Siddharth Sankaran

siddharth2996@gmail.com | +1 (716) 515-8570 | LinkedIn | GitHub | Buffalo, NY

Summary

Data scientist with 3+ years of experience in designing, developing, testing, and deploying machine learning solutions for real world business problems across multiple verticals. An enthusiastic learner experienced in translating diverse business requirements into impactful deliverables and providing actionable insights.

Work Experience

Data Scientist, Mu Sigma Business Solutions- Bangalore, India Aug '18 – Aug '21

Customer Retention

- Designed a data-driven customer retention framework to help business determine who, when, why and how to target customers. Discovered actionable customer microsegments with potential losses of ~€140M
- Segmented customers as improving/declining using regression and performed hypothesis testing to determine reasons behind the behavior. Built an R-Shiny dashboard to help business generate insights

Customer Segmentation

- Bolstered business' customer targeting by clustering customers based on purchase behaviour and product preferences to mend a decline of ~€30M and helped business increase customer spend by ~7%
- Engineered features based on RFM analysis, performed dimensionality reduction, and ran multiple clustering algorithms (k-means, mean-shift, DBSCAN) to create five actionable customer segments

Anomaly Detection

- Constructed an ensemble anomaly detection model identify to spikes in weekly cost and forecast potential aberrations. Devised a heuristic algorithm to rank anomalies and helped business avoid a spike of $400,000
- Analysed key metrics to pinpoint the reason behind anomaly and constructed a Power BI Dashboard to showcase all results, enabling a business user to view results from this analysis in real time

Education

- **Master of Science in Computer Science and Engineering** Aug '21 – Jan '23
  SUNY at Buffalo, Buffalo, New York
- **Bachelor of Technology in Computer Science and Engineering** Jun '14 – May '18
  SASTRA University, Thanjavur, Tamil Nadu, India

Projects and Publications

- **Similarity Based Representation for Identifying Healthcare Anomalous Activities** - Built a **vision-based patient monitoring system** that can detect abnormal activities like **coughing, sneezing, vomiting, falling**, etc. (https://doi.org/10.1166/jmihi.2020.2903)
- **Engagement Recognition Using Video based micro expression tracking** – Built a model that can detect user's engagement (bored, engaged, frustrated, etc.) from videos by capturing micro expressions
- **GitHub Notification System** - Built a Publisher-Subscriber system to fetch and notify live GitHub updates and messages. Implemented this from scratch and deployed Kafka to be the message broker for version 2.0
- **Diabetes Prediction** – Built a Logistic Regression model with Gradient Descent from scratch and compared with a multi-layer regularized neural network model
- **Review of Vision-based Assistive Healthcare Monitoring** – Surveyed the latest works in the field of patient monitoring analytics and presented a paper at the 12th International IndiaCOM Conference

Skills

- Languages – Python (NumPy, Pandas, Scikit-learn, Matplotlib, Flask, Kafka-Python), R, SQL (Hive, SSMS)
- Data Science – EDA, Hypothesis Testing, Regression, Clustering, Tensorflow, OpenCV
- Business Intelligence – Power BI, R Shiny; Tools – Microsoft Office (Excel, PowerPoint, Word), JIRA, Git