

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

Problem 3: Churn Detection with Account's Activities and Transactions

No firms want customers to stop using their products or services. For such reason, identifying customers who are fading away has been one of more crucial tasks that every business aims to do. For the longest time, churn detection is done manually; however, this proves to consume significant amount of time and human resource. Nowadays, with fast-paced advancing technology, churn detection is starting to be done by computer. This results in firms being able to act in time to reduce churn rate considerably leading to significant increase in business value.

In TechJam, you are asked to automate churn detection for a financial institution. Account and transaction information is provided.

Data: 14,020 dummy accounts (account_no)

Provided Files:

- Filename: tj_03_deposit_txn.csv
 - Dummy deposit transaction data of all dummy accounts
 - Period: 8/2016 – 10/2016
 - 2,918,302 rows

Field name	Data Type	Description
account_no	string	Dummy account no
from_to_account_no	string	From to dummy account number
txn_amount	double	Net transaction amount
txn_dt	timestamp	Stamp date
txn_hour	int	Transaction hour
txn_type	string	Type of transaction

- Filename: tj_03_account_info.csv
 - Account information of all dummy accounts
 - Period: 8/2016 – 10/2016
 - 529,709 rows

Field Name	Date Type	Description
account_no	string	Dummy account number
txn_dt	timestamp	Transaction date
open_date	timestamp	Account opening date
customer_type	string	Type of customer
pos_dt	timestamp	Transaction date
last_active_date	timestamp	The most recent date in which the account is interacted to
dormant_days	string	Dormant days (the total number

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

		of days that a customer has not had any transaction since his or her last contact date)
compound_frq	string	Frequency by which interest is compounded in which the customer has to pay the bank
compound_frq_unit	string	Unit of measurement of compound frequency in which the customer has to pay the bank
eff_interest_rate	decimal	Effective interest rest

Training Data File:

- Filename: tj_03_training.csv
 - Training data with given label
 - 11,215 rows

Field Name	Data Type	Description
account_no	string	Dummy account number
label	int	0 : not closed in month 11 1 : closed in month 11

Test Data File:

- Filename: tj_03_test.csv
 - Account numbers to be predicted whether they will be closed in month 11 or not
 - 2,805 rows

Field Name	Data Type	Description
account_no	string	Dummy account number

Expected Output:

- Filename: 3.txt
 - Prediction result for each account from the test data file where each line is either 0 (non-churn) or 1 (churn) with no white space. Note that each line in this file must correspond to the same line in the test data file.

Examples:

- Filename: tj_03_deposit_txn.csv

3450016331,3450114404,250.0,2016-10-22 00:00:00,18,DR
3450004676,0000000000,750.0,2016-09-19 00:00:00,20,DR

- Filename: tj_03_account_info.csv

3450037335,2016-10-31 00:00:00,2014-02-07 00:00:00,0704,2016-10-31 00:00:00,0,6,M,0.50000
3450030294,2016-08-30 00:00:00,2011-01-17 00:00:00,0704,2016-08-29 00:00:00,1,6,M,0.50000

- Filename: tj_03_training.csv

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

3450034541,1 3450022684,1

- Filename: tj_03_test.csv

3450034541 3450022684

- Filename: 3.csv

0 1
