**Project Topic: Student performance prediction system.**

The goal of this project is to develop a web based student performance prediction system which will predict student performance using Scikit-learn, a popular machine learning framework. The system will prompt users to input student data like demographic data, previous records etc. All the steps are explained in detail below:

1. Data Collection:

* We collected student performance data from various sources, importing it into a Pandas DataFrame for further analysis.

*Libraries used:*

* Pandas - Used for reading and loading data into a DataFrame for manipulation and preprocessing.

1. Exploratory Data Analysis (EDA):

* We conducted exploratory data analysis to understand the dataset's characteristics and distributions, visualizing relationships between features and the target variable.

*Libraries used:*

* Pandas: Used for data exploration and summary statistics.
* Matplotlib and Seaborn: Employed for creating plots and visualizations to analyze data distributions and relationships.

1. Feature Engineering:

* We selected relevant features based on domain knowledge and performed necessary transformations on the data including label & one hot encoding as well as data scale standardization.

*Libraries used:*

* Pandas: Utilized for encoding categorical data into numerical data.
* Scikit-learn (sklearn): Optionally used for feature scaling or normalization.

1. Model Evaluation and Selection:

* We compared the performance of different models based on evaluation metrics and selected the best-performing model(s) for deployment.

*Algorithms used:*

* Scikit-learn (sklearn): Used for model evaluation, including accuracy score calculation and comparison.
* GridSearchCV: Employed to evaluate which model performs better on the selected data.

1. Model Development:

* We trained machine learning models to predict student performance, evaluating their performance using various metrics.

*Libraries used:*

* Scikit-learn (sklearn): Used for implementing machine learning algorithms, model training, evaluation, and hyper parameter tuning.

*Algorithms:*

* Support Vector Machines (SVM): Utilized for classification tasks, including student performance prediction.

1. Deployment:

* We deployed the model through a web application with frontend designed using HTML, CSS and JavaScript as well as Django – Python web framework for backend.