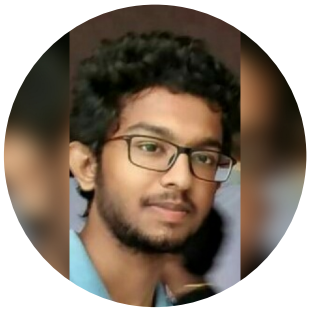
SHANKAR S

Data Science Analyst | Fin-Tech

Linkedin
Github
+91 6383 798 827
@ shankarselvaraj24@gmail.com
Chennai, Tamil Nadu



About Me - **Data Science Analyst with 2 years of experience. Core expertise in Statistics, Machine Learning & Analytics**

PROFESSIONAL EXPERIENCE

Data Science Analyst

Kaleidofin | Series A FinTech Startup serving the Underbanked Space, backed by BHARAT Inclusion, Omidyar Network & BLUME Ventures

Nov 2019 – Present Chennai, Tamil Nadu

- **Payments Prediction Model** - Developed end-to-end prediction model using machine learning techniques like LOFO importance for feature importance, lightGBM, Bayesian Optimization for hyperparameter tuning, SHAP plots for model interpretation by leveraging payments data of active Kaleidofin customer base.
- **Customer Profiling** - Categorized Low-income households from Hyderabad using Hierarchical Clustering for product design. This enables the organization to develop customized products for the underbanked sector.
- **Customer Transactions Prediction** - Implemented Bayesian Machine Learning model using PyMC3 to predict SEWA Bank customer transactions. Bayesian Multiple Linear Regression was used in factoring uncertainty of transactions as opposed to traditional ML models with deterministic output.
- **Geospatial Analysis** - Mapped entire Kaleidofin customer base using Folium, pgeocode libraries in Python to help visually comprehend the distribution of customers across India by making use of latitudinal & longitudinal data. Enabled local merchant payment options by mapping Kaleidofin customers to the nearest banking correspondents by making use of Haversine distance formula.
- **Enrolment Time Analysis** - Extracted event based real-time app data from Firebase DB using BigQuery to compute overall enrolment time of customers enrolled by Kaleidofin agents. Key enrolment time descriptive metrics helps in revamping enrolment process, and thus reducing the overall enrolment time.

Data Science Intern

Fiind Inc. | Predictive Analytics & AI Company

Jun 2018 – Nov 2018 Trichy, Tamil Nadu

- **Sales Forecasting** - Developed Decay Pattern Clustering technique in R to forecast sales by leveraging e-commerce sales data. Products are analysed to comprehend its life cycle and clustered based on their decaying patterns. Statistical techniques like Chi-square goodness of fit test has been used to compare distributions.

PROJECTS

Treatment Effectiveness Calculator

A medical application developed using R to calculate the treatment effectiveness of patients having higher cholesterol levels. The application uses real-world pre-treatment and post-treatment data as input and usage of statistical concepts like t-test to arrive at the effectiveness of the treatment.

SKILL-SETS

Database & ETL
MySQL | PostgreSQL | GCP BigQuery

Languages and Framework
Python | R | Apache Spark | Pandas | numpy | Sklearn | PyMC3 | Plotly | Folium

ML Ops
Git | Github | ShellScripting

Dashboarding
Tableau | AWS QuickSight

Tools & OS
Spyder | Jupyter | Microsoft Office Suite | BeautifulAI | Windows | Linux

EDUCATION

M.Sc Data Science (Integrated 5 year course)

PSG College of Technology

2015-2020 Coimbatore, Tamil Nadu

CERTIFICATION

Scalable ML using Apache Spark by IBM.

Advanced SQL for Data Scientists by Linkedin Learning.

Tableau Essential Training by Linkedin Learning.

REFEREES

Aaron Joseph Mathew
Data Scientist ex-Kaleidofin
@ aaronjosephmathew@kaleidofin.com
Bangalore
+91 984 434 8064