Bank of Baku – QNBWise

Advanced Scorecard Development
Training

December 2023

Table of Contents

1. End-to-end modeling flow

1. End-to-end modeling flow

Base model

Import libraries
- QNBAnalytics_ML

Read data
- Train
- Test

Create pipeline for base model
- Null feature elimination
- Constant feature elimination
- Low gini feature elimination
- Correlated feature elimination
- Binning
- Logistic regression

Save base model scores

End of Layer 1

1. End-to-end modeling flow

<u>Base model</u>

Import libraries
- QNBAnalytics_ML

Read data
- Train
- Test

Create pipeline for base model
- Null feature elimination
- Constant feature elimination
- Low gini feature elimination
- Correlated feature elimination
- Binning
- Logistic regression

Save base model scores

<u>End of Layer 1</u>

<u>Good and not good segment models</u>

Split the data
- Good segment
- Not good segment

Create pipeline for both segment models
- Null feature elimination
- Constant feature elimination
- Low gini feature elimination
- Correlated feature elimination
- Scaling
- Null imputation
- Encoding
- Logistic regression
- Random forest
- XGBoost
- LightGBM

Select best models

Save both segment model scores

<u>End of Layer 2</u>

1. End-to-end modeling flow

Base model

Import libraries
- QNBAnalytics_ML

Read data
- Train
- Test

Create pipeline for base model
- Null feature elimination
- Constant feature elimination
- Low gini feature elimination
- Correlated feature elimination
- Binning
- Logistic regression

Save base model scores

End of Layer 1

Good and not good segment models

Split the data
- Good segment
- Not good segment

Create pipeline for both segment models
- Null feature elimination
- Constant feature elimination
- Low gini feature elimination
- Correlated feature elimination
- Scaling
- Null imputation
- Encoding
- Logistic regression
- Random forest
- XGBoost
- LightGBM

Select best models

Save both segment model scores

End of Layer 2

Final (Meta) Model

Combine scores of these 3 models

Create pipeline for final (meta) model
- Logistic regression

Save final scores

End of modelling

2. Functions used in modeling

2.1. Data import
2.2. Data explore
2.3. Feature selection
2.4. Binning
2.5. Scaling
2.6. Null imputation
2.7. Encoding

2. Functions used in modeling

2.1. Data import

- Database

```
engine = connect_to_sql(db_username, db_password)
train = data_load(engine, sql=train_data_sql)
```

- File (csv, xlsx)

```
train = data_load(engine=None, data=train_data_path)
```

2. Functions used in modeling

2.2. Data explore

`data_explore()`

- Data types
  - Categoric
  - Numeric
  - Date
- Count, unique count
- Missing rate
- Mean, min, max, std

2. Functions used in modeling

2.2. Data explore

`data_explore()`

- Data types
  - Categoric
  - Numeric
  - Date
- Count, unique count
- Missing rate
- Mean, min, max, std

	datatypes	role	count	unique_count	missing_rate	mean	min	max	std	use
VAR1	numeric	input	10734	10	0.000000	0.933296	0.0	10.0	0.840168	True
VAR2	numeric	input	10734	5	0.000000	0.183622	0.0	4.0	0.435014	True
VAR3	numeric	input	10734	8	0.000000	0.415782	0.0	7.0	0.642262	True
VAR4	numeric	input	10734	7	0.000000	0.264021	0.0	7.0	0.498181	True
VAR5	numeric	input	10734	10	0.000000	0.679802	0.0	10.0	0.784676	True
VAR6	numeric	input	10734	5	0.000000	0.232159	0.0	4.0	0.470053	True
VAR7	numeric	input	10734	8	0.000000	0.49618	0.0	8.0	0.664786	True
VAR8	numeric	input	10734	10	0.000000	0.916806	0.0	10.0	0.834073	True
VAR9	categoric	input	10734	4	0.000000	NaN	NaN	NaN	NaN	True
VAR10	categoric	input	9708	4	9.558413	NaN	NaN	NaN	NaN	True

2. Functions used in modeling

2.3. Feature selection

- Null feature elimination

```
feature_elimination(eliminator='drop_null_features',
                    params={'threshold': 0.99})
```

2. Functions used in modeling

2.3. Feature selection

- Constant feature elimination

```
feature_elimination(eliminator='drop_constant_features')
```

2. Functions used in modeling

2.3. Feature selection

- Low gini feature elimination

```
feature_elimination(eliminator='drop_low_gini_features',
                    params={'threshold':0.05})
```

2. Functions used in modeling

Gini coefficient

A measure to assess
the predictive power of a feature and
the performance of a model.

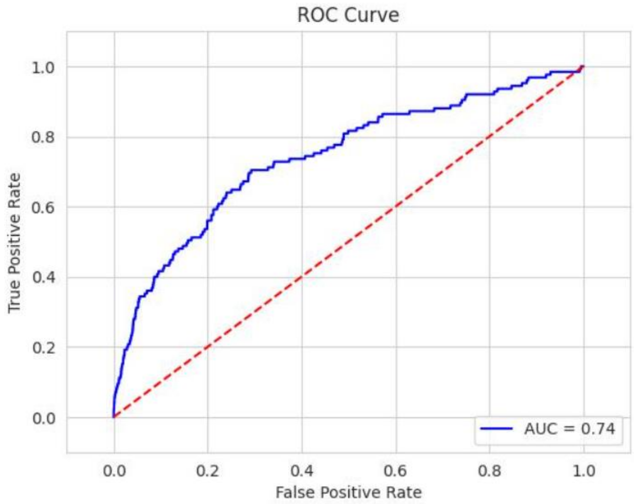
2. Functions used in modeling

Gini coefficient

A measure to assess
the predictive power of a feature and
the performance of a model.

Takes values between 0 and 1.

Calculated as
Gini = 2*AUC – 1

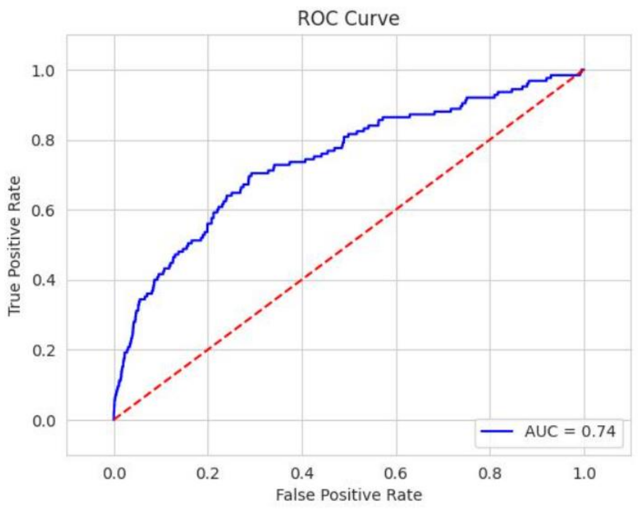


2. Functions used in modeling

AUC (Area Under the Curve)

A measure to assess
the predictive power of a feature and
the performance of a model.

Takes values between 0.5 and 1.



2. Functions used in modeling

2.3. Feature selection

- Correlated feature elimination

```
feature_elimination(
                eliminator='correlated_lower_gini_feature_elimination',
                params={'threshold’:0.90})
```

2. Functions used in modeling

2.4. Binning

`binning()`

Bin (AGE)	Non-event	Event	Event rate	WoE
(-inf, 26.50)	8170	427	5.0%	-0.99
[26.50, 36.50)	26672	682	2.5%	-0.27
[36.50, 49.50)	23341	307	1.3%	0.39
[49.50, inf)	26084	224	0.9%	0.82
Missing	48	0	0.0%	0
Totals	**84315**	**1640**	**1.9%**	

$$\mathrm{WOE} = \ln\left(\frac{\%\text{ of non-events}}{\%\text{ of events}}\right)$$

2. Functions used in modeling

2.5. Scaling

`scaling()`

StandardScaler
scales each feature to have a mean of 0 and a standard deviation of 1.

2. Functions used in modeling

2.5. Scaling

`scaling()`

StandardScaler
scales each feature to have a mean of 0 and a standard deviation of 1.

X	X_SCALED
54	0.97
57	1.21
24	-1.47
45	0.24
33	-0.74
62	1.62
48	0.48
40	-0.17
59	1.37
28	-1.14

2. Functions used in modeling

2.6. Null imputation

`null_imputation()`

imputes missing values

with the mean of the corresponding feature (if numeric)

with 'Null' (if categorical)

2. Functions used in modeling

2.7. Encoding

`encode_categoricals()`

Target Encoding

encodes categorical features as such

- calculate the mean of the target variable for each category
- replace the categorical values with the corresponding mean value

2. Functions used in modeling

2.7. Encoding

`encode_categoricals()`

Target Encoding
encodes categorical features as such
- calculate the mean of the target variable for each category
- replace the categorical values with the corresponding mean value

X	X_ENCODED
Credit Card	0.0157
Credit Card	0.0157
Credit Card	0.0157
Credit Card	0.0157
Loan	0.0087
Loan	0.0087
Loan	0.0087
Credit Card	0.0157
Credit Card	0.0157
Overdraft	0.0053

3. Algorithms used in modeling

3.1. Logistic Regression
3.2. Random Forest
3.3. XGBoost
3.4. LightGBM

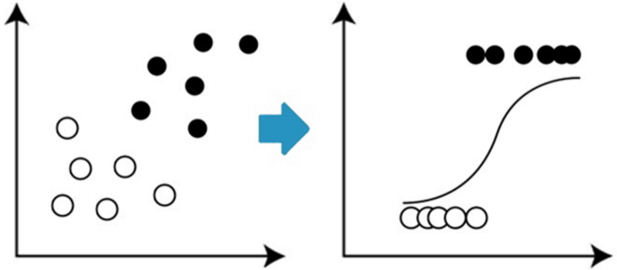
3. Algorithms used in modeling

3.1. Logistic regression

`LogisticRegression()`

Binary classification
Linear
Interpretability
- coefficients
- strength and direction

Regularization
- L1 (lasso)
- L2 (ridge)

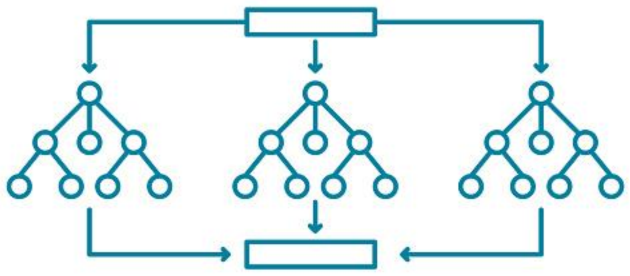


3. Algorithms used in modeling

3.2. Random forest

`RandomForest()`

Linear + nonlinear
Ensemble Learning
Decision Trees
Bootstrap Aggregating (Bagging)
Voting
Feature Importance

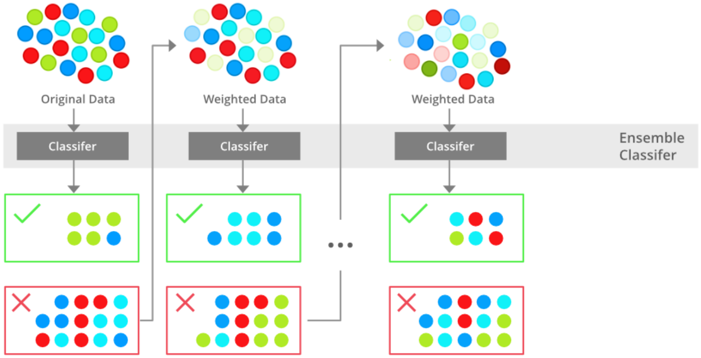


3. Algorithms used in modeling

3.3. XGBoost

`XGBoost()`

Linear + nonlinear
Ensemble Learning
Decision Trees
Gradient Boosting
Depth-wise tree growth
Regularization
Early stopping
Feature Importance

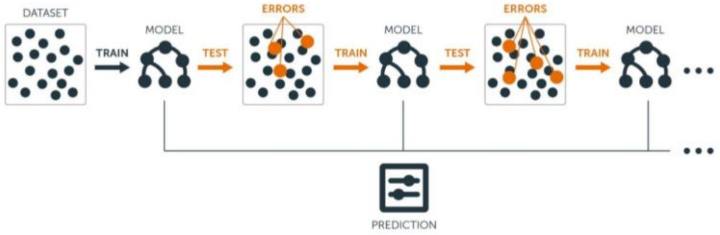


3. Algorithms used in modeling

3.4. LightGBM

LGBM()

Linear + nonlinear
Ensemble Learning
  Decision Trees
Gradient Boosting
  Leaf-wise tree growth
Regularization
Early stopping
Feature Importance



4. Algorithm - model - pipeline concepts

4. Algorithm - model - pipeline concepts

Algorithm = set of rules
(logistic regression algorithm)

Model = algorithm + data + learning = algorithm + learned parameters
(logistic regression model)

Pipeline = data processing and modeling steps
(scorecard pipeline)

5. Train - validation - test concepts

5. Train - validation - test concepts

Training

Train data = used to learn the model parameters

Validation data = used to tune hyperparameters / prevent overfitting

Testing

Test data = used to evaluate the final model's performance

6. Parameter optimization and model selection

6.1. Logistic Regression
6.2. Random Forest
6.3. XGBoost
6.4. LightGBM

6. Parameter optimization and model selection

```
'logistic_regression' : {
        'C': [0.001, …, 1000, log=True],
        'penalty': ['l1', 'l2'],
    }
```

6. Parameter optimization and model selection

```
'logistic_regression' : {
        'C': [0.001, …, 1000, log=True],
        'penalty': ['l1', 'l2'],
   }
```

In Grid Search,
7x2=14 combinations

	C	Penalty
1	0.001	L1
2	0.01	L1
...	...	...
...	...	...
...	...	...
13	100	L2
14	1000	L2

6. Parameter optimization and model selection

```
'random_forest_classifier' : {
        'max_depth': [2, …, 15, step=1],
        'max_features': ['sqrt'],
        'n_estimators': [100]
    }
```

6. Parameter optimization and model selection

```
'xgboost_classifier' : {
        'early_stopping_rounds': [10],
        'eval_metric': ["auc"],
        'n_estimators': [10000],
        'max_depth': [2, …, 9, step=1],
        'lambda': [1e-4, …, 10.0, log=True],
        'alpha': [1e-4, …, 10.0, log=True],
    }
```

6. Parameter optimization and model selection

```
'lightgbm_classifier' : {
        'early_stopping_round': [10],
        'n_estimators': [10000],
        'max_depth': [2, …, 25, step=1],
        'colsample_bytree': [0.6, …, 1.0, step=0.05],
        'reg_alpha': [1e-8, …, 10.0, log=True],
        'reg_lambda': [1e-8, …, 10.0, log=True]
   }
```

6. Parameter optimization and model selection

Model selection considerations;

Model complexity

- Overfitting
- Underfitting

Performance Metrics

Interpretability

7. Scoring Transformation

7. Scoring Transformation

Transform default probabilities into credit scores according to specified parameters

```
ref=200
odds_at_ref=100
points_to_double=20

odds = (1/default_prob)-1

credit_score =
 ref + points_to_double * (log(odds) - log(odds_at_ref)) / log(2)
```

7. Scoring Transformation

Transform default probabilities into credit scores according to specified parameters

```
ref=200
odds_at_ref=100
points_to_double=20

odds = (1/default_prob)-1

credit_score =
 ref + points_to_double * (log(odds) - log(odds_at_ref)) / log(2)
```

	A	B	C	D	E	F
1	**DEFAULT_PROB**	**ODDS**	**SCORE**		**ref**	200
2	0.0001	9999	332.9		**odds_at_ref**	100
3	0.0010	999	266.4		**points_to_double**	20
4	0.0100	99	199.7			
5	0.1000	9	130.5			
6	0.9900	0.01	-65.5			

8. Stacking

8. Stacking

1 Base Model

Logistic Regression

2 Split Data

Best Score → Worst Score

Good Segment | Not Good Segment

3 Segment Models

Good Segment: Logistic Regression, Random Forest, XGBoost, LightGBM

Not Good Segment: Logistic Regression, Random Forest, XGBoost, LightGBM

4 Combine Scores

- Base model
- Good segment model
- Not good segment model

5 Final Model

Logistic Regression

9. Outputs (tables and graphs)

9.1. Saved model (pipeline) pickle files
9.2. Eliminated features
9.3. Binning table
9.4. Gini table
9.5. Feature importance table
9.6. Feature importance graph
9.7. Probability and scores table

9. Outputs (tables and graphs)

9.1. Saved model (pipeline) pickle files

```
Models/
-base_model.pkl
-good_segment_model.pkl
-not_good_segment_model.pkl
-meta_model.pkl
```

9. Outputs (tables and graphs)

9.2. Eliminated features

```
Output/
-null_feature_elimination.xlsx
-constant_feature_elimination.xlsx
-low_gini_feature_elimination.xlsx
-correlated_feature_elimination.xlsx
```

	A	B
1		**NULL ELIMINATED FEATURES**
2	**0**	FEATURE_1324
3	**1**	FEATURE_1819
4	**2**	FEATURE_2130
5	**3**	FEATURE_2581
6	**4**	FEATURE_1718
7	**5**	FEATURE_1063
8	**6**	FEATURE_1993
9	**7**	FEATURE_1891
10	**8**	FEATURE_1771
11	**9**	FEATURE_614
12	**10**	FEATURE_297

9. Outputs (tables and graphs)

9.3. Binning table

Output/
-binning_table.xlsx

	A	B	C	D	E	F	G	H	I	J	K
1			Bin	Count	Count (%)	Non-event	Event	Event rate	WoE	IV	JS
2	AGE	0	(-inf, 22.50)	1014	1.2%	936	78	7.7%	-1.45496	0.05305	0.00610
3		1	[22.50, 23.50)	1103	1.3%	1040	63	5.7%	-1.13602	0.02963	0.00352
4		2	[23.50, 24.50)	1758	2.0%	1675	83	4.7%	-0.93514	0.02875	0.00347
5		3	[24.50, 25.50)	2160	2.5%	2063	97	4.5%	-0.88266	0.03061	0.00371
6		4	[25.50, 26.50)	2562	3.0%	2456	106	4.1%	-0.79701	0.02830	0.00345
7		5	[26.50, 28.50)	5655	6.6%	5463	192	3.4%	-0.59161	0.03093	0.00381
8		6	[28.50, 29.50)	2952	3.4%	2863	89	3.0%	-0.46887	0.00952	0.00118
9		7	[29.50, 30.50)	2923	3.4%	2857	66	2.3%	-0.17199	0.00109	0.00014
10		8	[30.50, 31.50)	2866	3.3%	2804	62	2.2%	-0.12820	0.00058	0.00007
11		9	[31.50, 35.50)	10406	12.1%	10185	221	2.1%	-0.10935	0.00153	0.00019
12		10	[35.50, 36.50)	2552	3.0%	2500	52	2.0%	-0.06706	0.00014	0.00002
13		11	[36.50, 37.50)	2304	2.7%	2263	41	1.8%	0.07101	0.00013	0.00002
14		12	[37.50, 42.50)	10277	12.0%	10132	145	1.4%	0.30686	0.00974	0.00121
15		13	[42.50, 43.50)	1684	2.0%	1662	22	1.3%	0.38487	0.00242	0.00030
16		14	[43.50, 44.50)	1660	1.9%	1642	18	1.1%	0.57343	0.00487	0.00060
17		15	[44.50, 45.50)	1592	1.9%	1575	17	1.1%	0.58893	0.00490	0.00060
18		16	[45.50, 48.50)	4592	5.3%	4544	48	1.0%	0.61050	0.01503	0.00185
19		17	[48.50, 49.50)	1539	1.8%	1523	16	1.0%	0.61599	0.00512	0.00063
20		18	[49.50, 50.50)	1689	2.0%	1673	16	0.9%	0.70992	0.00716	0.00088
21		19	[50.50, 52.50)	3415	4.0%	3384	31	0.9%	0.75296	0.01599	0.00195
22		20	[52.50, 58.50)	11205	13.0%	11107	98	0.9%	0.79050	0.05690	0.00693
23		21	[58.50, inf)	9999	11.6%	9920	79	0.8%	0.89300	0.06205	0.00751
24		22	Special	0	0.0%	0	0	0.0%	0	0	0
25		23	Missing	48	0.1%	48	0	0.0%	0	0	0
26		Totals		85955	100.0%	84315	1640	1.9%		0.39844	0.04813

9. Outputs (tables and graphs)

9.4. Gini table

```
Output/
-gini_table.xlsx
```

	A	B	C
1		**VARIABLE**	**GINI_SCORE**
2	**0**	FEATURE_1486	0.349
3	**1**	FEATURE_2108	0.173
4	**2**	FEATURE_2821	0.146
5	**3**	FEATURE_1062	0.094
6	**4**	FEATURE_717	0.086
7	**5**	FEATURE_2188	0.127
8	**6**	FEATURE_916	0.119
9	**7**	FEATURE_1206	0.138
10	**8**	FEATURE_113	0.113
11	**9**	FEATURE_1347	0.066
12	**10**	FEATURE_1976	0.127

9. Outputs (tables and graphs)

9.5. Feature importance table

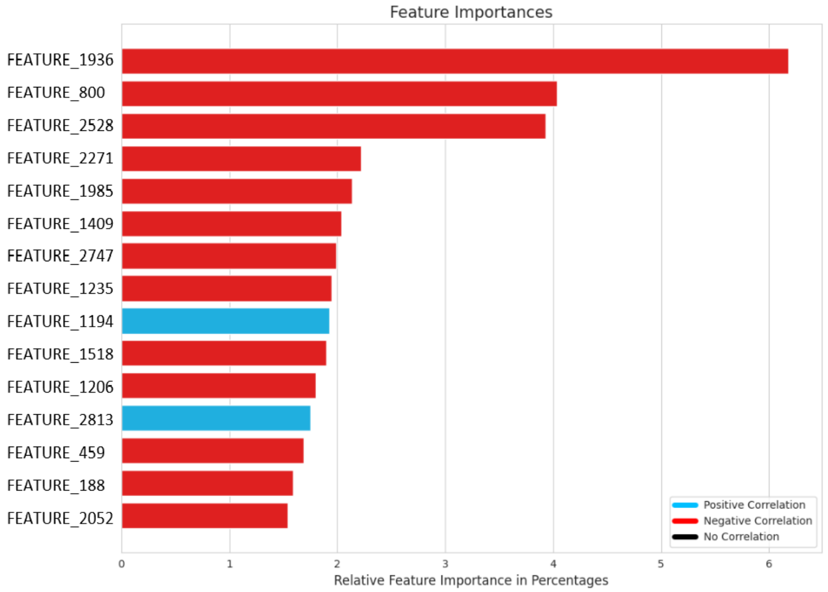
```
Output/
-feature_importance_table.xlsx
```

	A	B	C
1	**variable**	**feature_importance**	**correlation**
2	**FEATURE_1976**	6.18	-1
3	**FEATURE_1956**	4.04	-1
4	**FEATURE_1829**	3.93	-1
5	**FEATURE_788**	2.22	1
6	**FEATURE_237**	2.14	-1
7	**FEATURE_2007**	2.04	-1
8	**FEATURE_192**	1.99	-1
9	**FEATURE_738**	1.95	-1
10	**FEATURE_1198**	1.93	1
11	**FEATURE_2528**	1.90	-1
12	**FEATURE_1069**	1.80	-1

9. Outputs (tables and graphs)

9.6. Feature importance graph

```
Figures/
-feature_importance.png
```



9. Outputs (tables and graphs)

9.7. Probability and scores table

```
Output/
-scores.xlsx
```

	A	B	C	D	E	F
1	**ID**	**BASIC_PROBA**	**BASIC_SCORE**	**BOOSTED_PROBA**	**BOOSTED_SCORE**	**TARGET**
2	**APPL_331**	1.19%	194.6	1.08%	197.6	0
3	**APPL_387**	0.47%	221.4	0.92%	202.1	0
4	**APPL_1517**	3.36%	164.0	3.38%	163.9	0
5	**APPL_1288**	0.54%	217.7	1.04%	198.6	0
6	**APPL_1290**	0.32%	232.5	0.95%	201.2	0
7	**APPL_1908**	0.55%	217.2	1.00%	199.6	0
8	**APPL_1979**	1.43%	189.3	1.37%	190.5	0
9	**APPL_964**	0.39%	227.2	0.83%	205.1	0
10	**APPL_1131**	1.26%	193.0	1.11%	196.7	0
11	**APPL_480**	4.32%	156.5	3.05%	166.9	0
12	**APPL_1498**	2.05%	178.7	1.24%	193.4	0

10. Calibration

10. Calibration

Examples of calibration methods;

Isotonic calibration
Sigmoid calibration

11. Through the door analysis

11. Through the door analysis

The aim is to assess a new loan policy (e.g., a new scorecard model) in terms of;

the potential decrease in NPL rate and

the potential increase in the number of approved loans.

11. Through the door analysis

Score Band	Application Count	Application Count %	Approve Count	Approve Count %	
>200	10,506	23%	4,356	32%	=D2/$D$10
190-200	12,378	27%	4,294	32%	
180-190	7,438	16%	2,079	15%	
170-180	4,727	10%	1,070	8%	
160-170	3,055	7%	631	5%	
150-160	2,053	5%	380	3%	
140-150	1,334	3%	225	2%	
<140	3,844	8%	447	3%	
Grand Total	**45,335**	**100%**	**13,482**	**100%**	

*Applications: 202301 - 202304

11. Through the door analysis

	A	B	C	D	E	F	G
1	**Score Band**	**Application Count**	**Application Count %**	**Approve Count**	**Approve Count %**	**Approve Rate by Score Band**	
2	>200	10,506	23%	4,356	32%	41%	=D2/B2
3	190-200	12,378	27%	4,294	32%	35%	
4	180-190	7,438	16%	2,079	15%	28%	
5	170-180	4,727	10%	1,070	8%	23%	
6	160-170	3,055	7%	631	5%	21%	
7	150-160	2,053	5%	380	3%	19%	
8	140-150	1,334	3%	225	2%	17%	
9	<140	3,844	8%	447	3%	12%	
10	**Grand Total**	**45,335**	**100%**	**13,482**	**100%**	**30%**	

11. Through the door analysis

Score Band	Application Count	Application Count %	Approve Count	Approve Count %	Approve Rate by Score Band	12M NPL Rate Qnbwise Model Expectation
>200	10,506	23%	4,356	32%	41%	0.25%
190-200	12,378	27%	4,294	32%	35%	0.60%
180-190	7,438	16%	2,079	15%	28%	1.72%
170-180	4,727	10%	1,070	8%	23%	1.63%
160-170	3,055	7%	631	5%	21%	3.00%
150-160	2,053	5%	380	3%	19%	4.58%
140-150	1,334	3%	225	2%	17%	8.97%
<140	3,844	8%	447	3%	12%	9.13%
Grand Total	**45,335**	**100%**	**13,482**	**100%**	**30%**	

*Expected NPL rates obtained on the test data

11. Through the door analysis

Score Band	Application Count	Application Count %	Approve Count	Approve Count %	Approve Rate by Score Band	12M NPL Rate Qnbwise Model Expectation	Cumulative 12M NPL Rate Qnbwise Model Expectation	
>200	10,506	23%	4,356	32%	41%	0.25%	0.25%	=SUMPRODUCT(G$2:$G2;$D$2:D2)/SUM(D$2:$D2)
190-200	12,378	27%	4,294	32%	35%	0.60%	0.42%	
180-190	7,438	16%	2,079	15%	28%	1.72%	0.67%	
170-180	4,727	10%	1,070	8%	23%	1.63%	0.76%	
160-170	3,055	7%	631	5%	21%	3.00%	0.87%	
150-160	2,053	5%	380	3%	19%	4.58%	0.98%	
140-150	1,334	3%	225	2%	17%	8.97%	1.12%	
<140	3,844	8%	447	3%	12%	9.13%	1.39%	
Grand Total	**45,335**	**100%**	**13,482**	**100%**	**30%**			

11. Through the door analysis

	A	B	C	D	E	F	G	H	I	J
1	Score Band	Application Count	Application Count %	Approve Count	Approve Count %	Approve Rate by Score Band	12M NPL Rate Qnbwise Model Expectation	Cumulative 12M NPL Rate Qnbwise Model Expectation	Potential Count	
2	>200	10,506	23%	4,356	32%	41%	0.25%	0.25%	6,150	=B2-D2
3	190-200	12,378	27%	4,294	32%	35%	0.60%	0.42%	8,084	
4	180-190	7,438	16%	2,079	15%	28%	1.72%	0.67%	5,359	
5	170-180	4,727	10%	1,070	8%	23%	1.63%	0.76%	3,657	
6	160-170	3,055	7%	631	5%	21%	3.00%	0.87%	2,424	
7	150-160	2,053	5%	380	3%	19%	4.58%	0.98%	1,673	
8	140-150	1,334	3%	225	2%	17%	8.97%	1.12%	1,109	
9	<140	3,844	8%	447	3%	12%	9.13%	1.39%	3,397	
10	**Grand Total**	**45,335**	**100%**	**13,482**	**100%**	**30%**				

11. Through the door analysis

	A	B	C	D	E	F	G	H	I	J	K
1	Score Band	Application Count	Application Count %	Approve Count	Approve Count %	Approve Rate by Score Band	12M NPL Rate Qnbwise Model Expectation	Cumulative 12M NPL Rate Qnbwise Model Expectation	Potential Count	Adjusted Approve Count	
2	>200	10,506	23%	4,356	32%	41%	0.25%	0.25%	6,150	7,431	=D2 + I2/2
3	190-200	12,378	27%	4,294	32%	35%	0.60%	0.42%	8,084	8,336	
4	180-190	7,438	16%	2,079	15%	28%	1.72%	0.67%	5,359	4,759	
5	170-180	4,727	10%	1,070	8%	23%	1.63%	0.76%	3,657	2,899	
6	160-170	3,055	7%	631	5%	21%	3.00%	0.87%	2,424	1,843	
7	150-160	2,053	5%	380	3%	19%	4.58%	0.98%	1,673	1,217	
8	140-150	1,334	3%	225	2%	17%	8.97%	1.12%	1,109	780	
9	<140	3,844	8%	447	3%	12%	9.13%	1.39%	3,397	2,146	
10	**Grand Total**	**45,335**	**100%**	**13,482**	**100%**	**30%**					

*Potential Assessment procedure shows that
around 50% of the applications were rejected due to non-regulatory reasons.
So, they could potentially be approved depending on the risk appetite.

11. Through the door analysis

Score Band	Application Count	Application Count %	Approve Count	Approve Count %	Approve Rate by Score Band	12M NPL Rate Qnbwise Model Expectation	Cumulative 12M NPL Rate Qnbwise Model Expectation	Potential Count	Adjusted Approve Count	
>200	10,506	23%	4,356	32%	41%	0.25%	0.25%	6,150	7,431	=D2 + I2/2
190-200	12,378	27%	4,294	32%	35%	0.60%	0.42%	8,084	8,336	
180-190	7,438	16%	2,079	15%	28%	1.72%	0.67%	5,359	4,759	
170-180	4,727	10%	1,070	8%	23%	1.63%	0.76%	3,657	2,899	
160-170	3,055	7%	631	5%	21%	3.00%	0.87%	2,424	1,843	
150-160	2,053	5%	380	3%	19%	4.58%	0.98%	1,673	1,217	
140-150	1,334	3%	225	2%	17%	8.97%	1.12%	1,109	780	
<140	3,844	8%	447	3%	12%	9.13%	1.39%	3,397	2,146	
Grand Total	**45,335**	**100%**	**13,482**	**100%**	**30%**					

Filter: No Loan

	BGN	Other	Delinquent	Loan (Approved)	Grand Total	Potential %	
Count of ID	328	1161	626	184	2299	51%	=C3/F3

Filter: No Loan

Count of ID

Row Labels	BGN	Other	Delinquent	Loan (Approved)	Grand Total	Potential %
>200	98	318	86	60	562	57%
180-200	161	561	263	79	1064	53%
160-180	48	179	148	28	403	44%
140-160	14	60	62	11	147	41%
120-140	5	18	30	3	56	32%
<120	2	25	37	3	67	37%
Grand Total	**328**	**1161**	**626**	**184**	**2299**	**51%**

*Potential Assessment procedure shows that around 50% of the applications were rejected due to non-regulatory reasons. So, they could potentially be approved depending on the risk appetite.

11. Through the door analysis

Score Band	Application Count	Application Count %	Approve Count	Approve Count %	Approve Rate by Score Band	12M NPL Rate Qnbwise Model Expectation	Cumulative 12M NPL Rate Qnbwise Model Expectation	Potential Count	Adjusted Approve Count	Cumulative Adjusted Approve Count	
>200	10,506	23%	4,356	32%	41%	0.25%	0.25%	6,150	7,431	7,431	
190-200	12,378	27%	4,294	32%	35%	0.60%	0.42%	8,084	8,336	15,767	
180-190	7,438	16%	2,079	15%	28%	1.72%	0.67%	5,359	4,759	20,526	
170-180	4,727	10%	1,070	8%	23%	1.63%	0.76%	3,657	2,899	23,424	
160-170	3,055	7%	631	5%	21%	3.00%	0.87%	2,424	1,843	25,267	
150-160	2,053	5%	380	3%	19%	4.58%	0.98%	1,673	1,217	26,484	
140-150	1,334	3%	225	2%	17%	8.97%	1.12%	1,109	780	27,263	
<140	3,844	8%	447	3%	12%	9.13%	1.39%	3,397	2,146	29,409	=SUM(J$2:$J9)
Grand Total	**45,335**	**100%**	**13,482**	**100%**	**30%**						

11. Through the door analysis

Score Band	Application Count	Application Count %	Approve Count	Approve Count %	Approve Rate by Score Band	12M NPL Rate Qnbwise Model Expectation	Cumulative 12M NPL Rate Qnbwise Model Expectation	Potential Count	Adjusted Approve Count	Cumulative Adjusted Approve Count	Cumulative Adjusted 12M NPL Rate Qnbwise Model Expectation	
>200	10,506	23%	4,356	32%	41%	0.25%	0.25%	6,150	7,431	7,431	0.25%	
190-200	12,378	27%	4,294	32%	35%	0.60%	0.42%	8,084	8,336	15,767	0.43%	
180-190	7,438	16%	2,079	15%	28%	1.72%	0.67%	5,359	4,759	20,526	0.73%	
170-180	4,727	10%	1,070	8%	23%	1.63%	0.76%	3,657	2,899	23,424	0.84%	
160-170	3,055	7%	631	5%	21%	3.00%	0.87%	2,424	1,843	25,267	1.00%	
150-160	2,053	5%	380	3%	19%	4.58%	0.98%	1,673	1,217	26,484	1.16%	
140-150	1,334	3%	225	2%	17%	8.97%	1.12%	1,109	780	27,263	1.39%	
<140	3,844	8%	447	3%	12%	9.13%	1.39%	3,397	2,146	29,409	1.95%	=SUMPRODUCT(G$2:$G9;$J$2:J9)
Grand Total	**45,335**	**100%**	**13,482**	**100%**	**30%**							/SUM(J$2:$J9)

11. Through the door analysis

	A	B	C	D	E	F	G	H	I	J	K	L
1	**Score Band**	**Application Count**	**Application Count %**	**Approve Count**	**Approve Count %**	**Approve Rate by Score Band**	**12M NPL Rate Qnbwise Model Expectation**	**Cumulative 12M NPL Rate Qnbwise Model Expectation**	**Potential Count**	**Adjusted Approve Count**	**Cumulative Adjusted Approve Count**	**Cumulative Adjusted 12M NPL Rate Qnbwise Model Expectation**
2	>200	10,506	23%	4,356	32%	41%	0.25%	0.25%	6,150	7,431	7,431	0.25%
3	190-200	12,378	27%	4,294	32%	35%	0.60%	0.42%	8,084	8,336	15,767	0.43%
4	180-190	7,438	16%	2,079	15%	28%	1.72%	0.67%	5,359	4,759	20,526	0.73%
5	170-180	4,727	10%	1,070	8%	23%	1.63%	0.76%	3,657	2,899	23,424	0.84%
6	160-170	3,055	7%	631	5%	21%	3.00%	0.87%	2,424	1,843	25,267	1.00%
7	150-160	2,053	5%	380	3%	19%	4.58%	0.98%	1,673	1,217	26,484	1.16%
8	140-150	1,334	3%	225	2%	17%	8.97%	1.12%	1,109	780	27,263	1.39%
9	<140	3,844	8%	447	3%	12%	9.13%	1.39%	3,397	2,146	29,409	1.95%
10	**Grand Total**	**45,335**	**100%**	**13,482**	**100%**	**30%**						

*With the total number of approvals is constant (13.4K),
the expected NPL rate decreases from 1.39% to less than 0.43%.

11. Through the door analysis

	A	B	C	D	E	F	G	H	I	J	K	L
1	**Score Band**	**Application Count**	**Application Count %**	**Approve Count**	**Approve Count %**	**Approve Rate by Score Band**	**12M NPL Rate Qnbwise Model Expectation**	**Cumulative 12M NPL Rate Qnbwise Model Expectation**	**Potential Count**	**Adjusted Approve Count**	**Cumulative Adjusted Approve Count**	**Cumulative Adjusted 12M NPL Rate Qnbwise Model Expectation**
2	>200	10,506	23%	4,356	32%	41%	0.25%	0.25%	6,150	7,431	7,431	0.25%
3	190-200	12,378	27%	4,294	32%	35%	0.60%	0.42%	8,084	8,336	15,767	0.43%
4	180-190	7,438	16%	2,079	15%	28%	1.72%	0.67%	5,359	4,759	20,526	0.73%
5	170-180	4,727	10%	1,070	8%	23%	1.63%	0.76%	3,657	2,899	23,424	0.84%
6	160-170	3,055	7%	631	5%	21%	3.00%	0.87%	2,424	1,843	25,267	1.00%
7	150-160	2,053	5%	380	3%	19%	4.58%	0.98%	1,673	1,217	26,484	1.16%
8	140-150	1,334	3%	225	2%	17%	8.97%	1.12%	1,109	780	27,263	1.39%
9	<140	3,844	8%	447	3%	12%	9.13%	1.39%	3,397	2,146	29,409	1.95%
10	**Grand Total**	**45,335**	**100%**	**13,482**	**100%**	**30%**						

*With the total number of approvals is constant (13.4K), the expected NPL rate decreases from 1.39% to less than 0.43%.

*With the NPL rate constant (1.39%), the total number of approvals increases (doubles) from 13.4K to around 27K.

12. NPL profit-loss scenario analysis

12. NPL profit-loss scenario analysis

The aim is to calculate

- the average profit or loss per loan, and
- the break-even NPL rate (NPL tolerance).

12. NPL profit-loss scenario analysis

The aim is to calculate

the average profit or loss per loan, and

the break-even NPL rate (NPL tolerance).

Taking account of

loan info,

loan commission,

demand deposit,

funding cost,

insurance,

early closure,

NPL Info.

12. NPL profit-loss scenario analysis

		Overall
Loan Info	*Ticket Size*	10,000TL
	Interest Rate Simple %	2.00%
	# of Installments	12
	Installment Amount	946TL
Loan Commission	*File Fee*	0.5%
	Penetration	100%
Demand Deposit	*Demand Deposit Monthly Pool*	35.0%
	Demand Deposit Duration (Day)	3
	Demand Deposit Income	29TL
Funding Cost	*Funding Cost %*	1.00%
Insurance	*Penetration*	80.0%
	Premium Rate	3.0%
	Share	40%
Early Closure	*Early Closure Rate %*	50.0%
	Early Closure Avg. Inst. Paid	9
NPL Info	*Avg. Inst. Paid*	5
	Avg. NPL Month	9
	NPL Amt	6,120TL
	PV of NPL	5,596TL
	NPV of Recovery	60.0%
	Interest Income of Default (NPV)	407TL
	Loss without OPEX	1,832TL
Result	*Net Interest Income*	643TL
	Loan Commission Income	50TL
	Insurance Income	96TL
	Early Closure Interest Loss	25TL
	Total Demand Deposit Income	29TL
	Opex	0TL
	Net Income	**793TL**
	NPL Loss (/w Opex)	**1,832TL**
NPL Tolerance	***NPL Tolerance***	**30.22%**
PD	*Origination PD (24M) - limit pd*	1.50%
VNB		754TL

13. Modeling with WiseMiner