

MANDALIKA

Dokumen Laporan Final Project Stage 3

Anggota:

- Ardilla Safitri
- Jonse Kennedy
- Fakhry Husnul
- Lise Listianti
- Riska Diah N
- Arni Cici Suryani



MODEL EVALUATION

Model yang digunakan:

- a. Logistic Regression
- b. kNN
- c. Decision Tree
- d. Random Forest
- e. AdaBoost
- f. XgBoost

LOGISTIC REGRESSION

Metrics	Train	Test
Accuracy	0.86	0.84
Precision	0.86	0.48
Recall	0.85	0.77
F1	0.85	0.59
AUC-ROC	0.93	0.89

Confusion Matrix		Predictive	
		True	False
Actual	True	1092	196
	False	53	178

KNN

Metrics	Train	Test
Accuracy	0.94	0.85
Precision	0.90	0.50
Recall	1.00	0.82
F1	0.95	0.62
AUC-ROC	1.00	0.90

Confusion Matrix		Predictive	
		True	False
Actual	True	1095	193
	False	41	190

DECISION TREE

Metrics	Train	Test
Accuracy	1.00	0.91
Precision	1.00	0.68
Recall	1.00	0.77
F1	1.00	0.73
AUC-ROC	1.00	0.86

Confusion Matrix		Predictive	
		True	False
Actual	True	1205	83
	False	52	179

RANDOM FOREST

Metrics	Train	Test
Accuracy	1.00	0.95
Precision	1.00	0.85
Recall	1.00	0.80
F1	1.00	0.82
AUC-ROC	1.00	0.97

Confusion Matrix		Predictive	
		True	False
Actual	True	1255	33
	False	47	184

ADABOOST

Metrics	Train	Test
Accuracy	0.89	0.85
Precision	0.88	0.50
Recall	0.89	0.69
F1	0.89	0.58
AUC-ROC	0.96	0.89

Confusion Matrix		Predictive	
		True	False
Actual	True	1127	161
	False	71	160

XGBOOST

Metrics	Train	Test
Accuracy	1	0.95
Precision	1	0.87
Recall	1	0.81
F1	1	0.84
AUC-ROC	1	0.98

Confusion Matrix		Predictive	
		True	False
Actual	True	1261	27
	False	43	188

CONCLUSION

1. Metric yang digunakan ialah Recall karena ingin menurunkan nilai false negatif.
2. Model yang digunakan ialah XGBoost karena memiliki nilai recall tertinggi dibandingkan modeling yang lain dan memiliki false negatif terendah dibandingkan model yang lain..