Python Developer

SAI SUNANDH RAMAYANAM
Phone: +1 9285502812
Email: ramayanamsaisunandh@gmail.com
LinkedIn : https://www.linkedin.com/in/sunandh-r/

PROFESSIONAL SUMMARY:

- Highly skilled and results-driven Python Developer and Data Engineer with over six years of experience in building scalable ETL pipelines, AI/ML applications, and cloud-based solutions. Specializing in big data processing, deep learning, and automation, with expertise in Python, SQL, PySpark, AWS, and distributed computing.
- Adept at designing end-to-end data pipelines, managing cloud infrastructure, and developing AI-powered solutions that enhance business operations and decision-making.
- Proficient in developing ETL workflows using Apache Airflow, extracting and processing raw data from AWS S3, and transforming it for storage in SQL (RDS, PostgreSQL, MySQL) and NoSQL databases.
- Hands-on experience in PySpark for large-scale data processing and implementing feature engineering techniques to extract valuable insights for AI/ML models. Strong expertise in AWS services, including IAM for access control, S3 for scalable storage, Lambda for automation, SNS for event-driven notifications, and CloudWatch for monitoring and performance optimization.
- Deep learning expert with experience in PyTorch, TensorFlow, and Scikit-learn, developing and fine-tuning neural networks for predictive modeling, NLP, and time-series forecasting. Built and optimized CNNs, RNNs, Transformers, and generative AI models for various AI-driven applications, improving accuracy through hyperparameter tuning, transfer learning, and real-time inference. Skilled in deploying AI models using Flask, FastAPI, and ONNX, ensuring seamless integration into production environments.
- Extensive experience in SQL database design and query optimization, creating schemas, managing table structures, and executing complex queries for analytics. Implemented big data frameworks such as Apache Spark, Hadoop, and Kafka for real-time data streaming, ensuring efficient handling of structured and unstructured data.
- Strong cloud computing background, leveraging AWS, Azure, and GCP for developing cloud-native applications. Implemented CI/CD pipelines using Jenkins, GitHub Actions, and Terraform to automate testing, deployment, and infrastructure provisioning. Expertise in containerization (Docker, Kubernetes) and serverless computing (AWS Lambda, Azure Functions) for scalable AI deployments.
- Experienced in SQL Server Integration Services (SSIS) for designing and managing ETL workflows, automating data extraction, transformation, and loading processes for large-scale enterprise applications. Proficient in SQL Server Reporting Services (SSRS) for creating interactive reports, dashboards, and data visualizations to support business intelligence and decision-making. Skilled in SQL Server Analysis Services (SSAS) for building and optimizing multidimensional and tabular data models, enabling efficient data analysis and performance tuning for complex analytical queries.
- Passionate about AI model validation and testing, using cross-validation techniques, A/B testing, and performance metric evaluations to ensure accuracy and reliability. Conducted extensive monitoring and debugging using AWS CloudWatch, X-Ray, and logging frameworks to optimize pipeline performance and model efficiency.
- Adaptable and proactive in Agile and DevOps environments, collaborating with cross-functional teams, participating in Scrum meetings, sprint planning, and retrospectives, and managing workflows using Jira and Confluence. Committed to continuous learning, actively exploring Generative AI, fine-tuning LLMs, and integrating AI-powered features into business applications.
- Seeking opportunities to leverage expertise in Python, AI/ML, ETL, cloud engineering, and full-stack development to build intelligent, data-driven solutions that drive business efficiency, innovation, and technological advancements.

TECHNICAL SKILLS

Skillset	Skills
Programming Languages	Python (3.12), C, C++, Java
Machine Learning & Deep Learning	Scikit-learn, PyTorch, TensorFlow, XGBoost, LightGBM, YOLO, EfficientNet, OpenCV
Generative AI & LLMs	GPT, BERT, LLaMA, T5, Diffusion Models
Cloud Computing & MLOps	Google Cloud Platform (GCP), Amazon Web Services (AWS), Microsoft Azure, Vertex AI, BigQuery, SageMaker, MLflow, Kubeflow, TensorBoard
Data Engineering & ETL	Apache Spark, Kafka, Airflow, Pandas, NumPy, SciPy

DevOps & Automation	Docker, Kubernetes, Terraform, CI/CD (Cloud Build, GitHub Actions, Jenkins)
Database Management	SQL, MySQL, PostgreSQL, BigQuery, MongoDB, Firebase, Apache Cassandra
Statistical Analysis & Data Science	Data wrangling, Hypothesis testing, Feature engineering, Time-series forecasting
API Development & Backend	Flask, FastAPI, Django, RESTful APIs, OAuth2, JWT authentication
Visualization & Reporting	Matplotlib, Seaborn, Plotly, Tkinter, Google Data Studio, R's Shiny
Security & System Monitoring	IAM roles, Access control, GCP Cloud Monitoring, Logging tools
Agile Development & Collaboration	Agile methodologies, Jira, Confluence, Git
Emerging AI Trends & Innovation	Federated learning, AutoML, Reinforcement learning, Responsible AI

CERTIFICATIONS

- AWS Certified Data Engineer - Associate: Demonstrated expertise in architecting, deploying, and managing scalable data solutions using AWS services like S3, Redshift, Glue, and RDS.
- Microsoft Power Platform Fundamentals: Proficient in designing business solutions and workflows using Power Apps, Power BI, and Power Automate for actionable insights and process automation.
- Google Cloud Data Engineer Certification (Target Certification): Skilled in utilizing GCP tools such as BigQuery, Cloud Dataflow, and Vertex AI for scalable and efficient data processing and analytics.
- DevOps Specialist: Advanced proficiency in CI/CD pipeline creation using Jenkins, Cloud Build, and Infrastructure as Code (IaC) practices to ensure seamless application deployment and scalability.
- Python Data Science & Machine Learning Certification: Deep expertise in Python for data manipulation, statistical analysis, and machine learning, including libraries like Scikit-learn, Pandas, and NumPy.
- PySpark Specialist Certification: Advanced skills in distributed data processing, building scalable ETL pipelines, and executing large-scale data transformations using PySpark.
- Machine Learning & AI Specialist: Proven ability in designing, training, and deploying machine learning models for classification, regression, and clustering, with experience in hyperparameter tuning and evaluation.
- SQL Certification: Mastery in querying, managing, and analyzing structured data in relational databases, ensuring high performance and data integrity.
- JupyterLab Data Analytics & Workflow Automation Certification (Self-taught): Expert in using JupyterLab for data analysis, pipeline development, and visualizing workflows for reproducibility and collaboration.
- Robotic Process Automation (RPA) Practitioner: Skilled in automating repetitive tasks and improving operational efficiency with RPA tools.
- IoT & Data Analytics: Hands-on experience with IoT sensor integration, real-time data analysis, and designing solutions for interconnected systems.
- Web Development (Django): Expertise in developing scalable web applications using Django, with strong backend and front-end integration capabilities.

WORK EXPERIENCE

PYTHON DEVELOPER || Averosoft June 2023 – Present

- Designed, developed, and deployed end-to-end machine learning models for real-world applications such as fraud detection, recommendation systems, predictive analytics, and NLP-driven automation.
- Collaborated with stakeholders, including business leaders, medical professionals, and engineers, to understand and define objectives for AI-driven solutions such as disease prediction, facial recognition, and structural similarity analysis.
- Translated business and healthcare needs into AI models, determining the feasibility of different computer vision approaches. Worked across multiple domains, including medical imaging, fraud detection, quality control, and business analytics, ensuring solutions align with industry requirements.
- Gathered large-scale image datasets from various sources, including medical scans, surveillance footage, and business-related images. Cleaned and processed images using OpenCV, scikit-image, and NumPy, applying transformations such as noise reduction, image segmentation, morphological operations, histogram equalization, edge detection, and feature extraction.
- Used data augmentation techniques like flipping, rotation, scaling, and brightness adjustments to improve model generalization. Applied dimensionality reduction techniques such as PCA to enhance computational efficiency.

- Designed and implemented machine learning models using supervised learning algorithms, including Support Vector Machines (SVM), Random Forests, Decision Trees, and Gradient Boosting for classification and regression tasks in business and medical applications.
- Developed deep learning architectures using Convolutional Neural Networks (CNNs) and Transformer-based models for complex image processing tasks such as object detection, segmentation, and disease classification.
- Utilized unsupervised learning methods such as K-Means clustering, Autoencoders, and Isolation Forest for anomaly detection, fraud prevention, and pattern recognition. Conducted structural similarity analysis using SSIM and perceptual hashing to compare images for quality control, duplicate detection, and validation in different industries.
- Optimized model performance using hyperparameter tuning techniques, including Grid Search, Random Search, and Bayesian Optimization, ensuring high accuracy and efficiency.
- Leveraged distributed computing frameworks such as PySpark to handle large-scale datasets and accelerate training processes. Developed and deployed machine learning models as APIs using Flask and FastAPI, integrating them with real-world business and healthcare applications for seamless real-time inference.
- Monitored model performance using MLOps tools such as MLflow and TensorBoard, tracking key metrics and detecting model drift for continuous model improvement. Designed automated retraining workflows to ensure AI solutions remain accurate and relevant.
- Implemented explainability frameworks such as SHAP and LIME to enhance transparency and trust in AI-driven decision-making, particularly in critical applications like healthcare diagnostics and financial fraud detection.
- Researched and integrated Generative AI models, including GANs, VAEs, and Diffusion Models, to improve data augmentation, enhance model accuracy, and expand AI capabilities. Explored optimization techniques such as model quantization, pruning, and compression to improve inference speed and enable AI deployment across cloud, edge, and mobile environments.
- Engineered scalable data pipelines and ETL workflows using Apache Spark, Kafka, and Airflow, ensuring seamless data ingestion, transformation, and model retraining for big data applications.
- Developed AI-powered automation solutions for customer support, finance, and healthcare, including speech-to-text processing, sentiment analysis, and autonomous decision-making.
- Leveraged transfer learning and fine-tuned pre-trained models to accelerate model deployment for specialized applications in medical imaging, geospatial analysis, and industrial defect detection.
- Automated AI/ML deployment workflows using CI/CD pipelines with GitHub Actions, Jenkins, and Docker, ensuring fast and reliable updates to production environments.
- Worked on federated learning architectures to train ML models across decentralized data sources while preserving privacy in industries like healthcare, fintech, and cybersecurity.
- Created interactive AI-driven analytics dashboards using Streamlit and Dash, enabling business users to visualize model predictions and insights in real-time.
- Designed multi-modal AI solutions integrating text, image, and audio processing for smart assistants, AR/VR applications, and AI-powered content creation.
- Actively engaged in AI research, open-source contributions, and community knowledge sharing, staying ahead of emerging trends in AI ethics, responsible AI, and next-gen deep learning architectures.

SOFTWARE ENGINEER || Fresenius Kabi (Accenture) February 2020 – December 2022

- Developed and maintained automated ETL pipelines using Python, SQL, and Apache Airflow to efficiently process and manage large-scale datasets. Established secure connections to AWS S3 using Python for data extraction, leveraging Pandas for transformation and storing processed data in Excel, CSV, or SQL/RDS databases.
- Designed and implemented SQL databases, creating schemas and tables for structured data storage and retrieval.
- Utilized AWS IAM to manage secure access to AWS resources, ensuring proper authentication and authorization for data pipelines. Integrated AWS Lambda and AWS SNS to automate notifications upon new data availability in S3, triggering necessary ETL workflows in real time.
- Employed AWS CloudWatch for monitoring and logging data pipeline performance, setting up alerts for failures or performance bottlenecks.
- Processed and queried large datasets using PySpark, optimizing performance and scalability for big data workloads. Integrated Apache Airflow for automated scheduling, monitoring, and execution of data transformation and loading tasks.
- Conducted feature engineering and correlation analysis to extract meaningful insights, improving the predictive accuracy of machine learning models.
- Developed and trained deep learning models using PyTorch, leveraging neural networks for advanced predictive analytics. Fine-tuned machine learning models through hyperparameter optimization and efficient training strategies.

- Ensured robust and scalable AI model deployment by validating accuracy and optimizing performance for real-world applications.
- Optimized data pipelines for efficiency, ensuring reduced latency and improved data processing reliability. Implemented AWS Step Functions to coordinate multi-step workflows and ensure seamless execution of ETL processes.
- Validated and tested AI models using cross-validation techniques and performance metrics to guarantee accuracy. Conducted extensive debugging and monitoring using AWS CloudWatch and AWS X-Ray to enhance operational visibility and performance tuning.

Research Assistant || Undergraduate December 2018 – January 2020

- Preprocessed the health dataset by implementing data cleaning and transformation workflows in JupyterLab using Python, ensuring the dataset was accurate, consistent, and ready for machine learning applications.
- Conducted exploratory data analysis (EDA) in JupyterLab using Python libraries such as Pandas, Matplotlib, and Seaborn to uncover trends and correlations within the dataset, providing a foundation for model development.
- Designed and implemented multiple machine learning models using Scikit-learn, including logistic regression, decision trees, random forests, and support vector machines, to predict diabetes based on patient health data.
- Performed hyperparameter tuning using grid search and cross-validation techniques in Scikit-learn to improve model accuracy and reduce overfitting, ensuring reliable predictions.
- Evaluated model performance by calculating key metrics such as precision, recall, F1-score, and ROC-AUC, providing a comprehensive understanding of model effectiveness.
- Automated the end-to-end machine learning pipeline in JupyterLab, integrating steps such as data preprocessing, model training, validation, and evaluation into a seamless workflow.
- Visualized model predictions and feature importance using Python libraries like Matplotlib and Plotly, creating intuitive charts and graphs to effectively communicate findings to the professor and research team.
- Implemented real-time error tracking and debugging for the machine learning workflow in JupyterLab, ensuring the reliability of scripts and minimizing computational errors during execution.
- Collaborated with the professor to document all stages of the research process, including dataset preparation, model development, evaluation, and result interpretation, enabling reproducibility of the research.
- Regularly tested and iterated on machine learning models, refining the pipeline to ensure optimal accuracy and performance for diabetes prediction using Scikit-learn and other essential Python libraries.

DATA ANALYST INTERN || COSMO FERRITES April 2018 – October 2018

- Collaborating with stakeholders to understand their needs around financial data, including revenue, purchase orders, scheduled deliveries, and completed orders.
- With clear requirements in hand, I dive into data extraction, ensuring seamless integration from various sources like databases, spreadsheets, and external APIs using Python tools like Pandas and SQL.
- Once the data is consolidated, I preprocess and clean it, resolving missing values, detecting anomalies, and standardizing formats to ensure accuracy and consistency. Leveraging libraries like NumPy and Pandas, I implement advanced calculations and derive key performance metrics such as monthly revenue growth, delivery efficiency, and order fulfillment rates.
- Moving forward, I use Matplotlib and Seaborn to visualize insights. For instance, I create dashboards highlighting revenue trends, delivery timelines, and order statuses, allowing stakeholders to identify bottlenecks and areas of opportunity. These insights empower teams to make data-driven decisions, improving operational efficiency.
- To ensure long-term scalability, I design and implement automation workflows, integrating data pipelines to schedule regular updates and processing of financial and operational data. Tools like Apache Airflow or Python's schedule library help automate these processes.
- Additionally, I write reusable code modules and document the workflows thoroughly, enabling seamless handoffs or future enhancements by other teams.
- Through these efforts, I deliver robust, scalable, and insightful software solutions that enable organizations to unlock the full potential of their data.