parents are with each school of study. * Classification Model with accuracy : 0.6666 * KNN model with accuracy : 0.625 * random forest Model with Accuracy : 0.71875   Link   * <https://www.kaggle.com/tuli09/student-s-performance-analysis> |
| 59 | Methodology   * EDA   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * plotly ( For data visualization ) * Seaborn ( For EDA )   Findings   * Plotting on Number of Students per Course * Plotting on Number of Boys and Girls per Course * Plotting on ParentschoolSatisfaction by topics * Plotting on Average active participation in the classroom * Plotting on Absences of School Satisfaction * Plotting of discussion and School satisfaction   Link   * <https://www.kaggle.com/lariss4vitoria/dados-edu-india-larissa> |
| 60 | Methodology   * EDA * LogisticRegression * RandomForestClassifier * DecisionTreeClassifier * XGBClassifier   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * plotly ( For data visualization ) * Seaborn ( For EDA )   Findings   * Plotting on categorical data * Plotting of categorical features with numerical features * Plotting on relationships between numerical features. * Plotting on Gender Comparison With Parents Relationship * Logistic Regression Model with score : 0.8020833333333334 * XgBoost Model with Accuracy : 0.84375 * RandomForestClassifier with Accuracy : 0.84375   Link   * <https://www.kaggle.com/rupampatil/student-s-academic-performance-with-ml-eda> |
| 61 | Methodology   * EDA   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * TensorFlow (For Machine Learning ) * Seaborn ( For EDA )   Findings   * Plotting on categorical data * Plotting on Numerical data * Plotting on [Marks by gender, Marks by semester, Marks by absence] * Plotting on [raisedhands, AnnouncementsView]   Link   * <https://www.kaggle.com/stanislauhlebik/students-academic-performance-eda> |
| 62 | Methodology   * KNN   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Seaborn ( For EDA )   Findings   * KNN model with accuracy : 0.71875 * Plotting on Error Rate vs. K Value * Confusion Matrix   Link   * <https://www.kaggle.com/pradeeppradee/model-knn-to-predict-student-performance> |
| 63 | Methodology   * SVM * Logistic Regression * KNN * Decision Tree * Random Forest * Naive Bayes * XGBoost   Technique   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Sklearn ( For Machine Learning Algorithm )   Findings   * Plotting on categorical data * Plotting on Numerical data * SVM model with Accuracy : 0.7358490566037735 * Logistic Regression Model with Accuracy : 0.7735849056603774 * KNN model with Accuracy : 0.6289308176100629 * Decision Tree model with Accuracy : 0.7295597484276729 * Random Forest model with Accuracy : 0.8301886792452831 * Naive Bayes model with Accuracy : 0.49056603773584906 * XgBoost Model with Accuracy : 0.7924528301886793   Link   * <https://www.kaggle.com/ozkanbanu/matematik-m-hendisli-inde-tasar-m-uygulamalar-ml> |
| 64 | Methodology :   * Data Exploration * Machine Learning   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Sklearn ( For Machine Leaning Algorithms )   Findings :   * Count Plot for the [ Class (Respect to the L,H,M) , Gender , Gender ( with respect to the L,M,H ) , Nationality , Nationality (With respect to the L,M,H) , Place of Birth ( With Respect to the L,M,H ) ,Stage ID , Stage ID ( with Respect to the L,M,H ) , Grade ID , Grade ID ( with respect to the L,M,H ) , Section ID , Section ID (with respect to the L,M,H)] * Maximum Score , Minimum Score , Average Score , Minimum Prediction , Maximum Prediction , Average Prediction for the perceptron Classifier * RBF [ Minimum , Maximum , Average ] Score , Linear [ Minimum , Maximum , Average ] , Polynomial [ Minimum , Maximum , Average ] Score , Sigmoid [ Minimum , Maximum , Average ] Score , Prediction Misses [ RBF , Linear , Polynomial , Sigmoid ] * 3-Depth Tree [ Minimum , Maximum , Average ] Scores and 5-Depth Tree [ Minimum , Maximum , Average ] Score * The Final Result Average Accuracy Score of all the Algorithms [ Perceptron , SVM(rbf) , SVM(linear) , SVM(polynomial(1)) , SVM(polynomial(2)) , SVM(polynomial(3)) , SVM(polynomial(5)) , SVM(Sigmoid) , Random Forest (depth=3) , Random Forest (depth=5)]   Link :  <https://www.kaggle.com/entropicinertia/predicting-student-academic-performance> |
| 65 | Methodology :   * Data Analysis * Machine Learning   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Seaborn ( For EDA ) * Sklearn ( For Machine Learning Algorithms )   Findings :   * BoxPlot and Swarn Plot for [ Class -> Raise Hand , Class -> Visit Resources , Class -> Announcement , Class -> Discussion] * Factor plot for [ Class -> Parent Answering Survey , Class -> Parent School Satisfaction , Class -> Student Absence Days ] * F1 Score , Accuracy Score and Classification Report for the Linear Regression and Random Forest Classifier   Links :   * <https://www.kaggle.com/chengpp/notebook1dd4a74262> |
| 66 | Methodology :   * Exploratory Data Analysis * Data Cleaning * Data Visualization * Feature Engineering   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Seaborn ( For EDA ) * Sklearn ( For Machine Learning Algorithms )   Findings :   * The histogram of No. of Student vs No. of student Raised their hands on particular time. * The graphs based on [ Gender (High , Low , Medium) , Students from the country , Stage ID (High , Low , Medium) , Grade ID (High , Low , Medium) , Topics , Topic (High , Low , Medium) , Relation With Parents , Parent School Satisfactory , Parent Answering Survey ] * The graph of the Feature Importance with F Score. * The Kdeplot for topic with respect to the High , Low , Medium * The Co-relation Graph for the every features with each other. * Accuracy score and classification Report for the [ SVC , Random Forest Classifier , Logistic Regression , Multi-Layer Perceptron , XGBoost ]   Link :   * <https://www.kaggle.com/nasirislamsujan/an-eda-on-sap> |
| 67 | Methodology :   * Data Visualization * Exploratory Data Analysis * XGBoost * Deep Learning   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Seaborn ( For EDA ) * Sklearn ( For Machine Learning Algorithms ) * Plotly ( For data visualization )   Findings :   * The graphs on [ Gender , Nationality , Place of Birth , Stage ID ( Middle school , lower level , High level ) , Grade ID , Topic (with respect to vote count) , Percentage and Semester , Relation of Mother and Father , Raise Hand and Percentage , Parent School Satisfaction and Percentage , Parent Answering Survey and Percentage , Student Absence Day and Percentage , Class and Percentage ] * The count Graph of the [ Class , Gender , Stage ID (Low , Medium , High) , Semester , Topic , Nationality , Topic (M,F) , Nationality (M,F) , Nationality (Father , Mother) , Nationality (Above-7 and Under-7) ] * The Subplot on [ Class and Visit Resources (L,M,H) , Class and Announcement(L,M,H) , Class and Raise Hand (L,M,H) , Class and Discussion (L,M,H) , Gender and Raise Hands , Gender and Discussion ] * The Swarm Plot on [ Gender and Announcement , Gender and Raise Hands ] * The Box Plot on [ Class and Discussion , Class and Visit Resources ] * The Point Plot on [ Semester and Visit Resource , Semester and Announcement ] * The regression Plot [ Raise Hand and Visit Resource , Announcement View and Discussion ] * Graph Analysis of Gender Vs Place of Birth * Accuracy Score and Classification Report of [ Logistic Regression and XGBoost ] * Feature Importance graph to the Features in the dataset   Link :   * <https://www.kaggle.com/harunshimanto/student-s-academic-performance-with-ml-eda> |
| 68 | Methodology :   * Data Visualization   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * MatplotLIb ( For data Visualization ) * Seaborn ( For EDA )   Findings :   * The count graph on [ Class (with respect to the M,L,H) , Topic (With respect to the M,L,H) , Class (With respect to the gender) , Place Of Birth (with respect to the Pass , Fail) , Stats (With respect to the Father and Mother) , Parent Answering Survey (With respect to the Pass and Fail) , Student Absence Days (With respect to the Pass and Fail)]   Link :   * <https://www.kaggle.com/sid321axn/data-visualization-my-1st-kernel> |
| 69 | Methodology :   * EDA (Exploratory Data Analysis) * Clustering   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * MatplotLIb ( For data Visualization ) * Seaborn ( For EDA ) * Sklearn ( For Clustering )   Findings :   * The visualization of data in the different type of graphs with varying from the BoxPlot , FactorPlot , CountPlot , SwarnmPlot. * The heat map of correlation matrix with features [ Raise Hand , Visit Resources , Announcement View , Discussion ] * The accuracy score and Classification Report for the Random Forest Classifier , SVC . * The cluster using KMean Clustering Algorithm. * Feature importance graph with all the features.   Link :   * <https://www.kaggle.com/bossimuimu/exploratory-data-analysis-and-clustering> |
| 70 | Methodology :   * Data Visualization * Machine Learning (SVM)   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * MatplotLIb ( For data Visualization ) * Seaborn ( For EDA ) * Sklearn ( For SVM)   Findings :   * The count graph of the [ Class (with respect to the M,L,H) , Topic , Gender , Nationality ]. * The heat map of all the features with respect to each other. * The Accuracy Score and Classification Report for the SVC (Support Vector Classifier) , Linear SVC , Non-Linear SVC with rbf. * The Data Frame which describe all the Model’s score.   Link :   * <https://www.kaggle.com/rangarirb/svm-classification> |
| 71 | Methodology :   * Data Visualization * Machine Learning (SVC)   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * MatplotLIb ( For data Visualization ) * Sklearn ( For SVM , Data pre-processing )   Findings :   * The count graph on [ Class (with respect to the L,M,H) , Topic , Gender , Nationality ] * The Accuracy score before and after scaling the data using SVC ( Support Vector Classifier )   Link :   * <https://www.kaggle.com/ahsanmemon/predicting-performance> |
| 72 | Methodology :   * EDA ( Exploratory Data Analysis ) * Machine Learning   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * MatplotLIb ( For data Visualization ) * Sklearn ( For Machine Learning , Data pre-processing )   Findings :   * Cor-relation heatmap of features [ Raise Hand , Announcement View , Visit Resources , Discussion ] * The Line Plot on [ Raise Hand , Raise Hand and Visit Resources , Discussion and Raise Hand ] * Subplot on Raise Hand and Visit Resource * Histogram Plot on [ Raise Hand ] * Histogram Subplot with non cumulative and cumulative frequency * Scatter Plot on [ Raise Hand and Discussion , Raise Hand and Announcement , Raise Hand (in terms of gender) , Discussion (in terms of gender) ] * Bar Plot on [ Raise Hand (With respect to the topic) , Topic (with respect to the Raise Hand , Discussion , Announcement View) , Discussion and Raise Hand (according to the Topic) , Raise Hand and Discussion (according to the Place Of Birth) ] * Point Plot on [ Raise Hand and Discussion , Raise Hand and Discussion (according to the Place Of Birth) ] * Count Plot on [ Gender , Stage ID (according to the L,M,H) ] * Pie Chart on [ Stage ID ] * The Accuracy Score on [ Logistic Regression Classifier , KNN Classifier , SVM Classifier , Naïve Bayes Classifier , Decision Tree Classifier , Random Forest Classifier ] * The Confusion matrix of [ Random Forest Classifier , KNN Classifier , Decision Tree Classifier ]   Link :   * <https://www.kaggle.com/sibelkcansu/machine-learning-classification> |
| 73 | Methodology :   * Data Exploration * Classification   Technique :   * MASS * ISLR * Tree * RandomForest * E1071   Findings :   * The Overview of data exploration * The Accuracy Score for the SVM (Support Vector Machine) , Decision Tree and Random Forest.   Link :   * <https://www.kaggle.com/sinnfashen/kernel90343248fe> |
| 74 | Methodology :   * ANN (Artificial Neural Network) * SVM (Support Vector Machine )   Technique :   * MASS * Rsamplr * Lattice * Klar * ElemStatLearn * Ggplot2 * Recipes * Keras * E1071   Findings :   * Logistic Regression Accuracy Score and Prediction Score * Artificial Neural Network Findings like Accuracy Score * The history graph of the ANN (Artificial Neural Network) training and Validation using epoch and loss * SVM Accuracy Score and Prediction Score * Algorithms Performance Comparison with all the 3   Link :   * <https://www.kaggle.com/urvimistry/hw4-urvi-mistry> |
| 75 | Methodology :   * Data Processing * ANN ( Artificial Neural Networks ) * Machine Learning ( Logistic Regression and SVM )   Techniques :   * Caret * Mlbench * Arules * Tidyverse * Rsample * MASS * Lattice * Klar * ElemStatLearn * Ggplot2 * Recipes * Kera * E1071 * Tensorflow   Findings :   * Accuracy Score for the Logistic Regression * Training and Validation Accuracy Score for ANN * System time taken by ANN and Logistic Regression   Link :   * <https://www.kaggle.com/mkadam26/homework4-madhavi-kadam> |
| 76 | Methodology :   * Data Analysis * Clustering * Machine Learning   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Sklearn ( For Machine Learning , Data pre-processing )   Findings :   * Cluster of Raise Hand and Visit Resources * Accuracy Score for the [ KNN , Decision Tree , SVM , Logistic Regression ]   Link :   * <https://www.kaggle.com/marlonferrari/data-analysis-student-s-behavior> |
| 77 | Methodology :   * Data Exploration * Data Visualization * Logistic Regression   Techniques :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Sklearn ( For Machine Learning , Data pre-processing ) * Matplotlib ( For data visualization )   Findings :   * The count graph on [ Gender , Gender (with respect to the under-7 and Above-7) , Gender (with respect to the Pass and Fail) , Gender (with respect to the L,M,H) , Gender (with respect to the Quiet and Active) , Parent School Satisfaction (with respect to the Bad and Good) ] * Accuracy Score for the Logistic Regression Model   Link :   * <https://www.kaggle.com/rashmiek99/student-performance> |
| 78 | Methodology :   * Data Visualization * Machine Learning * Feature Engineering   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Sklearn ( For Machine Learning , Data pre-processing ) * Matplotlib ( For data visualization )   Findings :   * Exploratory Data Analysis on [ Raise Hand , Visit Resources , Announcement View , Discussion ] * BoxPlot on [ class and (Raise Hand , Visit Resources , Announcement View , Discussion ) ] * Heat Map of features [ Raise Hand , Visit Resources , Announcement View , Discussion ] * Vertical Count Graph of Nationality , Place of Birth , Grade ID , Topic * Accuracy score , Classification Report and Confusion matrix of [ Logistic Regression , XGBClassifier , KNN , Decision Tree Classifier , Random Forest , SVC , LGMBClassifier ]   Link :   * <https://www.kaggle.com/sarzhan/students-perfomance-tutorial> |
| 79 | Methodology :   * Data Exploration * SVM ( Radial , Linear , Sigmoid , Polynomial )   Technique :   * Tidyverse * Random forest * Class * Rpart * E1071 * Caret * Corrplot * caTools * party * DataExplorer   Findings :   * SVM prediction of data based on [ Linear , Radial , Sigmoid , Polynomial ] * The graph on Performance of SVM with epsilon and cost * Fine Tuning Model   Link :   * <https://www.kaggle.com/robertkibet/using-svm-to-predict-student-performance-class> |
| 80 | Methodology :   * Data Exploration * Machine Learning   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Sklearn ( For Machine Learning , Data pre-processing ) * Matplotlib ( For data visualization )   Findings :   * The countPlot on [ Gender , Nationality , Place of Birth , Stage ID , Grade ID , Section ID , Topic , Semester , Relation , Parent Answer Survey , Parent School Satisfaction , Student Absence Days , Class ] * The countPlot on [ Gender (with respect of L,M,H) , Nationality (with respect of L,M,H) , Place of Birth (with respect of L,M,H) , Stage ID (with respect of L,M,H) , Grade ID (with respect of L,M,H) , Section ID (with respect of L,M,H) , Topic (with respect of L,M,H) , Semester (with respect of L,M,H) , Relation (with respect of L,M,H) , Parent Answer Survey (with respect of L,M,H) , Parent School Satisfaction (with respect of L,M,H) , Student Absence Days (with respect of L,M,H) , Class (with respect of L,M,H) ] * The BoxPlot on [ Class and Raise Hand , Class and Visit Resource , Class and Announcement View , Class and Discussion ] * Heat Map on all the important Features * Cross Validation score for the [ Logistic Regression . SVM , KNN , Decision Tree , Random Forest , Gradient Boosting ] * The confusion matrix for the [ Logistic Regression and Gradient Boosting ] * The hyper parameter and AUC Score of the Logistic Regression   Link :   * <https://www.kaggle.com/kvsivasankar/students-academic-performance-model> |
| 81 | Methodology :   * Exploratory Data Analysis   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Seaborn ( For EDA )   Findings :   * The Pie Chart of the [ Nationality , Place of Birth , Nationality and Place of Birth ] * Histogram on the [ Announcement View , Discussion , Raise Hand , Visit Resources ] * Box Plot on [ Gender and (Discussion , Visit Resources , Announcement View , Raise Hand) ] * Sworm Plot on [ Gender and (Discussion , Visit Resources , Announcement View , Raise Hand) ] * Sub Plot on [ Gender and (Discussion , Visit Resources , Announcement View , Raise Hand) ] * Heat Map of [ Raise Hand , Visit Resources , Announcement View , Discussion , Score ]   Link :   * <https://www.kaggle.com/tommzzhou/exploratory-data-analysis> |
| 82 | Methodology :   * Data Exploration * Machine Learning   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Sklearn ( For Machine Learning )   Findings :   * The Subplot of [ Gender , Gender (with respect to L,M,H) , Relation , Relation (with respect to L,M,H)] * The Pie Chart on [ Yes and No , Good and Bad , Nationality ] * The count Plot on [ Topic (with respect to gender) , Nationality (with respect to gender) , Grade ID (with respect to the L,M,H) ] * The Accuracy score of the [ SVM , Logistic Regression , Random Forest , Naïve bayes , Decision Tree ]   Link :   * <https://www.kaggle.com/caglatoprak/academic-performace-analysis> |
| 83 | Methodology :   * EDA   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization )   Findings :   * The count Plot on [ Class (with respect to L,M,H) ]   Link :   * <https://www.kaggle.com/manish1984/seaborn-tutorial-count-plots> |
| 84 | Methodology :   * Data Exploration * PCA (Principal Component Analysis) * Machine Learning   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Sklearn ( For Machine Learning )   Findings :   * Histogram on different types of data visualization * PCA Graph * K-Mean Cluster graph * Histogram of cluster count * Graph of GMM ( Gaussian Mixture ) * Data Frame of Cluster   Link :   * <https://www.kaggle.com/haochengmei/hw4-meihaocheng-tom> |
| 85 | Methodology :   * Exploratory Data Analysis * Classification * Feature Engineering * Machine Learning   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Sklearn ( For Machine Learning ) * Seaborn ( For EDA )   Findings :   * The count graph of [ Nationality , Class (with respect to gender) ] * Accuracy Score for the [ KNN , Naïve bayes , SGD (stochastic Gradient Descent) , SVM , Decision Tree , Random Forest , XGBoost , Logistic Regression , ] * The graph of Error rate and K value of KNN.   Link :   * <https://www.kaggle.com/obirgul/understanding-classification-for-beginners> |
| 86 | Methodology :   * Data Analysis * Data Visualization   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization )   Findings :   * Box Plot on Class and (Discussion and Visit Resources) * Bar Plot on Visit Resource and Announcements View   Link :   * <https://www.kaggle.com/ghydaa/project> |
| 87 | Methodology :   * Data Visualization * Exploratory Data Analysis * Classification   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Seaborn ( For EDA ) * Sklearn ( For Machine Learning )   Findings :   * The count plot of [ Nationality , Class (with respect to gender)] * The scatter plot of Raise Hand and Visit Resources * The Accuracy score and Confusion matrix of [ Logistic Regression , KNN , SVM , Naïve Bayes , Decision Tree , Random Forest ]   Link :   * <https://www.kaggle.com/canayakincansel/students-academic-performans> |
| 88 | Methodology :   * Machine Learning * Data Cleaning   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Matplotlib ( For data visualization ) * Sklearn ( For Machine Learning )   Findings :   * The Accuracy Score for the [ Naïve Bayes , Decision Tree , Random Forest ] * The heat map of confusion matrix [Naïve Bayes , Decision Tree , Random Forest ] * The Data Frame with Model name and Score for that model   Link :   * <https://www.kaggle.com/rahaf1418/project-513> |
| 89 | Methodology :   * Classification * Data Cleaning   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Sklearn ( For Machine Learning )   Findings :   * The Precision Score , Recall Score and F1 Score for the [ SGDClassifier with cross validation , Random Forest ] * Heat Map of the confusion matrix with the parameters of [ Middle Level , Lowe Level , High Level and Middle Level , Lowe Level , High Level ]   Link :   * <https://www.kaggle.com/gopalkrishnanayak/students-academic-performance-classification> |
| 90 | Methodology :   * Feature Engineering * PCA ( Principal Component Analysis ) * Data Cleaning   Technique :   * NumPy ( For Linear Algebra ) * Pandas ( For data processing ) * Sklearn ( For PCA)   Findings :   * Plotting of 3 attribute in a 3D scatter plot [ Visit Resources , Discussion , Raise Hand ] * Histogram on the importance of each feature vector using explained variance * Projection Matrix of the Features.   Link :   * <https://www.kaggle.com/siddalore/academic-performance-study> |