
FRAUD/ FINANCIAL CRIME PREVENTION SOLUTIONS FOR BANKS(CREDIT CARD) WITH REINFORCEMENT LEARNING

— Suraj Kashozhala —
Zahra Qureshi

In my use case, I need to resolve an entity (personal or corporate name) that has encountered credit card transaction .

Since the KYC data cannot be obtained, there is not much data (account number, customer number, etc.) and it depends on the name and address of the party.

Goal is to find accuracy, f1 score, precision, sensitivity, specificity . for the following modeles

- *K-NN

- *Decision tree

- *SVM

Later we are going to use reinforcement learning to improve the model and improve it.

Solution Design

ML model used (K-NN, SVM, Decision tree, DQN)

Metrics – Confusion matrix, Mean reward per thousand episodes

Data(Train) – 85%

Data(Test) - 15%

- **Data download**
 - **Data Analysis**
 - **Pre-processing**
 - **Model training**
 - **Validation**
-

Implementation

***Data :**

<https://raw.githubusercontent.com/MLWave/Black-Boxxy/master/credit-card-default.csv>

***Data Analysis :** Check Null values & correlation.

***Pre-processing :** Train/Test splits , use robust scaler for standardization.

***Model training :** K-NN, SVM, Decision tree, DQN.

***Validation :** Confusion matrix (accuracy, f1 score, precision, sensitivity, specificity), Mean reward per thousand episodes (RL).

Pre-processing : Correlation

limit_bal	1.00	0.14	-0.16	-0.20	-0.19	-0.19	-0.17	-0.17	0.29	0.28	0.28	0.29	0.30	0.29	0.20	0.18	0.21	0.20	0.22	0.22	-0.15	0.26	-0.15	-0.14	-0.02	0.10
age	0.14	1.00	-0.00	-0.01	-0.02	-0.01	-0.02	-0.02	0.06	0.05	0.05	0.05	0.05	0.05	0.03	0.02	0.03	0.02	0.02	0.02	0.01	-0.10	-0.08	0.23	0.09	0.45
pay_1	-0.16	-0.00	1.00	0.67	0.43	0.37	0.35	0.31	-0.06	-0.05	-0.04	-0.03	-0.02	-0.02	-0.09	-0.06	-0.07	-0.07	-0.07	-0.06	0.37	-0.05	0.02	0.04	0.02	0.02
pay_2	-0.20	0.01	0.67	1.00	0.63	0.48	0.44	0.40	0.01	0.01	0.02	0.03	0.04	0.04	-0.10	-0.05	-0.06	-0.05	-0.05	-0.04	0.34	-0.08	0.06	0.04	0.04	0.01
pay_3	-0.19	-0.02	0.43	0.63	1.00	0.63	0.48	0.43	-0.02	-0.00	0.00	0.02	0.03	0.03	-0.04	-0.08	-0.06	-0.05	-0.05	-0.05	0.29	-0.07	0.05	0.04	0.04	0.00
pay_4	-0.19	-0.01	0.37	0.48	0.63	1.00	0.66	0.50	-0.02	-0.01	0.01	0.02	0.04	0.04	-0.05	-0.04	-0.08	-0.05	-0.05	-0.05	0.28	-0.07	0.05	0.04	0.04	0.01
pay_5	-0.17	-0.02	0.35	0.44	0.48	0.66	1.00	0.66	-0.01	0.00	0.01	0.04	0.05	0.06	-0.05	-0.04	-0.04	-0.07	-0.05	-0.05	0.27	-0.05	0.04	0.03	0.03	0.00
pay_6	-0.17	-0.02	0.31	0.40	0.43	0.50	0.66	1.00	-0.01	0.00	0.01	0.03	0.06	0.06	-0.05	-0.04	-0.04	-0.03	-0.07	-0.04	0.25	-0.04	0.04	0.01	0.03	-0.00
bill_amt1	0.29	0.06	-0.06	0.01	-0.02	-0.02	-0.01	-0.01	1.00	0.95	0.89	0.86	0.83	0.80	0.14	0.10	0.16	0.16	0.17	0.18	-0.02	-0.02	0.03	-0.02	0.03	0.03
bill_amt2	0.28	0.05	-0.05	0.01	-0.00	-0.01	0.00	0.00	0.95	1.00	0.93	0.89	0.86	0.83	0.28	0.10	0.15	0.15	0.16	0.17	-0.01	-0.02	0.03	-0.02	0.03	0.02
bill_amt3	0.28	0.05	-0.04	0.02	0.00	0.01	0.01	0.01	0.89	0.93	1.00	0.92	0.88	0.85	0.24	0.32	0.13	0.14	0.18	0.18	-0.01	-0.01	0.02	-0.02	0.02	0.03
bill_amt4	0.29	0.05	-0.03	0.03	0.02	0.02	0.04	0.03	0.86	0.89	0.92	1.00	0.94	0.90	0.23	0.21	0.30	0.13	0.16	0.18	-0.01	-0.00	0.02	-0.03	0.02	0.02
bill_amt5	0.30	0.05	-0.02	0.04	0.03	0.04	0.05	0.06	0.83	0.86	0.88	0.94	1.00	0.95	0.22	0.18	0.25	0.29	0.14	0.16	-0.01	0.00	0.02	-0.03	0.02	0.03
bill_amt6	0.29	0.05	-0.02	0.04	0.03	0.04	0.06	0.06	0.80	0.83	0.85	0.90	0.95	1.00	0.20	0.17	0.23	0.25	0.31	0.12	-0.01	-0.00	0.02	-0.03	0.02	0.02
pay_amt1	1.00	0.03	-0.09	-0.10	-0.04	-0.05	-0.05	-0.05	0.14	0.28	0.24	0.23	0.22	0.20	1.00	0.29	0.25	0.20	0.15	0.19	-0.07	0.05	-0.03	-0.02	0.00	0.01
pay_amt2	0.18	0.02	-0.06	-0.05	-0.08	-0.04	-0.04	-0.04	0.10	0.10	0.32	0.21	0.18	0.17	0.29	1.00	0.24	0.18	0.18	0.16	-0.06	0.04	-0.03	-0.02	0.00	0.01
pay_amt3	0.21	0.03	-0.07	-0.06	-0.06	-0.08	-0.04	-0.04	0.16	0.15	0.13	0.30	0.25	0.23	0.25	0.24	1.00	0.22	0.16	0.16	-0.06	0.06	-0.04	-0.03	0.01	0.01
pay_amt4	0.20	0.02	-0.07	-0.05	-0.05	-0.05	-0.07	-0.03	0.16	0.15	0.14	0.13	0.29	0.25	0.20	0.18	0.22	1.00	0.15	0.16	-0.06	0.05	-0.03	-0.02	0.00	0.01
pay_amt5	0.22	0.02	-0.07	-0.05	-0.05	-0.05	-0.05	-0.07	0.17	0.16	0.18	0.16	0.14	0.31	0.15	0.18	0.16	0.15	1.00	0.15	-0.06	0.05	-0.02	-0.03	0.00	0.00
pay_amt6	0.22	0.02	-0.06	-0.04	-0.05	-0.05	-0.05	-0.04	0.18	0.17	0.18	0.18	0.16	0.12	0.19	0.16	0.16	0.16	0.15	1.00	-0.05	0.05	-0.03	-0.03	0.00	0.01
default	-0.15	0.01	0.37	0.34	0.29	0.28	0.27	0.25	-0.02	-0.01	-0.01	-0.01	-0.01	-0.01	-0.07	-0.06	-0.06	-0.06	-0.06	-0.05	1.00	-0.05	0.04	0.03	0.04	0.03
grad_school	0.26	-0.10	-0.05	-0.08	-0.07	-0.07	-0.05	-0.04	-0.02	-0.02	-0.01	-0.00	0.00	-0.00	0.05	0.04	0.06	0.05	0.05	0.05	-0.05	1.00	-0.69	-0.33	0.02	-0.15
university	-0.15	-0.08	0.02	0.06	0.05	0.05	0.04	0.04	0.03	0.03	0.02	0.02	0.02	0.02	-0.03	-0.03	-0.04	-0.03	-0.02	-0.03	0.04	-0.69	1.00	-0.41	-0.03	0.06
high_school	-0.14	0.23	0.04	0.04	0.04	0.04	0.03	0.01	-0.02	-0.02	-0.02	-0.03	-0.03	-0.03	-0.02	-0.02	-0.03	-0.02	-0.03	-0.03	0.03	-0.33	-0.41	1.00	0.01	0.11
male	-0.02	0.09	0.02	0.04	0.04	0.04	0.03	0.03	0.03	0.03	0.02	0.02	0.02	0.02	0.00	0.00	0.01	0.00	0.00	0.00	0.04	0.02	-0.03	0.01	1.00	-0.03
married	0.10	0.45	0.02	0.01	0.00	0.01	0.00	-0.00	0.03	0.02	0.03	0.02	0.03	0.02	0.01	0.01	0.01	0.01	0.00	0.01	0.03	-0.15	0.06	0.11	-0.03	1.00
limit_bal	age	pay_1	pay_2	pay_3	pay_4	pay_5	pay_6	bill_amt1	bill_amt2	bill_amt3	bill_amt4	bill_amt5	bill_amt6	pay_amt1	pay_amt2	pay_amt3	pay_amt4	pay_amt5	pay_amt6	default	grad_school	university	high_school	male	married	

Q-Learning Model Algorithm

Training data $D = (x_1, p_1), (x_2, p_2), \dots, (x_n, p_n)$

Number Of Samples N

Memory M , Capacity N

Activation Function Q

Random Weights as W

$y(i) = r(i) + \gamma \max_{a'} Q(s(i), a'; W)$

$L(w) = (y(i) - Q(s(i), a(i); W))^2$

THE Q TABLE I IMPLEMENTED:

$Q_table[current_state, action] = (1 - lr) * Q_table[current_state, action] + lr * (reward + \gamma * \max(Q_table[next_state, :]))$

$total_episode_reward = total_episode_reward + reward$

Performance -1

	Accuracy	F-1 Score	Precision	Sensitivity	Specificity
Decision Tree	0.722667	0.395349	0.381665	0.811412	0.41005
SVM	0.795111	0.395349	0.551773	0.909843	0.390955
k-nn	0.781111	0.395335	0.507886	0.910984	0.323618

Performance - 2

Mean reward per thousand episodes

1000 mean episode reward: 0.052

2000 mean episode reward: 0.218

3000 mean episode reward: 0.435

4000 mean episode reward: 0.602

5000 mean episode reward: 0.665

6000 mean episode reward: 0.684

7000 mean episode reward: 0.667

8000 mean episode reward: 0.677

9000 mean episode reward: 0.701

10000 mean episode reward: 0.683

Deliverables

1. We are going to submit our model using jupyter notebook

THANK YOU