Bank Customer Segmentation

Function : #Customer Marketing Strategies | **Industry :** #Banking Sector

rubiscape©
make sense

Goal

• To identify spending and investment habits of the customers and segment them corresponding to age group and locations.
• Utilize these segments for marketing strategies.
• To plan cross-sell and up-sell segment wise.

Technique

• Statistical Analysis
• Clustering Algorithms
• Classification Algorithms
• Visualization

Impact

• Streamlined marketing efforts: Targeting specific customer segments, tailored messaging and promotions
• Increased efficiency : Automating customer segmentation
• Improved decision-making: Gaining insights into customer behavior and preferences
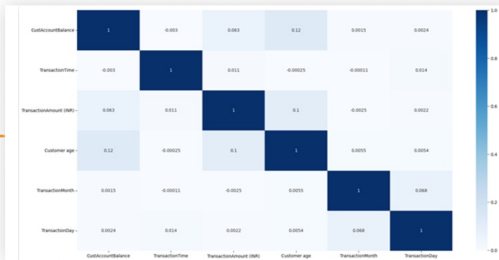
Result

Bank Transaction Analysis

MODELS	Runtime	Accuracy	Sensitivity	Specificity	F-Score	AUC
Extreme_Gradient_Boost	133.7	58.9%	0.6018	0.5768	0.5945	0.6312
Random_Forest	591.3651	57.4%	0.6264	0.5209	0.5952	0.6078
k_Nearest_Neighbor	2784.4404	56.3%	0.4246	0.7023	0.4931	0.5892
Decision_Tree	46.1035	56.0%	0.6438	0.4766	0.5943	0.5803
Adaboost	109.2078	54.5%	0.6256	0.4645	0.5791	0.5568
Binomial_Logistic_Regression	51.4397	53.0%	0.7524	0.3066	0.6154	0.5306

Value Points

Understand the **what, why, when, where** & **how**
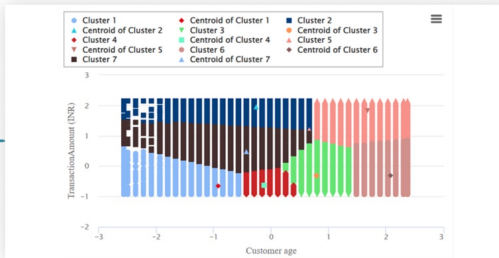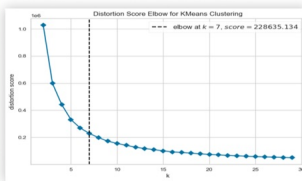
Exploratory Analysis

Exploratory data analysis on the preprocessed data to derive meaningful data insights

The maximum features affecting transactions using correlation map are considered

Exploratory data analysis enables business owners to derive meaningful insights and making better data-driven decisions as opposed to intuitive ones.

Clustering

WCSS is used to obtain optimal clusters and then k-means, and hierarchical clustering are implemented.

Cluster analysis provides valuable insights about the dynamics in different bank transactions, aiding decision-making in the financial domain.
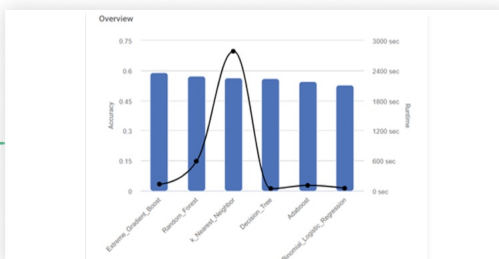
Classification

```
Confusion Matrix :
 [[63820 29083]
  [43836 49357]]
Accuracy for Logistic Regression model is : 60.816460%
Classification Report
             precision  recall  f1-score  support

         0      0.59      0.69      0.64     92903
         1      0.63      0.53      0.58     93193

  accuracy                          0.61    186096
 macro avg       0.61      0.61      0.61    186096
weighted avg     0.61      0.61      0.61    186096
```

Applied multiple classification algorithms and then based on accuracy metric best fit is identified
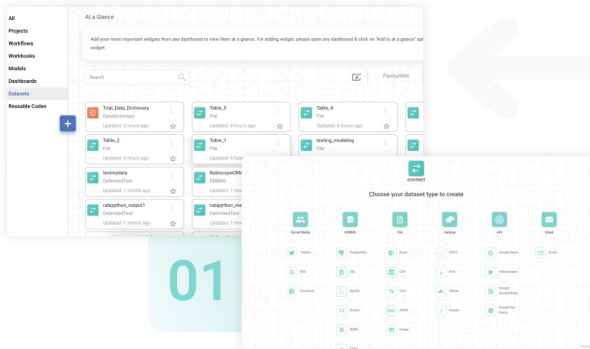
Classification analysis in bank transactions, particularly for fraud detection, has wide-ranging business impacts including operational efficiency and regulatory compliance,

Multi Persona DSML Platform

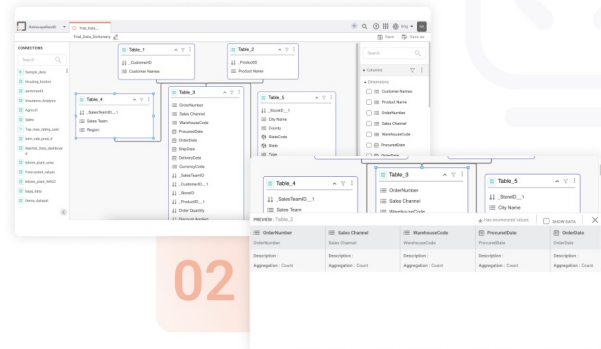For all your data needs- Data Engineering, Data Science, Data Visualisation, IoT

rubiscape © make sense

Data Connect

01

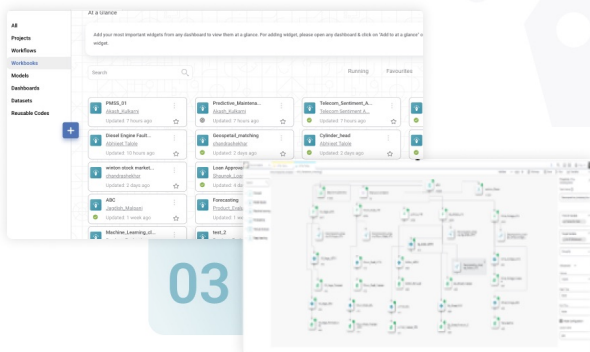Data Source: Bank_Transcation_Analysis

Metadata Manger

02

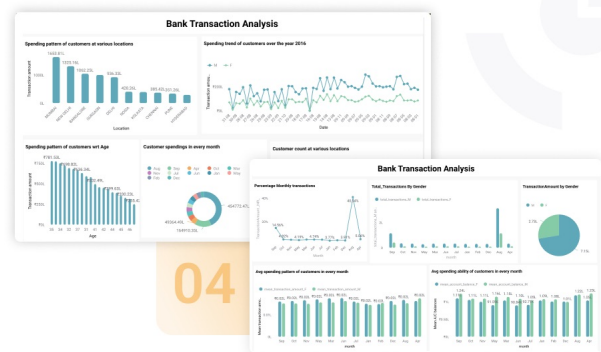Comprehensive Data Operations, encompassing Metadata Management

Model Studio

03

Modelling, encompassing the selection and configuration of models

Visualisation

04

Viz Ops, Illustrating The Core Trends And Graphical Representations

Agile Data Science

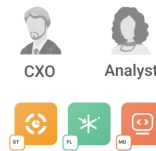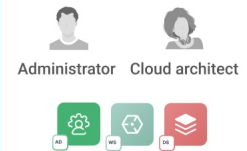Encapsulating **best practices**, tools and **methods**

Data Ops	ML Ops	Viz Ops	Tech Ops
Data Architect Data Manager	ML Experts Data Scientist	CXO Analyst	Administrator Cloud architect

Ideate	Acquire	Explore	Model	Present	Deploy
• What is the **goal**? • How can you leverage the **data**? • What do you want to **predict**?	• How is data **sampled**? • Which data is **relevant**? • Any data **privacy** issue?	• Plot the **data** • Are there **anomalies**? • Are there **patterns**?	• **Build** a model • **Fit** the model • **Validate** the model	• What did we **learn**? • Do the results make **sense**? • Can we tell a **story**?	• Where to **Deploy**? • What is the **Structure** of Pipeline? • How to **Optimise** and **Scale**?