

Training run "#172_congo_final_model"

Date: 2024-12-12 16:04
Author: Mark K.
Partner: finca-congo
Target mean encoding: Disabled
Number of features: 16
Calibration: Disabled
Macro enrichment: Disabled
Validation: Out-of-sample

Data overview

| Samples | Train size | Test size | Total good | Total bad | Train good | Train bad | Test good | Test bad | Train NPL | Test NPL |
|---------|------------|-----------|------------|-----------|------------|-----------|-----------|----------|-----------|----------|
| 12489 | 9991 | 2498 | 11420 | 1069 | 9136 | 855 | 2284 | 214 | 0.0856 | 0.0857 |

| id | Feature | Type | Missing | Min | Max | Mode |
|----|------------------|---------|---------|---------|-----------|------------|
| 1 | disb_amt_usd | float64 | 0.0 | 43.2255 | 195660.79 | 1422.25 |
| 2 | installment_num | int64 | 0.0 | 6 | 24 | 10 |
| 3 | eir | float64 | 0.0 | 0.0 | 97.16 | 68.66 |
| 4 | age | float64 | 0.00088 | 3.0 | 89.0 | 34.0 |
| 5 | collateral_balan | float64 | 0.0 | 0.0 | 39110.8 | 0.0 |
| 6 | product_id | int64 | 0.0 | 3 | 66 | 8 |
| 7 | marital_status | object | 0.0 | - | - | m |
| 8 | gender | object | 0.0 | - | - | m |
| 9 | branch_name | object | 0.0 | - | - | kalemie |
| 10 | branch_city | object | 0.0 | - | - | lubumbashi |
| 11 | cust_sector | object | 0.0 | - | - | 10003 |
| 12 | usd_exchange | float64 | 0.0 | 1953.7 | 2720.0 | 1995.14 |
| 13 | consumer_price_i | float64 | 0.0 | 150.422 | 262.47 | 159.749 |
| 14 | food_inflation | float64 | 0.0 | 4.48 | 23.71 | 12.22 |
| 15 | inflation_rate | float64 | 0.0 | 5.13 | 46.8 | 5.85 |
| 16 | interest_rate | float64 | 0.0 | 7.5 | 25.0 | 7.5 |

Model performance and hyperparameters

| Train AUC | Test AUC | Train Gini | Test Gini |
|-----------|----------|------------|-----------|
| 0.7805 | 0.7551 | 0.562 | 0.51 |

| iterations | depth | learning_rate |
|------------|-------|---------------|
| 300 | 4 | 0.03 |

Threshold breakdown

| Threshold | Accuracy | Model NPL | Approval rate | Good found | Bad found | TP | TN | FP | FN |
|-----------|----------|-----------|---------------|------------|-----------|-----|------|------|-----|
| 0.01 | 0.098 | 0.0 | 0.013 | 0.014 | 1.0 | 214 | 32 | 2252 | 0 |
| 0.02 | 0.204 | 0.01 | 0.12 | 0.13 | 0.986 | 211 | 298 | 1986 | 3 |
| 0.03 | 0.292 | 0.01 | 0.211 | 0.228 | 0.977 | 209 | 521 | 1763 | 5 |
| 0.04 | 0.353 | 0.017 | 0.277 | 0.298 | 0.944 | 202 | 681 | 1603 | 12 |
| 0.05 | 0.416 | 0.02 | 0.344 | 0.369 | 0.921 | 197 | 843 | 1441 | 17 |
| 0.06 | 0.479 | 0.023 | 0.412 | 0.44 | 0.888 | 190 | 1006 | 1278 | 24 |
| 0.07 | 0.548 | 0.026 | 0.488 | 0.52 | 0.85 | 182 | 1187 | 1097 | 32 |
| 0.08 | 0.608 | 0.036 | 0.562 | 0.593 | 0.766 | 164 | 1354 | 930 | 50 |
| 0.09 | 0.667 | 0.037 | 0.628 | 0.661 | 0.729 | 156 | 1510 | 774 | 58 |
| 0.1 | 0.716 | 0.043 | 0.689 | 0.722 | 0.654 | 140 | 1648 | 636 | 74 |
| 0.11 | 0.755 | 0.049 | 0.741 | 0.771 | 0.579 | 124 | 1762 | 522 | 90 |
| 0.12 | 0.79 | 0.052 | 0.786 | 0.815 | 0.523 | 112 | 1862 | 422 | 102 |
| 0.13 | 0.821 | 0.06 | 0.836 | 0.859 | 0.416 | 89 | 1963 | 321 | 125 |
| 0.14 | 0.838 | 0.066 | 0.866 | 0.885 | 0.336 | 72 | 2021 | 263 | 142 |
| 0.15 | 0.86 | 0.067 | 0.894 | 0.912 | 0.304 | 65 | 2084 | 200 | 149 |
| 0.16 | 0.872 | 0.069 | 0.912 | 0.929 | 0.266 | 57 | 2121 | 163 | 157 |
| 0.17 | 0.88 | 0.07 | 0.924 | 0.94 | 0.243 | 52 | 2146 | 138 | 162 |
| 0.18 | 0.89 | 0.071 | 0.937 | 0.952 | 0.224 | 48 | 2175 | 109 | 166 |
| 0.19 | 0.898 | 0.072 | 0.948 | 0.962 | 0.206 | 44 | 2198 | 86 | 170 |
| 0.2 | 0.9 | 0.074 | 0.956 | 0.968 | 0.173 | 37 | 2212 | 72 | 177 |
| 0.21 | 0.906 | 0.076 | 0.966 | 0.977 | 0.145 | 31 | 2231 | 53 | 183 |
| 0.22 | 0.907 | 0.078 | 0.973 | 0.981 | 0.117 | 25 | 2241 | 43 | 189 |
| 0.23 | 0.909 | 0.079 | 0.978 | 0.985 | 0.093 | 20 | 2250 | 34 | 194 |
| 0.24 | 0.909 | 0.081 | 0.984 | 0.988 | 0.065 | 14 | 2257 | 27 | 200 |
| 0.25 | 0.91 | 0.084 | 0.99 | 0.992 | 0.033 | 7 | 2265 | 19 | 207 |
| 0.26 | 0.912 | 0.084 | 0.993 | 0.995 | 0.023 | 5 | 2272 | 12 | 209 |
| 0.27 | 0.913 | 0.084 | 0.995 | 0.996 | 0.019 | 4 | 2276 | 8 | 210 |
| 0.28 | 0.913 | 0.084 | 0.996 | 0.997 | 0.019 | 4 | 2277 | 7 | 210 |
| 0.29 | 0.913 | 0.085 | 0.998 | 0.998 | 0.005 | 1 | 2279 | 5 | 213 |
| 0.3 | 0.914 | 0.085 | 0.998 | 0.999 | 0.005 | 1 | 2281 | 3 | 213 |

Optimal threshold (Youden's J statistic): 0.089

Train/test split validations

| Kolmogorov-Smirnov test | | |
|-------------------------|---------|------------|
| Feature | P-value | Similarity |
| disb_amt_usd | 0.2861 | High |
| eir | 0.4189 | High |
| usd_exchange | 0.5676 | High |

| | | |
|--------------------------|--------|------|
| inflation_rate | 0.7302 | High |
| consumer_price_index_cpi | 0.7541 | High |
| collateral_balance | 0.7952 | High |
| product_id | 0.8624 | High |
| food_inflation | 0.9094 | High |
| interest_rate | 0.9153 | High |
| age | 0.9764 | High |
| installment_num | 0.9806 | High |

| | |
|------------------------|-------|
| Adversarial validation | |
| AUC | 0.507 |

PSI

| product_id | population | sample | psi |
|------------|------------|--------|-----|
| 3 | 0.01 | 0.04 | Low |
| 8 | 64.51 | 63.17 | Low |
| 12 | 0.91 | 1.04 | Low |
| 13 | 17.74 | 18.9 | Low |
| 15 | 11.78 | 12.21 | Low |
| 27 | 0.09 | 0.04 | Low |
| 42 | 0.01 | 0.04 | Low |
| 56 | 4.43 | 3.88 | Low |
| 57 | 0.06 | 0.12 | Low |
| 58 | 0.05 | 0.12 | Low |
| 62 | 0.37 | 0.4 | Low |
| 66 | 0.05 | 0.04 | Low |

| marital_status | population | sample | psi |
|----------------|------------|--------|-----|
| d | 1.31 | 1.4 | Low |
| m | 69.9 | 69.02 | Low |
| null_value | 2.3 | 2.32 | Low |
| other | 0.06 | 0.04 | Low |
| s | 24.94 | 25.42 | Low |
| w | 1.49 | 1.8 | Low |

| gender | population | sample | psi |
|--------|------------|--------|-----|
| f | 27.98 | 27.82 | Low |
| m | 72.02 | 72.18 | Low |

| branch_name | population | sample | psi |
|--------------|------------|--------|----------|
| bel_air | 2.66 | 3.2 | Low |
| boma | 3.94 | 4.56 | Low |
| bukavu_1 | 3.24 | 2.72 | Low |
| bukavu_2 | 3.96 | 4.08 | Low |
| durba | 4.76 | 4.52 | Low |
| goma | 6.35 | 5.84 | Low |
| gombe | 5.42 | 5.0 | Low |
| kalemie | 13.97 | 15.41 | Moderate |
| kananga | 1.83 | 1.36 | Moderate |
| kanangatshik | 0.26 | 0.2 | Low |
| katuba | 5.69 | 5.48 | Low |
| kintambo | 2.57 | 2.48 | Low |
| kisangani | 2.67 | 2.6 | Low |
| kisanganibut | 0.21 | 0.28 | Low |

| | | | |
|------------|------|------|-----|
| kolwezi | 4.57 | 4.2 | Low |
| likasi | 4.88 | 5.32 | Low |
| limete | 2.51 | 2.6 | Low |
| lubumbashi | 8.55 | 8.49 | Low |
| masina | 2.59 | 2.32 | Low |
| matadi | 5.0 | 5.2 | Low |
| mbuji_mayi | 1.82 | 2.04 | Low |
| ndjili | 2.28 | 2.28 | Low |
| ngaba | 2.31 | 2.44 | Low |
| other | 0.33 | 0.28 | Low |
| upn | 5.52 | 5.16 | Low |
| victoire | 2.1 | 1.88 | Low |

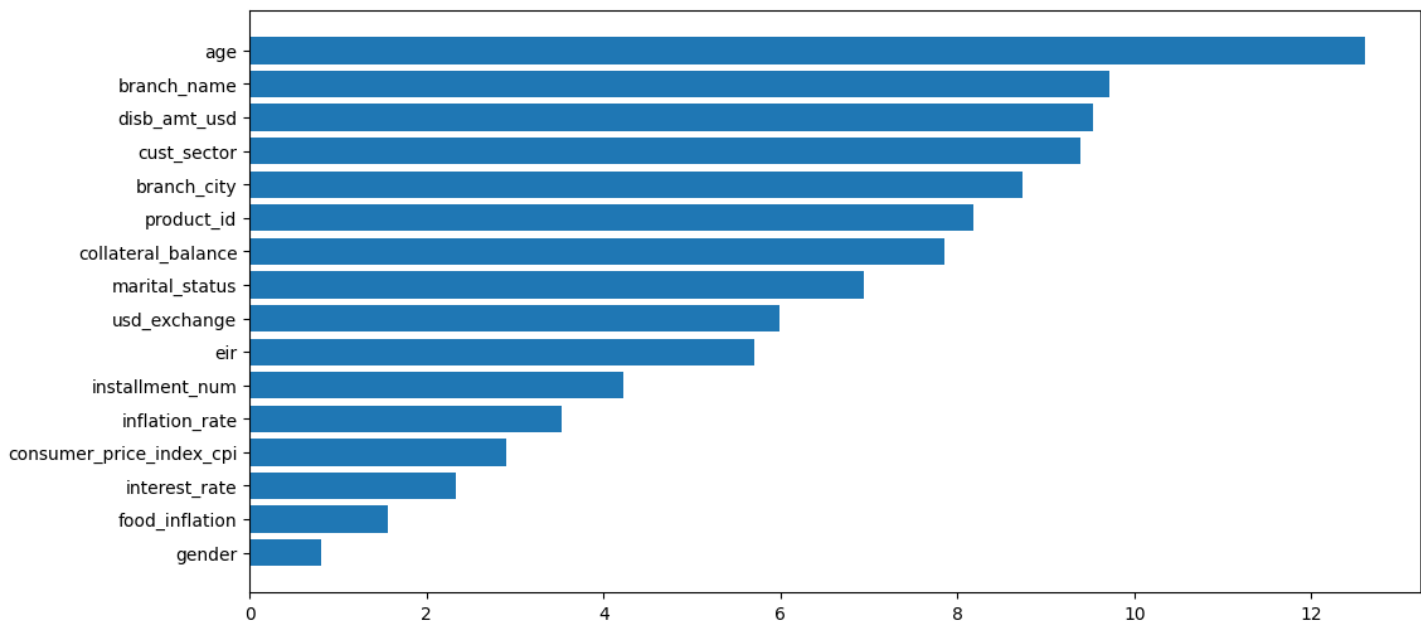
| branch_city | population | sample | psi |
|-------------|------------|--------|----------|
| boma | 3.94 | 4.56 | Low |
| bondo | 4.76 | 4.52 | Low |
| bukavu | 7.2 | 6.81 | Low |
| butembo | 0.21 | 0.28 | Low |
| goma | 6.35 | 5.84 | Low |
| gombe | 5.42 | 5.0 | Low |
| kalamu | 2.1 | 1.88 | Low |
| kalemie | 13.97 | 15.41 | Moderate |
| kananga | 1.83 | 1.36 | Moderate |
| kisangani | 2.67 | 2.6 | Low |
| kolwezi | 4.57 | 4.2 | Low |
| lemba | 2.31 | 2.44 | Low |
| likasi | 4.88 | 5.32 | Low |
| limete | 2.51 | 2.6 | Low |
| lubumbashi | 16.89 | 17.17 | Low |
| masina | 2.59 | 2.32 | Low |
| matadi | 5.0 | 5.2 | Low |
| mbujimayi | 1.82 | 2.04 | Low |
| ndjili | 2.28 | 2.28 | Low |
| ngaliema | 8.1 | 7.65 | Low |
| other | 0.33 | 0.28 | Low |
| tshikapa | 0.26 | 0.2 | Low |

| cust_sector | population | sample | psi |
|-------------|------------|--------|-----|
| 10003 | 63.46 | 62.33 | Low |
| 10010 | 2.29 | 2.4 | Low |
| 11001 | 0.52 | 0.4 | Low |
| 11002 | 2.73 | 2.48 | Low |

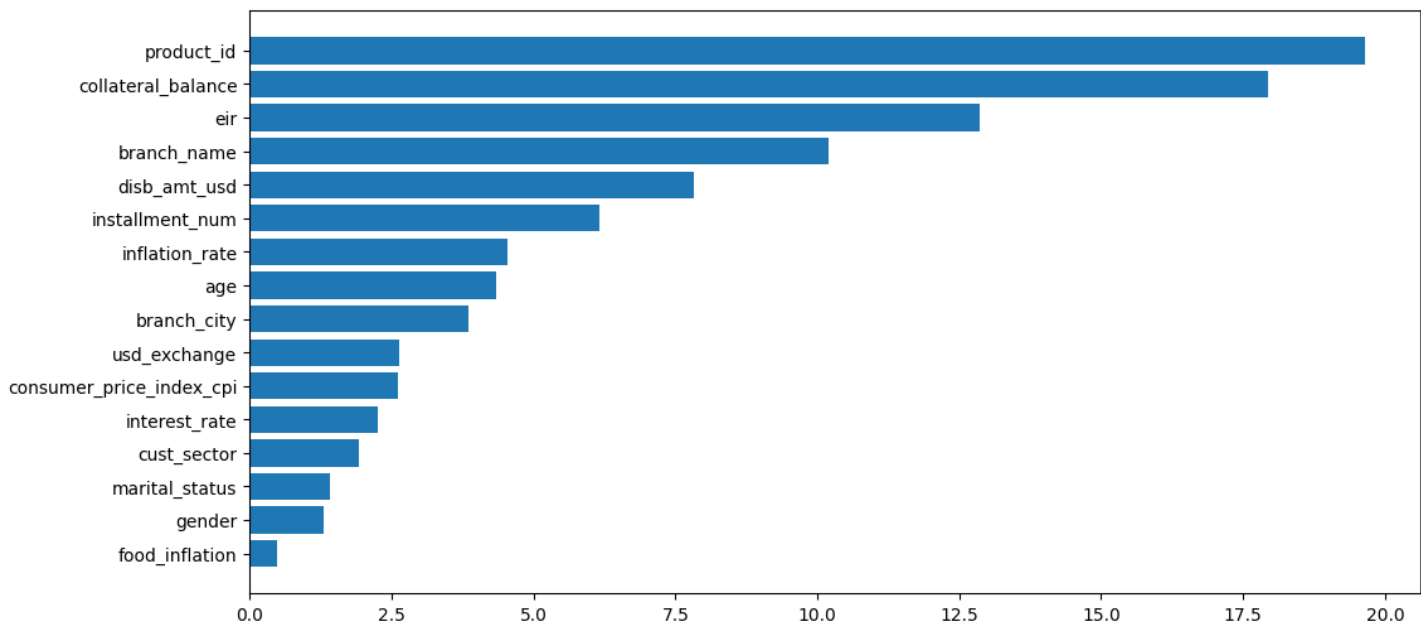
| | | | |
|-------|-------|-------|----------|
| 21003 | 0.34 | 0.2 | Low |
| 21004 | 0.14 | 0.08 | Low |
| 22001 | 0.1 | 0.08 | Low |
| 23001 | 0.08 | 0.08 | Low |
| 24001 | 0.18 | 0.2 | Low |
| 25001 | 0.06 | 0.12 | Low |
| 25004 | 0.07 | 0.08 | Low |
| 26002 | 0.06 | 0.04 | Low |
| 26004 | 0.04 | 0.04 | Low |
| 30001 | 0.27 | 0.44 | Low |
| 30004 | 0.24 | 0.24 | Low |
| 30006 | 0.26 | 0.2 | Low |
| 3030 | 0.18 | 0.04 | Moderate |
| 3040 | 0.29 | 0.4 | Low |
| 40001 | 0.19 | 0.2 | Low |
| 40002 | 0.18 | 0.2 | Low |
| 40003 | 2.4 | 2.0 | Low |
| 40004 | 0.35 | 0.52 | Low |
| 40005 | 1.47 | 1.36 | Low |
| 40006 | 0.27 | 0.32 | Low |
| 40007 | 0.36 | 0.48 | Low |
| 40009 | 11.99 | 12.85 | Low |
| 40010 | 1.52 | 1.48 | Low |
| 40011 | 0.1 | 0.12 | Low |
| 40012 | 0.1 | 0.04 | Low |
| 40013 | 0.45 | 0.52 | Low |
| 40014 | 0.02 | 0.04 | Low |
| 40017 | 1.12 | 1.28 | Low |
| 40018 | 0.82 | 0.96 | Low |
| 40019 | 0.31 | 0.44 | Low |
| 40021 | 0.14 | 0.12 | Low |
| 40022 | 0.22 | 0.4 | Moderate |
| 40023 | 0.89 | 1.16 | Low |
| 50006 | 0.07 | 0.08 | Low |
| 50007 | 0.22 | 0.24 | Low |
| 60001 | 0.38 | 0.36 | Low |
| 70002 | 0.28 | 0.28 | Low |
| 70003 | 0.27 | 0.24 | Low |
| 80002 | 0.38 | 0.64 | Moderate |
| 80004 | 0.36 | 0.24 | Low |

| | | | |
|-------|------|------|-----|
| 90002 | 0.21 | 0.16 | Low |
| 90003 | 0.1 | 0.08 | Low |
| 90005 | 0.31 | 0.44 | Low |
| 90011 | 0.03 | 0.08 | Low |
| 90012 | 0.14 | 0.16 | Low |
| 90013 | 0.06 | 0.08 | Low |
| f3082 | 0.02 | 0.04 | Low |
| fsc10 | 0.42 | 0.44 | Low |
| fsc77 | 1.08 | 1.04 | Low |
| fsc80 | 0.07 | 0.08 | Low |
| other | 1.16 | 0.96 | Low |

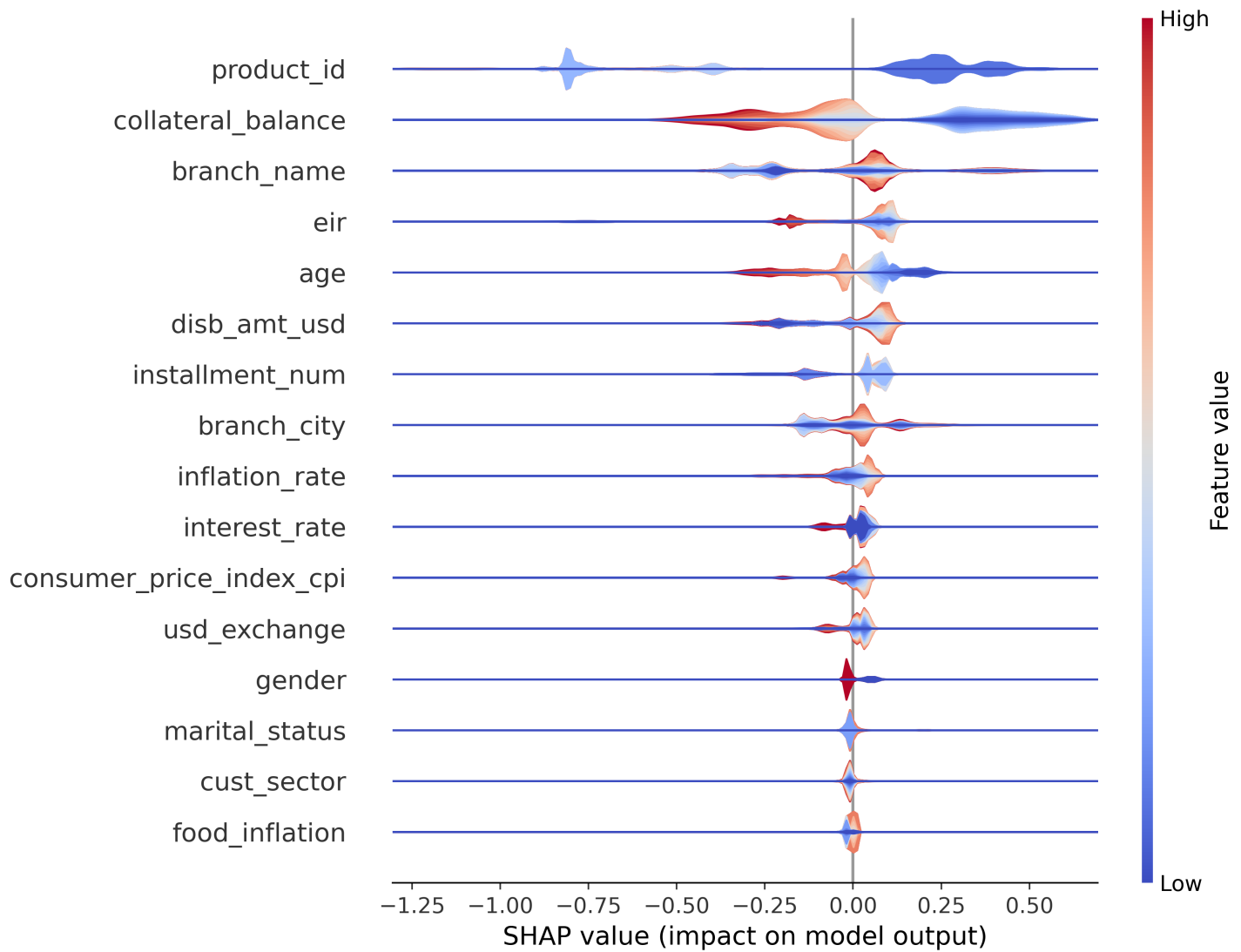
Adversarial validation feature importance



Model feature importance



Shapley feature importance



Expected calibration error: 0.005

PD Buckets and Default Rates

