Veda Jenetty Immaraj
Ph: +1 (469) 536 8320
Email: vedajenettyimmaraj@gmail.com
LinkedIn: linkedin.com/in/jenetty-immaraj-951883124
Texas, USA

PGP in Data Science
from Carnegie Mellon University

B.Tech in Computer Science
from JNTU-Hyderabad

Professional Summary

- Seasoned Data Scientist with over 8 years of experience, seeking dynamic career opportunities leveraging technical expertise and creativity to contribute significantly to an organization's growth and development while advancing professional growth.
- Expertise includes end-to-end data science solution development, with a particular focus on architecting, developing, and deploying ML solutions on leading cloud platforms such as AWS, GCP, and Azure ML Studio.
- Experienced in building machine learning, and deep learning models for B2B, B2C customer analytics, supply chain, marketing, and finance use cases.
- Passionate about solving challenging business problems and presenting data-driven business decisions to stakeholders.
- Proficient in data preprocessing techniques, exploratory data analysis (EDA), hands-on deep learning model building, and report generation.
- Skilled in working with NoSQL, SQL, Shell Scripting, Automation scripts, and data visualization tools such as TABLEAU.
- Possess an AWS Associate Architect Certification and an Airflow Astronomer Certification.
- Successfully built applications from scratch and worked on multiple use cases on AWS.
- Developed a churn prediction model with 82% accuracy that helped in retaining customers and saving ~.9M $ annually.
- Led the conversion of a 1.2 M $ contract from a Hi-Tech customer.
- Led the development of a Proof-of-Concept project which unlocked an engagement of ~300K $ within 6 months.
- Contributed to converting a ~500K $ RFP to further penetrate the account using existing business knowledge.
- Experience driving strategic business decisions driven by quantitative data with great attention to detail.
- Built Call-to-action based employee churn to reduce attrition at Atlassian, built an HR BOT to help the HRBP's in call-to-action for HR attrition, and built self-serve dashboards.
- Designed, developed & deployed custom Deep learning models (LSTM & Encoder Decoder) for a Sales Recommendation Engine (Upsell\Cross-sell) which helped increase the YoY targets by 20 %.
- Orchestrated ML pipelines using Google Pub-Sub and dockized them to deploy on GPU servers.
- Proposed pricing analytics solution for an e-commerce giant with more than 50+ million users and 500k SKU which Uncovered a lift in demand 1.08X and saw an average 4% increase in revenue in A/B Tested groups.
- Developed an ML WebApp using Flask API backed by the Naïve Bayes model with 75% AUC for product marketing and customer success team, developed a chatbot for the Service desk manager to automate L1 & L2service incidents which enhanced the business process by 40%.

Technical Skills

- **Tools**: Jupyter Notebook, Pycharm, SAS, Tableau, MySQL, PostgreSQL, Teradata, MLOps, GCP cloud, AWS, Microsoft Azure ML Studio, Airflow, Snowflake, PowerPoint, Excel, MS SQL Server.
- **Programming:** Python, R, SQL, Spark, Hadoop, Hive, Numpy, Pandas, Matplotlib, SciPy, Scikit-Learn, Kera's, TensorFlow.
- **Analytics**: Regression, Classification, Clustering, Hypothesis Testing, A/B Testing, T-Test, F Test, ANOVA, Forecasting, statistical methods, supervised learning, unsupervised learning, cross-validation, statistical modeling, Predictive Analytics.
- **Machine Learning**: Principal Component Analysis, Natural Language Processing, Deep Learning, AI/ML algorithms, Autoencoders, Text Analytics

Certifications

- Data Science & Big Data Analytics - Certified by LTI of Carnegie Mellon University certification in 2016 - ranked 3rd worldwide in.

Professional Experience: 7+ Years

- Worked as a Sr. Data Scientist from March 2021 to August 2022.
- Worked as a Data Scientist from January 2019 to February 2021.
- Worked as a Software Engineer from August 2015 to January 2019.

Projects

Genpact, Client Deutsche Bank, New York, NY **Nov 2021–Aug 2022**

Title: Revolutionizing Production Support System for Top-Tier Banking Enterprise
Objective: The project aimed to design and develop a cutting-edge Production Support System powered by Artificial Intelligence that would significantly reduce Mean Time to Recovery (MTTR), minimize manual effort, And mitigate service impact.
Outcome: The project yielded the following outcomes: a 15% improvement in Mean Time to Recovery (MTTR) Through the implementation of self-service, a 45% reduction in call center hold times, and an anticipated annual Savings of $2M.

- Developed an AI Powered Production Support system using custom-tailored ML models such as Doc2Vec,
- SBert, and Word2Vec models for the NLP engine.
- Developing a state-of-the-art AI Powered Production Support system, utilizing custom-tailored machine learning models such as Doc2Vec, SBert, and Word2Vec for the NLP engine.
- Designing the Technical Roadmap and Functional Architecture of the Solution.
- Playing a critical role in planning, architecting, and building end-to-end solutions.
- Collaborating with the client to understand their requirements and charting a path for project implementation.
- Developing the solution on the Google Cloud Platform, including the creation of a customized dashboard on Data Studio, based on the client's specifications.
- Developing the necessary APIs to support the solution.

Technology Stack: GCP, NLP, Jupyter Notebooks & Pycharm

Genpact, Hyderabad, India **Jan 2021 - Oct 2021**

Adidas - EU ECOM, Germany
Title: Football Jersey Forecasting
Problem: Due to the transfer of star players, such as C. Ronaldo moving from Juventus to Manchester United, there has been an unexpected increase in demand for club jerseys, resulting in missed opportunities for the sale of 50k units.
Objective: Forecast the demand of Jerseys for- 2022 Elite 5 Clubs and 8 World Cup Federations
Outcome: I was a vital contributor in the management of a €3.2M project aimed at enhancing operational efficiency.

- Defined the solution roadmap and architected and developed end-to-end solutions.
- Analyzed client requirements and devised strategies for project implementation.
- Built several multivariate time series forecasting models using key features such as historic demand, team performance, Google Trends, PDP visits, etc., to predict future demand for an upcoming season.
- Built an AutoML pipeline to enable the easy deployment, retraining, and reuse of the model in future projects.
- Developed the solution on Azure Machine Learning Studio in accordance with the client's specific needs.

Technology Stack: Azure Cloud, Multivariate Time Series Models, MySQL, Jupyter Notebooks & Pycharm

Quadratyx, Hyderabad, India **Feb 2019 - Dec 2021**

Title: Self-Learning Machines\Auto-Learning
Objective: Develop an AI platform that enables Sales and Product teams at Adidas to leverage Data Science as a Service for informed decision-making.
Responsibilities:

- Collaborated with fellow data scientists and engineers to build an AI platform for Click-Stream analytics to predict lead generation, which could be used by Adidas's Product teams to leverage data science as a service.
- Developed and implemented an in-house custom algorithm for automated binning of data to increase accuracy and actionability of insights.
- Utilized a stack of feature selection techniques and applied machine learning models such as Logistic Regression, Decision Trees, and Random Forest with Grid Search to develop predictive models for lead generation.
- Conducted statistical tests such as KS-Test, AUC-ROC, Correlation, Chi-square, and Anova Test to validate model performance and accuracy.
- Orchestrated the adoption of the product for North the American sales team and provided insights andrecommendations based on the models generated to increase lead generation.

Technology Stack: Python, SQL, Django, Teradata

Title: Project Evergreen Churn Analytics
Client: CA North America Sales Director, Product Marketing team, Customer Success team.
Objective: Identify churn patterns in CA's account base and develop an early warning system to reduce customerchurn. Communicate actionable insights to the business to mitigate churn risk for the current account base, specifically targeting those due for renewals.
Responsibilities:

- Led end-to-end delivery of the Early Warning System project, from planning to deployment, ensuring timely and high-quality delivery.
- Conducted extensive data preparation from multiple sources, including the digital presence of a customer, contract history, support experience, and adoption metrics, for a data volume of 500K records.
- Developed a custom feature selection algorithm to reduce the number of features from 150 to 10 actionable features, improving the model's interpretability and accuracy.
- Developed custom algorithms for automated binning of numeric variables, increasing the actionability of the insights derived from the models.
- Applied various machine learning models, including Logistic Regression, Random Forest, Decision Trees, KNN-Classifier, and Naïve Bayes', and selected Logistic Regression based on its actionability.
- Conducted statistical tests, including the AUC ROC test with Decile-wise lift, recall at the threshold with Out-of-timetesting, and K-Fold Cross-validation to evaluate the model's performance.
- Deployed the models for North America's Tier-2 and Commercial Accounts, providing a powerful tool to mitigate churn and improve customer retention.

Technology Stack: Python, SQL Server, Web Framework Django, Models Built**:** Logistic Regression, DecisionTrees, GBM, Random Forest, Boost

Title: Intelligent Voice Assistant Solution for Seamless Reporting from Any Relational Database
Objective: Create a dynamic chatbot that leverages conversational interactions to retrieve data from SQL/NoSQL databases, allowing for easy configuration and seamless information retrieval.

- Develop an advanced chatbot using Natural Language Processing to efficiently retrieve information from
- SQL/NoSQL databases based on user conversations, reducing response time and increasing customer satisfaction.
- Designed and developed a highly configurable product that enables querying any structured database using conversational sentences within 1-2 days.
- Created a program that can automatically generate SQL queries based on the database schema.
- Implemented a solution that maps natural language conversations to corresponding queries and retrieves the relevant data from the database, without requiring the user to know querying languages.
- Continuously enhanced the product to support integration with NoSQL databases for improved functionality.
- Configured the database to seamlessly integrate with the chatbot, resulting in a dynamic and personalized experience for end-users.

Technology Stack: Google Dialogflow, NLP, MySQL, Jupyter Notebooks & Pycharm

Title: Develop AI-Powered Voice Analytics Agent for a Prominent Microfinance Company
Client: Asirvad Microfinance
Objective: Minimize the net default rate on payments, predict the likelihood of an individual defaulting on a loan payment, and identify the underlying reasons for defaulting.

- Developing a solution to extract vital information from real-time voice recordings in HINDI/SPANISH.
- Planning, architecting, and building end-to-end solutions for the project.
- Collaborating with the client to understand their requirements and charting a path for project implementation.
- Developing a solution to extract critical information from real-time voice recordings.
- Building the solution on AWS cloud, utilizing S3, Lambda functions, and DynamoDB as per client specifications.
- Developing and implementing the required APIs, while also identifying any new API requirements using the Django framework

Technology Stack: AWS, NLP, MySQL, Jupyter Notebooks & Pycharm

Title: Product Tagging Classification
Client: Shopify
Objective: The goal is to improve the effectiveness of Project Evergreen by assessing the impact of community presence on customer churn. The classification engine will utilize NLP techniques, including both syntax and semantics, along with machine learning and deep learning models, to perform the classification.

- Developed and implemented a robust text classification engine to accurately tag untagged questions based on their associated products in communities.ca.com.
- Applied NLP approaches, including both syntax and semantics, to understand the context and meaning of the questions and effectively classify them.
- Built and trained text classification models using various machine learning algorithms, leveraging Python, Keras, and TensorFlow.
- Utilized deep learning techniques to enhance the classification accuracy and performance of the models.
- Collaborated with cross-functional teams, including data scientists, software engineers, and product managers, to integrate the classification engine into Project Evergreen.
- Conducted thorough testing and validation of the classification engine to ensure its accuracy and reliability in real-world scenarios.

Technology Stack: Python, Machine Learning Frameworks: Keras, TensorFlow, NLP Libraries: NLTK, SpaCy, Jupyter Notebook, PyCharm, Version Control: Git, GitHub.

Virtusa Pvt Ltd, Hyderabad, India **July 2015- Dec 2018**

Title: GTM Models Development
Client: VMWare PVT LTD
Objective: The objective of the project is to define data-driven cross-sell prospects for existing customers of VMware, aiming to improve the conversion rate from opportunity to contract.
Responsibilities:

- Perform time-based cross-validation to divide the data into training and testing sets.
- Generate features from various data sources to enhance the model's predictive capabilities.
- Conduct exploratory data analysis (EDA) to gain insights and identify patterns in the data.
- Apply custom feature selection techniques to identify actionable features that have a significant impact on the cross-sell prospects.
- Utilize multiple machine learning algorithms and ultimately select logistic regression as the final model.
- Evaluate and measure the business impact of the GTM models, specifically in the NA region, by comparing the conversion rates from an opportunity to contract.

Technology Stack: Python, Postgres SQL

Other Projects:

- **ML WEB APP:** Developed an ML WebApp using Flask API backed by the Naïve Bayes model with 75% AUC for product marketing and customer success team which helped the team to understand the customer personas.

- **Chatbot Development:** Leveraged RASA and chatbot LUMA framework to develop a bot and deployed them on AWS EC2 instance for the Service desk manager to automate L1 & L2 service incidents which enhanced the business process by 40%

- **Active Learning in Churn propensity Model:** Built a churn propensity model for" Big 6 and Little 3" products at CA, which is actionable to prevent product churn, developed a custom ML pipeline that automated feature engineering, feature creation, and selection tasks with 100% accuracy and a real-time dashboard to track overfitting, underfitting, model and dataset drifts & convey causal features over time for product churn.

- **Statistical Analysis:** Performed various descriptive and inferential statistics and presented insights to the team the S&M team.

- **Customer acquisition modeling:** Built a propensity to MQL, SQL for B2B sales, leveraged campaigns & external intent signals from Bombora and zoom info data, Deployed the model with an actionable pitch to prospects which saw an incline in lead to conversion rate by 19%.

Leadership Experience

Shareholder's Value Creation

- Led the product strategy, roadmap, and implementation of analytics dashboards, providing a 360-degree view ofthe company KPI metrics in product, growth, operations, and sales, built a bespoke solution on KPI led B2B Subscription.
- Contributed to converting a ~500K $ RFP to further penetrate the account using existing business knowledge.
- Experience in driving strategic business decisions driven by quantitative data with great attention to detail.

Team Leadership

- Detail-oriented leader-built ramp-up roadmap for junior/senior data scientists to catalyst the onboarding time to the project.
- Collaborated with cross-functional leaders to design the plan of attack for multi-million $ RFPs on Analytics,AI\ML.
- Mentored 4 data scientists and analysts to increase productivity and unlock their potential. Led multiple cross-functional teams of Developers, Quality, and business Analysts.
- Implemented MLOPs for ml model monitoring, followed production-ready clean coding standards, with industrybest practices, active contributor to organizational tech events.

Achievements & Awards

- Received SPOT award from the ELT for demonstrating DS solutions for B2B SaaS subscription business.
- Built a B2B SaaS S&M data-driven transformational framework to elevate growth metrics by providing analytics-led interventions, this framework is followed by the data science practice for project execution & RFPs.

Educational Background

- Master's in data science from the University of North Texas.
- Bachelor of Technology in Computer Science & Engineering from JNTU-Hyderabad.