Connect & explore how I can be of an asset to you as I have been for others

OpenToOpportunities

I Implement **Data** and **AI** solutions for Your Use Cases

★ Implemented **14+** use cases
★ Consulted **4+** Startups
★ Regularly invited as panelist
★ Speaker at Tech Talks

Data Analyst --> **ML Modeling** --> **AI Engineer**
Design Thinking + Rapid Prototyping + Productionize + Story Teller

Data Scientist | Time Series Analytics, Recommendation Systems, Python, SQL, Power BI | I Help companies leverage AI, ML & Big Data to Identify Anomalies, Mitigate Risk & Personalize offerings

aws certified
Cloud Practitioner
FOUNDATIONAL

dbt
Learn
Fundamentals

ORACLE Certified
Professional
Oracle Cloud Infrastructure Generative AI
2024

pandas

git

Power BI

SQL

databricks

LangChain

Firebase

Gemini

Streamlit

C3.ai

Google coursera
Google **Project Management Certificate**
Certificate of Completion

Azure

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

Talk at Microsoft office NYC
during
AI Camp Meetup
link

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

Big Data Text Analysis - YouTube Channel Video Classification

[Project Slides]   **[DEMO]**

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

Technology
- **PYSPARK**
- **PYTHON**
- **FLASK**
- **PANDAS**
- **HTML/CSS**
- **JOBLIB**
- **LIGHTGBM**

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

Workflow

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

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

Fraud Detection -
Paycheck Protection Program

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.
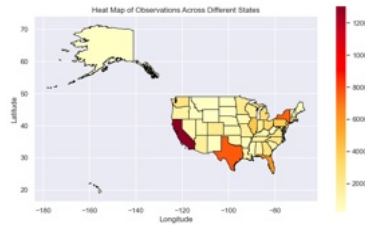- The primary objective is to reduce frauds in the future by applying anomaly detection methods to identify outliers and building machine learning models to potentially detect possible frauds.

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 Sklearn
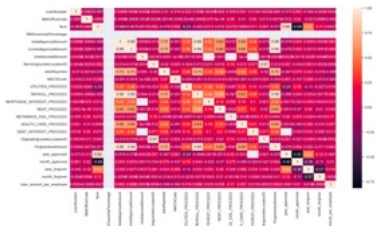
Exploratory Analysis

- Investigated data to understand distribution, anomaly with heat map and charts
- Others were specific to data exploration with Python

Some examples of this effort can be seen below.
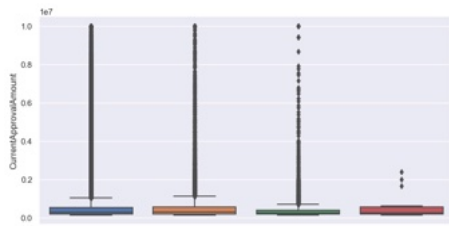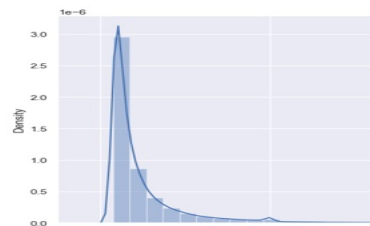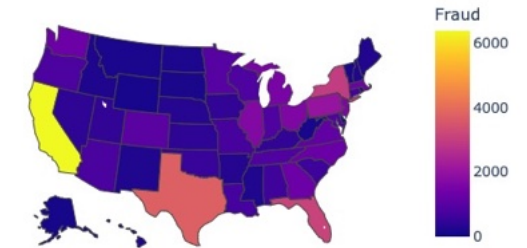

Approval Amount Outlier Analysis

Heatmap of number of loans approved across all the states

Correlation heat-matrix to identify key dependencies

Loan Approval Amounts Density Plot

Anomaly Detection

- Data is Standardization using standard scalar
- Findings Higher the anomaly score, higher the probability of loan being fraud
- Used Isolation forest to calculate anomaly scores
- Calculated Anomaly score for features:
  - Initial Approval Amount
  - Payroll proceed
  - Jobs Reported
- Identified loan being fraud biases of anomaly score top 95% = anomaly (i.e. fraud)

State-wise probable frauds in USA
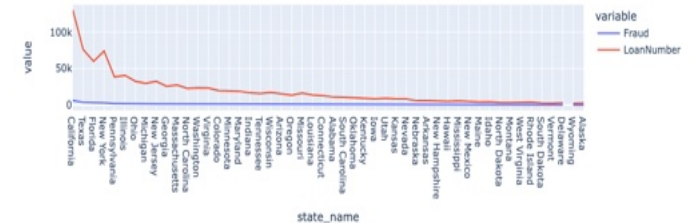
State-wise Loans vs Frauds

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

Tableau Viz on Profit VS Sales of Coffee in Different States of USA [LINK]
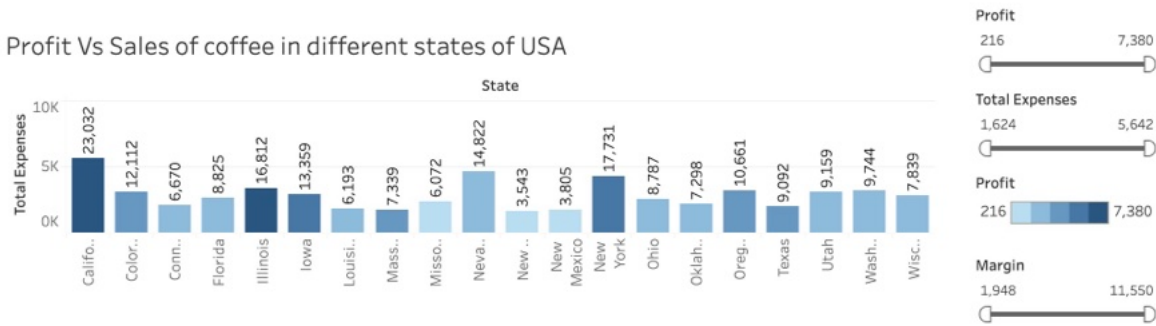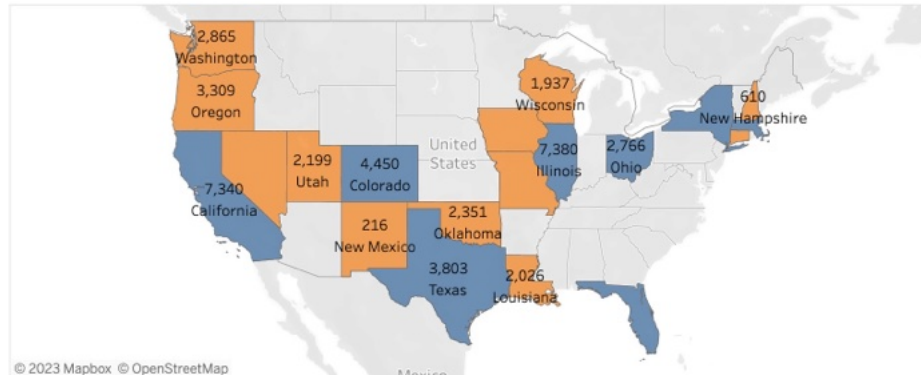
Electric Vehicle Analysis [LINK]
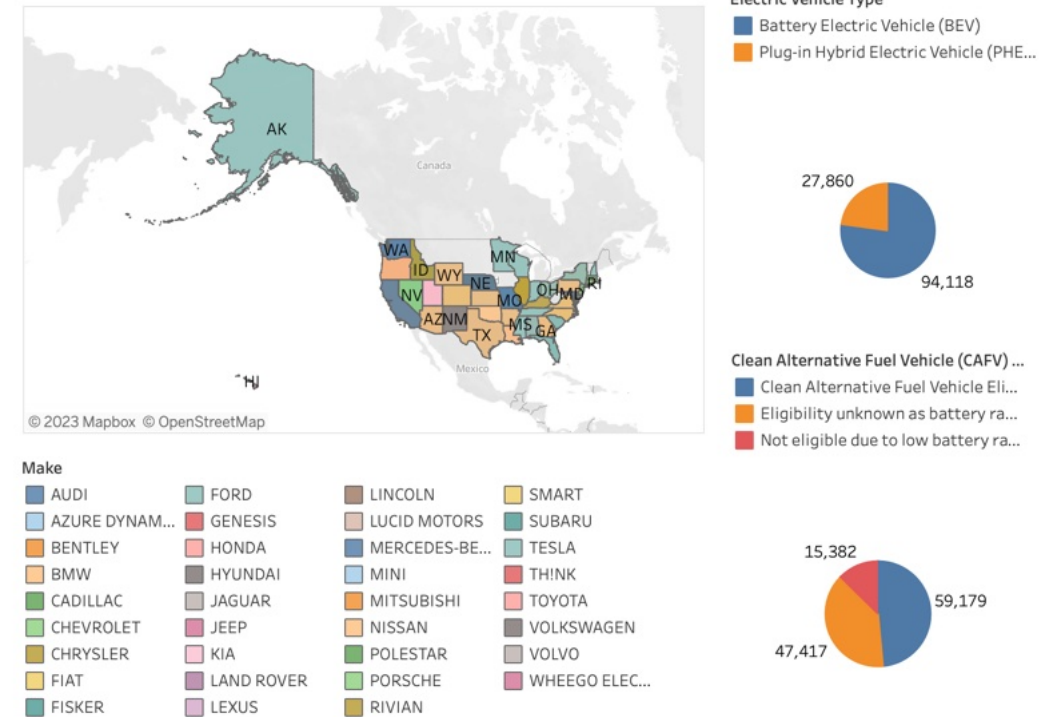
Goal Based Student Diet Personalization
(Interviewed students using **Design Thinking** tools (Empathy), analyzed data to understand user needs to develop product to save busy student's deteriorating health)
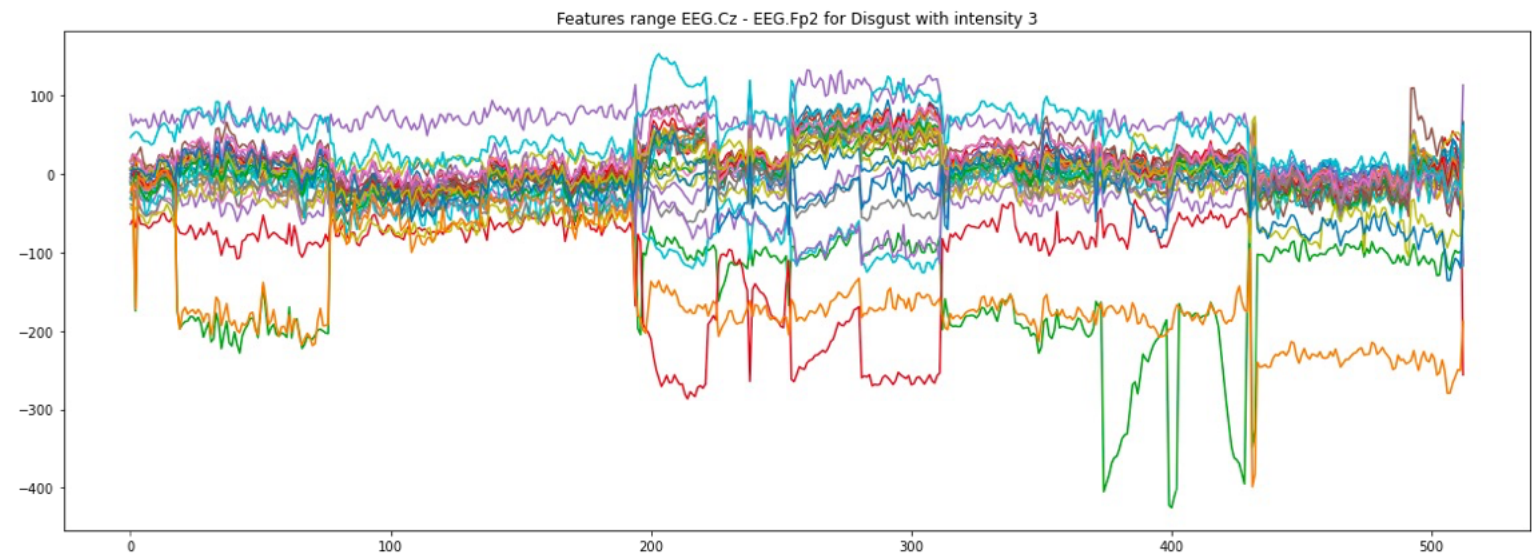
[**Prototype**]   [**Project Slides**]

[**Explainer Video**]

EEG Emotion Detection
Deep Learning Hard Parameter Sharing

Features range EEG.Cz - EEG.Fp2 for Disgust with intensity 3

	precision	recall	f1-score	support
0	0.80	0.77	0.78	13005
1	0.81	0.81	0.81	40448
2	0.81	0.80	0.80	16179
3	0.79	0.81	0.80	54272
4	0.76	0.79	0.77	22733
5	0.80	0.80	0.80	33280
6	0.78	0.74	0.76	17408
accuracy			0.80	197325
macro avg	0.79	0.79	0.79	197325
weighted avg	0.80	0.80	0.80	197325
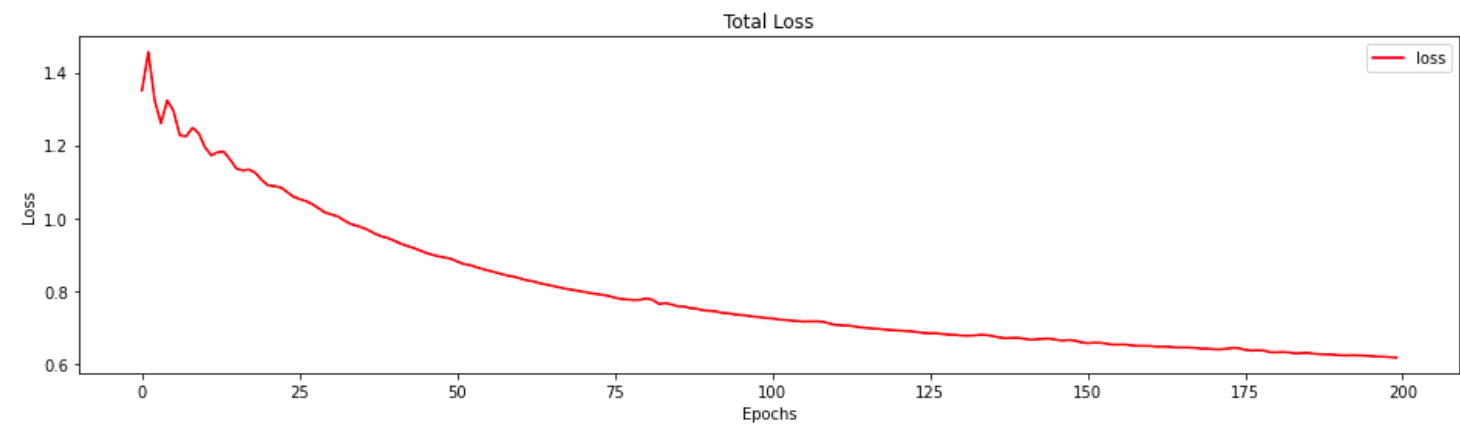
Amogh Mahadev Kokari -- linkedin.com/in/amoghkokari

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 .....

References available on request

Michael Washington
Director of Finance

Dr Thomas Lechlar
Project Management Professional

Dr Edward Stohr
BIA Program Director

RJ Lehman
Founder Bullwhip

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