VALLI RAJA SEKAR M
| +91 9600231572 | sekarvalliraja@gmail.com
Linkedin | Portfolio | Github

Model Building | Predictive Modeling | Machine Learning | Regression | Classification | Forecasting |Langchain | Tableau | SQL | NLP | Statistics| Generative- AI |Deep Learning|Data Mining|Python | AZURE |AWS|GCP

Experienced Data Scientist with a comprehensive background covering different aspects of Machine Learning, NLP, and Generative AI. I have successfully designed and implemented projects in various domains, showcasing a hands-on approach and a track record of delivering impactful results. My expertise extends to the financial sector, specifically in transaction banking and the management of High Net Worth (HNW) client holdings

Holding an MBA from ICFAI Business School and recently completing a Post Graduate Program in Data Science and Business Analytics from Texas McCombs, University of Texas, I bring a unique blend of business acumen and technical skills. This combination positions me as a versatile professional capable of delivering data-driven solutions that align with business objectives.

My banking experience has played a crucial role in shaping my understanding of financial processes, such as Know Your Customer (KYC) procedures, transaction analysis, and risk management. This deep understanding of banking operations has been instrumental in my data science career, enabling me to leverage data effectively to drive strategic decision-making and innovation

Proficient in cutting-edge technologies, including Generative AI, OpenAI, Langchain, and Hugging Face, I employ innovative approaches to solve complex problems. By integrating these advanced tools, I provide actionable insights that enhance decision-making processes in intricate business scenarios..

P R O F E S S I O N A L- E X P ER I EN C E

GWC Data.Ai,Data scientist Bengaluru(Mar 2023-Present)

DKSH (FINANCE SECTOR)

Overview :

Developed a comprehensive sales forecasting system for various country regions at SKU level, using time series analysis and regression models. Leveraged Azure Machine Learning and Azure Synapse Pipelines to handle 25 million data points. Achieved a model accuracy of 90% through advanced encoding techniques and model optimization.

Role:

- Conducted time series forecasting for regions with sufficient data, ensuring robust predictions for sales at SKU levels.
- Built regression models using XGBoost and Random Forest for regions lacking comprehensive data, focusing on achieving high accuracy and generalizability.
- Designed and implemented end-to-end machine learning pipelines in Azure Synapse, automating data ingestion, preprocessing, model training, and deployment.
- Fine-tuned models to achieve a 90% accuracy score, utilizing cross-validation and hyperparameter tuning to enhance predictive performance.
- Collaborated with stakeholders to interpret model results, providing actionable insights that improved inventory management and sales strategies, ultimately driving a 15% increase in sales efficiency.
- Forecasted time series data from January 2021 to December 2023 to predict the next year's data, aiming to reduce MAPE to less than 30. Evaluated different combinations of models, with tasks performed in Azure Synapse and Azure ML service.

Machine Bringup Application(Manufacturing Sector)-TATA ELCTRONICS

Overview:
In Machine Bring-Up Application project at Tata Electronics, focusing on the Classification phase to determine the eligibility of the machine tuning model. Successfully gathered data from diverse sources, performed comprehensive data exploration using Jupyter notebooks, and implemented a Logistic Regression model with a notable 95% accuracy. The project's impact included a 20% improvement in efficiency and a 25% reduction in errors, demonstrating significant strides in eligibility assessments.

- Led the end-to-end execution of the Machine Bring-Up Application project, overseeing data collection from PostgreSQL, SQL Server, and APIs.
- Utilized Jupyter notebooks for thorough data exploration, applying statistical concepts for effective Exploratory Data Analysis (EDA).
- Rigorously tested various models and selected Logistic Regression, showcasing expertise in model tuning and optimization.
- Implemented modular coding practices for enhanced code structure, readability, and maintainability, promoting seamless integration within the team.
- Executed the deployment phase using internal servers and SQL Server, ensuring the Logistic Regression model's efficient integration for eligibility assessments.
- Demonstrated proficiency in database management, specifically leveraging SQL Server during the deployment phase.
- Actively contributed to continuous improvement, collaborating with cross-functional teams to refine the Classification phase and overall project.
- Successfully implemented the Machine Bring-Up Application, handling a substantial dataset of 6 crores data points, resulting in a 20% improvement in efficiency and a 25% reduction in errors.

Predictive Modeling for Personal Loan Eligibility (Banking Sector)

- Gathered comprehensive customer data including KYC details, banking transactions, customer segment, past loans, company employed, and CIBIL score. Performed data cleaning, normalization, and feature engineering to prepare the dataset for modeling.
- Conducted EDA to identify key features impacting loan eligibility. Visualized distributions and correlations to understand patterns and relationships in the data, guiding feature selection and model development.
- Utilized CatBoost for its robustness in handling categorical features and building an initial predictive model. Implemented Randomized Search Regressor, Grid Search CV, and Bayesian Search CV for hyperparameter tuning to optimize model performance.
- Evaluated model performance using cross-validation and various metrics. Fine-tuned the CatBoost model to achieve an optimal accuracy score of 89%, ensuring the model's generalizability and reliability in predicting loan eligibility.
- Deployed the final model into the bank’s decision-making system. Collaborated with stakeholders to provide actionable insights based on model predictions, improving the bank's loan approval process and customer satisfaction

LGC,TEXT CONVERSION USING LANGCHAIN AND OPEN-AI

- Collaborated with LGC to convert German-based text into English using the Helsinki model from Hugging Face, facilitating language translation tasks.
- Implemented Regex techniques to extract essential data points, ensuring accuracy and relevance in the converted text for effective data science analysis
- Contributed expertise in leveraging advanced NLP technologies to optimize data processing pipelines, enhancing efficiency and precision in data-driven insights.
- Played a pivotal role in bridging linguistic barriers, enabling seamless cross-linguistic data analysis and interpretation for comprehensive data science solutions
- Demonstrated proficiency in integrating language translation tools and Regex methodologies to support robust data science workflows, empowering informed decision-making processes.

Customer Feedback Sensitivity Classification (Financial Sector)

- Spearheaded a multi-channel complaint classification project aimed at categorizing customer feedback across various mediums such as smile score reviews, social media, and comments.
- Applied advanced natural language processing techniques, including TF-IDF transformation, to convert text data into a format suitable for machine learning analysis.

- Implemented a logistic regression model for precise classification of complaints into three sensitivity categories: More Sensitive, Medium Sensitive, and Less Sensitive.
- Successfully deployed the classification model on an internal server, ensuring real-time analysis and proactive measures to avoid regulatory actions
- Played a key role in establishing an early warning system, leveraging machine learning insights to take preemptive steps based on the sensitivity of customer feedback.

HDFC BANK -CHENNAI-AUG 2020 -AUG2022

Manager | Aug 2021 – Aug 2022

- Managed High Net Worth (HNW) client portfolios, overseeing important portfolios with assets worth 60 crores by generating business from existing and new customers.
- Coordinated with various business groups across multiple product lines, including third-party products, to bring in new business and enhance customer relationships.
- Strategized and implemented business processes to increase sales within the branch and across the portfolio, focusing on whole sales products, forex, bank guarantees, and other offerings
- Developed customized reports using Tableau for branch, cluster, and portfolio analysis, presenting key performance indicators and insights to top officials
- Successfully built strong relationships with HNW clients, understanding their financial goals and providing tailored solutions to meet their investment and wealth management needs
- Consistently exceeded sales targets and contributed to the bank's growth, recognized for exceptional performance and contribution to the organization's success

Management Trainee | Aug 2020 – Aug 2021

- Utilized SQL to access and extract data from various databases, ensuring accurate and timely retrieval of information for analysis and reporting purposes.
- Successfully built comprehensive reports using Tableau, providing key insights on customer behavior, sales performance, and branch metrics to support strategic decision-making.
- Efficiently managed data allocation across branches in the region, optimizing resource utilization and enhancing operational efficiency.
- Spearheaded presentations to top management officials using PowerPoint and Tableau, effectively communicating complex data-driven analyses and performance evaluations.
- Demonstrated exceptional sales acumen while working in an assigned branch, consistently meeting and exceeding sales targets set by the manager.
- Collaborated with cross-functional teams to identify opportunities for process improvent, resulting in enhanced productivity and streamlined operations
- Promoted to Manager for outstanding performance and dedication during the management trainee program

E D U C A T I O N

Degree	Institution	GPA/CPGA /PERCENTAGES	Year of Completion
PG Diploma in Data Science and Business Analytics	University of Texas and Great Lakes	3.25	2023
MBA	ICFAI Business School, Hyderabad	7.16	2020
B.Sc. Mathematics	Loyola College,Chennai	6.89	2016
Class XII	Bharath M. M.H.H.S Matriculation	86.91	2013
Class X	Bharath M. M.H.H.S Matriculation	90.6	2011