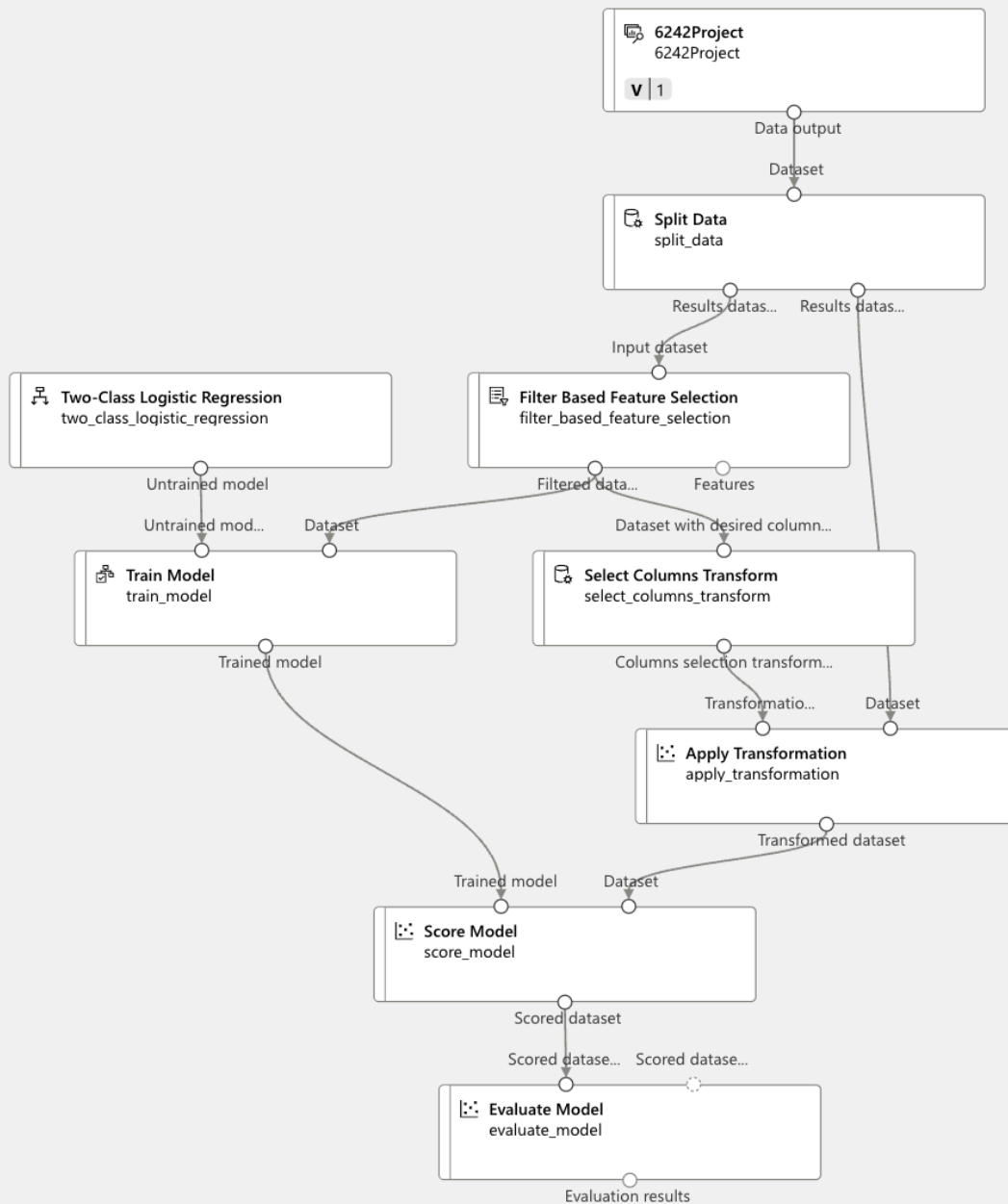


Columns of Interest - 8 feature selection using chisquare



Filtered_dataset

×

Rows ?

Columns ?

556,222

9

ne	loan_term	loan_amount	derived_ethnicity	derived_race	state_code	derived_sex	loan_type	applicant_age
								
	0.349015	-0.068587	Not Hispanic or Latino	White	TX	Male	3	25-34
	0.349015	0.424572	Not Hispanic or Latino	Asian	WA	Joint	1	45-54
	0.349015	-0.008105	Not Hispanic or Latino	White	WA	Female	1	25-34
	-2.601487	-0.101154	Hispanic or Latino	White	TX	Joint	1	45-54
	0.349015	0.005852	Hispanic or Latino	White	NY	Female	2	35-44
	-0.634485	0.038419	Ethnicity Not Available	Race Not Available	WA	Sex Not Available	1	55-64
	0.349015	0.736285	Not Hispanic or Latino	White	NY	Male	1	45-54
	0.349015	-0.03602	Not Hispanic or Latino	Black or African American	TX	Male	3	25-34
	0.349015	-0.022063	Not Hispanic or Latino	Asian	TX	Male	1	25-34
	0.349015	0.015157	Not Hispanic or Latino	White	NY	Joint	1	25-34
	0.349015	-0.143026	Not Hispanic or Latino	Black or African American	TX	Joint	2	55-64
	0.349015	0.101154	Hispanic or	White	TX	Male	1	25-44

Features

×

Filtered_dataset









×

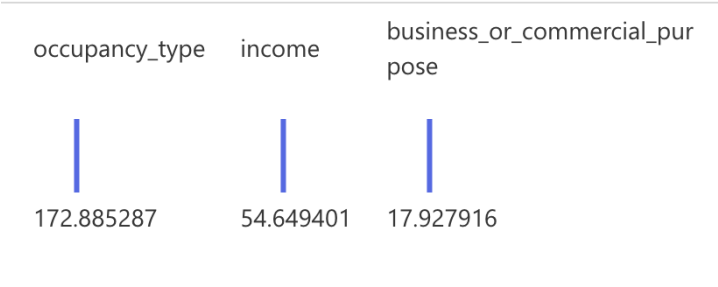
Rows ?










Columns ?

1

12

loan_term	loan_amount	derived_ethnicity	derived_race	state_code	derived_sex	loan_type	applicant_ag
							
92477.59312	27093.030968	7372.932699	4138.372793	3761.172333	3074.777832	1709.0506	1330.116676



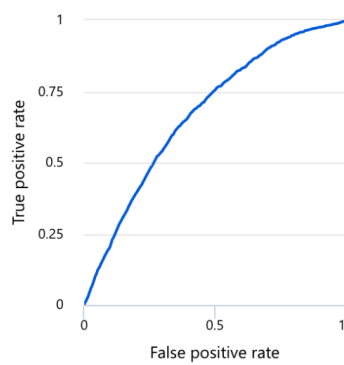
Features	×	Scored_dataset	×	Filtered_dataset	×	Untrained_model	×	Trained_model
Rows ? Columns ?								
238,381 11								
state_code	derived_ethnicity	derived_race	derived_sex	loan_type	loan_amount	loan_term	applicant_age	ou
								
TX	Ethnicity Not Available	Race Not Available	Male	1	-0.138374	0.349015	25-34	1
TX	Ethnicity Not Available	Race Not Available	Female	1	-0.059282	0.349015	45-54	1
NY	Not Hispanic or Latino	White	Joint	1	0.038419	0.349015	25-34	1
NY	Not Hispanic or Latino	White	Joint	1	-0.105807	0.349015	25-34	1
WA	Ethnicity Not Available	Race Not Available	Sex Not Available	1	0.126815	0.349015	25-34	1
WA	Not Hispanic or Latino	White	Joint	1	0.0012	0.349015	<25	1
TX	Not Hispanic or Latino	White	Male	1	-0.087197	-1.617986	25-34	0
TX	Not Hispanic or Latino	White	Joint	1	0.094248	0.349015	25-34	1
WA	Not Hispanic or Latino	White	Female	1	0.215212	0.349015	25-34	1

outcome	Scored Labels	Scored Probabilities
1	1	0.818019
1	1	0.777043
1	1	0.930677
1	1	0.930673
1	1	0.890291
1	1	0.907332
0	1	0.796457
1	1	0.912387
1	1	0.91096

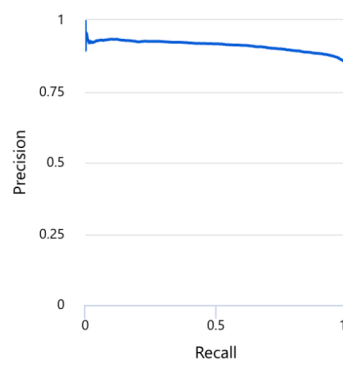
Features ☒ Evaluation_results ☒ Filtered_dataset ☒ Untrained_model ☒ Trained_model ☒ Scored_dataset

● Scored dataset (left port)

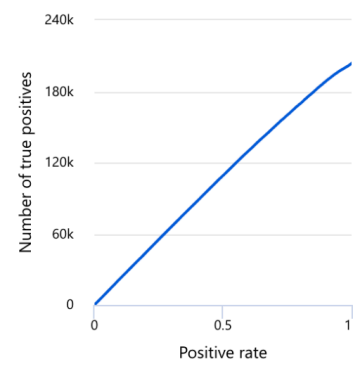
ROC curve



Precision-recall curve



Lift curve



Threshold 0.5

Accuracy 0.851
Precision 0.855
Recall 0.995
F1 Score 0.92
AUC 0.674

		Actual	
		1	0
Predicted	1	202 598	34 376
	0	1 053	354

Score bin ↓	Positive exam...	Negative exam...	Fraction above thres...	Accura...	F1 Score	Precisi...	Recall	Negative precis...	Negative recall	Cumulative A...
(0.900,1.000]	74159	6248	0.337	0.431	0.522	0.922	0.364	0.180	0.820	0.034
(0.800,0.900]	96333	15257	0.805	0.771	0.862	0.888	0.837	0.285	0.381	0.316
(0.700,0.800]	26344	8344	0.951	0.846	0.915	0.868	0.967	0.417	0.141	0.535
(0.600,0.700]	4223	3604	0.984	0.849	0.918	0.857	0.987	0.330	0.037	0.637
(0.500,0.600]	1539	923	0.994	0.851	0.920	0.855	0.995	0.252	0.010	0.663
(0.400,0.500]	604	260	0.998	0.853	0.921	0.854	0.998	0.173	0.003	0.671
(0.300,0.400]	364	83	1.000	0.854	0.921	0.854	1.000	0.115	0.000	0.673
(0.200,0.300]	82	11	1.000	0.854	0.921	0.854	1.000	0.000	0.000	0.673
(0.100,0.200]	3	0	1.000	0.854	0.921	0.854	1.000	1.000	0.000	0.673
(0.000,0.100]	0	0	1.000	0.854	0.921	0.854	1.000	1.000	0.000	0.673

Columns of Interest - 6 feature selection using chisquare

Filtered_dataset						
<div> <div>Rows ?</div> <div>Columns ?</div> </div>						
556,2227						
outcome	loan_term	loan_amount	derived_ethnicity	derived_race	state_code	derived_sex
1	0.349015	-0.068587	Not Hispanic or Latino	White	TX	Male
1	0.349015	0.424572	Not Hispanic or Latino	Asian	WA	Joint
1	0.349015	-0.008105	Not Hispanic or Latino	White	WA	Female
1	-2.601487	-0.101154	Hispanic or Latino	White	TX	Joint
1	0.349015	0.005852	Hispanic or Latino	White	NY	Female
1	-0.634485	0.038419	Ethnicity Not Available	Race Not Available	WA	Sex Not Available
1	0.349015	0.736285	Not Hispanic or Latino	White	NY	Male
0	0.349015	-0.03602	Not Hispanic or Latino	Black or African American	TX	Male
1	0.349015	-0.022063	Not Hispanic or Latino	Asian	TX	Male
1	0.349015	0.015157	Not Hispanic or Latino	White	NY	Joint
1	0.349015	-0.143026	Not Hispanic or Latino	Black or African American	TX	Joint
1	0.349015	-0.101154	Hispanic or Latino	White	TX	Male
1	0.349015	0.052376	Not Hispanic or Latino	Black or African American	TX	Female
...	Black or African

Features							
<div> <div>Rows ?</div> <div>Columns ?</div> </div>							
112							
outcome	loan_term	loan_amount	derived_ethnicity	derived_race	state_code	derived_sex	loan_type
1	92477.59312	27093.030968	7372.932699	4138.372793	3761.172333	3074.777832	1709.0506

Features



Scored_dataset



Filtered_dataset











Rows ?

238,381

Columns ?

9

state_code	derived_ethnicity	derived_race	derived_sex	loan_amount	loan_term	outcome	Scored Labels	Scored Probabilitie
								
TX	Ethnicity Not Available	Race Not Available	Male	-0.138374	0.349015	1	1	0.820995
TX	Ethnicity Not Available	Race Not Available	Female	-0.059282	0.349015	1	1	0.798099
NY	Not Hispanic or Latino	White	Joint	0.038419	0.349015	1	1	0.929208
NY	Not Hispanic or Latino	White	Joint	-0.105807	0.349015	1	1	0.929202
WA	Ethnicity Not Available	Race Not Available	Sex Not Available	0.126815	0.349015	1	1	0.890022
WA	Not Hispanic or Latino	White	Joint	0.0012	0.349015	1	1	0.939102
TX	Not Hispanic or Latino	White	Male	-0.087197	-1.617986	0	1	0.78744
TX	Not Hispanic or Latino	White	Joint	0.094248	0.349015	1	1	0.913313
WA	Not Hispanic or Latino	White	Female	0.215212	0.349015	1	1	0.908463
NY	Hispanic or Latino	White	Male	0.029114	0.349015	1	1	0.835991
WA	Not Hispanic or Latino	Asian	Joint	0.177992	0.349015	1	1	0.937567
TX	Not Hispanic or Latino	White	Joint	0.210559	0.349015	1	1	0.913319



Score bin ↓	Positive exam...	Negative exam...	Fraction above thres...	Accura...	F1 Score	Precisi...	Recall	Negative precis...	Negative recall	Cumulative A...
(0.900,1.000]	70916	5916	0.322	0.418	0.506	0.923	0.348	0.178	0.830	0.032
(0.800,0.900]	104392	16853	0.831	0.786	0.873	0.885	0.861	0.297	0.344	0.348
(0.700,0.800]	22108	8229	0.958	0.844	0.914	0.864	0.969	0.374	0.107	0.567
(0.600,0.700]	3545	2757	0.985	0.847	0.917	0.856	0.987	0.266	0.028	0.645
(0.500,0.600]	1567	669	0.994	0.851	0.919	0.855	0.994	0.214	0.009	0.664
(0.400,0.500]	645	210	0.998	0.853	0.920	0.854	0.998	0.167	0.003	0.670
(0.300,0.400]	426	86	1.000	0.854	0.921	0.854	1.000	0.161	0.000	0.672
(0.200,0.300]	51	10	1.000	0.854	0.921	0.854	1.000	0.000	0.000	0.673
(0.100,0.200]	1	0	1.000	0.854	0.921	0.854	1.000	1.000	0.000	0.673
(0.000,0.100]	0	0	1.000	0.854	0.921	0.854	1.000	1.000	0.000	0.673






Columns of Interest - 4 feature selection using chisquare

Rows ?








556,222

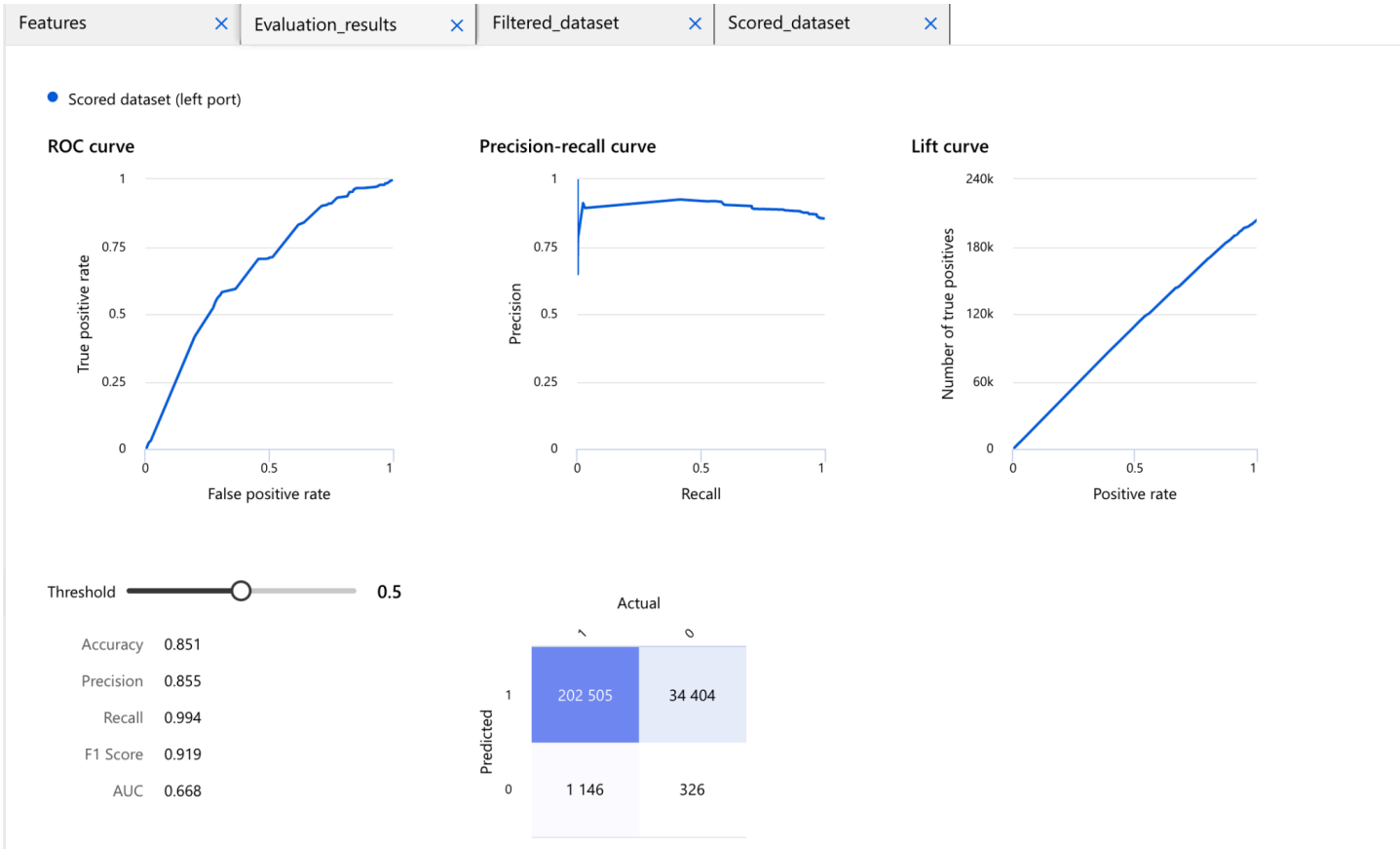
Columns ?

5

outcome	loan_term	loan_amount	derived_ethnicity	derived_race
				
1	0.349015	-0.068587	Not Hispanic or Latino	White
1	0.349015	0.424572	Not Hispanic or Latino	Asian
1	0.349015	-0.008105	Not Hispanic or Latino	White
1	-2.601487	-0.101154	Hispanic or Latino	White
1	0.349015	0.005852	Hispanic or Latino	White
1	-0.634485	0.038419	Ethnicity Not Available	Race Not Available
1	0.349015	0.736285	Not Hispanic or Latino	White
0	0.349015	-0.03602	Not Hispanic or Latino	Black or African American
1	0.349015	-0.022063	Not Hispanic or Latino	Asian
1	0.349015	0.015157	Not Hispanic or Latino	White
1	0.349015	-0.143026	Not Hispanic or Latino	Black or African American
1	0.349015	-0.101154	Hispanic or Latino	White
1	0.349015	0.052376	Not Hispanic or Latino	Black or African American
1	0.349015	-0.124416	Hispanic or Latino	Black or African American
1	0.349015	-0.03602	Hispanic or Latino	Race Not Available
1	0.349015	-0.091849	Not Hispanic or Latino	White
1	-2.601487	-0.101154	Not Hispanic or Latino	Asian
0	0.349015	-0.138374	Hispanic or Latino	Black or African American
0	0.349015	-0.133721	Not Hispanic or Latino	Asian
1	0.349015	-0.091849	Not Hispanic or Latino	White
1	0.349015	0.545535	Not Hispanic or Latino	Asian

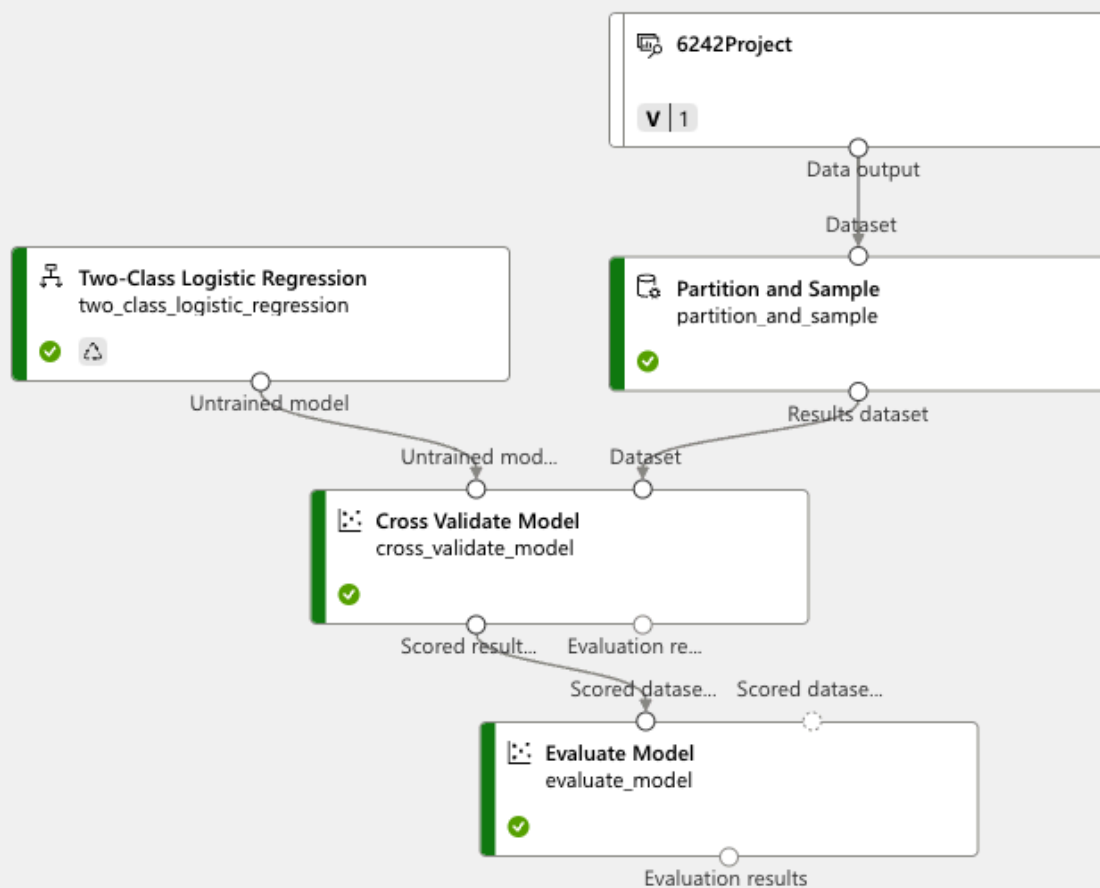
Features		Filtered_dataset					
Rows ?	Columns ?						
1	12						
outcome	loan_term	loan_amount	derived_ethnicity	derived_race	state_code	derived_sex	loan_type
1	92477.59312	27093.030968	7372.932699	4138.372793	3761.172333	3074.777832	1709.0506

Features		Scored_dataset		Filtered_dataset		
Rows ?	Columns ?					
238,381	7					
derived_ethnicity	derived_race	loan_amount	loan_term	outcome	Scored Labels	Scored Probabilities
						
Ethnicity Not Available	Race Not Available	-0.138374	0.349015	1	1	0.852411
Ethnicity Not Available	Race Not Available	-0.059282	0.349015	1	1	0.852424
Not Hispanic or Latino	White	0.038419	0.349015	1	1	0.906169
Not Hispanic or Latino	White	-0.105807	0.349015	1	1	0.906153
Ethnicity Not Available	Race Not Available	0.126815	0.349015	1	1	0.852455
Not Hispanic or Latino	White	0.0012	0.349015	1	1	0.906165
Not Hispanic or Latino	White	-0.087197	-1.617986	0	1	0.81614
Not Hispanic or Latino	White	0.094248	0.349015	1	1	0.906175
Not Hispanic or Latino	White	0.215212	0.349015	1	1	0.906188
Hispanic or Latino	White	0.029114	0.349015	1	1	0.818508
Not Hispanic or Latino	Asian	0.177992	0.349015	1	1	0.900636
Not Hispanic or	White	0.210559	0.349015	1	1	0.906188



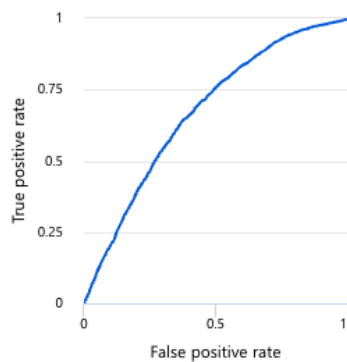
Score bin ↓	Positive exam...	Negative exam...	Fraction above thres...	Accura...	F1 Score	Precisi...	Recall	Negative precis...	Negative recall	Cumulative A...
(0.900,1.000]	106221	9452	0.485	0.552	0.665	0.918	0.522	0.206	0.728	0.075
(0.800,0.900]	77365	15337	0.874	0.812	0.891	0.881	0.901	0.331	0.286	0.391
(0.700,0.800]	13745	6428	0.959	0.843	0.913	0.863	0.969	0.357	0.101	0.565
(0.600,0.700]	3250	2557	0.983	0.845	0.916	0.856	0.985	0.237	0.028	0.636
(0.500,0.600]	1924	630	0.994	0.851	0.919	0.855	0.994	0.221	0.009	0.654
(0.400,0.500]	560	247	0.997	0.852	0.920	0.854	0.997	0.119	0.002	0.661
(0.300,0.400]	525	71	1.000	0.854	0.921	0.854	1.000	0.116	0.000	0.664
(0.200,0.300]	61	8	1.000	0.854	0.921	0.854	1.000	1.000	0.000	0.664
(0.100,0.200]	0	0	1.000	0.854	0.921	0.854	1.000	1.000	0.000	0.664
(0.000,0.100]	0	0	1.000	0.854	0.921	0.854	1.000	1.000	0.000	0.664

Columns of Interest - using 10-fold cross-validation

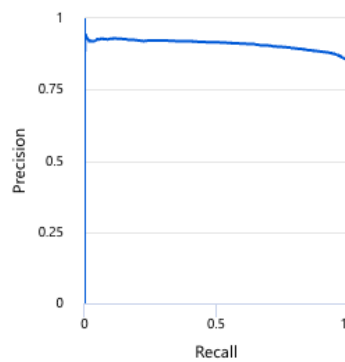


● Scored dataset (left port)

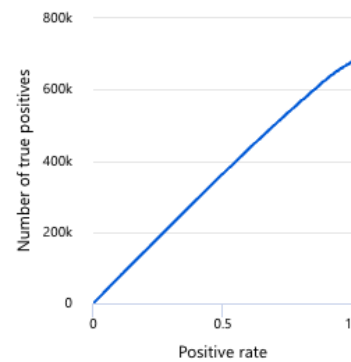
ROC curve



Precision-recall curve



Lift curve



Threshold 0.5

Accuracy 0.852

Precision 0.855

Recall 0.995

F1 Score 0.92

AUC 0.672

	Actual	
	1	0
Predicted	675 727	114 173
	3 533	1 170

Score bin ↓	Positive exam...	Negative exam...	Fraction above thres...	Accura...	F1 Score	Precisi...	Recall	Negative precis...	Negative recall	Cumulative AUC
(0.900,1.000]	242331	21197	0.332	0.423	0.514	0.920	0.357	0.177	0.816	0.034
(0.800,0.900]	330252	50229	0.810	0.776	0.865	0.889	0.843	0.292	0.381	0.314
(0.700,0.800]	81750	26367	0.947	0.846	0.914	0.870	0.963	0.413	0.152	0.523
(0.600,0.700]	16511	12793	0.983	0.850	0.919	0.858	0.988	0.361	0.041	0.631
(0.500,0.600]	4883	3587	0.994	0.852	0.920	0.855	0.995	0.249	0.010	0.662
(0.400,0.500]	1985	866	0.998	0.853	0.921	0.855	0.998	0.164	0.003	0.669
(0.300,0.400]	1217	256	1.000	0.854	0.922	0.855	1.000	0.127	0.000	0.671
(0.200,0.300]	320	45	1.000	0.855	0.922	0.855	1.000	0.214	0.000	0.672
(0.100,0.200]	11	3	1.000	0.855	0.922	0.855	1.000	1.000	0.000	0.672
(0.000,0.100]	0	0	1.000	0.855	0.922	0.855	1.000	1.000	0.000	0.672

Rows 12

Columns 11

Fold Number	Number of examples in fold	AUC	Accuracy	Precision	Recall	F1	True Negative	False Positive	False Negative	True Positive
0	79461	0.674379	0.851814	0.855549	0.994613	0.919855	113	11409	366	67573
1	79460	0.674097	0.851045	0.85498	0.994284	0.919385	131	11448	388	67493
2	79460	0.668707	0.853033	0.856373	0.995162	0.920565	114	11349	329	67668
3	79460	0.673822	0.850076	0.853601	0.99482	0.918816	133	11562	351	67414
4	79461	0.670383	0.851726	0.855205	0.994964	0.919806	111	11440	342	67568
5	79460	0.670461	0.850944	0.85445	0.99487	0.919329	129	11496	348	67487
6	79460	0.6775	0.854103	0.85764	0.994965	0.921213	92	11250	343	67775
7	79461	0.666178	0.854016	0.857419	0.995109	0.921147	106	11267	333	67755
8	79460	0.67336	0.850289	0.854055	0.994559	0.918968	109	11527	369	67455
9	79460	0.671027	0.851636	0.855314	0.994639	0.91973	132	11425	364	67539
Mean	794603	0.671991	0.851868	0.855459	0.994799	0.919881	117	11417.3	353.3	67572.7
Standard Deviation	794603	0.003268	0.001425	0.001342	0.000273	0.000843	13.711309	103.277458	18.245547	123.438739