DW\_Results

Ashish Kulkarni

Friday, February 20, 2015

This document presents results of running the recurring transactions model on the raw transactions in "Export Data Checking - DW.csv" and "Export Data Card- DW.csv".

## Checking Data

A summary of the transactions is shown below -

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| description | amount | category | date | transaction\_id |
| USAA FUNDS TRANSFER DB | 500 | Transfers | 2015-1-30 | 1 |
| CHECK # xxxxxx1059 | 150 | Checks | 2015-1-29 | 2 |
| YODLEE.COM INC DIRECT DEP \*\*\*\* | 1000 | Paychecks/Salary | 2015-1-29 | 3 |
| ATM SURCHARGE REBATE | 3 | Deposits | 2015-1-26 | 4 |
| INTEREST PAID | 0.02 | Interest | 2015-1-26 | 5 |
| USAA FUNDS TRANSFER DB | 800 | Transfers | 2015-1-26 | 6 |

The model first filters in potentially similar transactions based on their description. It then runs a rules-based binary classifier to label these transactions as 'Valid' (recurring) or 'Invalid' (non-recurring) transactions.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| transaction\_id | similar\_transactions | desc | date\_diff | date\_var | amt | cat | discret | sub | pred |
| 43 | 44|40|38|39|27|28|30|37|10|6|1|16 | usaafundstransferdb | 0.65 | 5.3 | 119.3 | Transfers | FALSE | FALSE | Invalid |
| 2 |  | checkxxxxxx | 0 | 0 | 0 | Checks | FALSE | FALSE | Invalid |
| 32 |  | checkxxxxxx | 0 | 0 | 0 | Home Improvement | TRUE | FALSE | Invalid |
| 48 | 24|11 | checkxxxxxx | 5.07 | 6.38 | 0 | Personal Care | TRUE | FALSE | Invalid |
| 46 | 34|15 | checkxxxxxx | 4.64 | 0.5 | 0 | Rent | FALSE | FALSE | Invalid |
| 45 | 31|17|12|3 | yodleecomincdirectdep | 2.75 | 0.9 | 0 | Paychecks/Salary | FALSE | FALSE | Valid |
| 23 | 4 | atmsurchargerebate | 4.71 | 0 | 0.01 | Deposits | FALSE | FALSE | Invalid |
| 22 | 5 | interestpaid | 4.71 | 0 | 0 | Interest | FALSE | FALSE | Invalid |
| 19 | 7 | usaapcautopay | 4.43 | 0 | 0 | Insurance | FALSE | FALSE | Invalid |
| 20 | 8 | pacificgaselpayment | 4.43 | 0 | 1.02 | Utilities | FALSE | FALSE | Invalid |
| 18 | 9 | concurepense | 3.57 | 0 | 0 | Business Miscellaneous | FALSE | FALSE | Invalid |
| 41 | 47|25|35|13 | usaacompaymntcreditcrd | 2 | 19.85 | 230.7 | Credit Card Payments | FALSE | FALSE | Invalid |
| 26 | 36|21 | targetonlinepmt | 0.71 | 14.73 | 17.45 | Home Improvement - Cleaning | FALSE | FALSE | Invalid |

The final list of recurring transactions -

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Transation ID | Similar Transactions | Category | Subscription | Amount | Description |
| 45 | 31|17|12|3 | Paychecks/Salary | FALSE | 1000 | YODLEE.COM INC DIRECT DEP \*\*\*\* |

## Card Data

A summary of the transactions is shown below -

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| description | amount | category | date | mem\_id | transaction\_id |
| CVS/PHARMACY #x9370 SAN ANSELMO CA | 45 | Healthcare/Medical | 2015-1-31 | 1 | 1 |
| MOLLIE ST GREEN GREENBRAE CA | 6.99 | Groceries | 2015-1-31 | 1 | 2 |
| xx xxxx4008 CORTE MADERA CA | 40.93 | Uncategorized | 2015-1-30 | 1 | 3 |
| MARIN JOE'S RESTAURANT CORTE MADERA CA | 134.8 | Restaurants/Dining | 2015-1-30 | 1 | 4 |
| AMAZON MKTPLACE PMTS AMZN.COM/BILLWA | 26.03 | Gifts | 2015-1-29 | 1 | 5 |
| SQ \*FALAFEL TAZAH REDWOOD CITY CA | 9.76 | Restaurants/Dining | 2015-1-28 | 1 | 6 |

The model first filters in potentially similar transactions based on their description. It then runs a rules-based binary classifier to label these transactions as 'Valid' (recurring) or 'Invalid' (non-recurring) transactions.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| mem | transaction\_id | similar\_transactions | desc | date\_diff | date\_var | amt | cat | discret | sub | pred |
| 1 | 62 | 2 | molliestgreengreenbraeca | 6.14 | 0 | 27.3 | Groceries | TRUE | FALSE | Invalid |
| 1 | 69 | 70|66|5 | amazonmktplacepmtsamzncombillwa | 2.24 | 12.33 | 0.67 | Gifts | TRUE | FALSE | Invalid |
| 1 | 101 | 6 | sqfalafeltazahredwoodcityca | 10 | 0 | 0.48 | Restaurants/Dining | TRUE | FALSE | Invalid |
| 1 | 64 | 59|8|36 | fastrakcscxxxxxxxca | 1 | 14.94 | 0.07 | Automotive Expenses - Fast track | FALSE | FALSE | Invalid |
| 1 | 117 | 94|128|9 | nobhillredwoodcityca | 3.67 | 59.07 | 2.53 | Groceries | TRUE | FALSE | Invalid |
| 1 | 93 | 58|10 | comcastcaliforniacomcastca | 4.36 | 0.01 | 0 | Cable/Satellite Services | FALSE | FALSE | Valid |
| 1 | 15 | 16 | smartecartestpaulmn | 0 | 0 | 0 | Travel | TRUE | FALSE | Invalid |
| 1 | 104 | 97|98|17 | unitedxxxxxxxxxxxxxxxxxxxxt | 3.19 | 21.97 | 69.69 | Travel | TRUE | FALSE | Invalid |
| 1 | 114 | 47|19 | eatcomsanbrunoca | 5 | 6.9 | 7.38 | Restaurants/Dining | TRUE | FALSE | Invalid |
| 1 | 100 | 61|21 | sprintwirelessxxxxxxks | 4.36 | 0.01 | 0 | Telephone Services | FALSE | FALSE | Valid |
| 1 | 107 | 129|65|85|28 | usaacompmtthankyousanantoniot | 2 | 20.61 | 230.7 | Credit Card Payments | FALSE | FALSE | Invalid |
| 1 | 31 | 40 | safewaystorexxxxmillvalleyca | -1 | 0 | 0.74 | Groceries | TRUE | FALSE | Invalid |
| 1 | 126 | 35 | chevronxxxqsancarlosca | 9.29 | 0 | 0.4 | Automotive Expenses - Wash | FALSE | FALSE | Invalid |
| 1 | 95 | 127|34 | chevronxxxsancarlosca | 3.07 | 77.23 | 2.33 | Gasoline/Fuel | TRUE | FALSE | Invalid |
| 1 | 106 | 81|38 | traderjoesqpssanfranciscoca | 3.57 | 2.61 | 16.24 | Groceries | TRUE | FALSE | Invalid |
| 1 | 123 | 80|39 | netflicomnetflicomca | 4.36 | 0.01 | 0 | Entertainment | TRUE | TRUE | Valid |
| 1 | 87 | 72 | travisbelmontca | 1.14 | 0 | 1.45 | Gasoline/Fuel | TRUE | FALSE | Invalid |
| 1 | 76 | 77 | safewaystorexxxxredwoodcityca | 0 | 0 | 3.95 | Groceries | TRUE | FALSE | Invalid |
| 1 | 110 | 86|79 | walgreenssanfranciscoca | 1.43 | 2 | 7.2 | Personal Care | TRUE | FALSE | Invalid |
| 1 | 131 | 83 | amazoncomamzncombillwa | 4.72 | 0 | 13.61 | Gifts | TRUE | FALSE | Invalid |
| 1 | 103 | 121 | chevronxxxsanfranciscoca | -1.57 | 0 | 0.09 | Gasoline/Fuel | TRUE | FALSE | Invalid |

Final list of recurring transactions

## Warning: longer object length is not a multiple of shorter object length

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Transation ID | Similar Transactions | Category | Subscription | Amount | Description |
| 93 | 58|10 | Cable/Satellite Services | FALSE | 7.99 | NETFLIX.COM NETFLIX.COM CA |
| 100 | 61|21 | Telephone Services | FALSE | 7.99 | NETFLIX.COM NETFLIX.COM CA |
| 123 | 80|39 | Entertainment | TRUE | 7.99 | NETFLIX.COM NETFLIX.COM CA |