

# 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()
```

	id	Marital status	Application mode	Application order	Course \
0	0	1	1	1	9238
1	1	1	17	1	9238
2	2	1	17	2	9254
3	3	1	1	3	9500
4	4	1	1	2	9500

	Daytime/evening attendance	Previous qualification \
0	1	1
1	1	1
2	1	1
3	1	1
4	1	1

	Previous qualification (grade)	Nacionality	Mother's qualification
...	\		
0	126.0	1	1
...			
1	125.0	1	19
...			
2	137.0	1	3
...			
3	131.0	1	19
...			
4	132.0	1	19
...			

	Curricular units 2nd sem (credited)	Curricular units 2nd sem (enrolled) \
0	0	
6		
1	0	
6		
2	0	
6		
3	0	

```

8
4
7
0

Curricular units 2nd sem (evaluations) \
0 7
1 9
2 0
3 11
4 12

Curricular units 2nd sem (approved) Curricular units 2nd sem
(grade) \
0 6
12.428571
1 0
0.000000
2 0
0.000000
3 7
12.820000
4 6
12.933333

Curricular units 2nd sem (without evaluations) Unemployment
rate \
0 0 11.1
1 0 11.1
2 0 16.2
3 0 11.1
4 0 7.6

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

[5 rows x 38 columns]

train_data.info
<bound method DataFrame.info of id Marital status
Application mode Application order Course \
0 0 1 1 1

```

9238				
1	1	1	17	1
9238				
2	2	1	17	2
9254				
3	3	1	1	3
9500				
4	4	1	1	2
9500				
...	...	...	...	...
...				
76513	76513	1	17	1
9254				
76514	76514	1	1	6
9254				
76515	76515	5	17	1
9085				
76516	76516	1	1	3
9070				
76517	76517	1	1	1
9773				

	Daytime/evening attendance	Previous qualification \
0	1	1
1	1	1
2	1	1
3	1	1
4	1	1
...	...	...
76513	1	1
76514	1	1
76515	1	1
76516	1	1
76517	1	1

	Previous qualification (grade)	Nacionality	Mother's qualification \
0	126.0	1	
1			
1	125.0	1	
19			
2	137.0	1	
3			
3	131.0	1	
19			
4	132.0	1	
19			
...	...	...	
...			

76513	121.0	1
19		
76514	125.0	1
1		
76515	138.0	1
37		
76516	136.0	1
38		
76517	133.1	1
19		

	...	Curricular units 2nd sem (credited)	\
0	...	0	
1	...	0	
2	...	0	
3	...	0	
4	...	0	
...	...	...	
76513	...	0	
76514	...	0	
76515	...	0	
76516	...	0	
76517	...	0	

	Curricular units 2nd sem (enrolled)	\
0	6	
1	6	
2	6	
3	8	
4	7	
...	...	
76513	6	
76514	6	
76515	5	
76516	6	
76517	6	

	Curricular units 2nd sem (evaluations)	\
0	7	
1	9	
2	0	
3	11	
4	12	
...	...	
76513	8	
76514	9	
76515	8	
76516	0	
76517	6	

	Curricular units 2nd sem (approved)	Curricular units 2nd sem
(grade) \		
0	6	
12.428571		
1	0	
0.000000		
2	0	
0.000000		
3	7	
12.820000		
4	6	
12.933333		
...	...	
...		
76513	5	
10.600000		
76514	6	
13.875000		
76515	5	
11.400000		
76516	0	
0.000000		
76517	6	
13.666667		

	Curricular units 2nd sem (without evaluations)	Unemployment
rate \		
0	0	
11.1		
1	0	
11.1		
2	0	
16.2		
3	0	
11.1		
4	0	
7.6		
...	...	.
...		
76513	0	
13.9		
76514	0	
9.4		
76515	1	
9.4		
76516	0	
7.6		
76517	0	
15.5		

	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
...	...	...	...
76513	-0.3	0.79	Graduate
76514	-0.8	-3.12	Graduate
76515	-0.8	-3.12	Enrolled
76516	2.6	0.32	Dropout
76517	2.8	-4.06	Graduate

[76518 rows x 38 columns]>

```
train_data.isnull().sum()
```

```
id 0
Marital status 0
Application mode 0
Application order 0
Course 0
Daytime/evening attendance 0
Previous qualification 0
Previous qualification (grade) 0
Nationality 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
Target                                0
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
0	76518	1	1	1	9500
1	76519	1	1	1	9238
2	76520	1	1	1	9238
3	76521	1	44	1	9147
4	76522	1	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
0	141.0	1	3
1	128.0	1	1
2	118.0	1	1
3	130.0	1	1
4	110.0	1	1

Curricular units 1st sem (without evaluations) \	
0	0
1	0
2	0
3	0
4	0

Curricular units 2nd sem (credited) (enrolled) \	Curricular units 2nd sem
0	0
8	
1	0
6	
2	0
6	
3	3
8	
4	0
6	

Curricular units 2nd sem (evaluations) \	
0	0
1	6
2	11
3	14
4	9

Curricular units 2nd sem (approved) (grade) \	Curricular units 2nd sem
0	0
0.000000	
1	6
13.500000	
2	5
11.000000	
3	5
11.000000	
4	4
10.666667	

Curricular units 2nd sem (without evaluations) rate \	Unemployment	
0	0	13.9
1	0	11.1
2	0	15.5
3	0	8.9



4 2 7.6

	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'>

RangeIndex: 51012 entries, 0 to 51011

Data columns (total 37 columns):

#	Column	Non-Null Count
0	id	51012 non-null
1	Marital status	51012 non-null
2	Application mode	51012 non-null
3	Application order	51012 non-null
4	Course	51012 non-null
5	Daytime/evening attendance	51012 non-null
6	Previous qualification	51012 non-null
7	Previous qualification (grade)	51012 non-null
8	Nacionality	51012 non-null
9	Mother's qualification	51012 non-null
10	Father's qualification	51012 non-null
11	Mother's occupation	51012 non-null
12	Father's occupation	51012 non-null
13	Admission grade	51012 non-null

14	Displaced	51012 non-null
int64		
15	Educational special needs	51012 non-null
int64		
16	Debtor	51012 non-null
int64		
17	Tuition fees up to date	51012 non-null
int64		
18	Gender	51012 non-null
int64		
19	Scholarship holder	51012 non-null
int64		
20	Age at enrollment	51012 non-null
int64		
21	International	51012 non-null
int64		
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		
25	Curricular units 1st sem (approved)	51012 non-null
int64		
26	Curricular units 1st sem (grade)	51012 non-null
float64		
27	Curricular units 1st sem (without evaluations)	51012 non-null
int64		
28	Curricular units 2nd sem (credited)	51012 non-null
int64		
29	Curricular units 2nd sem (enrolled)	51012 non-null
int64		
30	Curricular units 2nd sem (evaluations)	51012 non-null
int64		
31	Curricular units 2nd sem (approved)	51012 non-null
int64		
32	Curricular units 2nd sem (grade)	51012 non-null
float64		
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		
36	GDP	51012 non-null
float64		

dtypes: float64(7), int64(30)  
memory usage: 14.4 MB

```
test_data.isnull().sum()
```

id	0
Marital status	0
Application mode	0
Application order	0
Course	0
Daytime/evening attendance	0
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	

---