

**"NOTE: All data and labels given in this problem are synthesized and not related to any customer or any real person."**

### **Problem 1: Non-Performing Loan Flag Prediction from Credit Card Information**

Each year, non-performing loans cost the economy millions of baht. Predicting whether a customer will default on their credit card loans is a non-trivial task, which if solved, can be highly beneficial to both KBank and KBank's customers. If default could be predicted in advance, loans could be restructured, this would save customers from bankruptcy, and prevent KBank from losing revenue.

Your task is to use the data given to predict if customers defaulted on their credit card loans in a certain period of time.

#### **Provided Files (for both Training and Testing) :**

- ◆ File name: tj\_01\_creditcard\_transaction.csv
  - Credit card transaction
  - Period: 01/01/2016 - 31/08/2016
  - Size: ~ 600,000 rows

Field Name	Data Type	Description
card_no	STRING	Credit Card Number
txn_date	DATETIME	Transaction Date
txn_hour	TIMESTAMP	Transaction Time ( only Hour )
txn_amount	INT	Transaction amount (THB)
mer_cat_code	STRING	VISA Merchant category code
mer_id	STRING	Merchant ID <u>Note</u> 0 is unidentified

**"NOTE: All data and labels given in this problem are synthesized and not related to any customer or any real person."**

- ◆ File name: tj\_01\_creditcard\_card.csv
  - Dummy credit card information
  - Size: 11,875 rows

Field Name	Data Type	Description
card_no	STRING	Dummy credit card number.
bill_cyc	SMALL INTEGER	Bill cycle date.
pos_dt	TIMESTAMP	Date information was updated.
cst_id	STRING	Dummy Customer ID.
open_dt	TIMESTAMP	Card issue date.
exp_dt	STRING	Card expiry date.
cr_lmt_amt	DECIMAL	Credit limit amount.
prev_cr_lmt_amt	DECIMAL	Previous credit limit.

- ◆ File name: tj\_01\_creditcard\_customer.csv
  - Dummy customer information
  - Size: 9,544 rows

Field Name	Data Type	Description
cst_id	STRING	Dummy Customer ID.
incm_amt	INTEGER	income amount per month.
age	INTEGER	Age.
main_zip_cd	STRING	Dummy zip code of home address.
cr_line_amt	DECIMAL	Total available credit amount.
pos_dt	TIMESTAMP	Date information was updated.

**"NOTE: All data and labels given in this problem are synthesized and not related to any customer or any real person."**

### Training Data File:

- ◆ File name: tj\_01\_training.csv
  - Description: dummy credit card number and its corresponding NPL flag
  - Size: 9500 rows

Field Name	Data Type	Description
card_no	STRING	Dummy credit card number.
npl_flag	STRING	Whether this card is subject to NPL condition 0: not NPL 1: subject to NPL

### Testing Data File

- ◆ File name: tj\_01\_test.csv
  - Description: the set of dummy credit card to be predicted whether they are subject to NPL or not.
  - Size: 2375 rows

Field Name	Data Type	Description
card_no	STRING	Dummy credit card number.

### Expected Output

- ◆ File name: l.txt
  - Description: Prediction result whether the corresponding credit card number in the test data file is subject to NPL or not.
  - Each line contains either 0(not NPL) or 1(subject to NPL)
  - Each line must correspond to the credit card number in the same line from the tj\_01\_test.csv data file

Field Name	Data Type	Description
npl_flag	STRING	Whether the corresponding card number is subject to NPL condition 0: not NPL 1: subject to NPL

**"NOTE: All data and labels given in this problem are synthesized and not related to any customer or any real person."**

### Examples

File: tj\_01\_credit\_card\_transaction.csv

```
1234000000009154,2016-12-07 00:00:00.0,22,1550,4511,0
12340000000017165,2016-12-07 00:00:00.0,22,250,6011,0
```

File: tj\_01\_creditcard\_card.csv

```
12340000000023349,20,2017-02-23 00:00:00,1000004674,2010-12-20
00:00:00,1220,70000.00,0.00
12340000000007292,17,2017-02-23 00:00:00,1000010084,2014-06-04
00:00:00,0619,69000.00,60000.00
```

File: tj\_01\_creditcard\_customer.csv

```
1000001115,47676,41,11110,100000.00,2017-02-23 00:00:00
1000004693,27157,52,10170,80000.00,2017-02-23 00:00:00
```

File: tj\_01\_training.csv

```
12340000000011560,0
12340000000022416,1
```

File: tj\_01\_test.csv

```
12340000000024065
12340000000007629
```

Output File: 1.csv

```
1
0
```