A MINOR PROJECT REPORT ON

"PREDICTION OF STUDENT PERFORMANCE ANALYSIS"

Submitted

In the partial fulfilment of the requirements for
Data Mining Techniques Course

By

K. MAHESWARI	**(171FA04147)**
U. MOUNIKA	**(171FA04181)**
G. AISHWARYA	**(171FA04204)**

Under the esteemed guidance of S Shivaprasad, Asst Professor

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
VIGNAN'S FOUNDATION FOR SCIENCE, TECHNOLOGY AND RESEARCH
(Accredited by NAAC "A" grade)
Vadlamudi , Guntur.

VIGNAN'S
Foundation for Science, Technology & Research
(Deemed to be University)
-Estd. u/s 3 of UGC Act 1956

CERTIFICATE

This is to certify that the report entitled "**Prediction of Student Performance Analysis**" is being submitted by **K Maheswari(171FA04147), U Mounika(171FA04181)** and **G Aishwarya (171FA04204)** in partial fulfilment of course work of **Data Mining Techniques** as a Minor Project, carried in Department of CSE, Vignan's Foundation for Science, Technology and Research, Deemed to be University.

S Shivaprasad
Asst Professor

Head Of The Department

External Examiner

DECLARATION

We hereby declare that the project entitled "**PREDICTION OF STUDENT PERFORMANCE ANALYSIS**" submitted for the DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING. Thisdissertation is our original work and the project has not formed the basis for the award of any degree, associate-ship and fellowship or any other similar titles and no part of it has been published or sent for publication at the time of submission.

By

K Maheswari (171FA04147)
U Mounika(171FA04181)
G Aishwarya (171FA04204)

Date:

ACKNOWLEDGEMENT

We are very grateful to our beloved Chairman **Dr. Lavu Rathaiah**, and Vice Chairman **Mr. Lavu Sri Krishna Devarayalu,** for their love and care.

It is our pleasure to extend our sincere thanks to Vice-Chancellor **Dr. M.Y.S. Prasad** for providing an opportunity to do my academics in VFSTR.

It is a great pleasure for me to express my sincere thanks to **Dr. D. Venkatesulu HOD, CSE** of VFSTR**,** for providing me an opportunity to do my Mini Project.

We feel it our responsibility to thank **Mr. S. Siva Prasad** under whose valuable guidance that the project came out successfully after each stage.

We extend our whole hearted gratitude to all our faculty members of Department of Computer Science and Engineering who helped us in our academics throughoutcourse.

Finally we wish to express thanks to our family members for the love and affection overseas and forbearance and cheerful depositions, which are vital for sustaining effort, required for completing this work.

With Sincere regards,

K Maheswari (171FA04147)

U Mounika (171FA04181)

G Aishwarya (171FA04204)

CONTENTS:

ABSTRACT

Our project presents learning analytic of the students data for measuring the impact of students performance. For this purpose, a specifically designed Google Sheet questionnaire was developed and distributed across several undergraduate students of engineering colleges and universities. Data Mining is the most prevalent techniques to evaluate students performance and is extensively used in educational sector known as Educational Data Mining(EDM). The prime motto of our study is to discover the performance of students using some classification techniques and discovering the best one which yields optimal results.The dataset is pre-processed to filter duplicate records, missing fields are identified and filled with the destined data. Data Mining technique Gaussian Naive Bayes is employed on the dataset using Weka, Jupyter and minor tools. The data will be useful for institutions dealing with prevalent decision in society to find new hidden patterns from students data. The hidden patterns then are discovered can be used to understand the problem arise in the educational field.

Keywords: Data Mining, Educational Data Mining, Naive Bayes, student performance, hidden patterns.

INTRODUCTION

Nowadays, data mining is playing a vital role in educational institutions and one of the most important areas of research with the objective of finding meaningful information from the data stored in huge dataset. Educational data mining(EDM) is avery important research area which helpful to predict useful information from educational database to improve educational performance, better understanding and to have better assessment of the students learning process. Data Mining or knowledge discovery has become the area of growing significance because it helps in analyzing data from different perspectives and summarizing it into useful information.Hence, the main aim of this project is to establish a prediction model with credible accuracy rate by training a small dataset i.e. lesser number of records. Using visualization, classification and clustering algorithms the best attributes that are important in the prediction model are obtained. These attributes are fed to machine learning algorithms to evaluate and obtain an accurate model. In this project, the dataset is collected focusing on several factors that affects a student performance such as the students social life, education support, grades, prior academic performance, skills, interests etc likewise 36 attributes are considered.

LITERATURE SURVEY

An association mining technique as Apriori that uses the partial information about the contents of a shopping cart for the prediction of products that the

customers wish to buy or are more likely to buy along with the already brought products [1].

The proposed knowledge based decision technique will guide the student for admission in proper branch of engineering. Another approach is also developed to analyze the accuracy rate of for decision [2].

Previous research has focused predominantly on how to obtain exhaustive lists of associations. However, users often prefer a quick response to targeted queries. For instance, they may want to learn about the buying habits of customers. For this, paper proposes an approach that converts the market-basket database into an item set tree [3].

This paper proposes the use of a recommendation system based on data mining techniques to help students to take decisions on their academics itineraries. More specifically, it provides support for the student to better choose how many and which courses to enrol on, having as basis the experience of previous students with similar academic achievements [4].

Right from making test papers, conducting test and giving results; everything has to be done manually. A lot of paper work is involved. A lot of time and energy is required to get results, but still no accurate suggestion could be provided to the student to choose the stream.

So far, no system has been developed to help the college to analyse its admissions so that it can develop its branches on the basis of preference's submitted by students,.

Proposed Methodology

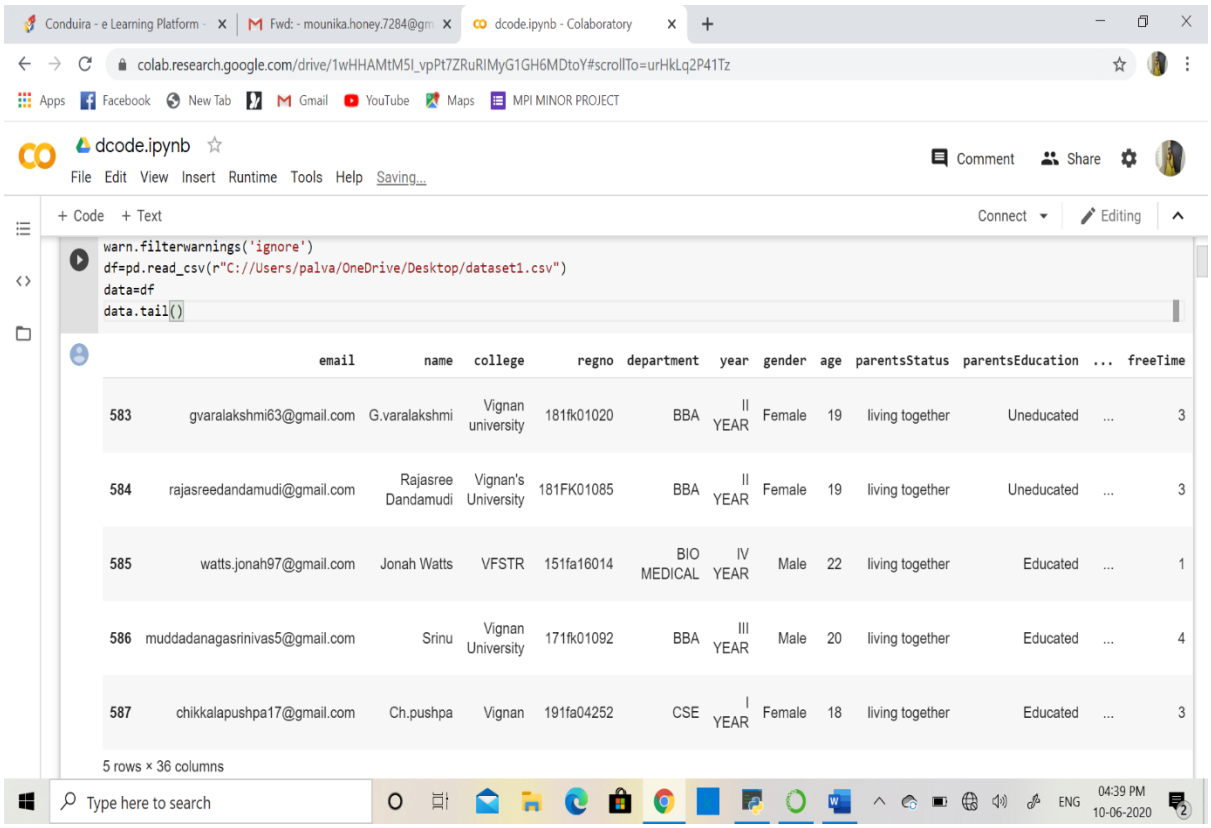
In order to analyze students performance with more accuracy, it requires more number of attributes to be considered. Here, the dataset comprises of 32 attributes which corresponds around students family support, financial support, skill set, time to study, educational support, social life, personal characteristics, social skills, problems, activities etc. The data is collected from college and university students of Andhra Pradesh. In this paper, we propose a classification algorithm of data mining i.e. Naive Bayesian algorithm to classify the data which is useful for the educational institutes. The entire dataset in analyzed by using clustering, visualization and also analysis between various attributes so that a clear picture of students performance can be known. Some association rules are also found using Apriori algorithm of Association Rule Mining(ARM) to know the dependency between the attributes. It greatly helps in understanding the common problems facing by the students. The Naive Bayesian algorithm here gives an accuracy rate of 72%.

Proposed Methodology

1. Data mining concepts
2. Choosing Classification technique
3. Building a model
4. Implementation of Naive Bayes
5. Choosing Apriori Algorithm for Association Rule Mining
6. Generating Association rules
7. Visualization of the results

Database Creation

The dataset used in this project is based on the data given by students of various colleges and universities in Andhra Pradesh. It is CSV file consists of 478 rows with 33 attributes.



Data preprocessing

Datasets in any data mining application can have should be noticed missing values. These missing values can get propagated due to lack of communication among the parameters in data collection system. These missing values can affect the performance of a data mining system, and it.

Proposed methods in our project

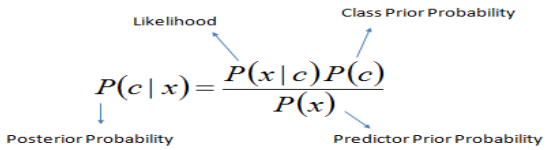
Classification algorithm seems the most appropriate type of algorithms to implement our proposed solution. Classification refers to the task giving a machine learning algorithm features and having the algorithm put the instance/data points into one of many discrete classes. Classes are categorical in nature, it isn't possible for an instance to be classified as partially another. A classic example of a classification task is classifying whether as it is either sunny or rainy. Here in classification we are using Naive Bayes algorithm for predicting performance of students based on attendance , cgpa ,extra curricular activities etc.

Steps followed in Architectural design

1. The data must be preprocessed in order to remove errors in the data. The replacing of missing values is done with default value. The preprocessing operations include data cleaning, data integration, data transformation, data reduction and data discretization.

2. The higher priority attributes are selected by which the student performance is measured. The lower priority attributes such as student names, age etc are ignored.

3. Now we use Naive Bayes algorithm to classify the data using the attribute probabilities. The Naive Bayes algorithm is based on Bayesian probability function which is used to predict the students performance.



$$P(c \mid X) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

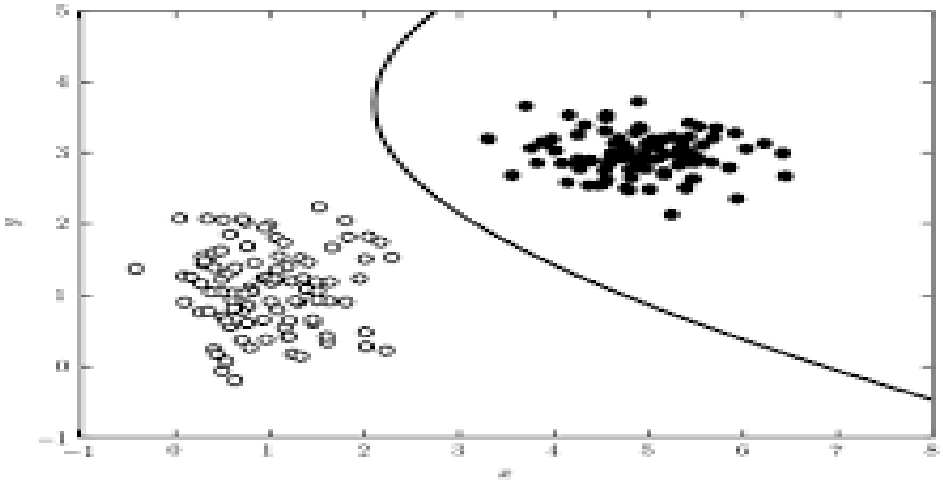
4. Now we use Apriori Algorithm for Association Rule Mining using WEKA tool.

5. Generate valid association rules basing on the parameters like minimum support count, confidence etc.

6. Visualization of results using graphs. This can be done using various libraries in python such as pandas, seaborn , matplotlib etc.

A Gaussian Naive Bayes algorithm is a special type of NB algorithm. Bayes theorem is based on conditional probability. The conditional probability helps us calculating the probability that something will happen, given that something else has already happened.

Gaussian Naive Bayes algorithm is specifically used when the features have continuous values. It's also assumed that all the features are following a Gaussian distribution i.e., normal distribution.



Data visualization is the act of taking information (data) and placing it into a visual context, such as a map or graph. Data visualizations make big and small data easier for the human brain to understand, and visualization also makes it easier to detect patterns, trends, and outliers in groups of data.

We need data visualization because a visual summary of information makes it easier to identify patterns and trends than looking through thousands of rows on a spreadsheet. It's the way the human brain works. Charts and graphs make communicating data findings easier even if you can identify the patterns without them.

Common general types of data visualization:

- Charts
- Tables
- Graphs
- Maps
- Info graphics
- Dashboards

Code:

Database Creation:

In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import warnings as warn
```

In [2]:
```
warn.filterwarnings('ignore')
df=pd.read_csv(r"C://Users/palva/OneDrive/Desktop/dataset1.csv")
data=df
data.tail()
```

Classification using Visualization:

In [3]:
```
fractions = data["onlineLearningRate"].value_counts().values.tolist()
plt.xlabel("onlineLearningRate")
plt.ylabel("percentage")
plt.title("attendance")
plt.hist(fractions,bins=5)
plt.show()
```

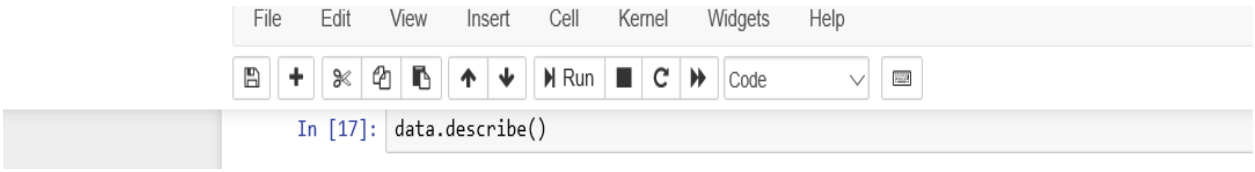
Run Code

In [7]:
```
sns.pairplot(data)
```

In [10]:
```
iF=data[data["internetFacility"]=="Yes"]
sns.pairplot(iF)
```

In [12]:
```
sns.distplot(data["onlineLearningRate"])
```

In [16]:
```
plt.plot(data["weeklyStudyTime"],data["backlogs"])
```



In [17]:
```
data.describe()
```

In [20]:
```
import pandas as pd
from sklearn import datasets, linear_model
from sklearn.model_selection import train_test_split
from matplotlib import pyplot as plt
import warnings
warnings.filterwarnings("ignore")
```

In [24]:
```
dat=pd.read_csv(r'C://Users/palva/OneDrive/Desktop/sampledataset1.csv')
dat.head()
```

Out[24]:

In [25]:
```
from sklearn import preprocessing
#creating labelEncoder
le = preprocessing.LabelEncoder()
# Converting string labels into numbers.
cgp=le.fit_transform(dat['CGPA'])
print(cgp)
```

Classification using Naive Bayes:

In [29]:
```
olr=le.fit_transform(dat['onlineLearningRate'])
print(olr)
```

```
[2 4 4 4 2 1 0 1 0 2 2 1 2 3 0 3 4 2 0 0 2 0 4 4 3 2 2 3 3 2 3 1 3 4 0 2 3
```

In [31]:
```
from sklearn.metrics import classification_report, confusion_matrix
from sklearn.model_selection import cross_val_score

def fit_predict_show_performance(classifier, X_train, y_train):

    y_test=np.reshape(cgp,(-1,1))
    prediction = classifier.predict(y_test)


    ten_performances = cross_val_score(estimator=classifier,X=X_train,y=y_train,cv=10)

    k_fold_performance = ten_performances.mean()

    print("====== Classification Report ======= ")

    cr = classification_report(y_test,prediction)

    print(cr)

    print("====== Confusion matrix ======= ")

    cm = confusion_matrix(y_test,prediction)
```

In [32]:
```
from sklearn.naive_bayes import GaussianNB
from sklearn.metrics import accuracy_score

classifier = GaussianNB()
model=classifier.fit(np.reshape(cgp,(-1,1)),np.reshape(olr,(-1,1)))
print(model)
print(classifier.predict([[0, 0]]))    #Churned data ---> 0==yes, 1==no
print(classifier.predict([[0, 1]]))    #internet data ---> 0==DSL,  1==Fiber optic  ,2==No
print(classifier.predict([[0, 2]]))
print(classifier.predict([[1, 0]]))
print(classifier.predict([[1, 2]]))


fit_predict_show_performance(classifier,np.reshape(cgp,(-1,1)),np.reshape(olr,(-1,1)))
```

In [34]:
```
txtLabels = data["CGPA"].value_counts().keys().tolist()
fractions = data["CGPA"].value_counts().values.tolist()
plt.title("CGPA data")
plt.pie(fractions, labels=txtLabels,
autopct='%1.1f%%', shadow=True, startangle=90,
colors=sns.color_palette('muted') )
plt.axis('equal')
```

Out[34]:
```
(-1.1150864057241259,
 1.1246993383603994,
 -1.1296619869215583,
```

In [35]:
```
txtLabels = data["learnFrom"].value_counts().keys().tolist()
fractions = data["learnFrom"].value_counts().values.tolist()
plt.title("Learn From data")
plt.pie(fractions, labels=txtLabels,
autopct='%1.1f%%', shadow=True, startangle=90,
colors=sns.color_palette('muted') )
plt.axis('equal')
```

In [36]:
```
txtLabels = data["careerObjective"].value_counts().keys().tolist()
fractions = data["careerObjective"].value_counts().values.tolist()
plt.title("Career Objective Data")
plt.pie(fractions, labels=txtLabels,
autopct='%1.1f%%', shadow=True, startangle=90,
colors=sns.color_palette('muted') )
plt.axis('equal')
```

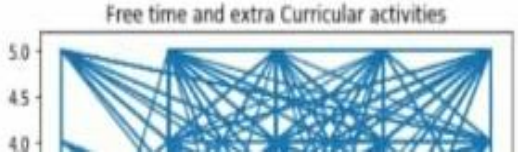
Out[36]:
```
(-1.1109050131289984,
 1.1024144262991773,
```

In [37]:
```
txtLabels = data["fears"].value_counts().keys().tolist()
fractions = data["fears"].value_counts().values.tolist()
plt.title("Fears Data")
plt.pie(fractions, labels=txtLabels,
autopct='%1.1f%%', shadow=True, startangle=90,
colors=sns.color_palette('muted') )
plt.axis('equal')
```

Out[37]:
```
(-1.1273201154442372,
 1.108850869006286,
 -1.1295619876303467,
```

In [40]:
```
plt.plot(data["freeTime"],data["extraCurricularActivities"])
plt.title("Free time and extra Curricular activities")
```

Out[40]:
```
Text(0.5, 1.0, 'Free time and extra Curricular activities')
```



```
In [41]: plt.plot(data["freeTime"],data["socialMediaUsage"])
         plt.title("Free time and Social Media Usage")

Out[41]: Text(0.5, 1.0, 'Free time and Social Media Usage')
```

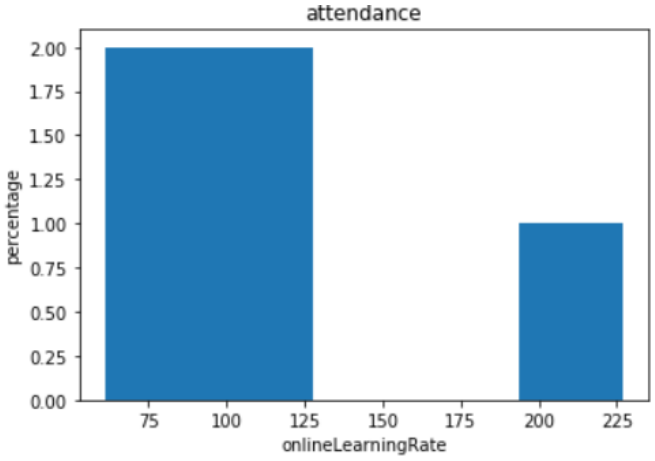
Results:

Out[2]:

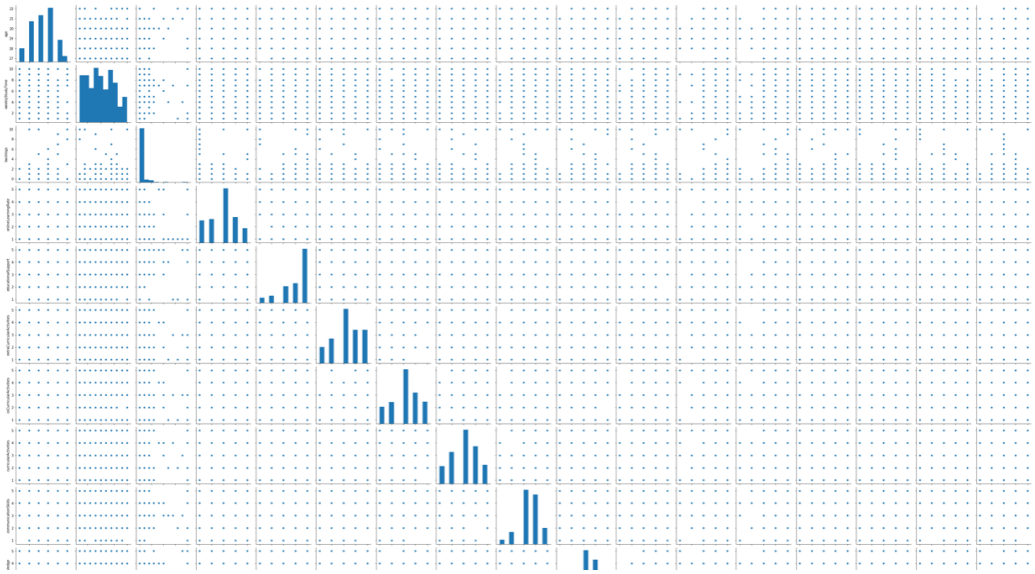
	email	name	college	regno	department	year	gender	age	parentsStatus	parentsEducation	...	freeTime	soc
583	gvaralakshmi63@gmail.com	G.varalakshmi	Vignan university	181fk01020	BBA	II YEAR	Female	19	living together	Uneducated	...	3	
584	rajasreedandamudi@gmail.com	Rajasree Dandamudi	Vignan's University	181FK01085	BBA	II YEAR	Female	19	living together	Uneducated	...	3	
585	watts.jonah97@gmail.com	Jonah Watts	VFSTR	151fa16014	BIO MEDICAL	IV YEAR	Male	22	living together	Educated	...	1	
586	muddadanagasrinivas5@gmail.com	Srinu	Vignan University	171fk01092	BBA	III YEAR	Male	20	living together	Educated	...	4	
587	chikkalapushpa17@gmail.com	Ch.pushpa	Vignan	191fa04252	CSE	I YEAR	Female	18	living together	Educated	...	3	

5 rows × 36 columns

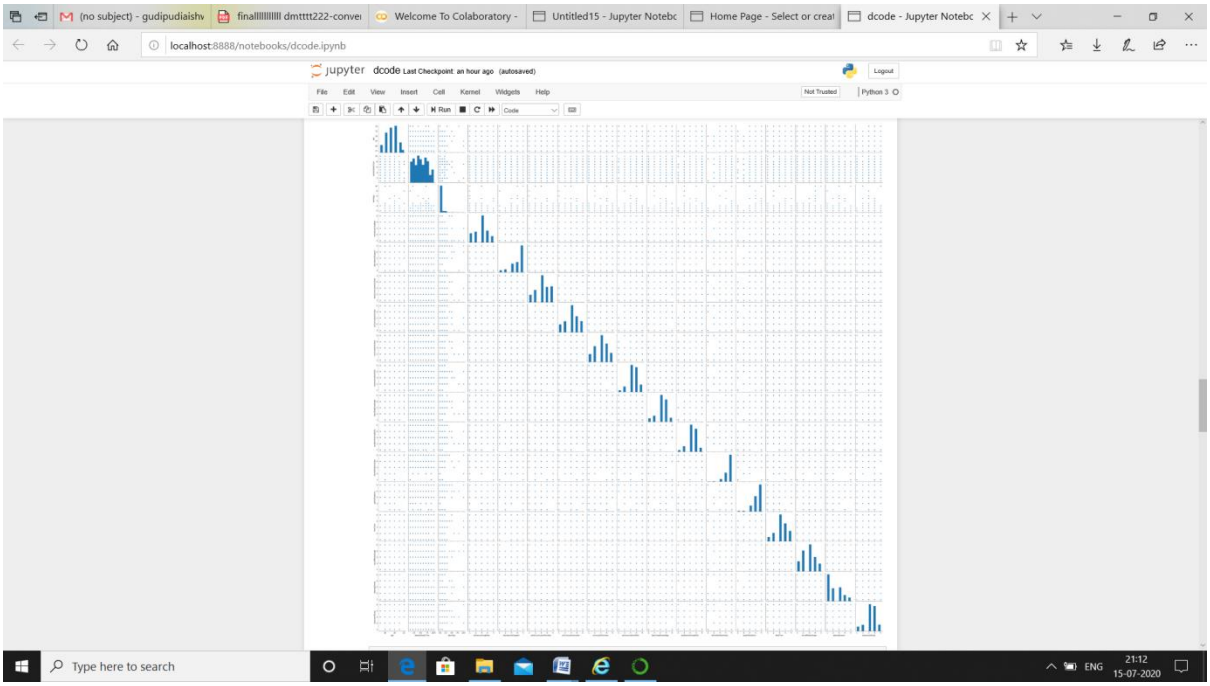
Out[3]:

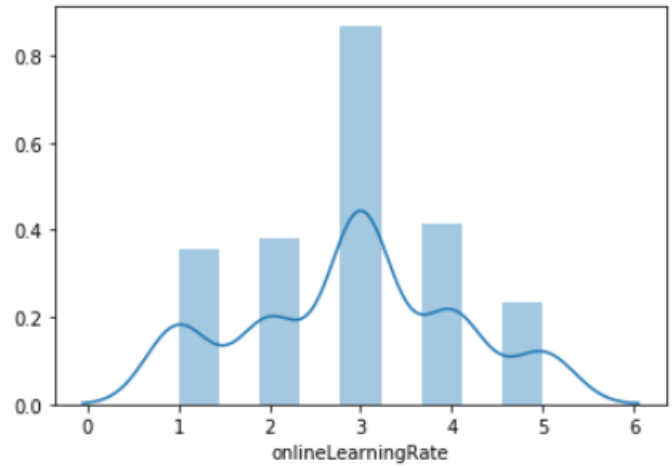


Out[7]: <seaborn.axisgrid.PairGrid at 0x191b7033630>

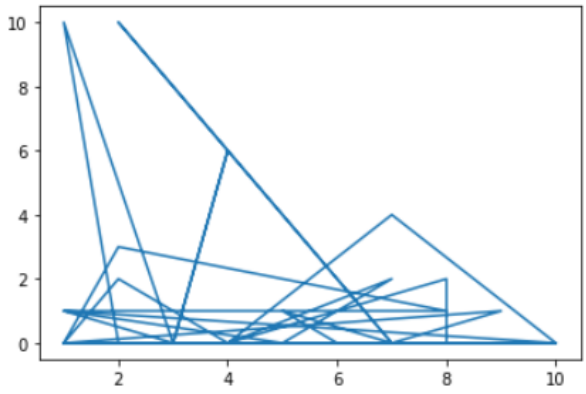


Out[10]:





Out[16]: [<matplotlib.lines.Line2D at 0x191d6e924e0>]



Out[17]:

	age	weeklyStudyTime	backlogs	onlineLearningRate	educationalSupport	extraCurricularActivities	coCurricularActivities	curricularActivities	cor
count	588.000000	588.000000	588.000000	588.000000	588.000000	588.000000	588.000000	588.000000	
mean	19.263605	4.984694	0.297619	2.906463	4.073129	3.263605	3.132653	3.049320	
std	1.234196	2.689136	1.188328	1.180373	1.189001	1.231432	1.185987	1.159782	
min	17.000000	1.000000	0.000000	1.000000	1.000000	1.000000	1.000000	1.000000	
25%	18.000000	3.000000	0.000000	2.000000	3.000000	2.000000	2.000000	2.000000	
50%	19.000000	5.000000	0.000000	3.000000	5.000000	3.000000	3.000000	3.000000	
75%	20.000000	7.000000	0.000000	4.000000	5.000000	4.000000	4.000000	4.000000	
max	22.000000	10.000000	10.000000	5.000000	5.000000	5.000000	5.000000	5.000000	

Out[24]:

	weeklyStudyTime	backlogs	onlineLearningRate	educationalSupport	extraCurricularActivities	coCurricularActivities	curricularActivities	communicationSkills
0	4	0	3	4	3	4	3	4
1	2	0	5	5	5	3	3	3
2	5	0	5	5	3	3	3	4
3	4	0	5	3	4	3	3	4
4	10	0	3	4	2	3	4	3

5 rows × 37 columns

Out[25]:

```
[2 1 2 0 1 1 0 1 2 0 2 1 1 2 0 1 2 1 1 2 2 2 2 2 2 2 0 0 1 2 1 2 0 1 1 1 1
 1 1 1 1 0 1 2 2 1 0 1 1 0 0 1 2 1 0 1 0 0 0 2 0 2 0 1 2 2 0 0 1 1 0 0 1 1
 2 2 1 2 2 1 0 2 2 0 2 2 2 0 1 0 0 0 2 0 1 1 0 2 0 2 2 0 0 1 0 2 2 1 1 1 1
 0 1 2 0 1 1 1 0 1 2 0 2 2 2 0 0 0 0 0 0 0 1 1 2 2 2 1 1 2 1 2 2 2 0 1 1 2 2
 0 0 2 1 0 0 0 0 0 1 1 2 2 1 1 1 1 1 1 1 2 1 0 2 2 2 2 1 1 1 1 1 1 2 0 1 1 2
 2 1 0 2 2 1 1 1 2 0 0 0 2 1 1 2 0 1 0 1 2 2 2 0 2 0 0 1 0 1 0 1 1 2 1 1 0
 2 1 1 2 1 2 2 1 2 0 1 1 2 1 1 2 0 0 0 0 0 2 2 2 0 2 2 2 2 1 2 2 2 1 1 2 2
 2 2 2 1 2 2 2 1 2 2 0 1 2 2 2 0 2 1 2 1 2 1 2 2 2 2 0 2 0 0 1 2 0 0 0 2 2
 2 2 1 1 2 1 2 1 1 2 0 2 2 0 2 0 0 1 0 2 2 2 2 2 0 1 2 0 0 1 0 1 0 0 1 2 2
 2 2 2 0 0 0 0 0 1 0 0 0 1 2 1 2 0 2 2 0 1 2 0 0 2 0 0 2 0 2 0 0 2 0 2 2 1
 1 2 1 2 0 2 0 1 0 1 2 1 1 2 2 0 1 2 1 1 2 1 2 1 1 1 2 2 2 1 2 1 0 2 1 0 2
 0 2 1 2 2 2 0 1 0 0 2 1 2 2 0 2 2 1 1 1 0 1 0 2 2 2 1 0 2 2 0 0 0 0 0 0 0
 2 0 0 2 2 0 0 0 2 0 0 0 0 0 0 0 1 0 0 0 1 2 0 0 0 0 0 0 2 1 1 0 1 0 0 0 0 2
 2 0 0 0 0 0 2 0 1 2 1 2 2 0 0 2 0 0 2 0 0 0 1 0 0 0 0 0 2 1 2 0 1 0 0 0 0 0
 2 0 0 2 0 0 0 0 0 0 2 0 0 0 2 0 0 0 0 2 0 0 2 2 0 2 0 0 0 0 0 2 2 0 0 0 0 2
 1 1 2 0 0 2 2 0 2 0 0 0 0 0 0 0 1 1 2 2 1 1 1 1 2 0 1 1 1 2 2 1 0]
```

Out[29]:

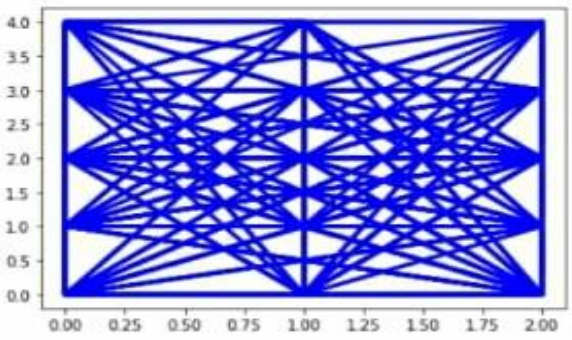
```
[2 4 4 4 2 1 0 1 0 2 2 1 2 3 0 3 4 2 0 0 2 0 4 4 3 2 2 3 3 2 3 1 3 4 0 2 3
 2 0 2 0 0 1 3 0 2 1 2 0 4 3 1 3 2 3 3 3 2 0 2 1 0 4 2 4 1 2 3 4 2 2 1 3 1
 2 4 1 3 2 1 2 0 2 2 0 0 0 2 3 2 0 2 2 2 1 0 2 1 2 2 3 2 2 1 2 0 2 4 1 4 4
 4 2 2 3 2 4 1 2 2 4 2 2 2 3 2 0 1 2 4 2 1 1 0 2 2 1 4 1 2 1 2 2 2 2 4 1 4
 3 4 2 2 2 3 2 2 1 3 2 3 0 2 2 1 2 2 2 1 3 3 2 2 4 2 0 1 1 0 1 1 2 1 3 3 2
 3 2 0 3 4 2 2 0 0 0 2 2 1 4 3 3 0 1 2 1 1 2 3 2 1 1 2 4 4 0 2 0 1 1 4 2 3
 3 2 3 0 2 2 2 4 1 1 4 3 0 1 4 2 3 3 3 0 2 0 3 2 2 0 2 1 1 3 0 3 3 0 1 3 0
 2 1 3 3 2 4 1 0 2 2 0 0 3 2 0 3 2 0 0 3 2 0 4 2 2 1 2 3 4 4 2 4 2 3 2 3 2
 0 4 3 4 2 3 2 2 1 0 2 2 2 2 2 2 0 0 2 2 1 2 1 0 1 2 1 3 2 1 2 1 2 2 1 4 0 2
 1 2 1 1 1 2 3 0 2 2 3 2 2 0 0 2 2 2 3 0 0 0 2 0 0 3 0 1 2 1 1 2 3 2 0 1 1
 2 0 2 2 3 3 3 3 3 1 2 4 2 1 2 2 3 0 4 4 0 0 3 0 3 3 2 2 4 0 0 0 3 3 0 0 2
 2 2 2 0 3 2 3 1 3 2 0 3 3 3 1 0 0 2 0 2 2 3 2 2 2 1 2 4 0 3 3 3 4 2 2 4 3
 2 1 1 2 1 2 2 2 1 3 2 2 4 2 2 4 4 2 2 0 3 2 0 2 0 2 2 4 3 4 2 2 1 1 2 2 2
 0 2 3 3 1 2 2 4 1 3 1 1 2 1 4 2 2 3 2 2 3 2 2 1 4 1 1 0 3 2 2 2 2 2 4 3 2 2
 2 1 4 2 2 1 2 2 2 2 4 3 0 2 2 1 0 3 4 2 2 2 2 2 1 3 1 2 3 3 2 1 3 3 3 2 2 1 1
 2 0 0 3 2 2 1 2 4 1 3 0 3 0 2 4 2 2 2 3 0 2 2 2 1 2 3 2 1 1 0 2 3]
```

Out[32]:

```
GaussianNB(priors=None, var_smoothing=1e-09)
[2]
[2]
[2]
[2]
[2]
====== Classification Report =======
              precision    recall  f1-score   support

           0       0.00      0.00      0.00       216
           1       0.00      0.00      0.00       164
           2       0.35      1.00      0.52       208

    accuracy                           0.35       588
   macro avg       0.12      0.33      0.17       588
weighted avg       0.13      0.35      0.18       588

====== Confusion matrix =======
[[  0   0 216]
 [  0   0 164]
 [  0   0 208]]
```

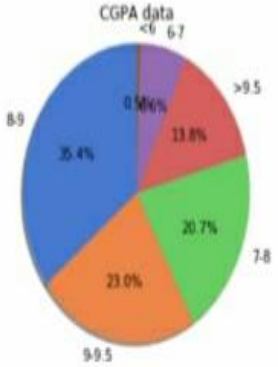
Out[33]:

```
Mean squared error: 3.15
accuracy:  0.23
```



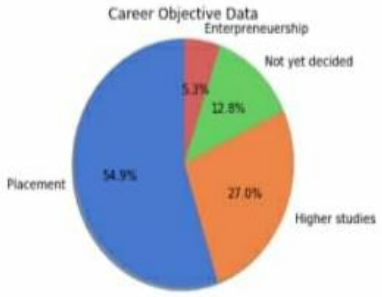
In [ ]:

Out[34]: (-1.1150864057241259,
1.1246993383603994,
-1.1296619869215583,
1.1014124755676933)



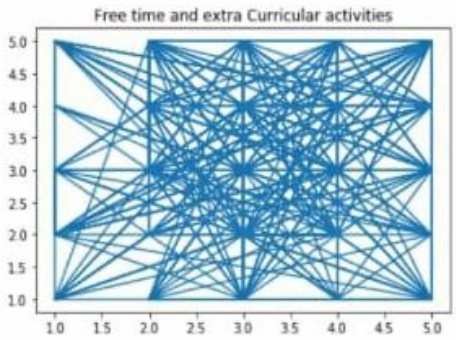
Out[35]: (-1.1255640118500934,
1.1235416488073766,
-1.1242669182936784,
1.1011555675377942)

```
Out[36]: (-1.1109050131289984,
          1.1024144262991773,
          -1.1106328283175384,
          1.1005063265125568)
```

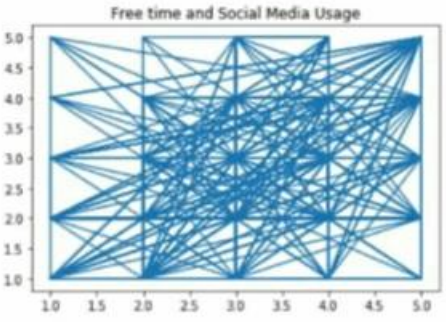



```
Out[37]: (-1.1273201154442372,
          1.108858869006286,
          -1.1295619876303467,
          1.1014077294936329)
```

Out[40]:




```
Out[41]: Text(0.5, 1.0, 'Free time and Social Media Usage')
```



Conclusion:

The performance of the student can be predicted based on the information about cgpa, attendance extra curricular activities etc. This study helps to identify the students performance and if they were poor in performance we can take acceptable action for career. This system will help the college to analyze the performance of students and take the necessary actions to improve students performance.

References:

https://ieeexplore.ieee.org/document/8697770

https://www.sciencedirect.com/science/article/pii/S1877050915036182

https://www.researchgate.net/publication/242341193_DATA_MINING_APPROACH_FOR_PREDICTING_STUDENT_PERFORMANCE