

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

Problem 2: Merchant Prediction

In this technological era, more than ever before, online shopping and social media sites are becoming more popular than ever. It is becoming a significant influence to the Thai retail trade. In order to make the interactions and transactions between merchants and buyers easier, people can now open bank accounts to reduce the use of cash during transactions.

Your job is to create a prediction model that finds individuals who are most likely merchants that opened bank accounts based on their credit card and deposit transactions.

Provided Files (for both Training and Testing) :

- ◆ File name: tj_02_account_transaction.csv
 - Description: Deposit transaction from all accounts
 - Period: 01/02/2016 - 30/05/2016
 - Size: 2,000,000 rows

Field Name	Data Type	Description
account_no	STRING	Account Number
form_to_account_no	STRING	From_To-Account Number
txn_amount	DECIMAL(18,2)	Net transaction amount
txn_dt	DATETIME	Transaction Date
txn_hour	INT	Transaction Time (only Hour)
txn_type	STRING	Type of Operation(CR: credit, DR: debit)

- ◆ File name: tj_02_credit_card_transaction.csv
 - Credit card transaction
 - Period: 01/01/2016 - 31/12/2016
 - Size: 500,000 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

- ◆ File name: tj_02_acc_x_card.csv

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

- Account – Credit Card Relation
- Period: 01/01/2016 - 31/12/2016
- Size: 11,500 rows
- Some accounts do not have accompanied by credit card

Field Name	Data Type	Description
account_no	STRING	Account Number
card_no	STRING	Credit Card Number

Training Data File:

- ♦ File name: tj_02_training.csv
 - Description: account no with label
 - Size: 9,321 rows

Field Name	Data Type	Description
account_no	STRING	Account Number
is_merchant	STRING	Is merchant label 0: non-merchant 1: merchant

Testing Data File

- ♦ File name: tj_02_test.csv
 - Description: input **ONLY account number** for submitting
 - Size: 2,331 rows

Field Name	Data Type	Description
account_no	STRING	Account Number

Expected Output

- ♦ File name: 2.txt
 - Description: Prediction result whether the corresponding account in test data file belonged to Merchant account or not
 - Each line contains either 0(not belong to a merchant) or 1(belonged to a merchant)
 - Each line must correspond to the account in the same line from the tj_02_test.csv data file

Field Name	Data Type	Description
------------	-----------	-------------

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

is_merchant	STRING	Is merchant label 0: non-merchant 1: merchant
-------------	--------	---

Examples

File: tj_02_account_transaction.csv

```
3450045224,0,50,2016-04-03,17,DR
3450034734,0,10000,2016-10-07,13,DR
```

File: tj_02_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_02_acc_x_card.csv

```
3450028461,12340000000026194
3450033777,12340000000010556
```

File: tj_02_training.csv

```
3450028461,1
3450033777,0
```

File: tj_02_test.csv

```
4313487888
3450033411
```

Output File: 2.csv

```
1
0
```