Data Pre-Processing

Handling Categorical Values

Date	4 september 2022
Team Id	PNT2022TMID46202
Project Name	Smart Lender- Applicant Credibility Prediction for Loan Approval

As we can see our dataset has categorical data we must convert the categorical data to integer encoding or binary encoding.

To convert the categorical features into numerical features we use encoding techniques. There are several techniques but in our project, we are using manual encoding with the help of list comprehension.

- In our project, Gender , married, dependents, self-employed, co-applicants income, loan amount , loan amount term, credit history With list comprehension encoding is done.

```
In [22]: from sklearn.preprocessing import LabelEncoder, OneHotEncoder
         le = LabelEncoder()
         oneh = OneHotEncoder()
         data['Education'] = le.fit_transform(data['Education'])
         data.head()
```

Out[22]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History
0	LP001002	1	No	0	0	No	5849	0.0	NaN	360.0	1.0
1	LP001003	1	Yes	1	0	No	4583	1508.0	128.0	360.0	1.0
2	LP001005	1	Yes	0	0	Yes	3000	0.0	66.0	360.0	1.0
3	LP001006	1	Yes	0	1	No	2583	2358.0	120.0	360.0	1.0
4	LP001008	1	No	0	0	No	6000	0.0	141.0	360.0	1.0