

CBU CODING CHALLENGE

Muhammadjafar Zokirov, Javoxir Mamatayev,
Ibrohim Ahmadjonov, Javohir Abdugafforov,
Sardor Nigmatov

Data Cleaning

Hamma sonli qator ma'lumotlari float formatga parse qilindi.

Quyidagi columnlar olib tashlandi:

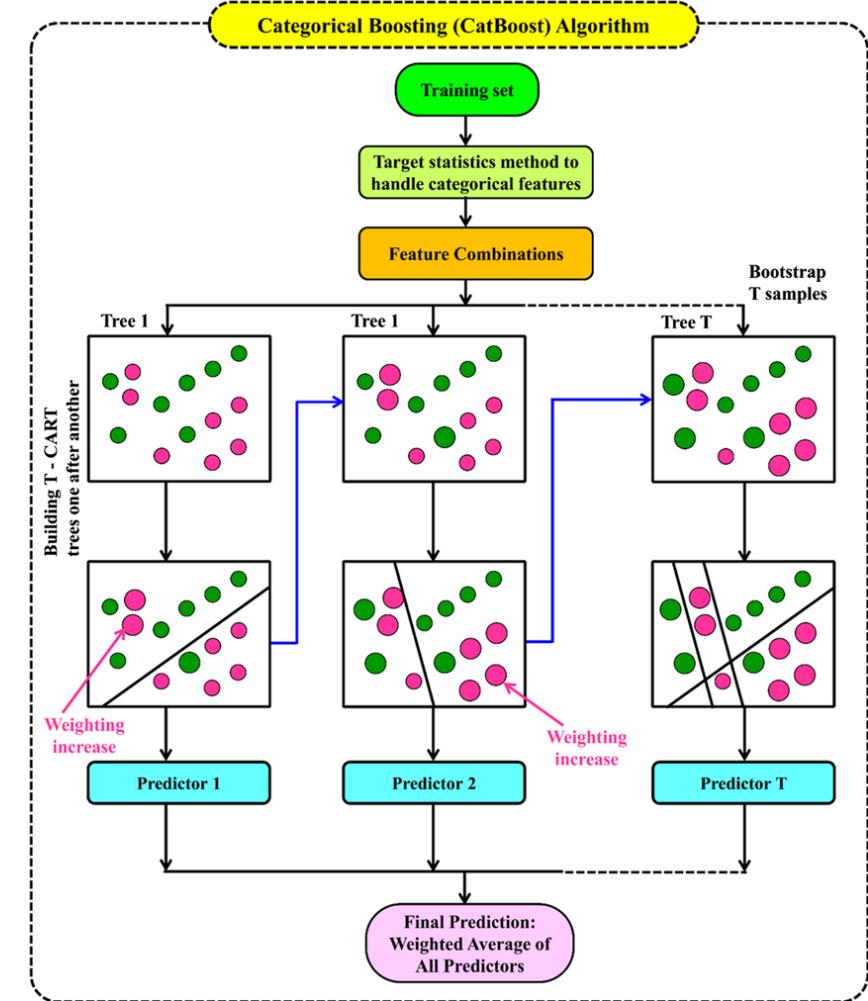
cust_id, customer_number,
cust_num, previous_zip_code, state,
marketing_campaign, loan_officer_id,
num_login_sessions,
application_hour,
total_monthly_debt_payment, credit_usage_amount, num_credit_accounts,
oldest_account_age_months, cost_of_living_index, total_monthly_debt_payment,

Dastlab Random Forest modeli tanlandi.

- ❑ One-hot encoding → feature sonini juda ko‘paytiradi
- ❑ Target encoding → leakage xavfi
- ❑ LabelEncoder → ordinal yaroqsiz

So‘ng yechim uchun CatBoost tanlandi.

- ❑ One-hot / Label encoding talab qilmaydi
- ❑ Overfittingga bardoshli
- ❑ Feature importance ishonchli



Optimization completed!

Best ROC-AUC: 0.8006

Best Parameters:

```
iterations          : 688
learning_rate       : 0.07671067229712569
depth              : 4
l2_leaf_reg         : 1.8585411871647346
min_data_in_leaf    : 26
border_count        : 197
bagging_temperature : 0.17150201140489996
random_strength     : 8.738456430998188
class_1_weight      : 18
```

STEP 6: Cross-Validation Training

```
=====
```

STEP 6: Cross-Validation Training

```
=====
```

```
Fold 1: ROC-AUC = 0.7644
Fold 2: ROC-AUC = 0.8008
Fold 3: ROC-AUC = 0.7863
Fold 4: ROC-AUC = 0.7839
Fold 5: ROC-AUC = 0.8079
```

...

```
account_diversity_index   3.683759
regional_median_rent      3.558911
total_monthly_debt_payment 3.339052
regional_unemployment_rate 3.024365
```

✓ Optimization completed!
Best ROC-AUC: 0.8006

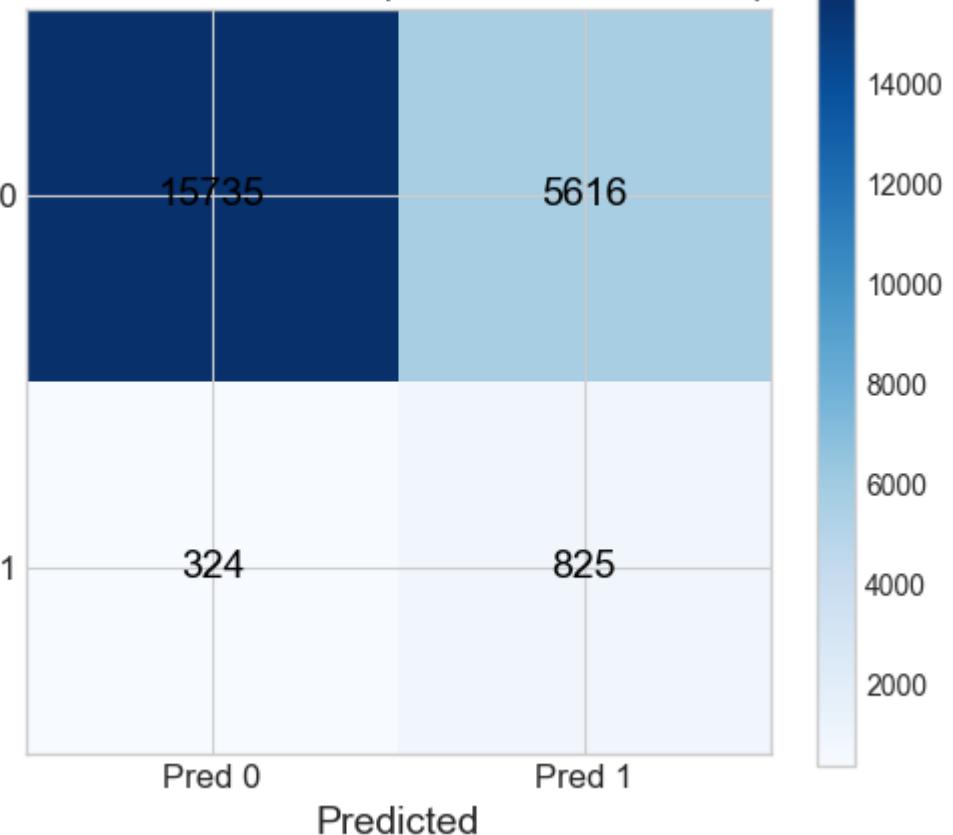
Best Parameters:
iterations : 688
learning_rate : 0.07671067229712569
depth : 4
l2_leaf_reg : 1.8585411871647346
min_data_in_leaf : 26
border_count : 197
bagging_temperature : 0.17150201140489996
random_strength : 8.738456430998188
class_1_weight : 18

STEP 6: Cross-Validation Training

STEP 6: Cross-Validation Training

=====
Fold 1: ROC-AUC = 0.7644
Fold 2: ROC-AUC = 0.8008
Fold 3: ROC-AUC = 0.7863
Fold 4: ROC-AUC = 0.7839
Fold 5: ROC-AUC = 0.8079

Confusion Matrix (Threshold = 0.490)



Top 15 Feature Importances

