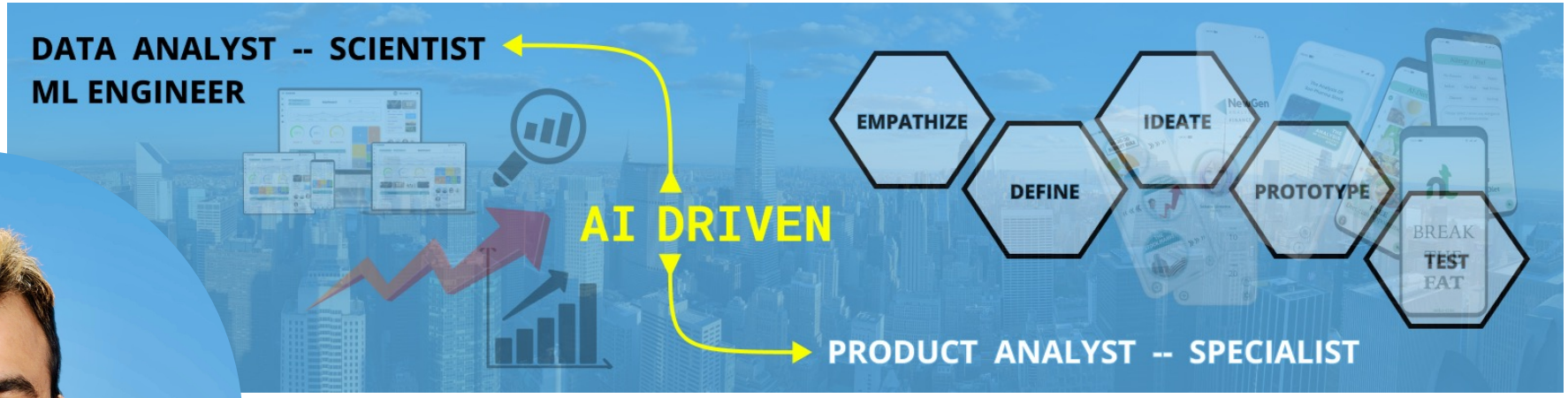


Data Scientist @ Con Edison | EX- MLE | EX-SDE | 3+ YOE

Forecasting, Recommendation | I find anomalies to mitigate risk & maximize efficiency

aws certified Cloud Practitioner FOUNDATIONAL

Google coursera Google Project Management Certificate Certificate of Completion

git databricks Azure

Figma SQL pandas

C3.ai

Power BI

AMOGH MAHADEV KOKARI -- LINKEDIN.COM/IN/AMOGHKOKARI

THE STORYTELLER

About Amogh

Hello, I'm an accomplished Analyst with passion for transforming complex datasets into actionable insights and strategic solutions across finance and various sectors. Holding a Master's in Information Systems, my expertise spans advanced statistical analysis, machine learning, AI technologies, and programming with a deep dive into **EXCEL**, **PYTHON**, **SQL**, **POWER BI** and more.

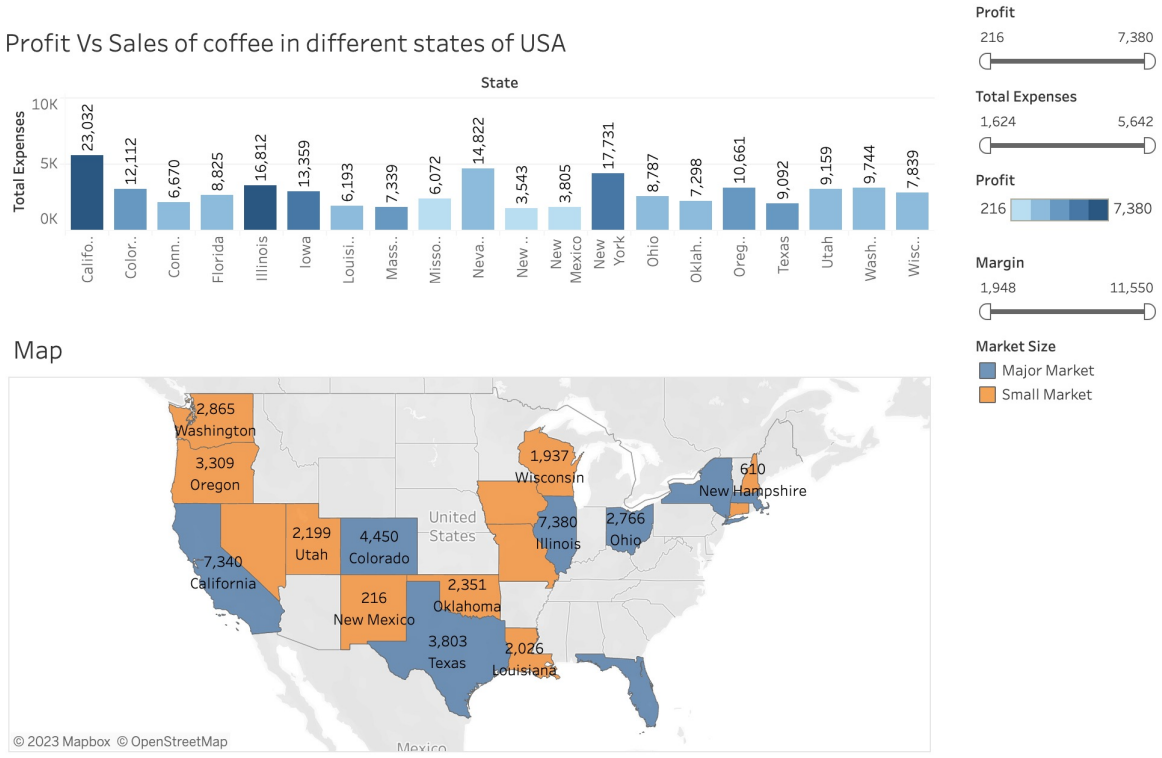
At the forefront of Data Science, I've engineered impactful models and systems, from **Forecasting** and **Recommendations** to optimizing operations and enhancing decision-making processes. My work with Con Edison and Bullwhip Technologies, among others, showcases my ability to leverage data for **predictive analytics**, **operational improvements**, and **insightful visualizations**, significantly affecting bottom lines and operational efficiency.

Certified in AWS Cloud, Google Business Intelligence, and as a Meta Database Engineer, I'm not just about data. I thrive in environments that challenge my problem-solving, agile methodologies, and stakeholder management skills. With a track record of leading projects from ideation through deployment, I excel in turning data into a compelling narrative for both technical and non-technical audiences.

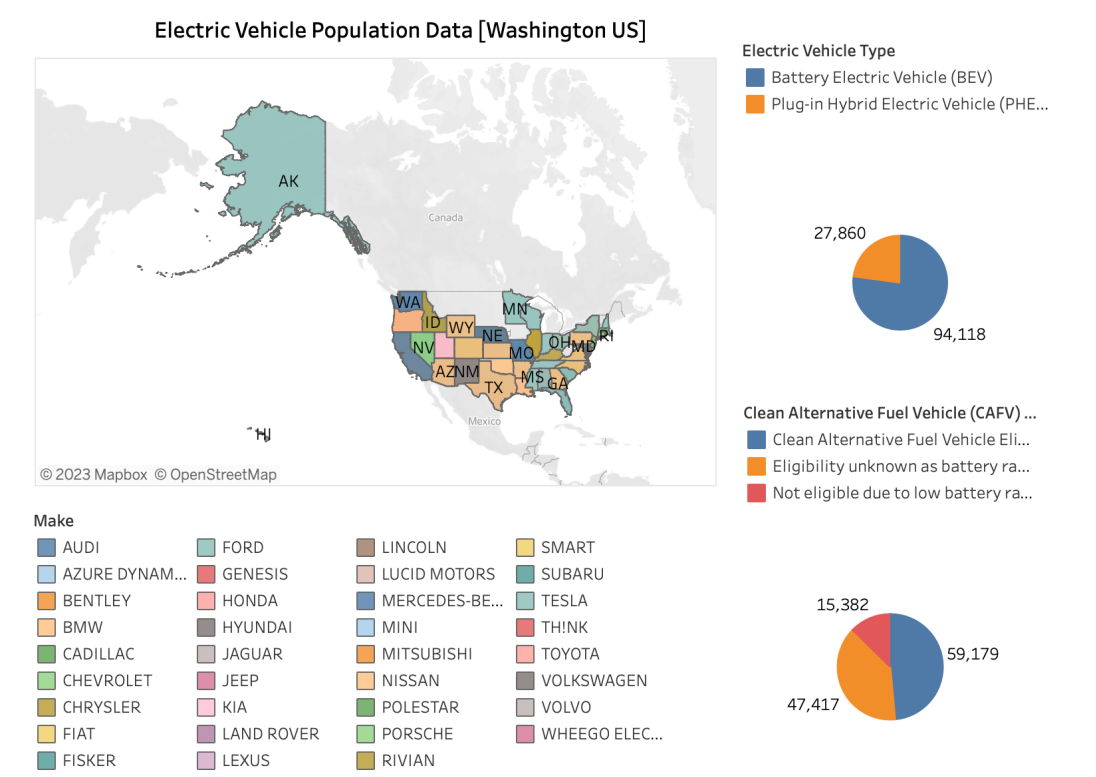
Whether streamlining processes, automating data collection, or driving user-centered solutions, I'm dedicated to pushing the boundaries of what data can do to solve real-world problems, making data-driven decision-making the cornerstone of business strategy and innovation.

Amogh Mahadev Kokari -- linkedin.com/in/amoghkokari

Tableau Viz on Profit VS Sales of Coffee in Different States of USA [LINK]



Electric Vehicle Analysis [LINK]



Amogh Mahadev Kokari -- linkedin.com/in/amoghkokari

Motivation

- Explore loan data from the Paycheck Protection Program (PPP), which provided relief to small and medium-sized businesses during the COVID-19 pandemic and identify probable fraudulent loans. The primary objective is to reduce frauds in the future by applying anomaly detection methods to identify outliers and building machine learning models to potentially detect possible frauds.

Technology

- Python was used for data pre-processing and cleaning
- Python libraries matplotlib, plotly, seaborn utilized for exploratory analysis and dashboarding
- Python to identify anomalies using pandas, matplotlib, IsolationForest in slkearn

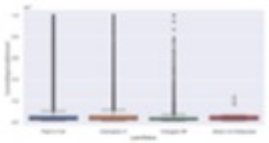
Fraud Detection - Paycheck Protection Program

Exploratory Analysis

Visualizations for this project are intended to aid the user get an essential understanding of the context of the subject and get a feel of what the dataset is trying to convey on the high level.

Supplementary intentions with the visualizations are to be informative for the user to know where to start looking for anomalies first. Ultimately, we want the the visualizations to aid the process of identifying anomalous phenomena by guiding the user's attention to some extreme unruly observations as a starting place for the outlier analysis.

Some examples of this effort can be seen below.



Approval Amount Outlier Analysis

Correlation heat-matrix to identify key dependencies

Heatmap of number of loans approved across all the states

Loan Approval Amounts Density Plot

Anomaly Detection

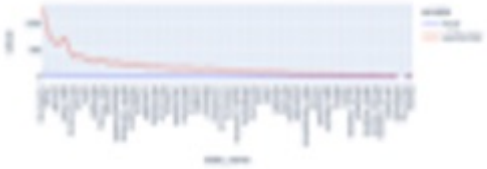
Standardization is done using standard scalar
Higher the anomaly score, higher the probability of loan being fraud

Isolation forest is being used to calculate anomaly scores
Anomaly score is calculated for features:
Initial Approval Amount, Payroll proceed and Jobs Reported

We then identify a loan being fraud if it lies in the top 95% of anomaly score calculated

State-wise probable frauds in USA

State-wise Loans vs Frauds



Current & Future Work

- Government agencies and policymakers may use the data to evaluate the effectiveness of the program and make necessary adjustments to future relief efforts
- Researchers may use data to study the economic impact of the pandemic on small businesses and to identify patterns and trends in the distribution of PPP loans
- Incorporate possible frauds with portal like Datamerch, Experian, Iapps to improve due diligence for debt-based Venture Capitalists

Amogh Mahadev Kokari -- linkedin.com/in/amoghkokari

Big Data Text Analysis - YouTube Channel Video Classification

[Project Slides] [DEMO]

Current & Future Work

- The current scope consists of analyzing data from one channel from YouTube to train on ML algorithms and Spark based to create predictions from user and display result on Frontend using Flask
- The current scope of deriving insights from one channel can be expanded to multiple channel input
- The preprocessing and ML modeling can be parallelized by running multiple instances
- Decentralization and guarding it from short term failure can be resolved using AWS ecosystem the entire project can be cloud based to handle high volume of requests and process them in real time paving the way of real time streaming analytics and prediction

Workflow

Data Pipeline

Conclusion

- The end-to-end analytics as a service based on decentralized microservices can be used in creating a robust real time prediction that can help content creators attract crowd on YouTube
- Although the limitation to the current development is on untuned algorithm but they tend to perform well in binary classification (0,1), auto tuning can be applied to make them better
- we believe the proposed big data architecture will be widely extensible and expandable that can be produced as industry grade service for everyone

Sample Data Pipeline build on MAGE AI
Don't forget to get in touch with presenters for a live demo
Scan Down for a video Preview and codebase

Technology

- PYSPARK
- PYTHON
- FLASK
- PANDAS
- HTML/CSS
- JOBLIB
- LIGHTGBM
- BEAUTIFULSOUP
- SKLEARN
- XGBOOST
- WORDCLOUD
- MATPLOTLIB
- NLTK
- JOBLIB
- REGEX

Motivation

- The motivation behind this is to empower the creation of high value content on streaming service like YouTube.
- The underlying fundamental is to build a scalable decentralized framework for ingesting data from different sources, training ML algorithms and predicting user content
- Scope is extended to show the effectiveness of Spark framework on large datasets

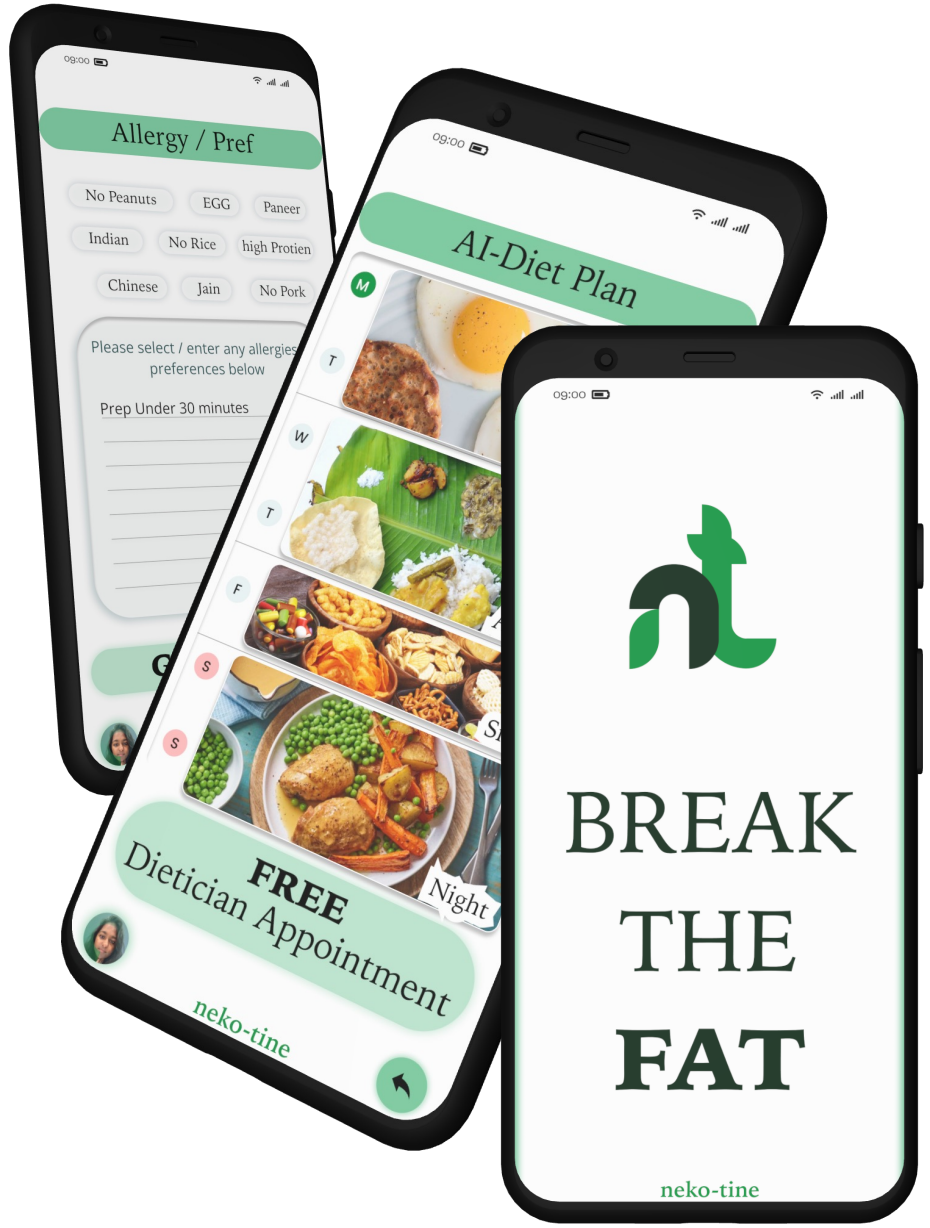
Amogh Mahadev Kokari -- linkedin.com/in/amoghkokari

Goal Based Student Diet Personalization

(Interviewed students using **Design Thinking** tools (Empathy), analyzed data to understand user needs to develop product to save busy student's deteriorating health)

[Prototype] [Project Slides]

[Explainer Video]



Amogh Mahadev Kokari -- linkedin.com/in/amoghkokari

Key Accomplishments

Power Grid Optimization at Con Edison: Mapped 1M+ NYC grid components, developing a machine learning model that predicted EV charging with 96% accuracy, optimizing charging times to reduce grid stress and cut electricity bills.

Keyword Strategy Innovation at Bullwhip Technologies: Created a time series forecasting model and a recommendation system, boosting keyword prediction accuracy and increasing user engagement and revenue by 15%.

Operational Efficiency Boost at Build Health International: Developed a financial oversight dashboard, significantly streamlining project management and reducing employee transition time from one day to one hour.

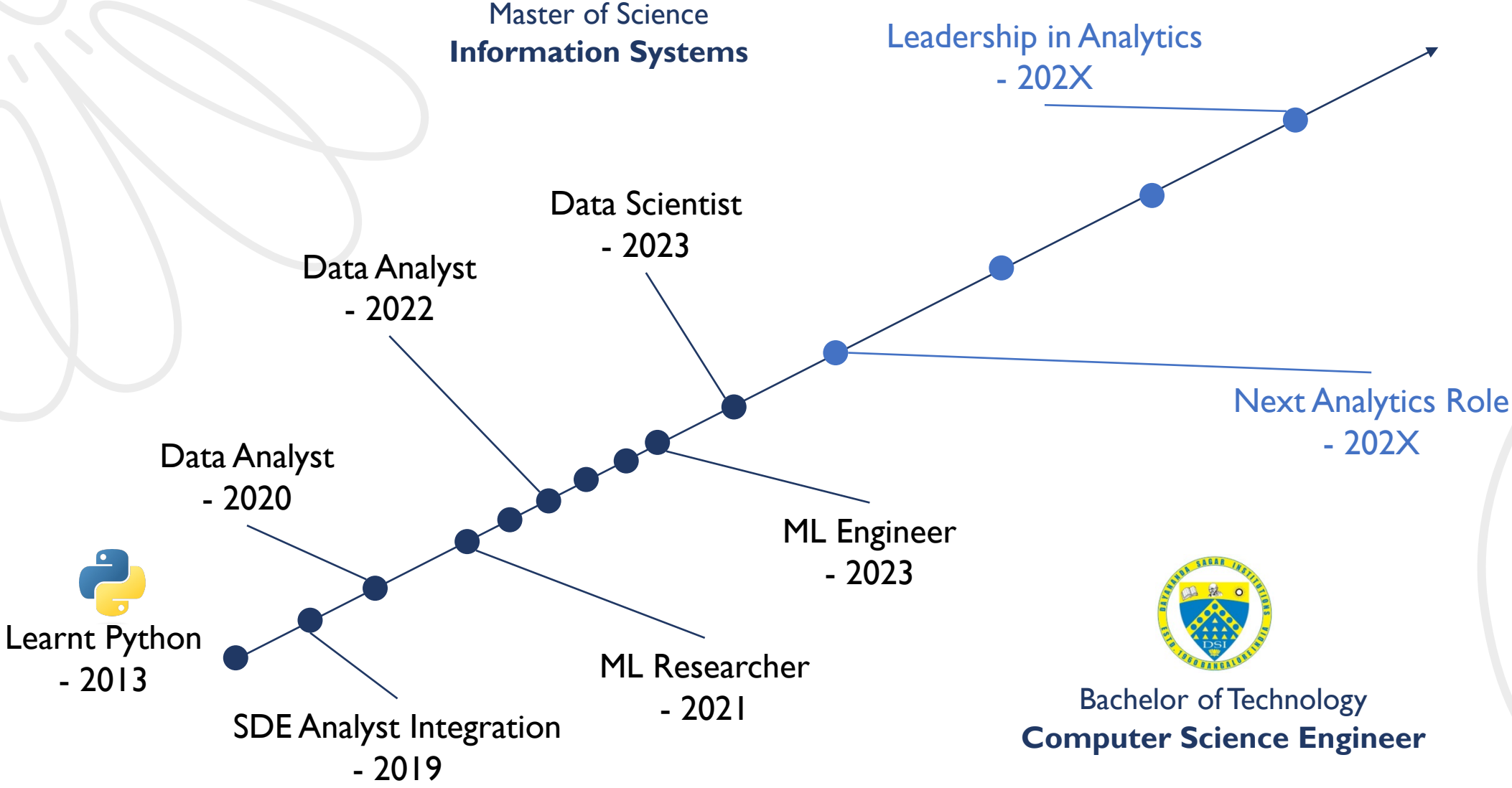
Payment Security Enhancement at Razorpay, Nioneer: Led fraud detection on 1M+ transactions and automated payment processes, saving $60,000 monthly, while improving digital engagement by 20% through targeted marketing strategies.

Amogh Mahadev Kokari -- linkedin.com/in/amoghkokari

CAREER TIMELINE

Master of Science
Information Systems

Bachelor of Technology
Computer Science Engineer



Testimonials

Chinmay Rathod (Co-Founder, Lead Data Scientist)

I had the pleasure of working with Amogh on a beta testing project that involved deploying code on AWS. His strong knowledge of cloud computing, software development, and data science was incredibly helpful for the project …..

Balaji Rao (PhD, Blockchain Researcher)

I worked closely with Amogh for over a year as a research assistant at the Stevens Institute of Technology. Amogh is a highly skilled individual with expertise in software development, data engineering, and data science. He built the Human-Computer Interaction Lab's official website, collaborating with a team of 5 researchers and developers to create an interactive user experience …..

Referances available on request

Michael Washington
Director of Finance

Dr Thomas Lechlar
Project Management Professional

Dr Edward Stohr
BIA Program Director

RJ Lehman
Founder Stealth Startup

Dr Joseph Morabito
Business Intelligence Professional

Amogh Mahadev Kokari -- linkedin.com/in/amoghkokari

Thank You

LINKEDIN.COM/IN/AMOGHKOKARI/

GITHUB.COM/AMOGHKOKARI

GETFOODRECIPE.STREAMLIT.APP

Amogh Mahadev Kokari -- linkedin.com/in/amoghkokari