Academic Success Classification Model

Understanding both Train and Test Data

```
# Importing Needed Libraries
import pandas as pd
import numpy as np
train_data = pd.read_csv("train.csv")
train data.head()
       Marital status Application mode Application order
   id
                                                                Course \
0
    0
                                                                  9238
                                                            1
1
    1
                     1
                                       17
                                                            1
                                                                  9238
2
    2
                     1
                                                            2
                                       17
                                                                  9254
3
                                        1
                                                            3
    3
                     1
                                                                  9500
4
    4
                                        1
                                                            2
                                                                  9500
   Daytime/evening attendance Previous qualification \
0
1
                             1
                                                       1
2
                             1
                                                       1
3
                             1
                                                       1
4
   Previous qualification (grade)
                                     Nacionality Mother's qualification
    \
                             126.0
                                                                         1
0
. . .
                             125.0
                                                                        19
1
. . .
                             137.0
                                                                         3
2
                             131.0
                                                                        19
3
                             132.0
                                                                        19
4
   Curricular units 2nd sem (credited) Curricular units 2nd sem
(enrolled) \
0
                                       0
6
1
                                       0
6
2
6
3
                                       0
```

```
8
4
                                     0
7
   Curricular units 2nd sem (evaluations) \
0
                                        9
1
2
                                        0
3
                                       11
4
                                       12
   Curricular units 2nd sem (approved) Curricular units 2nd sem
(grade) \
12.428571
                                     0
0.000000
0.000000
                                     7
12.820000
                                     6
12.933333
   Curricular units 2nd sem (without evaluations) Unemployment
rate \
                                                0
                                                                11.1
                                                0
1
                                                                11.1
2
                                                                16.2
                                                                11.1
                                                                 7.6
   Inflation rate
                  GDP
                           Target
0
                  2.02 Graduate
              0.6
              0.6 2.02
1
                        Dropout
2
              0.3 -0.92
                          Dropout
3
                   2.02
                         Enrolled
              0.6
              2.6 0.32 Graduate
[5 rows x 38 columns]
train_data.info
<bound method DataFrame.info of</pre>
                                          id Marital status
Application mode Application order
                                     Course \
                           1
                                                                1
```

9238	_	_			
1 9238	1	1	17		1
2	2	1	17		2
9254	2	1			2
3 9500	3	1	1		3
4	4	1	1		2
9500					
76513	76513	1	17		1
9254					
76514	76514	1	1		6
9254 76515	76515	5	17		1
9085	70313	3	17		
76516	76516	1	1		3
9070	76517	1	1		1
76517 9773	76517	1	1		1
3773					
0 1 2 3 4 76513 76514 76515 76516	Day cime, ever	1 1 1 1 1 1 1	Previous qualifi	1 1 1 1 1 1 1	
76517		1		1	
	Previous quaication \	alification (grade		Mother's	
0 1 1		120	0 1		
		125	0 1		
19		127	0 1		
2 3 3		137	0 1		
3		131	0 1		
19					
4 19		132	0 1		
13					
		1-1			

```
121.0
76513
                                                      1
19
76514
                                   125.0
                                                      1
76515
                                   138.0
37
76516
                                   136.0
                                                      1
38
76517
                                   133.1
                                                      1
19
             Curricular units 2nd sem (credited)
0
1
                                                   0
2
                                                   0
3
                                                   0
4
                                                   0
. . .
76513
                                                   0
76514
                                                   0
76515
                                                   0
76516
                                                   0
76517
       Curricular units 2nd sem (enrolled)
0
                                             6
1
                                             6
2
                                             6
3
                                             8
4
                                             7
76513
                                             6
76514
                                             6
76515
                                             5
76516
76517
       Curricular units 2nd sem (evaluations) \
0
1
                                                 9
2
                                                 0
3
                                               11
4
                                                12
76513
                                                 8
76514
                                                 9
76515
                                                 8
76516
                                                 0
76517
```

(Curricular	units	2nd	sem	(approved	d)	Curricular	units	2nd	sem
(grade) 0						6				
12.42857 1	71					0				
0.000000)					0				
0.000000)									
3 12.82000	90					7				
4 12.93333	33					6				
	,5									
76513						5				
10.60006 76514	90					6				
13.87500 76515	00					5				
11.40000	90									
76516 0.000000)					0				
76517 13.66666	57					6				
	Curricular	units	2nd	sem	(without	eva	luations)	Unempl	ovme	ent
rate \	Curricular	units	2nd	sem	(without	eva	luations)	Unempl	oyme	ent
rate \ 0 11.1	Curricular	units	2nd	sem	(without	eva	Θ	Unempl	oyme	ent
rate \ 0 11.1 1 11.1	Curricular	units	2nd	sem	(without	eva		Unempl	oyme	ent
rate \ 0 11.1 1 11.1 2	Curricular	units	2nd	sem	(without	eva	Θ	Unempl	Loyme	ent
rate \ 0 11.1 1 11.1 2 16.2 3	Curricular	units	2nd	sem	(without	eva	0 0	Unempl	oyme	ent
rate \ 0 11.1 1 11.1 2 16.2 3 11.1 4	Curricular	units	2nd	sem	(without	eva	0 0 0	Unempl	oyme	ent
rate \ 0 11.1 1 11.1 2 16.2 3 11.1	Curricular	units	2nd	sem	(without	eva	0 0 0	Unempl	oyme	ent
rate \ 0 11.1 1 11.1 2 16.2 3 11.1 4 7.6	Curricular	units	2nd	sem	(without	eva	0 0 0 0	Unempl	Loyme	ent .
rate \ 0 11.1 1 11.1 2 16.2 3 11.1 4 7.6 76513 13.9	Curricular	units	2nd	sem	(without	eva	0 0 0 0	Unempl	Loyme	ent .
rate \ 0 11.1 1 11.1 2 16.2 3 11.1 4 7.6 76513 13.9 76514 9.4	Curricular	units	2nd	sem	(without	eva	0 0 0 0 	Unempl	Loyme	ent .
rate \ 0 11.1 1 11.1 2 16.2 3 11.1 4 7.6 76513 13.9 76514 9.4 76515 9.4	Curricular	units	2nd	sem	(without	eva	0 0 0 0 9 0	Unempl	oyme	ent .
rate \ 0 11.1 1 11.1 2 16.2 3 11.1 4 7.6 76513 13.9 76514 9.4 76515 9.4 76516	Curricular	units	2nd	sem	(without	eva	0 0 0 0 	Unempl	oyme	ent .
rate \ 0 11.1 1 11.1 2 16.2 3 11.1 4 7.6 76513 13.9 76514 9.4 76515 9.4	Curricular	units	2nd	sem	(without	eva	0 0 0 0 9 0	Unempl	oyme	ent .

```
Inflation rate GDP
                                Target
0
                  0.6 2.02
                              Graduate
1
                  0.6 2.02
                               Dropout
2
                  0.3 - 0.92
                               Dropout
3
                  0.6 2.02
                              Enrolled
4
                  2.6 0.32
                              Graduate
                   . . .
                 -0.3
                       0.79
76513
                              Graduate
                 -0.8 -3.12
76514
                              Graduate
76515
                 -0.8 -3.12
                              Enrolled
                  2.6 0.32
76516
                             Dropout
76517
                  2.8 -4.06 Graduate
[76518 rows x 38 columns]>
train data.isnull().sum()
id
                                                    0
Marital status
                                                    0
Application mode
                                                    0
                                                    0
Application order
                                                    0
Course
Daytime/evening attendance
                                                    0
                                                    0
Previous qualification
Previous qualification (grade)
                                                    0
                                                    0
Nacionality
Mother's qualification
                                                    0
Father's qualification
                                                    0
Mother's occupation
                                                    0
Father's occupation
                                                    0
                                                    0
Admission grade
                                                    0
Displaced
                                                    0
Educational special needs
                                                    0
Debtor
Tuition fees up to date
                                                    0
Gender
                                                    0
                                                    0
Scholarship holder
                                                    0
Age at enrollment
International
                                                    0
                                                    0
Curricular units 1st sem (credited)
Curricular units 1st sem (enrolled)
                                                    0
Curricular units 1st sem (evaluations)
                                                    0
Curricular units 1st sem (approved)
                                                    0
Curricular units 1st sem (grade)
                                                    0
                                                    0
Curricular units 1st sem (without evaluations)
Curricular units 2nd sem (credited)
                                                    0
Curricular units 2nd sem (enrolled)
                                                    0
Curricular units 2nd sem (evaluations)
                                                    0
Curricular units 2nd sem (approved)
```

```
Curricular units 2nd sem (grade)
                                                    0
Curricular units 2nd sem (without evaluations)
                                                    0
                                                    0
Unemployment rate
                                                    0
Inflation rate
                                                    0
GDP
                                                    0
Target
dtype: int64
train data["Target"].value counts()
Target
Graduate
            36282
Dropout
            25296
Enrolled
            14940
Name: count, dtype: int64
test_data = pd.read_csv("test.csv")
test data.head()
      id Marital status Application mode Application order Course
  76518
                                           1
                                                               1
                                                                    9500
1 76519
                                           1
                                                               1
                                                                    9238
2 76520
                                           1
                                                               1
                                                                    9238
                                          44
3 76521
                                                               1
                                                                    9147
4 76522
                                          39
                                                               1
                                                                    9670
   Daytime/evening attendance Previous qualification \
0
                             1
                                                       1
1
                             1
                                                       1
2
                             1
                                                       1
3
                             1
                                                      39
4
                             1
                                                       1
   Previous qualification (grade) Nacionality Mother's qualification
    \
                                                                         3
0
                             141.0
. . .
1
                             128.0
                                                                         1
. . .
                             118.0
                                                                         1
2
. . .
                             130.0
                                                                         1
3
. . .
                             110.0
                                                                         1
4
. . .
```

0 1 2 3 4	Curricular	units	1st	sem	(without	eva	aluations	0 0 0	\		
(er 0 8 1 6 2 6 3 8 4 6	Curricular nrolled) \	units	2nd	sem	(credited	0 0 0 3	Curricul	ar.	units	2nd s	em
0 1 2 3 4	Curricular	units	2nd	sem	(evaluat:	1	6) \ 0 6 .1 .4				
0 0.0 1 13 2 11 3 11 4	Curricular rade) \ 000000 .500000 .000000 .000000	units	2nd	sem	(approved	d) 0 6 5 4	Curricul	ar.	units	2nd s	em
rat	Curricular te \	units	2nd	sem	(without	eva	aluations	0	Unempl	oymen	13.9
1								0			11.1
2								0			15.5
3								0			8.9

4	2	7.6
	_	710
Inflation rate GDP		
0 -0.3 0.79 1 0.6 2.02		
2 2.8 -4.06		
3 1.4 3.51		
4 2.6 0.32		
[5 rows x 37 columns]		
-		
test_data.info()		
<class 'pandas.core.frame.dataframe'=""></class>		
RangeIndex: 51012 entries, 0 to 51011		
Data columns (total 37 columns):		New No.11 Count
# Column		Non-Null Count
Dtype		
0 id		51012 non-null
int64		
1 Marital status		51012 non-null
int64		
2 Application mode		51012 non-null
int64		51010
<pre>3 Application order int64</pre>		51012 non-null
4 Course		51012 non-null
int64		JIOIZ HOH-HUCC
5 Daytime/evening attendance		51012 non-null
int64		
6 Previous qualification		51012 non-null
int64		
7 Previous qualification (grade)		51012 non-null
float64		51010 11
8 Nacionality		51012 non-null
int64 9 Mother's qualification		51012 non-null
9 Mother's qualification int64		JIUIZ HOH-HULL
10 Father's qualification		51012 non-null
int64		JIJII HUCC
11 Mother's occupation		51012 non-null
int64		
12 Father's occupation		51012 non-null
int64		
13 Admission grade		51012 non-null
float64		

14 Displaced	51012 non-null
int64	
15 Educational special needs	51012 non-null
int64 16 Debtor	51012 non-null
int64	JIVIZ HOH-HUCC
17 Tuition fees up to date	51012 non-null
int64	
18 Gender	51012 non-null
int64	
19 Scholarship holder	51012 non-null
int64	F101211
20 Age at enrollment	51012 non-null
int64 21 International	51012 non-null
int64	JIOIZ HOH-Hacc
22 Curricular units 1st sem (credited)	51012 non-null
int64	
23 Curricular units 1st sem (enrolled)	51012 non-null
int64	
24 Curricular units 1st sem (evaluations)	51012 non-null
int64	F1012 11
25 Curricular units 1st sem (approved)	51012 non-null
int64 26 Curricular units 1st sem (grade)	51012 non-null
float64	JIVIZ HOH-HUCC
27 Curricular units 1st sem (without evaluations)	51012 non-null
int64	SIGIL HON HALL
28 Curricular units 2nd sem (credited)	51012 non-null
int64	
29 Curricular units 2nd sem (enrolled)	51012 non-null
int64	51010
30 Curricular units 2nd sem (evaluations)	51012 non-null
int64	51012 non-null
31 Curricular units 2nd sem (approved) int64	51012 Holl-Hutt
32 Curricular units 2nd sem (grade)	51012 non-null
float64	31012 11011 11411
33 Curricular units 2nd sem (without evaluations)	51012 non-null
int64	
34 Unemployment rate	51012 non-null
float64	
35 Inflation rate	51012 non-null
float64	E1012 non null
36 GDP float64	51012 non-null
dtypes: float64(7), int64(30)	
memory usage: 14.4 MB	
•	
<pre>test_data.isnull().sum()</pre>	

id	0
Marital status	0
Application mode	0
Application order	0
Course	0
Daytime/evening attendance	Θ
Previous qualification	0
Previous qualification (grade)	0
Nacionality	0
Mother's qualification	0
Father's qualification	0
Mother's occupation	0
Father's occupation	0
Admission grade	0
Displaced	0
Educational special needs	0
Debtor	0
Tuition fees up to date	0
Gender	0
Scholarship holder	0
Age at enrollment	0
International	0
Curricular units 1st sem (credited)	0
Curricular units 1st sem (enrolled)	0
Curricular units 1st sem (evaluations)	0
Curricular units 1st sem (approved)	0
Curricular units 1st sem (grade)	0
Curricular units 1st sem (without evaluations)	0
Curricular units 2nd sem (credited)	0
Curricular units 2nd sem (enrolled)	0
Curricular units 2nd sem (evaluations)	0
Curricular units 2nd sem (approved)	0
Curricular units 2nd sem (grade)	0
Curricular units 2nd sem (without evaluations)	0
Unemployment rate	0
Inflation rate	0
GDP	0
dtype: int64	