Credit-Risk-Modelling-of-Mobile-Loan-Applicants-CBA-Safaricom- Public

Pin Unwatch 1 Fork 0 Star 0

main 1 branch 0 tags Go to file Add file Code

Tarachalk Update README.md ccecced 1 hour ago 20 commits

docs	first commit	17 hours ago
models	first commit	17 hours ago
notebooks	first commit	17 hours ago
references	first commit	17 hours ago
reports	naming convention changed	2 hours ago
src	first commit	17 hours ago
.gitignore	first commit	17 hours ago
LICENSE	first commit	17 hours ago
Loan_Database Raw.accdb	renamed database	16 hours ago
Makefile	first commit	17 hours ago
README.md	Update README.md	1 hour ago
requirements.txt	Modified file	15 hours ago
setup.py	first commit	17 hours ago
test_environment.py	first commit	17 hours ago
tox.ini	first commit	17 hours ago

About

This repository contains code and resources for building credit risk models for Safaricom customers. This repository may be useful for data scientists and analysts who are interested in developing and deploying credit risk models for Safaricom.

risk modelling loan default credit

Readme
View license
Activity
0 stars
1 watching
0 forks

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Languages

Jupyter Notebook 99.7% Other 0.3%

Suggested Workflows

Based on your tech stack

SLSA Generic generator — Configure
Generate SLSA3 provenance for your existing release workflows

Publish Python Package — Configure
Publish a Python Package to PyPI on release.

Python package — Configure
Create and test a Python package on multiple Python versions.

More workflows Dismiss suggestions

Credit-Risk-Modelling-of-Mobile-Loan-Applicants (CBA & Safaricom)

Author: Richard Taracha

Date: 23/10/2023

Table of Contents

Comprehending The Setting

Safaricom, a telecommunications behemoth with deep roots in Kenya, has carved a niche for itself in the realm of financial services through its M-PESA brand. M-PESA, transcending its origins as a mere mobile money transfer platform, has evolved into a comprehensive financial ecosystem catering to the diverse needs of millions of Kenyans.

At the heart of M-PESA's success lies its ability to provide a secure and convenient means of transferring funds between individuals. This has not only revolutionized the way Kenyans conduct their daily financial transactions but has also played a pivotal role in promoting financial inclusion among the unbanked and underbanked segments of the population.

In 2013, Safaricom, in a strategic partnership with CBA, launched M-SHWARI, a savings and loan product designed to tap into the vast potential of the M-PESA user base. M-SHWARI has since grown into a formidable force in the Kenyan microfinance landscape, providing millions of Kenyans with access to affordable and readily available credit.

M-SHWARI loans are typically small in size and have a short repayment period of 30 days. This makes them ideal for meeting short-term financial needs such as medical emergencies, school fees, and business expenses. The convenience of accessing M-SHWARI loans through the M-PESA platform has made it a popular choice among Kenyans seeking quick and hassle-free access to credit.

In addition to M-SHWARI, Safaricom also offers Okoa Jahazi, a credit loan service for prepaid subscribers. Okoa Jahazi allows subscribers to access small loans that are repaid through automatic deductions from their airtime top-ups. This service has proven to be a lifeline for many Kenyans who find themselves in need of immediate cash during times of financial difficulty.

Safaricom's foray into the financial services sector has been nothing short of remarkable. Through M-PESA, the company has not only provided Kenyans with a convenient and secure means of conducting financial transactions but has also played a major role in promoting financial inclusion in the country. As Safaricom continues to innovate and expand its financial services offerings, it is poised to play an even greater role in shaping the future of finance in Kenya.

Business Case

As a Data Scientist working for the institution you have been tasked to identify factors that contribute to loan default. The findings of your analysis will be used to inform the creation of a credit scoring model that will be used to determine the creditworthiness of Okoa Jahazi loan applicants. The model will be used to predict whether or not an applicant will default on their loan based on their credit history and other relevant factors. This will help the institution make better lending decisions and reduce the number of loan defaults.

Data Pertinence and Attribution

In December 2021, a household survey conducted by Central Bank of Kenya (CBK), FSD Kenya and the Kenya National Bureau of Statistics (KNBS) revealed that 50.9% of mobile loan borrowers had defaulted on their loans. The survey also found that 12.5% of mobile loan borrowers had defaulted on their loans more than once. This is a worrying trend that needs to be addressed if Kenya is to achieve its goal of becoming a middle-income country by 2030.

Source: https://www.businessdailyafrica.com/bd/economy/half-mobile-phone-borrowers-default-3654550

Formulating the Benchmark of Success

The project's objective is to estimate when a person will default on their loan. We construct a model that incorporates all the relevant data to reliably predict the outcome.

Loan defaulting occurs when a person does not fulfill their debt obligation in the specified period. A loan is considered to be in default for risk modelling purposes if it is more than 90 days old.

We will have achieved our objective when we get at least one model with an accuracy score of around 80%.

Technologies and Tools: Python, Pandas, Numpy, Matplotlib, Scikit-Learn

Back to top

Project Structure & Deliverables

```
├── LICENSE
├── Makefile           <- Makefile with commands like `make data` or `make train`
├── README.md          <- The top-level README for developers using this project.
├── data
│   ├── external       <- Data from third party sources.
│   ├── interim        <- Intermediate data that has been transformed.
│   ├── processed      <- The final, canonical data sets for modeling.
│   └── raw            <- The original, immutable data dump.
│
├── docs               <- A default Sphinx project; see sphinx-doc.org for details
│
├── models             <- Trained and serialized models, model predictions, or model summaries
│
├── notebooks          <- Jupyter notebooks. Naming convention is a number (for ordering),
│                         the creator's initials, and a short `-` delimited description, e.g.
│                         `1.0-jqp-initial-data-exploration`.
│
├── references         <- Data dictionaries, manuals, and all other explanatory materials.
│
├── reports            <- Generated analysis as HTML, PDF, LaTeX, etc.
│   └── figures        <- Generated graphics and figures to be used in reporting
│
├── requirements.txt   <- The requirements file for reproducing the analysis environment, e.g.
│                         generated with `pip freeze > requirements.txt`
│
├── setup.py           <- makes project pip installable (pip install -e .) so src can be imported
├── src                <- Source code for use in this project.
│   ├── __init__.py    <- Makes src a Python module
│   │
│   ├── data           <- Scripts to download or generate data
│   │   └── make_dataset.py
│   │
│   ├── features       <- Scripts to turn raw data into features for modeling
│   │   └── build_features.py
│   │
│   ├── models         <- Scripts to train models and then use trained models to make
│   │   │                 predictions
│   │   ├── predict_model.py
│   │   └── train_model.py
│   │
│   └── visualization  <- Scripts to create exploratory and results oriented visualizations
│       └── visualize.py
│
└── tox.ini            <- tox file with settings for running tox; see tox.readthedocs.io
```

There are three deliverables for this project:

- A GitHub repository
- A Jupyter Notebook
- A non-technical presentation

Dataset Attributes:

CUST_TXN_DATE	Customer Transaction Date
CUST_ID_ACCT1	Account Identifier
DEPOSIT_AMNT_MPSA	Amount deposited on Mpesa
RCVD_AMNT_MPSA	Amount received on Mpesa
TRSF_FROM_BANK_TO_MPESA_AMNT	Amount on transfers from Bank on Mpesa
MPSA_CREDITS	Mpesa credits
DAYS_ARTM_LESS_2	Number of days airtime is less than 2
X_NR_TOT_LOAN_AMNT_OKOA_JAZI	Total loan amount on okoa jahazi
LOAN_AGE_DAYS	Loan age in days

Back to top

Recording the Experimental Design

The following steps were followed during the implementation of the project:

Step	Task
1	Define the Research Question.
2	Data Importation.
3	Data Exploration.
4	Data Cleaning.
5	Exploratory Data Analysis
6	Data Preparation
7	Data modelling.
8	Model Evaluation.
9	Summary of Findings
10	Recommendations / Conclusions
11	Challenge solution.

Data Prepration & Modeling (from point 5 to 8)

All details are in the notebook in the following directory `end-to-end-data-science-project/notebooks`

Back to top

Made with ❤️ by Richard Taracha