Suraj Karyamapudi

linkedin.com/in/suraj-karyamapudi | suraj.ai

Summary

Experienced software professional with 8+ years in artificial intelligence application development, including expertise in natural language processing (NLP), computer vision (CV), and machine learning (ML). Proven ability to design and implement predictive models, architect scalable data-driven solutions, and harness advanced analytics to drive impactful business insights. Adept at developing innovative AI-powered applications that transform data into strategic value, improve operational efficiencies, and deliver cutting-edge solutions in complex environments. Passionate about advancing AI technologies and dedicated to leading AI-driven strategies that drive growth, innovation, and meaningful change. Seeking a senior role to apply deep AI expertise in creating transformative, real-world solutions.

Technical Skills

Languages: Python, R, SQL, JavaScript
Developer and CI/CD Tools: VS Code, Jupyter, Docker, Jenkins
Libraries/Frameworks: Pandas, NumPy, SciPy, Matplotlib, TensorFlow, Keras, Scikit-learn, PyTorch
Cloud Computing: AWS, Azure
Databases: MySQL, PostgreSQL, OracleDB
APIs & Web Development: REST APIs, FastAPI, React, Flask
Data Visualization: PowerBI, Tableau

Certifications

- AWS Certified AI Practitioner
- AWS Certified Cloud Practitioner
- Microsoft 767: Implementing a Data Warehouse
- IBM Certified Academic Associate: DB2 Database and Application Fundamentals

Experience

Data Scientist Jul. 2020 – Oct. 2022

Prism Solutions International Pvt. Ltd. *Bengaluru, India*

- Developed and implemented machine learning models for sales forecasting, leveraging ARIMA and LSTM to predict demand trends and optimize inventory planning. Deployed on AWS using Lambda for real-time processing and RDS for scalable data storage, reducing operational inefficiencies.
- Built an anomaly detection system for financial transactions using Isolation Forests and Autoencoders, identifying irregularities in real-time transactions and reducing false positives for fraud detection, hence enabling timely identification of potential fraud.
- Developed customer behavior analysis models using Random Forests and Collaborative Filtering to predict churn, lifetime value, and purchase likelihood. These insights powered targeted marketing campaigns, improving customer retention and optimizing promotional strategies for higher ROI.
- Developed and implemented pattern recognition algorithms as part of a customer analytics project, using K-Means clustering and Apriori to analyze purchasing behavior. This identified key customer segments and product affinities, enabling cross-selling opportunities and personalized marketing strategies that increased customer engagement and boosted revenue.
- Established a CI/CD pipeline on AWS using Jenkins and Docker, enabling automated model retraining, testing, and deployment. Integrated FastAPI for backend services and JavaScript for the frontend, supporting real-time data processing and seamless user interaction. Leveraged MLOps practices, including monitoring and version control, to ensure efficient deployment, continuous improvement, and scalability of machine learning models.
- Developed a Streamlit dashboard as part of the sales forecasting and anomaly detection projects, providing real-time model insights and enabling non-technical teams to interact with predictive outputs. This tool enhanced decision-making by offering intuitive visualizations and streamlined access to actionable data.

- Deployed scalable AI workflows on AWS and Azure, leveraging AWS RDS for structured data storage and Azure Blob Storage for managing unstructured data. This multi-cloud approach enabled efficient handling of millions of data rows, ensuring scalability, seamless data processing, and high-performance analytics.
- Led cross-functional teams to provide strategic insights and advanced AI-driven data solutions, aligning machine learning and deep learning outcomes with client business goals.

Data Scientist Jul. 2017 – Jun. 2020

Abhyudaya HR Business Solutions *Hyderabad, India*

- Built a predictive employee attrition model using Random Forests and Gradient Boosting, analyzing factors like job satisfaction, engagement, and work environment to forecast potential turnover allowing HR teams to identify high-risk employees and implement targeted retention programs.
- Enhanced the candidate-matching system developed during my junior role by integrating additional data sources and more complex features, such as skills similarity and role progression. Utilized advanced NLP techniques like TF-IDF to analyze resumes and job descriptions, improving recruitment efficiency by over 30%.
- Upgraded the anomaly detection system for payroll using Autoencoders and Isolation Forests, optimized to handle large-scale payroll datasets and uncover complex irregularities. Improved detection accuracy by reducing false positives, enabling timely identification of errors and mitigating compliance risks.
- Conducted behavior analysis for employee retention and performance by segmenting employees based on behavioral patterns using clustering and classification algorithms. Created visualizations and dashboards to highlight these insights, allowing HR to implement data-backed retention strategies and targeted training programs.
- Built a performance prediction model using Random Forests, enabling HR to identify high-potential employees for promotions and training opportunities. Leveraged historical performance and behavioral data, achieving precision and recall metrics of 90%, and integrated predictions into HR dashboards for actionable insights.
- Created interactive dashboards using Tableau, visualizing workforce trends and key metrics for real-time HR decision-making.
- Worked with HR teams to implement automated solutions, streamlining recruitment, retention, and performance evaluation processes.

Junior Data Scientist Jul. 2016 – Jun. 2017

Abhyudaya HR Business Solutions *Hyderabad, India*

- Developed a foundational candidate-matching model using Logistic Regression and Random Forests, analyzing skills and historical job success rates to optimize role-to-candidate alignment. Enhanced recruitment efficiency, reducing time-to-hire by leveraging data-driven insights for better candidate selection.
- Built an interactive HR dashboard using Tableau to visualize key metrics such as employee turnover, recruitment efficiency, and departmental performance, enabling management to monitor HR performance in real time.
- Designed and implemented a payroll anomaly detection algorithm using statistical methods such as Z-Score Analysis and k-Means clustering to flag discrepancies and potential errors for review and ensure accurate payroll processing.
- Conducted data cleaning and pre-processing on large unstructured HR datasets, removing inconsistencies, handling missing values, and standardizing data formats, leading to high-quality structured datasets ready for analysis.
- Tested multiple machine learning models (Logistic Regression, Decision Trees) for use in recruitment analysis, recommending optimal approaches based on predictive accuracy.
- Collaborated with senior data scientists to enhance models and streamline HR operations, contributing to improvements in recruitment and payroll processes.

Software Developer - Full Stack Mar. 2014 – Jun. 2016

Wave9 IT Solutions *Hyderabad, India*

- Developed a high-traffic sales dashboard application, leveraging JavaScript for the frontend and Python with Flask API for backend services.
- Integrated Oracle Database for robust data storage solutions, handling large datasets efficiently to support business analytics and sales tracking.
- Utilized AWS for scalable deployment and maintenance, ensuring consistent application uptime and performance.
- Contributed to backend and frontend integration, ensuring seamless data flow and user experience across modules.
- Employed best practices in RESTful API design and implemented security standards for data protection.
- Collaborated closely with cross-functional teams to enhance features and ensure application reliability and scalability.

Research Freelance Projects

Customer Support Agent Using OpenAI and Azure ML | *LLMs, RAG, Azure ML, LangChain, FAISS, Streamlit*

- Developed a customer support agent using OpenAI's GPT-4 and Azure ML to handle ticket categorization, sentiment analysis, and automated response generation, addressing scalability and reducing operational costs for high-volume support environments.
- Configured Azure ML workspace and set up a vector database (FAISS) to store and retrieve embeddings efficiently, enabling real-time retrieval and response generation based on customer query embeddings.
- Leveraged graph based Retrieval-Augmented Generation (RAG) using Neo4j to combine retrieval-based and generation-based approaches, using OpenAI embeddings to generate contextually accurate responses for diverse customer support scenarios.
- Designed a feedback loop for ongoing prompt and response refinement, improving the model's performance over time based on real-world interactions.
- Built a modular Streamlit UI to demo the support agent's capabilities, allowing users to view multiple response options and select the best one, enhancing the end-user experience.
- Deployed the solution on Azure ML for scalable operation, ensuring seamless integration with the existing support infrastructure.

Creating An Image Search Engine Using Deep Learning | *Python, Computer Vision, CNN, Deeplearning*

- Built an Image Search Engine using VGG16 with Transfer Learning, extracting 512-dimensional feature vectors and applying Cosine Similarity to identify similar products from a 300+ image dataset.
- Streamlined feature extraction by reducing each image to a single 512-length vector using Global Average Pooling, improving search efficiency and accuracy.
- Implemented a scalable search solution using nearest neighbor algorithms, enabling dynamic updates to the feature store for new or removed products.
- Recommended model and metric enhancements like testing ResNet, Inception, and alternative distance metrics to further refine search relevance and performance.

Enhancing Targeting Accuracy Using ML | *Python, Machine Learning, Classification, Customer Targeting*

- Developed a targeted customer prediction model using Random Forests, achieving a classification accuracy of 93.5%, Precision of 88.7%, and F1-Score of 89.5%, enabling cost-effective marketing for a grocery retailer's delivery club campaign.
- Analyzed and pre-processed customer data from three tables, including demographic and transaction data, and handled class imbalance using Precision, Recall, and F1-Score for model evaluation, improving recall to 90.4%.
- Tested multiple classification algorithms (Logistic Regression, Decision Tree, Random Forest, KNN) and selected Random Forest based on consistent performance across metrics and feature importance insights.
- Proposed further improvements by exploring additional models like XGBoost and LightGBM, hyperparameter tuning, and enhancing feature engineering for deeper insights and increased predictive accuracy.

Education

National College of Ireland — Dublin, Ireland
Master of Science in Data Analytics — *October 2023*

Jawaharlal Technological University — Andhra Pradesh, India
Bachelor of Technology in Electronics & Communication Engineering — *November 2012*