**MACHINE LEARNING**

**PROJECT**

**HSE LEVEL STUDENTS GPA ANALYTICS**

**NAME:AJITH PRASANNA .R**

**ROLL NO :195214103**

**CLASS :1 MSC CS ‘A**

**HSE Level Student’s Gpa**

**Problem Statement**

Build a machine learning model to predict grade point average,the input variables are the marks in various subjects and output target variable is gpa.

**ML methodology**

The Machine Learning Methodology used here is **SGDRegressor** .SGD stands for stochastic gradient descent.In (stochastic gradient descent)we repeatedly run through the training set and each time we encounter a training example.we update the paramaters according to the gradient of the error with respect to that singl

**Data set description**

This dataset contains the marks of 526 students in 7 different subjects and their grade point averages (399 in train data,127 in test data).

This dataset have **126** rows **11** columns.

**Pre-processing**

**Load dataset**: Using pandas load csv file format of our data

**Shape of data**: Use .shape method get the no of rows and columns of dataset

**head()**: Use .head() method to show the first 5 data of dataset.

**tail():** Using .tail() method will show the last 5 data of dataset

**isnull()**: This method will shown any null values of the dataset if exists

**isnull.sum()**: Use to find the sum of null values of dataset.

**describe()**: This method will describe all the feautures of the dataset like mean count max std variance etc..

**info()**: This will return information about the dataframe including the data types of each column and memory usage

**corr()**: Used to find the pairwise correlation of all columns in the dataframe.

**dtypes**: This returns a Series with the data type of each column.

**value\_counts():** This method will be shown in count outcome values

**Building ML model(s)**

The following list  summarizes our approach to developing an application using ML model.

* Problem Definition
* Data Collection
* Data Preparation
* Data Analysis
* Feature Engineering
* Training and testing dataset

Created a ML model by gathering and preprocessing data followed by identifying dataset imbalances.

**Training and evaluation of all ML models.**

***from sklearn.model\_selection import train\_test\_split***

After importing split x and y

X- independent Y-dependent

Split the data around 20%-80% between testing and training stages.

**ML algorithom:**

***from sklearn.linear\_model import SGDRegressor***

Here using classification algorithm (SGDRegressor)

Then perform Fit and predict operation

**Classification Report:**

**Computed**

* **Mean squared error and variance**

**Conclusion**

This model gives key to predict grade point average, input variables as the marks in various subjects and output target variable at gpa.In future will develop one real time project on data analytics.