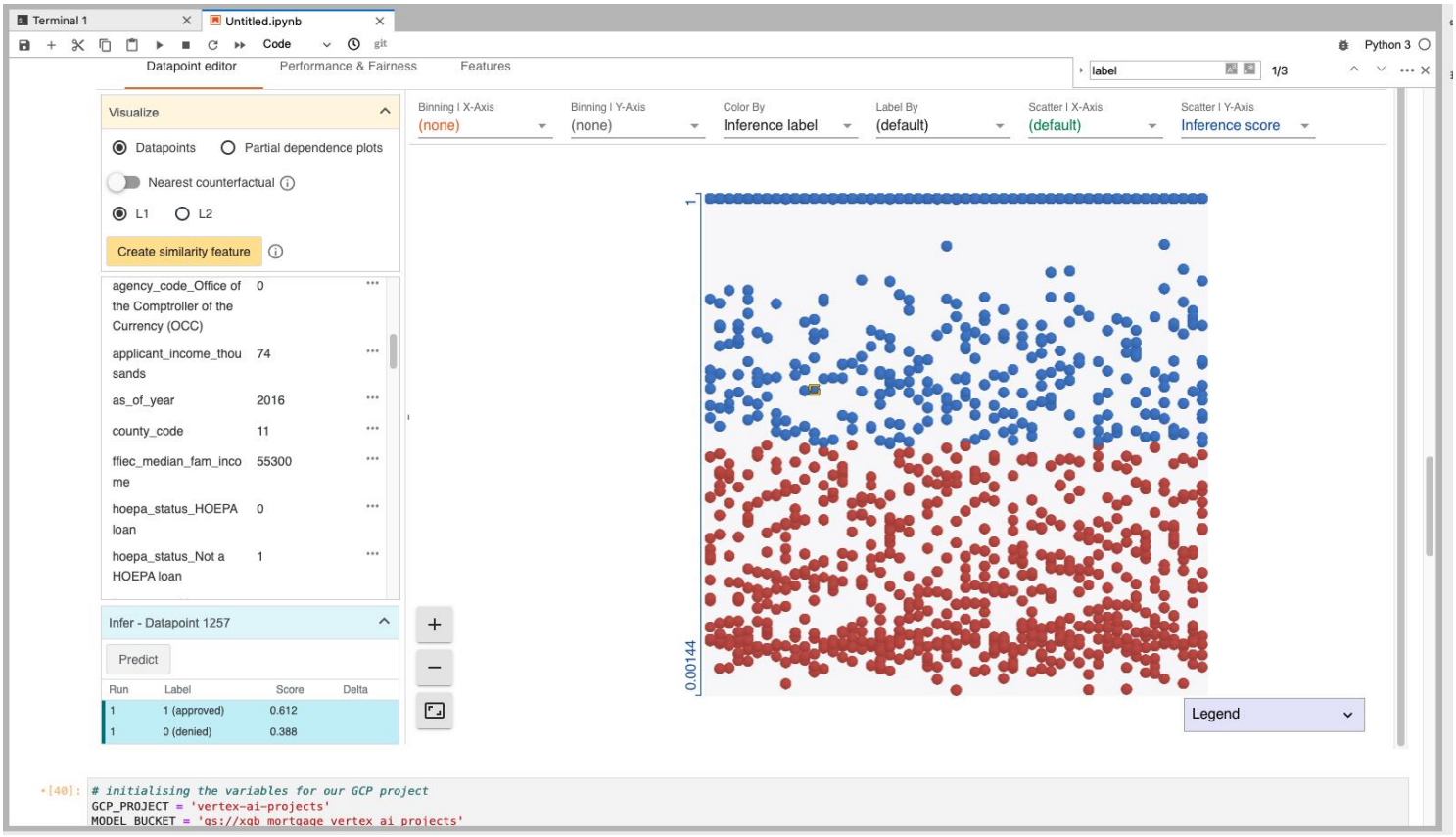
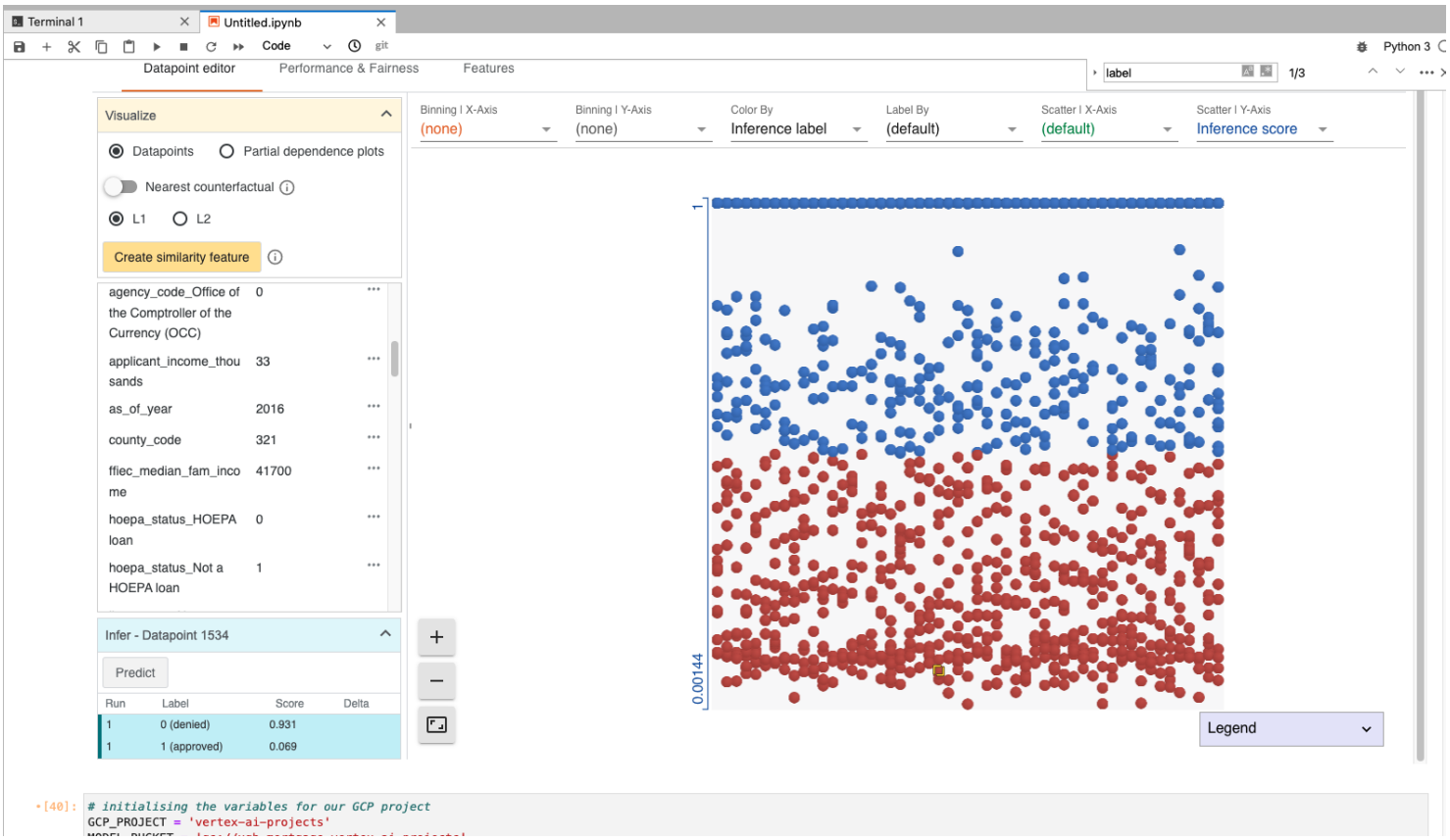
Explaining the model using What If Tool in GCP

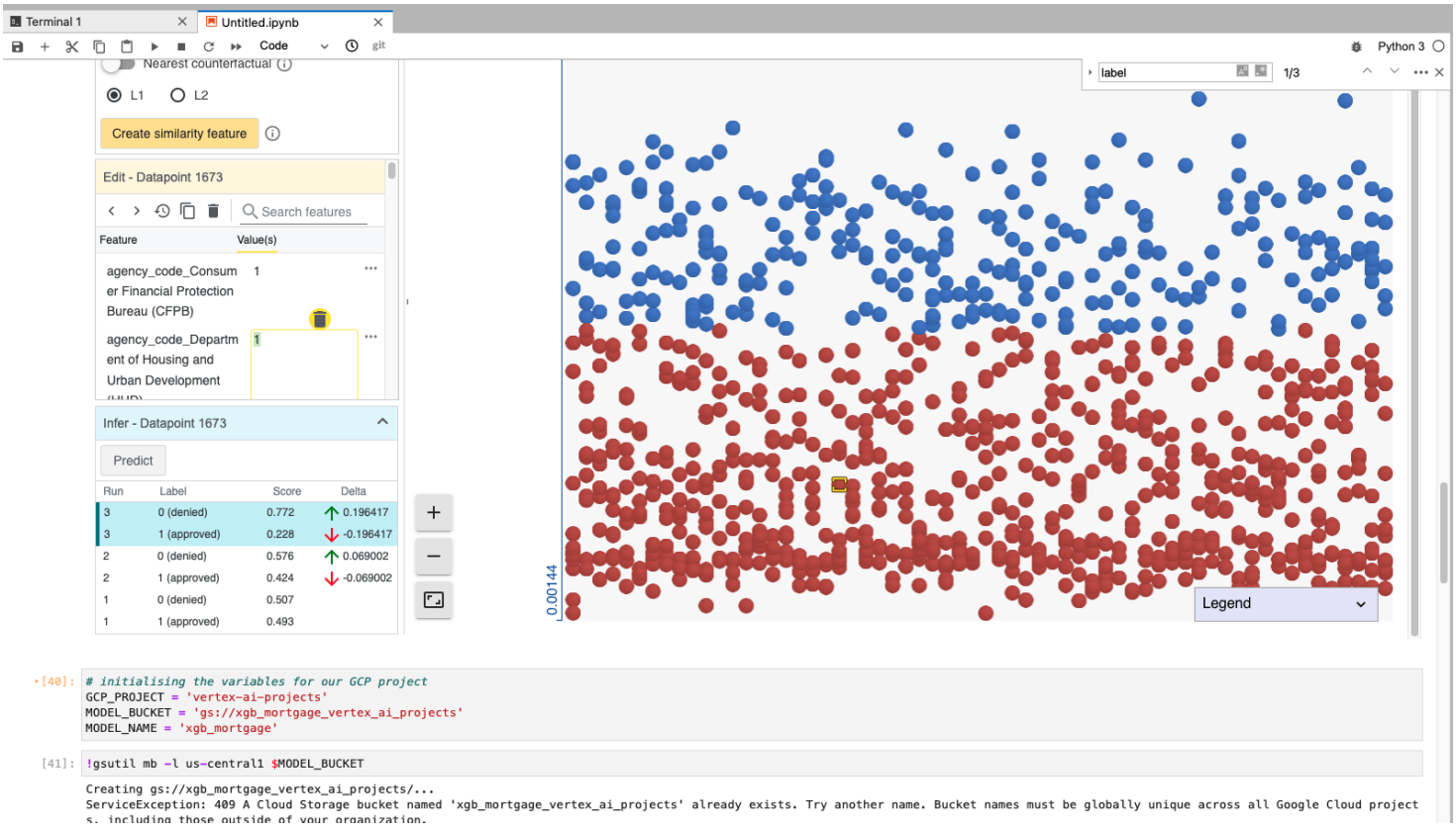
We try to analyse the data point which is closer to .5 threshold to get a better understanding of model.



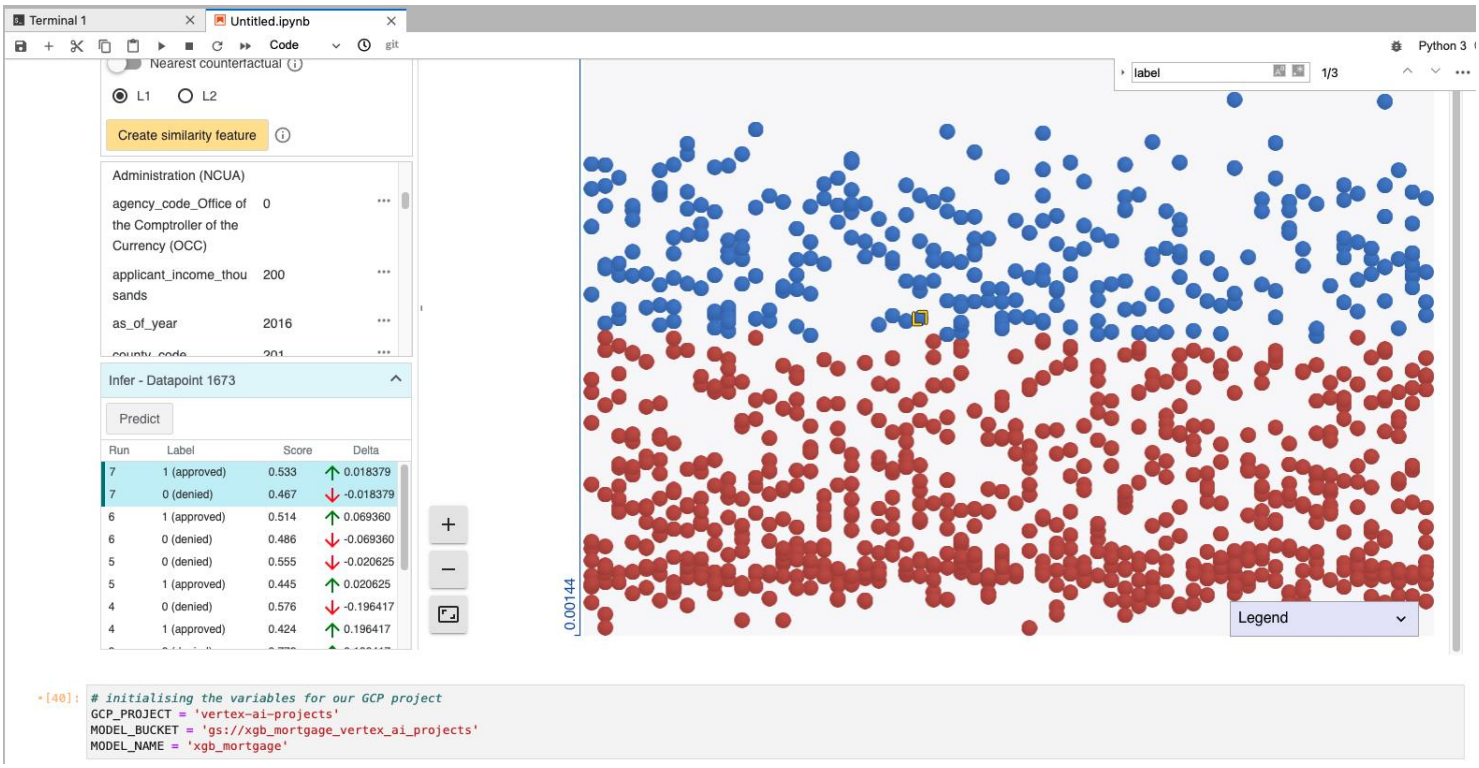
Now we choose a point at the extreme end to analyse which features are actually contributing more to the model.



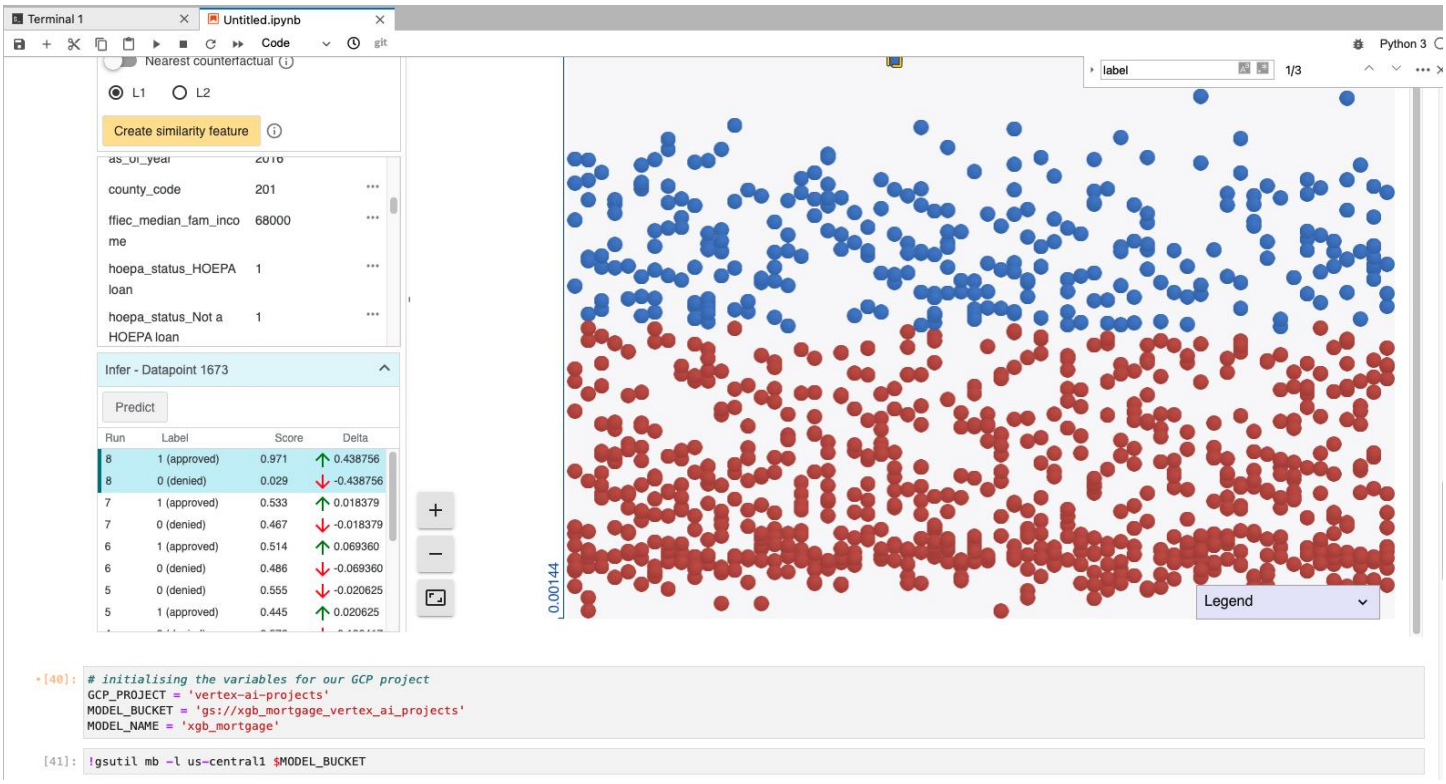
By changing the feature value we observe that there is a 20% change in the prediction probability of the target class indicating that the selected (changed) feature has a huge impact on the model and is important one.



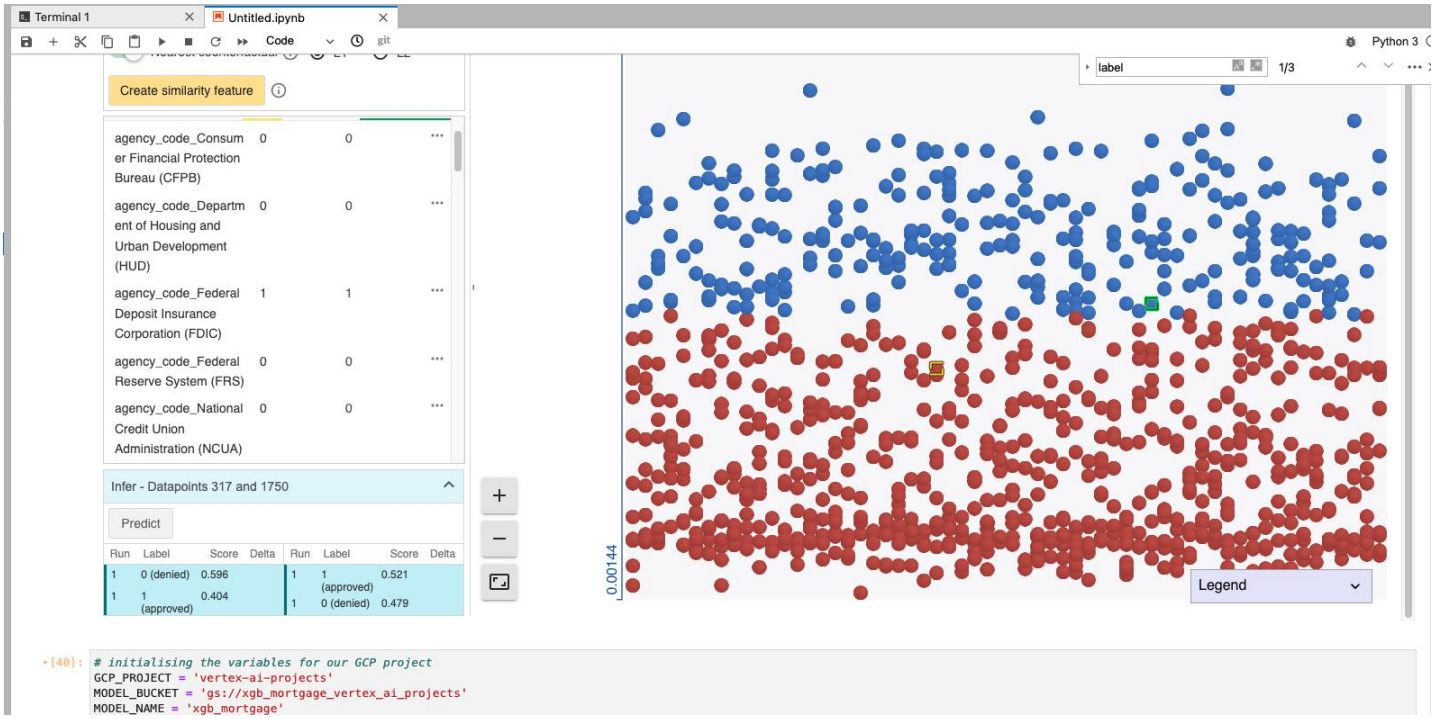
Changing the “applicant_income_thousands” value there is no significant change in the predictions. The delta which refers to the change is just 1.8%.



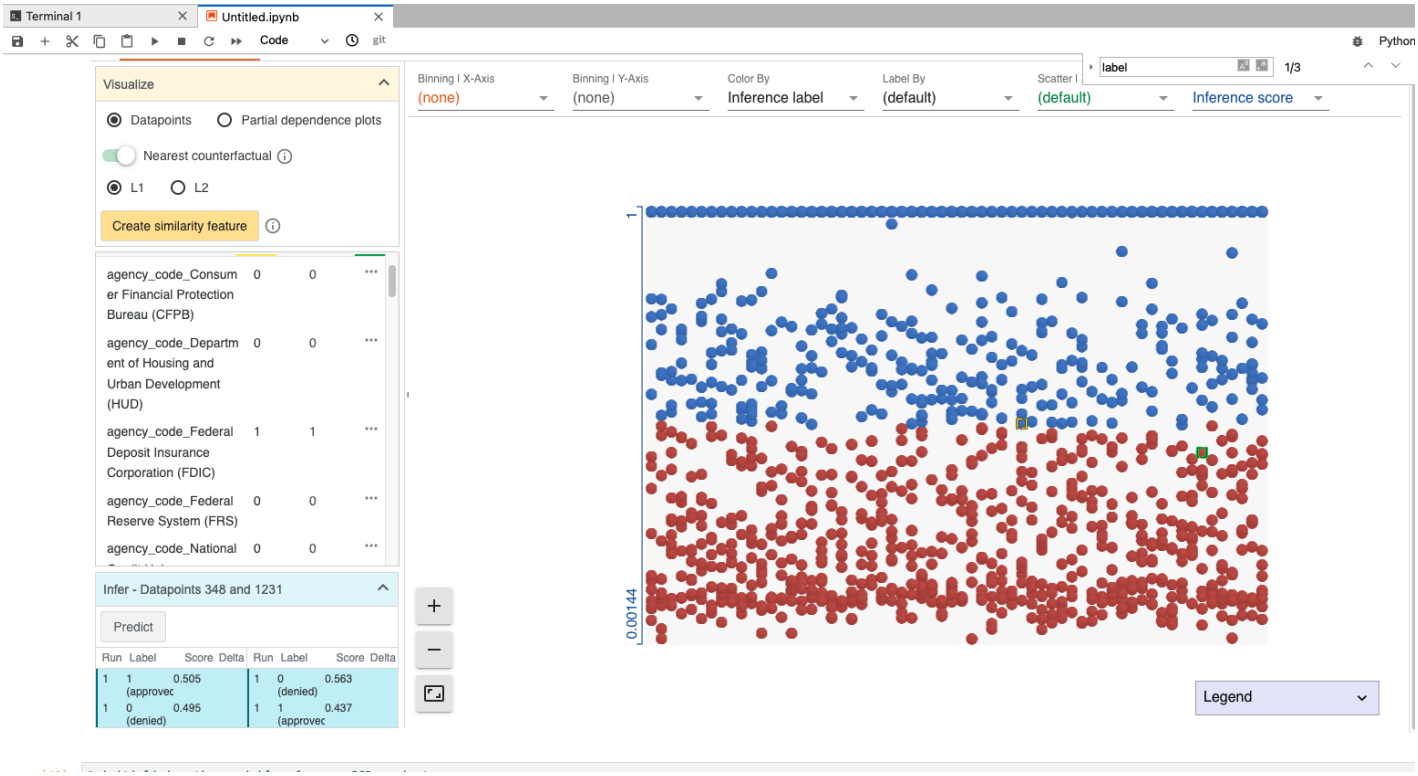
However, by changing the value of " hoepa_status_hoepa_loan" We observe a change in probability at 43%, indicating it is one of the most contributing factors for the model. Model has given huge importance to this feature column.



Now we try to find out nearest counterfactual data point w.r.t a single data point, basically data points which have similar feature values.

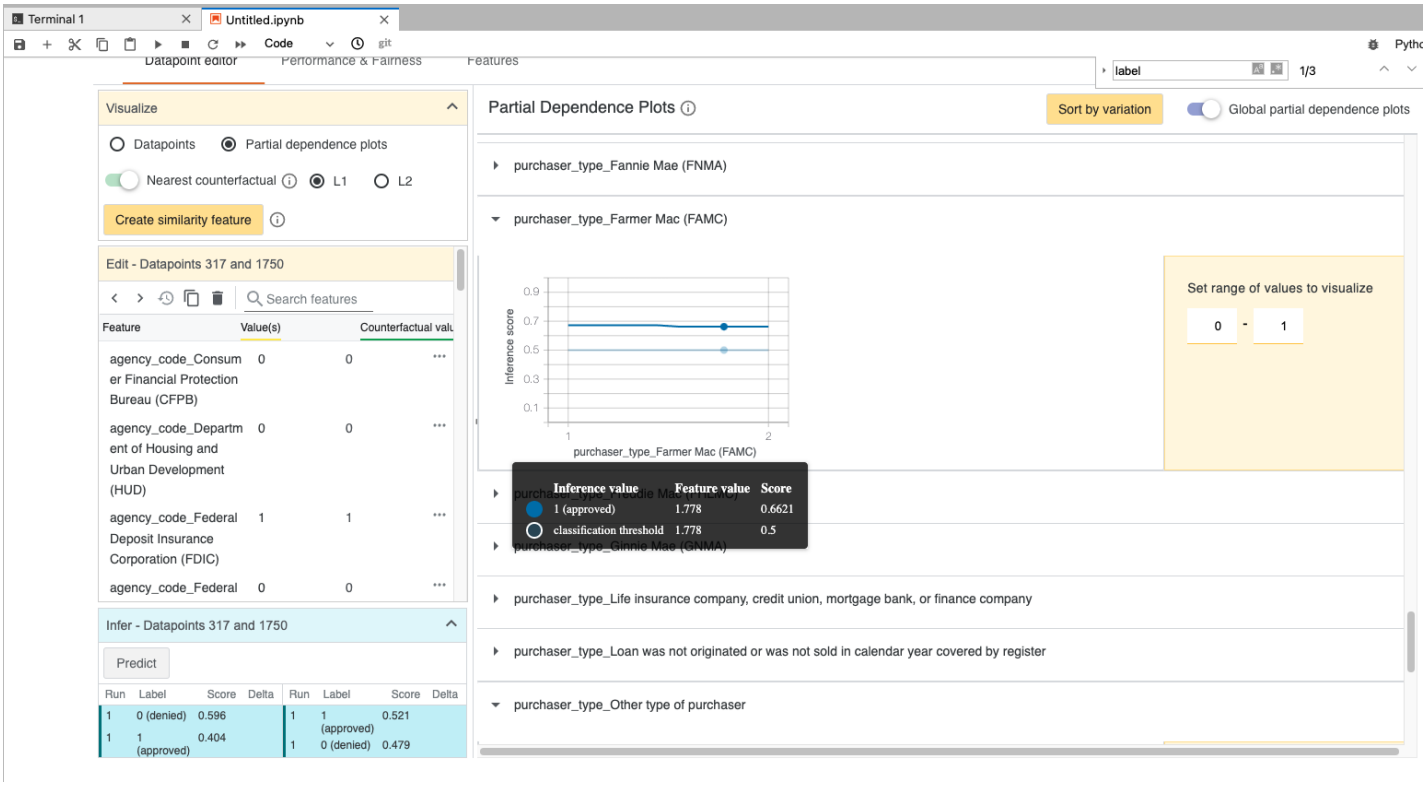


Note that there are 2 data point close to .5 threshold having similar feature values but are predicted in the opposite classes.

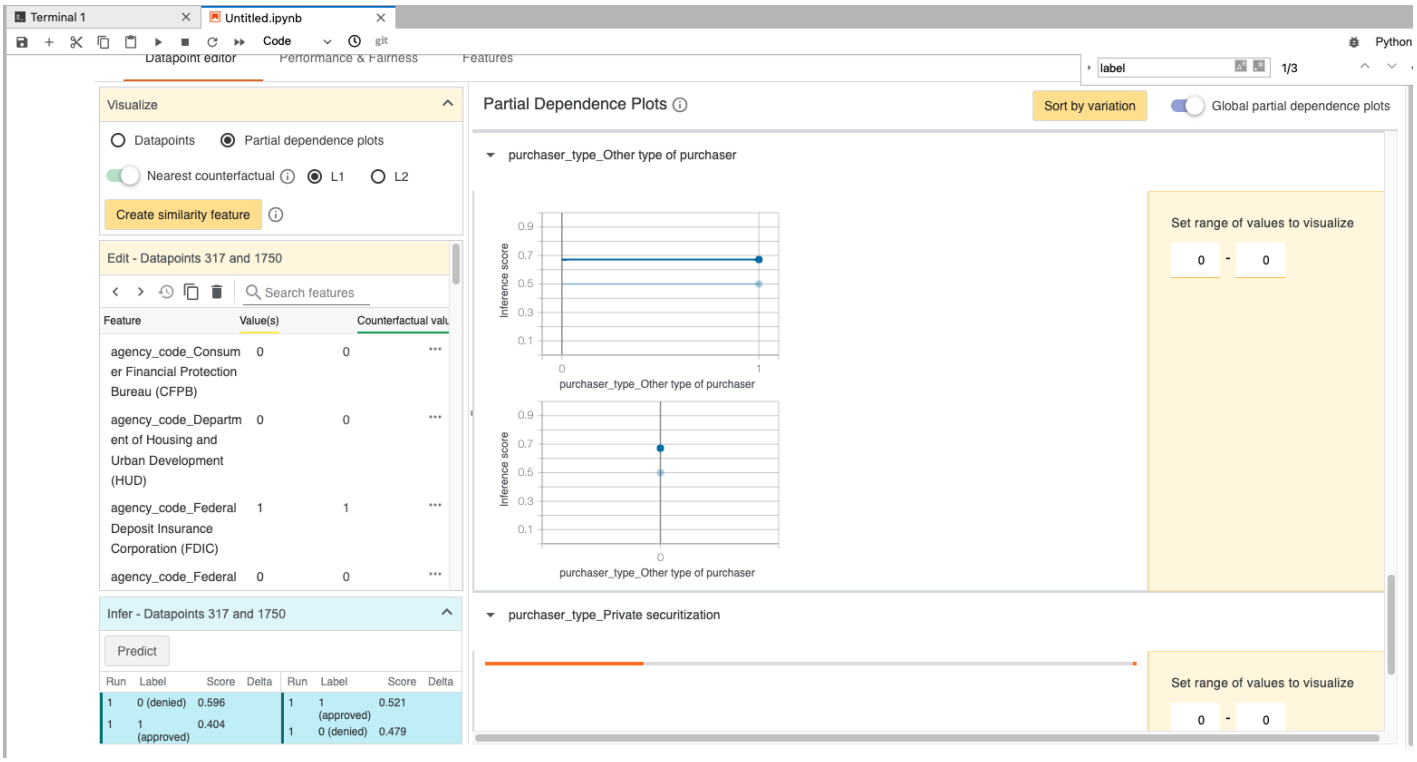


Here we explore the partial dependence plots for various feature columns.
We observe how each feature affects the model performance.

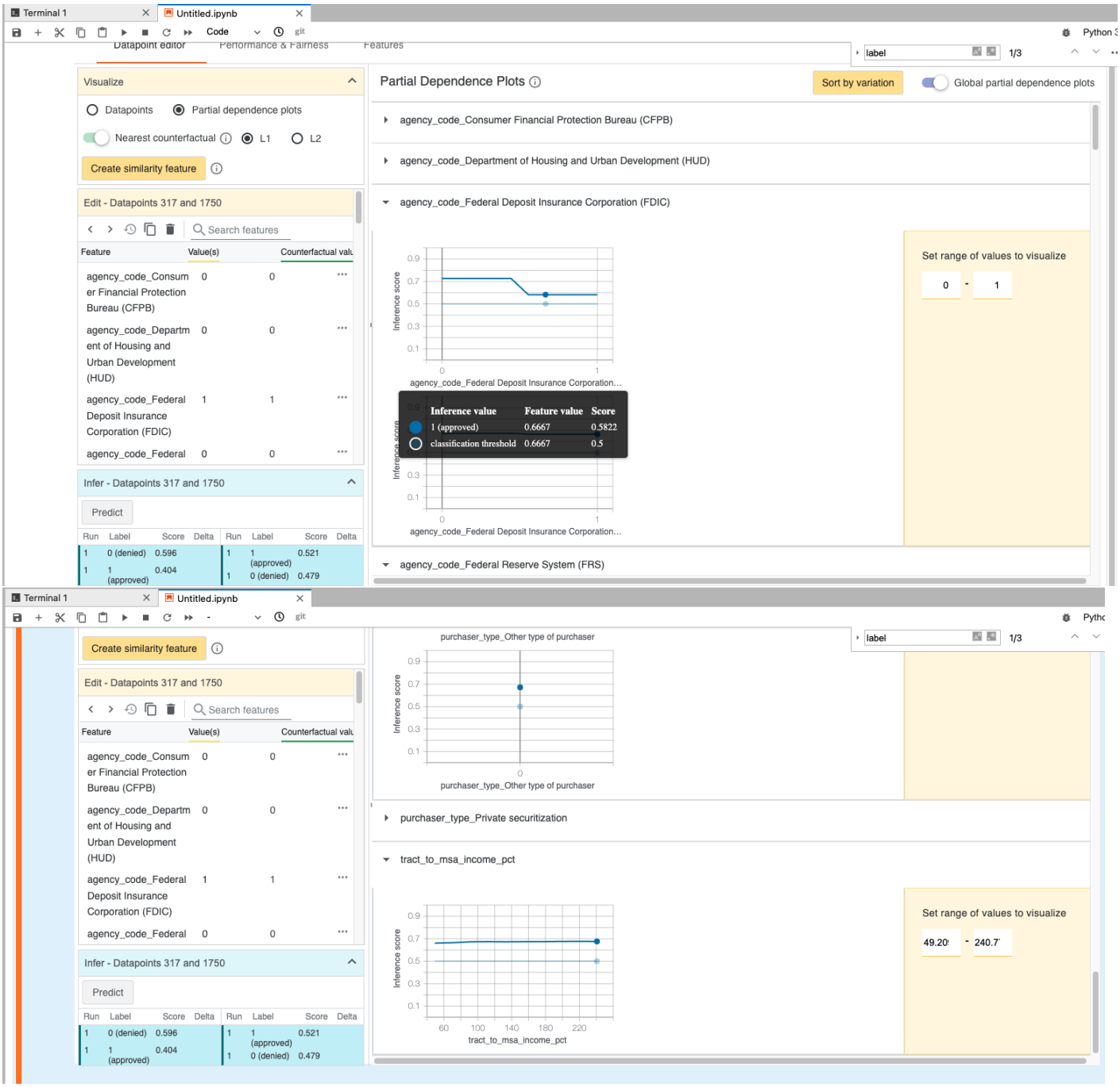
We notice that higher FAMC tends to lower likelihood of loan be approved.



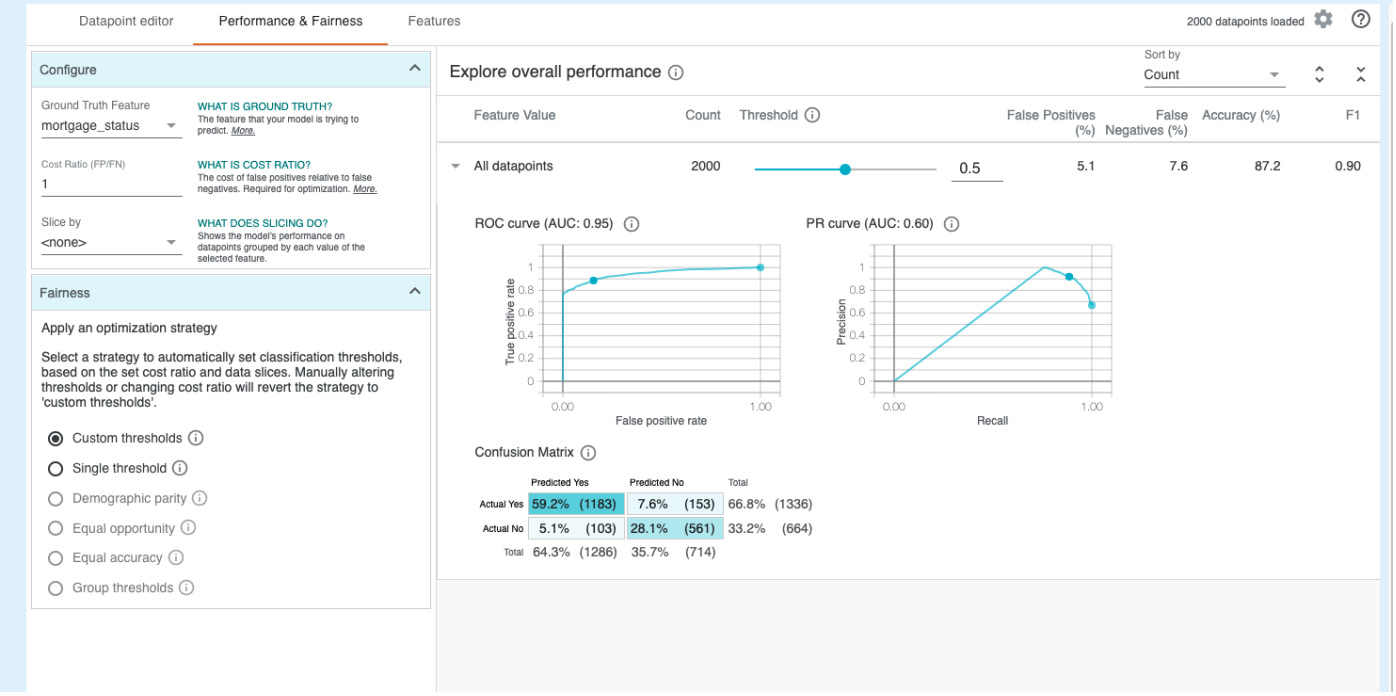
Feature “purchaser_type_other_type_of_purchaser” has little effect on the model decision.



Higher “FDIC” lowers the likelihood of loan being approved.



Now we explore the Overall performance and fairness of the model with all the metrics.



By changing the thresholds we can observe change in the values of metrics.
Precision and recall change significantly.

Highest precision of 0.82 is obtained at around threshold of 0.62 andrecall at 0.82

Datapoint editor | Performance & Fairness | Features

2000 datapoints loaded

Configure

Ground Truth Feature: mortgage_status

WHAT IS GROUND TRUTH?
The feature that your model is trying to predict. *More.*

Cost Ratio (FP/FN): 1

WHAT IS COST RATIO?
The cost of false positives relative to false negatives. Required for optimization. *More.*

Slice by: <none>

WHAT DOES SLICING DO?
Shows the model's performance on datapoints grouped by each value of the selected feature.

Fairness

Apply an optimization strategy

Select a strategy to automatically set classification thresholds, based on the set cost ratio and data slices. Manually altering thresholds or changing cost ratio will revert the strategy to 'custom thresholds'.

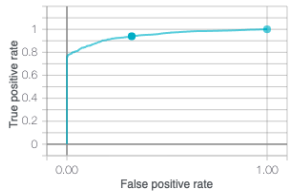
- (•) Custom thresholds
- ( ) Single threshold
- ( ) Demographic parity
- ( ) Equal opportunity
- ( ) Equal accuracy
- ( ) Group thresholds

Explore overall performance

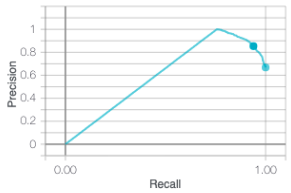
Sort by: Count

Feature Value	Count	Threshold	False Positives (%)	False Negatives (%)	Accuracy (%)	F1
All datapoints	2000	0.32	10.8	4.0	85.2	0.89

ROC curve (AUC: 0.95)



PR curve (AUC: 0.60)



Confusion Matrix

	Predicted Yes		Predicted No		Total	
Actual Yes	62.7%	(1255)	4.0%	(81)	66.8%	(1336)
Actual No	10.8%	(215)	22.4%	(449)	33.2%	(664)
Total	73.5%	(1470)	26.5%	(530)		

Datapoint editor | Performance & Fairness | Features

2000 datapoints loaded

Configure

Ground Truth Feature: mortgage_status

WHAT IS GROUND TRUTH?
The feature that your model is trying to predict. *More.*

Cost Ratio (FP/FN): 1

WHAT IS COST RATIO?
The cost of false positives relative to false negatives. Required for optimization. *More.*

Slice by: loan_amt_thousands | Buck… 2

WHAT DOES SLICING DO?
Shows the model's performance on datapoints grouped by each value of the selected feature.

Slice by (secondary): <none>

Fairness

Apply an optimization strategy

Select a strategy to automatically set classification thresholds, based on the set cost ratio and data slices. Manually altering thresholds or changing cost ratio will revert the strategy to 'custom thresholds'.

- (•) Custom thresholds
- ( ) Single threshold
- ( ) Demographic parity
- ( ) Equal opportunity
- ( ) Equal accuracy
- ( ) Group thresholds

Custom thresholds for 2 values of loan_amt_thousands

Sort by: Count

Feature Value	Count	Threshold	False Positives (%)	False Negatives (%)	Accuracy (%)	F1
[2, 1400)	1989	0.32	10.7	4.1	85.2	0.89
[1400, 2800]	11	0.32	18.2	0.0	81.8	0.89

Datapoint editor | Performance & Fairness | Features

2000 datapoints loaded

Configure

Ground Truth Feature: mortgage_status
WHAT IS GROUND TRUTH? The feature that your model is trying to predict. *More.*

Cost Ratio (FP/FN): 1
WHAT IS COST RATIO? The cost of false positives relative to false negatives. Required for optimization. *More.*

Slice by: loan_purpose_Home
Buck… 2
WHAT DOES SLICING DO? Shows the model's performance on datapoints grouped by each value of the selected feature.

Slice by (secondary): <none>

Fairness

Apply an optimization strategy

Select a strategy to automatically set classification thresholds, based on the set cost ratio and data slices. Manually altering thresholds or changing cost ratio will revert the strategy to 'custom thresholds'.

- (•) Custom thresholds
- ( ) Single threshold
- ( ) Demographic parity
- ( ) Equal opportunity
- ( ) Equal accuracy
- ( ) Group thresholds

Custom thresholds for 2 values of loan_purpose_Home purchase

Sort by: Count

Feature Value	Count	Threshold	False Positives (%)	False Negatives (%)	Accuracy (%)	F1
▸ 0	1078	0.32	12.8	3.9	83.3	0.87
▸ 1	922	0.32	8.4	4.2	87.4	0.92

Datapoint editor | Performance & Fairness | Features

2000 datapoints loaded

Configure

Ground Truth Feature: mortgage_status
WHAT IS GROUND TRUTH? The feature that your model is trying to predict. *More.*

Cost Ratio (FP/FN): 1
WHAT IS COST RATIO? The cost of false positives relative to false negatives. Required for optimization. *More.*

Slice by: loan_purpose_Home
Buck… 2
WHAT DOES SLICING DO? Shows the model's performance on datapoints grouped by each value of the selected feature.

Slice by (secondary): <none>

Fairness

Apply an optimization strategy

Select a strategy to automatically set classification thresholds, based on the set cost ratio and data slices. Manually altering thresholds or changing cost ratio will revert the strategy to 'custom thresholds'.

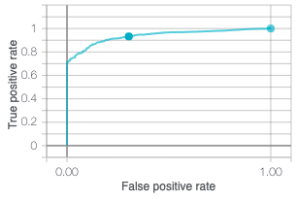
- (•) Custom thresholds
- ( ) Single threshold
- ( ) Demographic parity
- ( ) Equal opportunity
- ( ) Equal accuracy
- ( ) Group thresholds

Custom thresholds for 2 values of loan_purpose_Home purchase

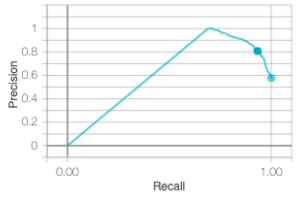
Sort by: Count

Feature Value	Count	Threshold	False Positives (%)	False Negatives (%)	Accuracy (%)	F1
▾ 0	1078	0.32	12.8	3.9	83.3	0.87

ROC curve (AUC: 0.94)



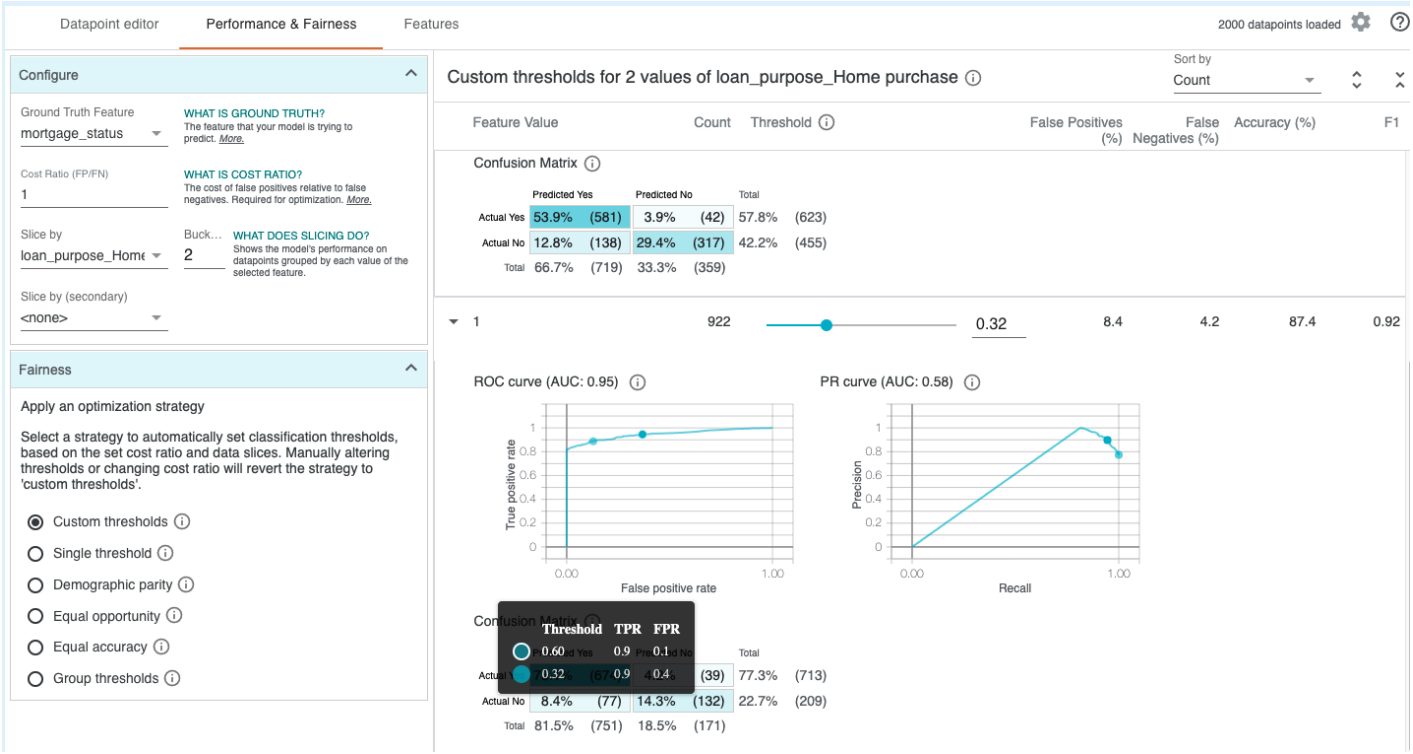
PR curve (AUC: 0.62)



Confusion Matrix

	Predicted Yes		Predicted No		Total	
Actual Yes	53.9%	(581)	3.9%	(42)	57.8%	(623)
Actual No	12.8%	(138)	29.4%	(317)	42.2%	(455)
Total	66.7%	(719)	33.3%	(359)		

▸ 1	922	0.32	8.4	4.2	87.4	0.92



Exploring feature distribution-

This can be used to ensure that the dataaset is balanced.
We observe that very few loans were approved on the basis of loan_type_conventional, GNMA , loan_purpose_improvement features etc.

Datapoint editor | Performance & Fairness | **Features**

2000 datapoints loaded

Sort by: Feature order | Reverse order | Feature search (regex enabled)

Features: int(45) float(2) string(3)

Numeric Features (47)

Chart to show: Standard | log | expand

	count	missing	mean	std dev	zeros	min	median	max
tract_to_msa_income_pct	2,000	0%	116.65	39.41	0%	24.42	111.07	293.28
ffiec_median_fam_income	2,000	0%	69.1k	14.8k	0%	14.4k	67.2k	124k
loan_type_Conventional (any loan other than FHA, VA, FSA, or RHS loans)	2,000	0%	0.75	0.43	24.8%	0	1	1
purchaser_type_Ginnie Mae (GNMA)	2,000	0%	0.09	0.28	91.5%	0	0	1
agency_code_Consumer Financial Protection Bureau (CFPB)	2,000	0%	0.3	0.46	69.55%	0	0	1
mortgage_status	2,000	0%	0.67	0.47	33.2%	0	1	1
loan_purpose_Home purchase	2,000	0%	0.46	0.5	53.9%	0	0	1

Datapoint editor | Performance & Fairness | Features

2000 datapoints loaded

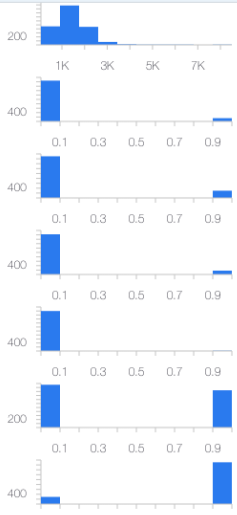
Sort by: Feature order | Reverse order | Feature search (regex enabled)

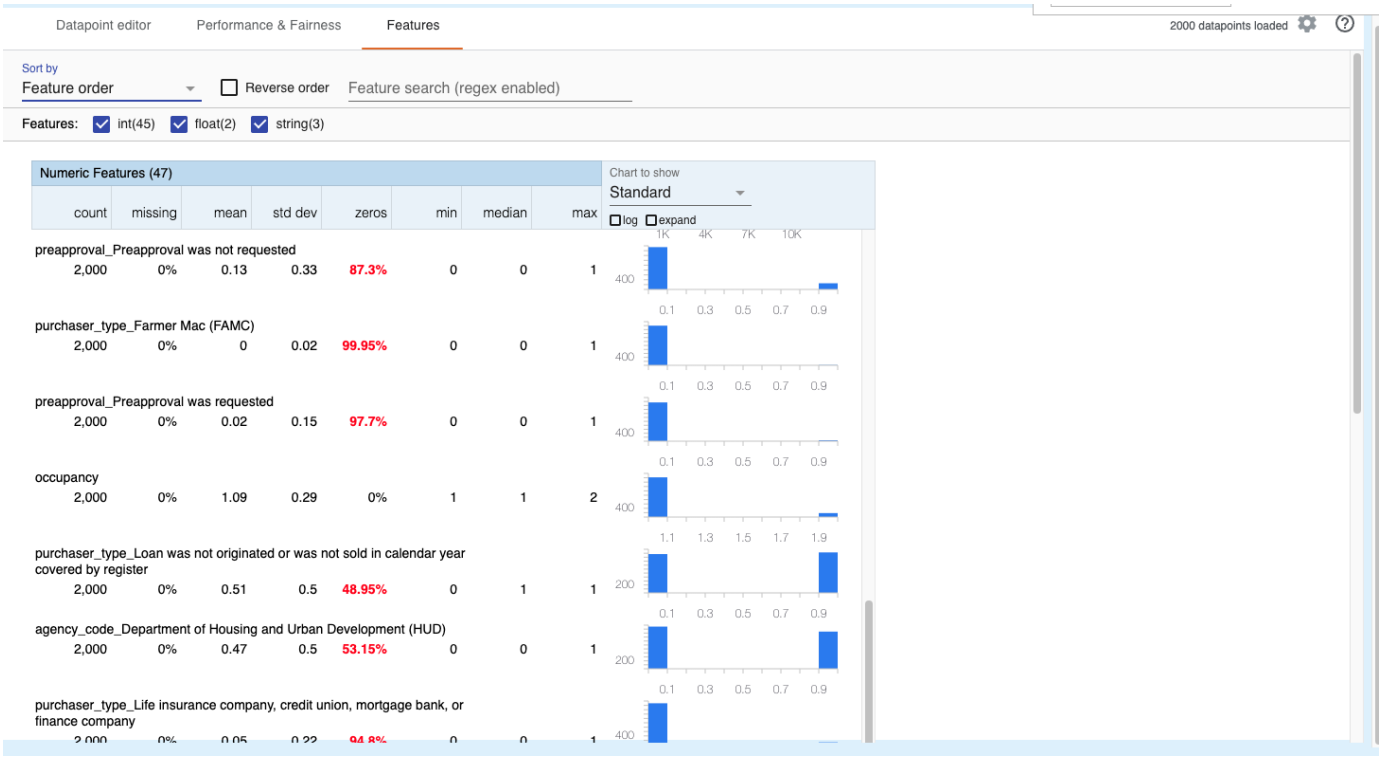
Features: int(45) float(2) string(3)

Numeric Features (47)

Chart to show: Standard | log | expand

	count	missing	mean	std dev	zeros	min	median	max
num_owner_occupied_units	2,000	0%	1,432.74	722.08	0%	71	1,327	8,503
loan_purpose_Home improvement	2,000	0%	0.07	0.26	92.65%	0	0	1
purchaser_type_Fannie Mae (FNMA)	2,000	0%	0.15	0.36	85.2%	0	0	1
agency_code_Federal Deposit Insurance Corporation (FDIC)	2,000	0%	0.09	0.28	91.3%	0	0	1
hoepa_status_HOEPA loan	2,000	0%	0	0.02	99.95%	0	0	1
loan_purpose_Refinancing	2,000	0%	0.47	0.5	53.45%	0	0	1
lien_status_Secured by a first lien	2,000	0%	0.85	0.35	14.6%	0	1	1




Datapoint editor
Performance & Fairness
Features
2000 datapoints loaded
Sort by
Feature order
Reverse order
Feature search (regex enabled)
Features:
int(45)
float(2)
string(3)
Numeric Features (47)
Chart to show
Standard
log
expand
count
missing
mean
std dev
zeros
min
median
max
preapproval_Preapproval was not requested
2,000 0% 0.13 0.33 87.3% 0 0 1
purchaser_type_Farmer Mac (FAMC)
2,000 0% 0 0.02 99.95% 0 0 1
preapproval_Preapproval was requested
2,000 0% 0.02 0.15 97.7% 0 0 1
occupancy
2,000 0% 1.09 0.29 0% 1 1 2
purchaser_type_Loan was not originated or was not sold in calendar year covered by register
2,000 0% 0.51 0.5 48.95% 0 1 1
agency_code_Department of Housing and Urban Development (HUD)
2,000 0% 0.47 0.5 53.15% 0 0 1
purchaser_type_Life insurance company, credit union, mortgage bank, or finance company
1K
4K
7K
10K
400
200
0.1
0.3
0.5
0.7
0.9
1.1
1.3
1.5
1.7
1.9