Khiran Kumar Chidambaram Sivaraman

+1(443)-788-2711 | sivkhiran3@gmail.com | Github

Experience

Data Analyst- Vanguard **Pennsylvania, USA | Mar 2023 - July 2023**

- Maintained test software for data collection and analysis using Python. Optimized the ETL pipelines to extract consumer data from multiple data sources (CSV, JSON, Excel) and load into MS SQL server using SSIS, saving up to 10+ hours weekly for data acquiring, parsing, entry, filter, etc. (data warehousing).
- Designed dashboard for data visualization using Tableau to monitor 15+ KPIs/metrics of a health insurance company enabling the management to identify the areas that require more attention. **JIRA** was used for issue and bug tracking maintaining the phase of the project.
- Designed and implemented a Predictive Model to facilitate resource allocation and investment decision making for an energy company using Python libraries like Sklearn, Pandas.Evaluated the migration of database changes and referenced data changes. **JIRA was linked to the Confluence** page for documentation of project report.

University of Maryland, Baltimore County -Graduate Teaching Assistant **Maryland, USA | July 2022 - Sept 2022**

- Thoroughly reviewed the code, graded the assignments, and provided feedback to 50 students in the **Intro to Data Analysis and Machine Learning** course.
- Prepared, executed, and shared the final solution key of the assignments to students, and responsibly completed multiple projects assigned by the professor.
- Mentored and clarified course-related questions.
- Worked diligently to complete the assigned projects within the deadlines while achieving the priorities and deliverables.

Quality Control Trainee/ Quality Test Engineer - HWASHIN Automotive India Private – Limited **Chennai, India | June 2018**

- Completed in-plant training at Hwashin Automotive, overseeing **product quality** through performance testing and inspections during development, encompassing **raw material, warranty claims, body parts, chassis quality, and new product** quality for clients including Hyundai and Honda.
- Performed Process capability study and risk analysis to reduce the process variation using **MINITAB** software and implemented containment actions while defining the **root cause analysis like fishbone diagrams and 8D methodology**.

Skills

Programming Languages	**Python-Jupyter Notebook, REST API, R**, C++, Scala, SQL, Linux, HTML, Bash(Scripting)
Data Visualization	Tableau, Power BI, Matlab, LabView, Microsoft Office Suite(Word, PowerPoint, Spreadsheet, Excel)
Cloud Computing	**AWS Cloud (Docker, EC2)**, Google Cloud Platform(GCP)
Python Libraries	ggplot, Matplotlib, seaborn, Streamlit, Pyplot,**Pandas**, PyTorch, PySpark, **NumPy**, **SciKit-Learn**, SciPy
AI and ML Library	**TensorFlow**, XGBoost, **Keras**, Natural Language Processing (NLP), ML Algorithms
Mechanical Tools	**Lathe Machine**, CNC Machine (G-Code), **Thermocouple, Strain Gauge**
Design Softwares	**CATIA V5**, SOLIDWORKS, AUTOCAD, ANSYS
Manufacturing Skills	MINITAB, Fishbone Diagrams, **FMEA, Root Cause Analysis**, Process Capability, DOE
Database	**MongoDB**, Hadoop, **Google Big Query**, MySQL
IDE	PyCharm, Eclipse
Collaboration tools	JIRA, Confluence

Education

Master of Professional Studies - Data Science **January 2021 - December 2022**
University of Maryland, Baltimore County
Courses: Data Structures & Algorithms, Intro to Data Analysis and Machine Learning, Project Management, Artificial Intelligence

Bachelor of Technology - Mechanical Engineering **May 2016 - May 2020**
SRM Institute of Science and Technology, India
Courses: Applied Thermal Engineering, **Elements of Mechatronics** - Intro to Sensors and Transducers, Heat & Mass Transfer, **Electric and Hybrid Vehicles**- Traction drive Power calculations for Electric Vehicles-Energy Storage Systems - Battery- Fuel Cell – Supercapacitors
Competition:Baja SAE, Developed an ATV as a team and designed test fixtures using CATIA V5 and involved in test equipment in collaboration with design engineering team.

Academic Projects

Neural Network Analysis of a heat pipe using hybrid nano fluids / Language: Python **January 2020 - May 2020**

- Conducted and automated a comprehensive analysis of a heat pipe charged with hybrid nano-fluids using machine learning techniques in Python.

- Employed **TensorFlow (Keras), Deep Convolution Neural Network (DCNN)** models in Python with **ReLU activation function** for the convolution layer to simulate inputs.
- Achieved an accuracy-correlation coefficient (R^2) of 0.991 Which is **99% accuracy** with the DCNN model.
- Increased heat pipe efficiency by decreasing other parameters (water inlet temperatures).

Chicago Taxi Usage Analysis / Language: Python | Database: MongoDB **September 2021 - December 2021**

- Explored large scale geospatial taxi usage dataset from Chicago spanning from 2014 to 2020.
- Investigated changes in taxi usage related to total trips, travel times, distances, and locations.
- Used Random Forests and Decision Trees to predict trip trends from 2014 to 2020.
- Observed an overall steady number of trips predicted for 2020, but the pandemic caused a sharp decrease compared to 2019's data.

Projects Completed

E-Commerce Churn Analysis / Language: Python **July 2023**

- Conducted comprehensive analysis of data structure, identified missing values, and implemented data cleaning strategies, and improved overall data quality.
- Employed one-hot encoding to accurately encode categorical variables, preserving data quality while enabling effective analysis.
- Employed various machine learning models for churn prediction, including **Logistic Regression, Random Forest, Gradient Boosting, SVM, and MLPClassifier**.
- Assessed models using metrics: accuracy, precision, recall, F1-score, and ROC-AUC. Identified Random Forest as the top-performing model (**Accuracy: 96%, ROC-AUC: 0.9957**).
- Assessed feature importances of the Random Forest model to identify churn factors.

Pizza Ordering System using Python (Data Structures and Algorithm) / Language: Python **January 2021 - May 2021**

- Developed a database that allows customers to customize and place pizza orders from their favourite store based on their preferences and analyzed the complexity of ordering the customized pizza using Binary search algorithms.

Prevalence of Tobacco Usage in United States from 1995 to 2010) / Language: R **January 2022 - May 2022**

- The goal was to predict how the prevalence of tobacco usage has changed in the U.S. over time and how the various measures taken by the U.S. government have affected people who smoke.
- Created a regression model to understand how the trends continued over the time period. A Multiple Linear Regression model was used to predict the trend.

Data Annotation Projects

A Comparative Evaluation of Sentiment Analysis in the case of Covid-19 Tweets / Language: Python **September 2021 - December 2021**

- Conducted COVID-19 sentiment analysis on **5825** public sentiments on Twitter reaction data points post-cleaning.
- Performed hypothesis testing, Trained and tested **Naïve Bayes, Logistic Regression, Long Short-Term Memory (LSTM), and Attention model** for predicting the outcome.
- Found Logistic Regression as the most efficient model, achieving an **accuracy of 86%**. Created a **Streamlit** app which shows the predicted output for the models used.

Customer Segmentation and Analysis / Language: Python **July 2023**

- Employed Customer Segmentation and Analysis to divide the customer base into distinct groups based on Product characteristics, behaviors, or attributes.
- Prepared **541,910** data points for analysis, reducing them to **406,830** by addressing missing values, removing outliers, and normalizing the data for accuracy and consistency.
- Investigated each customer segment to understand preferences, needs, and differences.
- Created customer profiles with specifications like demographics, preferences, buying behavior, schedule and relevant information for a better understanding of each segment.
- Conducted RFM analysis to identify high-value customers with consistent and substantial purchase behavior.

Certification

- **Career Essentials in Generative AI by Microsoft and LinkedIn (Pursuing)** **November 2023**
- **Career Essentials in Data Analysis by Microsoft and LinkedIn** **August 2023**
- **Power BI Essential training** **August 2023**
- **Learning Data Analytics Part 2: Extending and Applying Core Knowledge** **August 2023**
- **Regression Models** **July 2020**
- **Getting and Cleaning Data** **July 2020**
- **Introduction to Probability and Data with R** **July 2020**
- **R Programming** **May 2020**

Publications (Heat Transfer)

1. Kumararaja, K., Khiran Kumar, C. S. & Sivaraman, B. A convolutional neural network analysis of a heat pipe with Hybrid Nanofluids. International Journal of Ambient Energy **0,** 1–13. https://doi.org/10.1080/01430750.2021.2014959 (2021).