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

Problem 5 : Gender Prediction from Credit Card Transactions

Exploring the ways customers' use their credit cards will give banks, researchers, marketers and economists more information to understand their spending habits. With this kind of knowledge, banks or marketing strategists could tailor their marketing and communication programs according to customers' consumption behaviors or patterns.

This problem focuses on credit card spending transactions to answer a simple question: could you train an algorithm to predict the gender of the owner of a credit card?

Provided Files (for both Training and Testing) :

- File name: tj_05_credit_card_transaction.csv
  - Description: Dummy credit card transaction
  - Period: 01/01/2016 - 31/12/2016
  - Size: 900,000 rows

Field Name	Data Type	Description
card_no	STRING	Dummy 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 Note : 0 means unidentified

"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_05_training.csv
  - Description: dummy credit card number with the gender of its owner
  - Size: 18,486 rows

Field Name	Data Type	Description
card_no	STRING	Dummy Card Number
gender	STRING	Gender of the Owner 0: Female 1: Male

Testing Data File

- File name: tj_05_test.csv
  - Description: the list of credit card numbers to be predicted
  - Size: 4,623 rows

Field Name	Data Type	Description
card_no	STRING	Dummy Card Number

Expected Output

- File name: 5.txt
  - Description: Predicted gender of the owner of the corresponding credit card number in the test data file
  - Each line contains either 0 (Female) or 1 (Male)
  - Each line of prediction results must correspond to the card in the same line from the test data file (tj_05_test.csv)

Field Name	Data Type	Description
gender	STRING	Gender of the Owner 0: Female 1: Male

Examples

File: tj_05_credit_card_transaction.csv

```
1234000000010162,2016-12-07 00:00:00.0,22,200,4121,0
1234000000026524,2016-12-07 00:00:00.0,22,550,4899,0
```

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

File: tj_05_training.csv

```
1234000000022789, 0
1234000000024068, 1
```

File: tj_05_test.csv

```
1234000000010162
1234000000026524
```

Output File: 5.csv

```
1
0
```