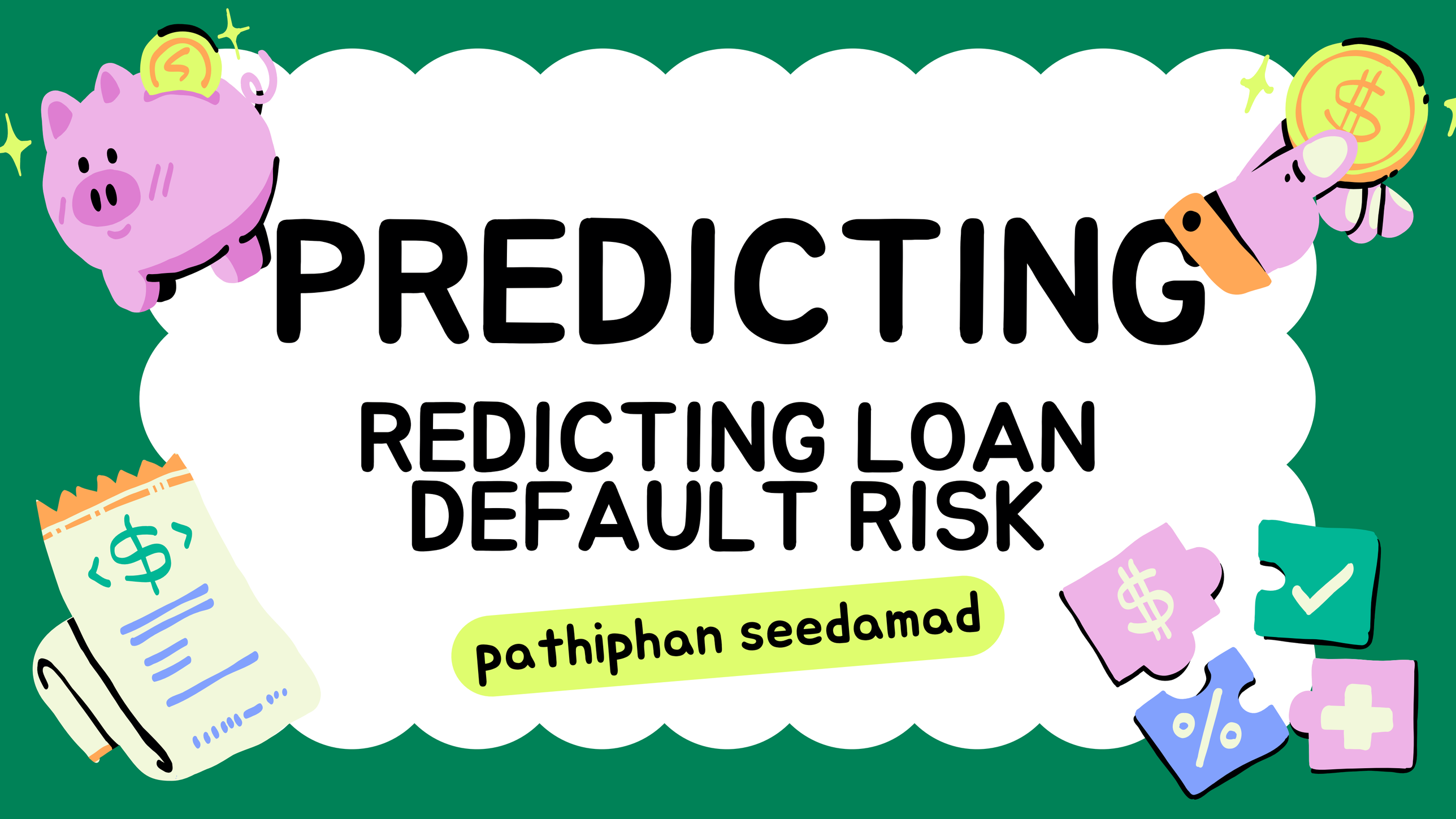
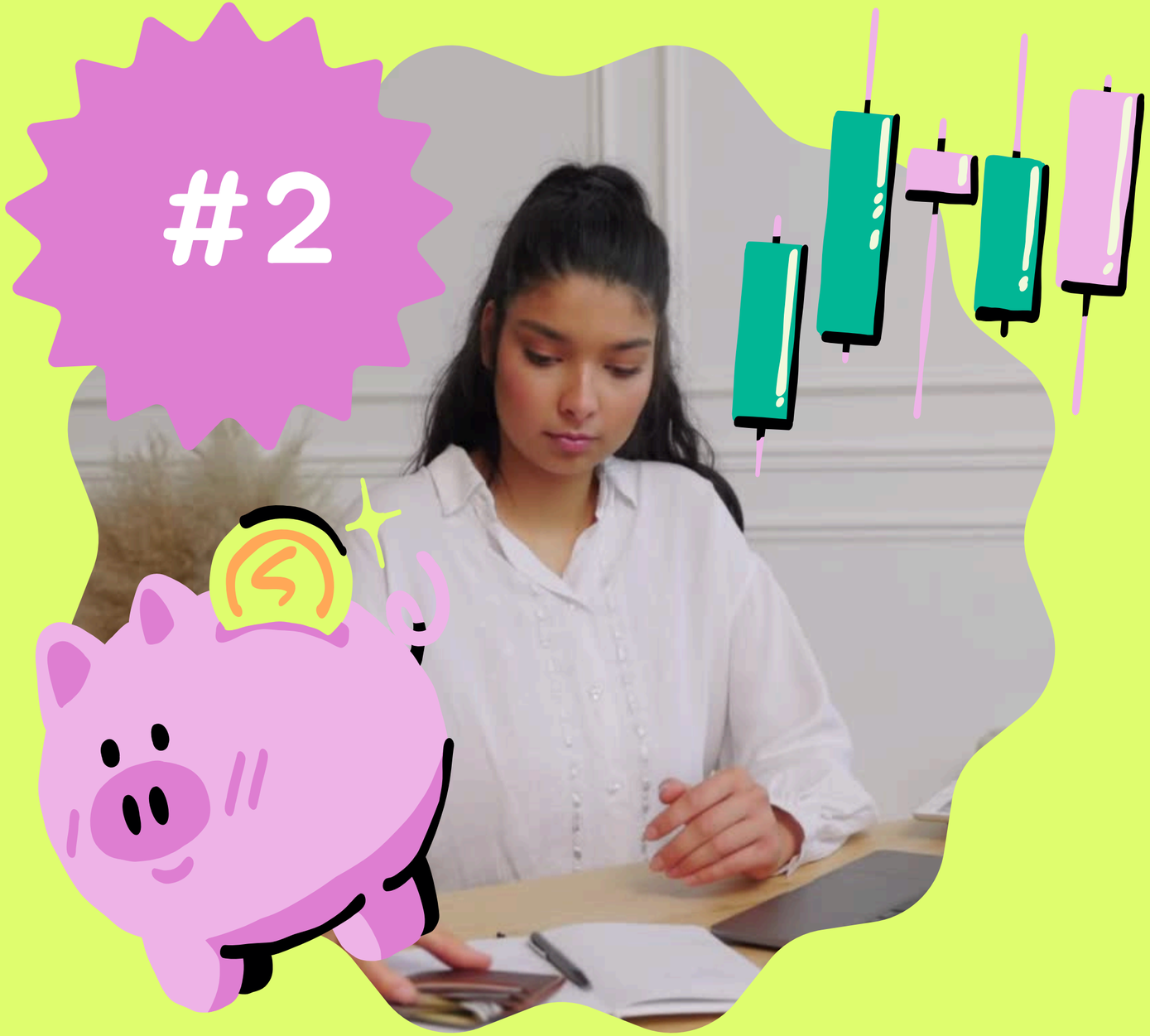

PREDICTING
REDICTING LOAN DEFAULT RISK
pathiphan seedamad

#1

OBJECTIVE

The inability of financial institutions to accurately assess the risk of loan applicants. This leads to frequent defaults and financial losses, so the project aims to use data from loan applicants to help them more effectively predict this risk.

#2

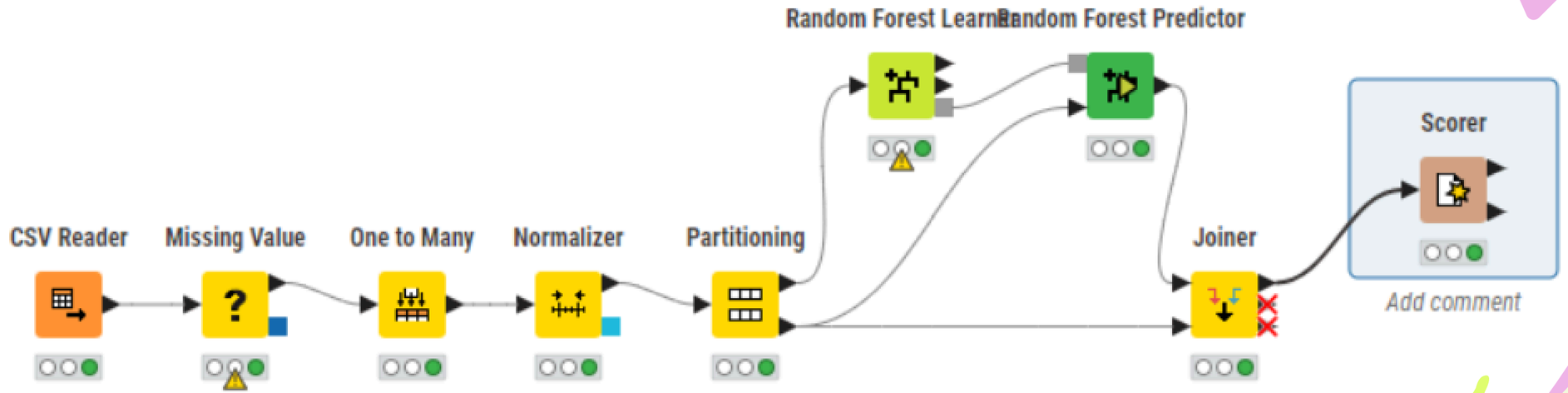


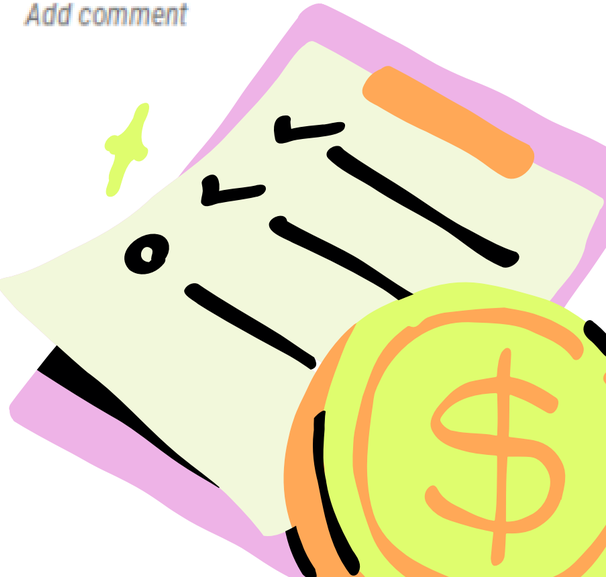
DATA USED

The data we used is the 'Loan Prediction' dataset from Kaggle, which contains important information such as income, loan application status, and more. This dataset has 614 rows and 13 columns

WORKFLOW & MODEL

#3





RESULT

#4

Rows: 3 | Columns: 11

Table | Statistics

#	RowID	TruePositives Number (integer)	FalsePositives Number (integer)	TrueNegatives Number (integer)	FalseNegatives Number (integer)	Recall Number (dou...	Precision Number (dou...	Sensitivity Number (dou...	Specificity Number (dou...	F-measure Number (dou...	Accuracy Number (dou...	Cohen's k... Number (dou...
1	Y	97	3	19	25	0.795	0.97	0.795	0.864	0.874	?	?
2	N	19	25	97	3	0.864	0.432	0.864	0.795	0.576	?	?
3	Overall	?	?	?	?	?	?	?	?	?	0.806	0.467

▶ 1: Confusion matrix ▶ 2: Accuracy statistics Flow Variables

Rows: 2 | Columns: 2

Table | Statistics

#	RowID	Y Number (integer)	N Number (integer)
1	Y	97	25
2	N	3	19

SUMMARY

FROM THE RESULTS, WE SEE THAT THE RANDOM FOREST MODEL CAN PREDICT LOAN DEFAULT RISK WITH AN ACCURACY OF 80.6%. THIS PREDICTION IS BENEFICIAL FOR FINANCIAL INSTITUTIONS IN REDUCING FINANCIAL RISK. FUTURE RECOMMENDATIONS INCLUDE ADDING NEW DATA AND EXPERIMENTING WITH OTHER MODELS FOR FURTHER IMPROVEMENT

THANK YOU

FOR WATCHING!