

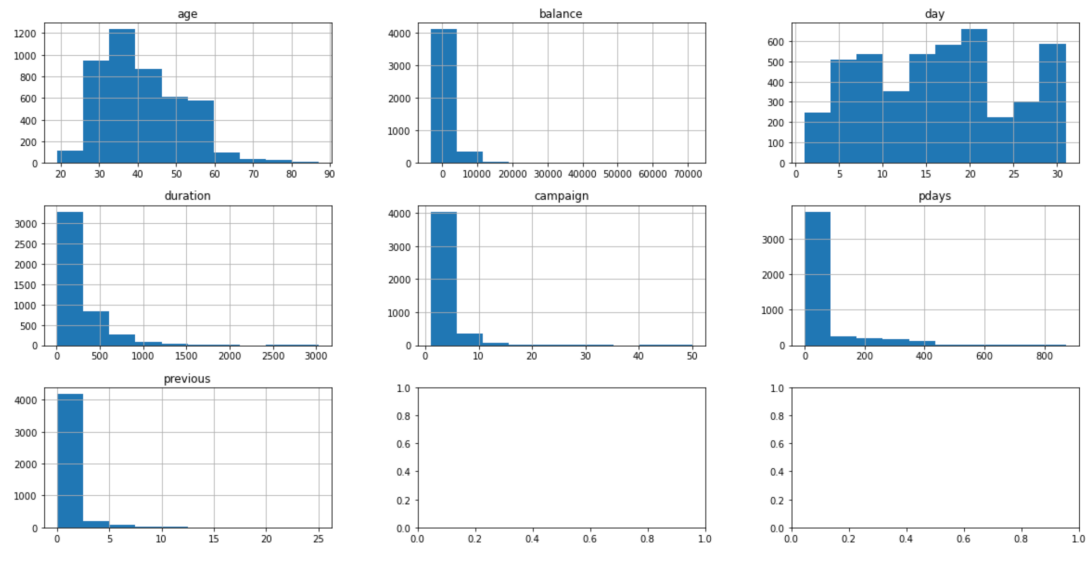
Bank Data Analytics and Modeling

Jim Woods, IXIS

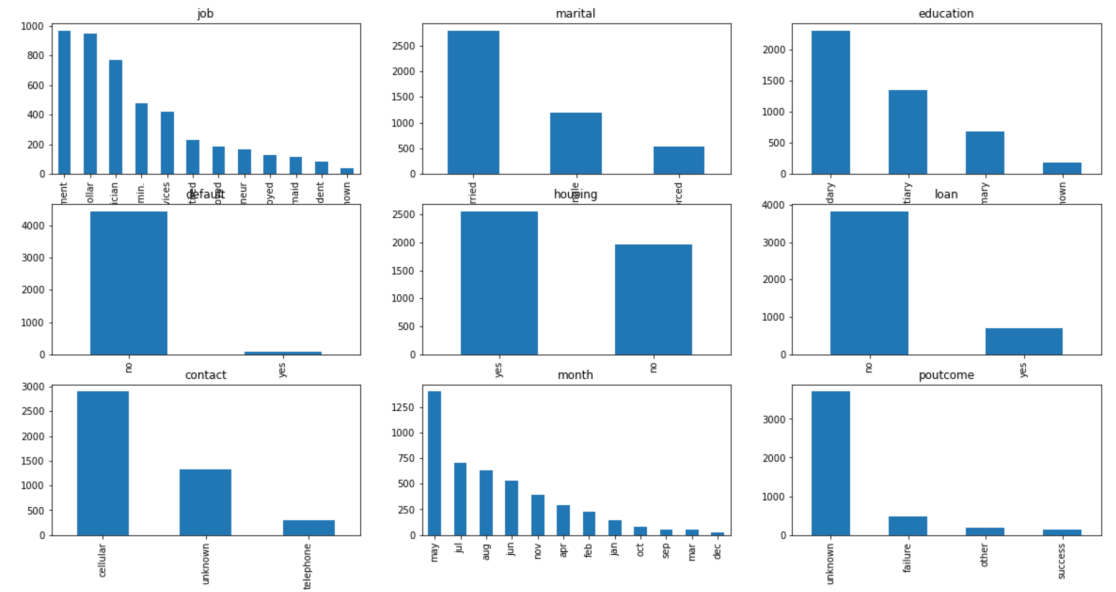
12/19/2022

Data Distributions

Distributions of Numeric Variables



Distribution of Categorical Variables



Check Feature Relevance One by One (Filtering Feature Selection)

Numeric Variables

	Feature	p-Value
3	duration	2.150445e-174
6	previous	3.477971e-15
5	pdays	2.287071e-12
4	campaign	3.885794e-05
0	age	2.424536e-03
1	balance	2.287158e-01
2	day	4.497348e-01

Categorical Variables

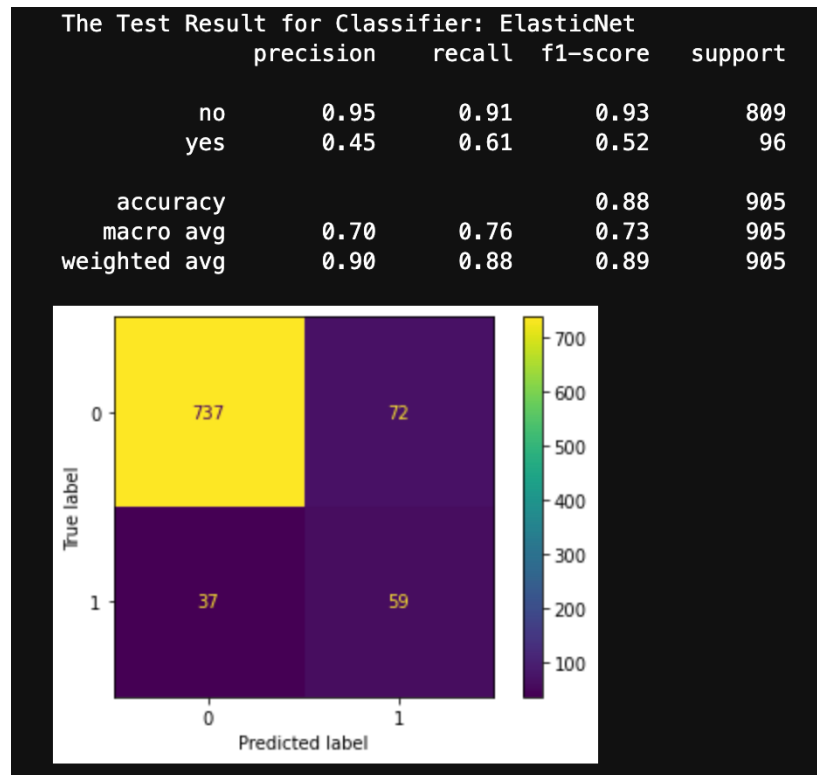
	Feature	p-Value
8	poutcome	1.539883e-83
7	month	2.195355e-47
6	contact	8.304301e-20
4	housing	2.714700e-12
0	job	1.901391e-10
5	loan	2.914829e-06
1	marital	7.373535e-05
2	education	1.625240e-03
3	default	1.000000e+00

Significant
Variables
(Keep)

Insignificant
Variables
(Filter Out)

Evaluation Results of 5 Models and the Chosen One - ElasticNet

ElasticNet – The Selected Model Performance



Comparison of 5 Models

	Classifier	Accuracy	Precision	Recall	F1Score
4	ElasticNet	0.879558	0.701289	0.762792	0.725484
1	LASSO	0.910497	0.778689	0.679111	0.714238
2	Support Vector Machine	0.791160	0.647285	0.814335	0.665999
0	Random Forest	0.841989	0.643008	0.709647	0.664781
3	k-Nearest Neighbor	0.790055	0.627085	0.754043	0.644573

	feature	coefficient
9	poutcome_success	143.094604
32	job_self-employed	-107.359792
28	job_entrepreneur	-114.641035
14	month_feb	-122.460917
8	poutcome_other	-124.025780
15	month_jan	-130.594623
44	education_unknown	-132.186477
11	month_apr	-134.578626
26	job_admin.	-158.395140
12	month_aug	-178.210062
17	month_jun	-218.428946
3	campaign	-221.058049
38	marital_divorced	-244.799838
33	job_services	-246.119392
20	month_nov	-267.663198
30	job_management	-288.004527
10	poutcome_unknown	-289.200984
43	education_tertiary	-289.543189
7	poutcome_failure	-331.813938
35	job_technician	-364.232770
16	month_jul	-394.570477
41	education_primary	-402.103740
40	marital_single	-405.150549
6	loan	-421.975122
27	job_blue-collar	-610.672467
25	contact_unknown	-720.213035
39	marital_married	-738.162187
19	month_may	-792.159410
42	education_secondary	-823.531055
5	housing	-1024.497520

Model Interpretation: Top Weighted Features in the Best Classifier

