### Venkata Sainath

**Email:** [**venkatasainathreddyp@gmail.com**](mailto:venkatasainathreddyp@gmail.com)

**Phone no: +1 (346) 409-5168**

**LinkedIn:** [**https://www.linkedin.com/in/venkatasainathp**](https://www.linkedin.com/in/venkatasainathp)

***With over six years of experience in machine learning, NLP, and data engineering, I specialize in deploying advanced AI models like GPT and BERT to deliver impactful solutions. I excel at building real-time ML models, scalable data pipelines, and AI-driven systems, leveraging AWS, Azure, Docker, and Kubernetes for optimal performance. Passionate about innovation, I bridge AI with business strategy, creating user-focused solutions that drive growth and solve complex challenges.***

# PROFESSIONAL SUMMARY:

* Developed and deployed real-time machine learning models using Python, TensorFlow, and PyTorch, optimizing fraud detection and recommendation systems.
* Built and implemented NLP pipelines with BERT, GPT, and Generative AI for enhanced sentiment analysis and text classification.
* Designed scalable backend systems using Python, Django, and microservices, along with frontend work using React, Redux, and TypeScript.
* Deployed AI models and microservices with Docker and Kubernetes on cloud platforms like AWS and Azure, ensuring high availability and performance.
* Created real-time data pipelines using Databricks, Apache Kafka, and PySpark for high-throughput applications.
* Implemented AI-driven recommendation engines and fraud detection systems, utilizing Cassandra, MongoDB, and Redis for real-time data processing.
* Automated CI/CD pipelines with Jenkins, GitLab, and GitHub Actions for seamless model deployment and zero-downtime updates.
* Developed secure RESTful APIs and gRPC services with OAuth2 for real-time applications.
* Applied data analysis and predictive modeling using Python, Pandas, and NumPy, and built interactive dashboards for business insights.
* Participated actively in Agile development processes, contributing to code reviews, pair programming, and ensuring efficient project delivery.

# TECHNICAL SKILLS:

|  |  |
| --- | --- |
| **Machine Learning/NLP** | GPT architectures, BERT, Natural Language Generation (NLG), Transfer Learning, Supervised/Unsupervised Learning, Reinforcement Learning, Generative AI, RAG models |
| **Data Handling** | Pandas, NumPy, Data Extraction, Transformation, Loading (ETL), Databricks, PySpark, Dask, AWS Lambda, Azure Data Factory, Cassandra, MongoDB |
| **Web Development Frameworks** | Django, Flask, Node.js, React (Hooks, Functional Components), Redux, TypeScript, Clojure, Rust, HTML5, CSS, Bootstrap, JavaScript, WebSockets |
| **Database Management** | MySQL, PostgreSQL, Oracle DB, SQL Queries, Stored Procedures, Cassandra, Redis, Azure Cosmos DB, AWS RDS |
| **API Development** | RESTful APIs, Django Rest Framework, Flask APIs, gRPC, OAuth2, Azure API Management |
| **Cloud Platforms** | Amazon Web Services (AWS), Microsoft Azure (Cosmos DB, Azure SQL, Azure DevOps, Azure App Services, Azure AKS), Docker, Kubernetes, AWS ECS |
| **Software Development Practices** | Test-Driven Development (TDD), Agile Methodology, Microservices Architecture, CI/CD Pipelines (Jenkins, GitHub Actions, GitLab CI, Azure DevOps), Scrum |
| **Scripting and Automation** | Python OS module, Bash/Shell Scripting, Linux/Unix Scripting, Ansible, Terraform, CloudFormation, AWS Lambda, APScheduler, Celery |
| **Data Analysis & Visualization** | Matplotlib, Seaborn, Bokeh, Elasticsearch, Kibana, OpenCV, RPA, Tableau, Predictive Analytics, Data Mining, Databricks |
| **Version Control** | Git, GitHub, GitLab, Bitbucket |
| **Security & Optimization** | Application Security Protocols (OAuth2, TinyML for Threat Detection), Performance Optimization (Redis Cache, Azure Active Directory), SEO |
| **UI/UX Development** | React, Redux, Bootstrap, jQuery UI, Responsive Web Design, User Interface Development |
| **Statistical Modeling** | Statistical Data Analysis and Modeling, Scikit-Learn, TensorFlow, PyTorch |
| **Testing & Automation** | PyTest, Selenium, JEST, Automated Testing (Unit, Integration, Functional, Regression, UAT, API Testing), CodeQL, Postman |
| **Real Time Processing** | Apache Kafka, Redis, Prometheus, Grafana, PySpark, OpenCV |
| **Containerization & Orchestration** | Docker, Kubernetes, AWS ECS, Azure AKS, AWS Lambda, Azure App Services |
| **CI/CD Tools** | Jenkins, GitHub Actions, GitLab CI |
| **Edge Computing** | TinyML, TensorFlow Lite, Edge-based Security Optimization, IoT Security |
| **Automation & Configuration Management** | Ansible, Terraform, AWS CloudFormation, Jenkins, GitLab CI, GitHub Actions |

**PROFESSIONAL EXPERIENCE:**

**Client: ABS wavesight April 2023-present**

**Location: Spring, Texas.**

## Role: Machine Learning Engineer

* Developed and deployed real-time machine learning models using **Python**, **TensorFlow**, **PyTorch**, and **Databricks**, optimizing data pipelines for high-throughput applications like fraud detection and recommendation systems, leveraging OpenCV for real-time image and video data analysis.
* Built and deployed advanced **NLP** pipelines with **BERT**, **GPT**, and **Generative** **AI** for real-time sentiment analysis and text classification on live platforms, significantly improving context-aware response generation.
* Architected Retrieval-Augmented Generation (**RAG**) models utilizing **LLMs** and **Gen AI** for real-time decision-making systems, enhancing the accuracy and relevance of natural language generation in live AI-driven applications.
* Integrated real-time **AI/ML** models with **RESTful APIs** and **gRPC**, reducing inference latency and utilizing **OpenCV** for real-time image processing and anomaly detection in production environments.
* Implemented scalable data pipelines with Databricks and **PySpark**, optimizing real-time data ingestion, transformation, and processing for machine learning workflows, improving efficiency in high-demand systems.
* Engineered end-to-end machine learning workflows using **Python, Node.js, Clojure, and Databricks,** deploying AI models Neo4j into production with improved response times and seamless integration in real-time operational systems.
* Developed real-time AI-based recommendation engines using deep learning models and collaborative filtering, optimizing user experience by processing high-frequency data streams in live e-commerce platforms.
* Optimized machine learning deployment pipelines using **Docker, Kubernetes**, and **Databricks,** ensuring scalable containerized environments for real-time model inference and continuous deployment with failover support.
* Designed secure, real-time **RESTful APIs** for AI services with **OAuth2,** enabling low-latency data processing for fraud detection and financial analytics in live environments.
* Developed scalable **data pipelines** in **AWS** **(S3, Redshift)**, integrating real-time machine learning models for chatbot applications and VR-based damage detection on marine ships, ensuring low-latency responses.
* Designed real-time data architectures using **Cassandra** and **MongoDB**, optimizing sharding and replication strategies to handle high-velocity distributed data with low-latency reads/writes.
* Integrated Redis for real-time caching and messaging, improving performance in high-traffic environments, supporting pub/sub messaging, and enabling instant access to data for real-time processing.
* Built real-time **CI/CD** pipelines with **Jenkins** and **GitLab**, automating the build, test, and deployment processes for machine learning models, ensuring rapid, zero-downtime deployments in production.
* Deployed and scaled containerized applications using **Docker** and **Kubernetes**, orchestrating real-time, zero-downtime deployments and implementing rolling updates to support mission-critical applications.
* Built and optimized real-time **data models** in Databricks and **PySpark**, handling large-scale data ingestion for location tracking devices using **TinyML** models in live environments.
* Implemented real-time security protocols using **Tiny-ML** algorithms for lightweight, on-device threat detection and anomaly detection in IoT environments, optimizing security for edge-based systems.
* Built data pipelines integrating **Snowflake** to handle high-volume IoT data, ensuring efficient data storage and retrieval for real-time analytics and recommendation systems.
* Designed and implemented advanced data models for IoT trackingdevices using **dimensional modeling** techniques to optimize real-time queries and reporting.
* Combined **LangChain** with **OpenCV** and **Kafka**, enriching image processing workflows by dynamically sourcing relevant data for real-time AI-driven decision-making, optimizing latency and accuracy in production environments.
* Optimized recommendation systems using Python with custom **Breadth-First Search** and **A\* algorithms**, Databricks, and Linux shell scripting for high-speed data processing.
* Enhanced MySQL performance with advanced indexing and **Linux-based automation** for database maintenance and backups.
* Architected and deployed distributed data pipelines using **Apache Spark** and **Databricks**, processing real-time IoT location tracking data with TinyML algorithms, and optimizing workflows for anomaly detection and live data streams.
* Integrated **Apache Spark** with **Airflow** for real-time VR data processing, automating defect detection in marine shipping containers, while ensuring seamless model retraining and scalable deployments across production environments.
* Integrated OCR to extract and classify text from scanned judiciary documents, fine-tuning models to reduce false positives and improve compliance.
* Integrated **GitHub** with **Python** **microservices**, automating CI/CD pipelines and enforcing code quality via **CodeQL** scans and custom build checks for real-time **AI/ML** model deployments.
* Actively participated in **agile** **scrum** ceremonies including PI planning, collaborating with the data team for code reviews and pair programming, ensuring efficient and cohesive development of AI/ML systems in production environments.

**Tools and Technologies:** Python, Node.js, Clojure, TypeScript, Bash/Shell Scripting, TensorFlow, PyTorch, Databricks, Airflow, OpenCV, BERT, GPT, Gen AI, RAG, PySpark, AWS (S3, Lambda, Redshift, Glue, Athena, SageMaker, EC2, EMR), Docker, Kubernetes, Snowflake, Dimensional Modeling, Databricks, GitHub Actions, Jenkins, Langchain, GitLab CI/CD, Apache Spark, Cassandra, MongoDB, Redis, RESTful APIs, gRPC, OAuth2, Jenkins, GitLab, GitHub Actions, CodeQL, Prometheus, Grafana, Tiny-ML, React, Redux, WebSockets, , React Dev Tools.

**Client: Navy Federal Credit union. Oct 2022 - April 2023**

**Location: Vienna, Virginia**

## Role: Sr. Python Developer.

* Built and maintained responsive UIs using **React** with **hooks, Redux, and TypeScript**, ensuring clean code, reusability, and efficient state management for real-time banking portals.
* Automated unit tests using **JEST** and integrated **React Dev Tools**, optimizing performance, debugging, and ensuring high test coverage for responsive financial applications.
* Engineered secure backend systems using **Python, Django**, and **Asyncio**, building scalable **REST** **APIs** to manage real-time transactions, user authentication, and regulatory compliance.
* Deployed microservices using **Rust**, **Clojure**, and **gRPC**, ensuring low-latency communication for payment gateways, fraud detection, and account management in banking systems.
* Orchestrated banking services with **Kubernetes** and **Docker**, automating container scaling during peak traffic, ensuring zero downtime for online banking and loan systems.
* Integrated real-time machine learning pipelines with **Python**, Kubernetes, and **MLOps**, continuously deploying updated models for fraud detection and credit scoring in production.
* Optimized financial data handling using **Azure Cosmos DB** and **SQL** Database, managing high-volume, low-latency transactions for real-time financial operations and reporting.
* Developed and optimized relational data models using **ER diagrams, UML,** and **normalization techniques** to ensure real-time data consistency and seamless transaction processing in dynamic, high-volume banking systems.
* Configured multi-region infrastructure with **Terraform**, **Ansible,** and **Azure DevOps**, automating real-time deployments for critical banking services with high availability.
* Built secure API management using **Azure API Management** and **Active Directory**, enabling real-time authentication, rate-limiting, and secure access for financial APIs.
* Automated API and end-to-end tests using **Postman**, **PyTest**, and **Selenium**, ensuring real-time verification and seamless deployments in production banking environments.
* Developed real-time ETL pipelines using **Azure Data Factory** and **Databricks**, transforming financial data for fraud monitoring and regulatory compliance.
* Developed fraud detection models using **A\*** search with Python on Linux, integrating with real-time financial data pipelines via **Apache Kafka**.
* Developed ETL workflows in **Azure Data Factory (ADF)** for financial data ingestion and transformation, improving fraud detection accuracy with real-time updates.
* Enhanced existing data models for financial transactions, ensuring schema normalization to maintain consistency and support real-time decision-making.
* Automated **ETL** processes on Linux with MySQL and shell scripting, achieving efficient financial data ingestion and processing.
* Leveraged machine learning frameworks like **Scikit-learn** and **TensorFlow**, alongside feature engineering and data schema design, to create instructed data models for real-time fraud detection and risk assessment in financial transactions.
* Utilized **OpenCV** and **RPA** for real-time document verification and fraud detection, automating identity checks in banking workflows.
* Built AI models to detect and anonymize **PII** in banking transactions, ensuring real-time compliance with data protection regulations.
* Developed real-time AI models using **Azure Text Classification** and **NER** for secure legal document processing, automating PII detection and redaction.
* Developed real-time fraud detection pipelines using **Apache Spark** on **Databricks** and **Azure Synapse**, processing millions of banking transactions, fine-tuning Spark jobs, and integrating systems to enhance detection speed and accuracy.
* Implemented **OCR** for loan document automation, streamlining data extraction and reducing manual processing times for faster approvals.
* Integrated **GenAI** and **Azure Data Factory** to deploy AI-driven personalized financial services, enhancing customer interactions with real-time data.
* Orchestrated **Spark-based ETL** workflows with **Apache Airflow** and **Azure Data Factory**, automating real-time data ingestion, model retraining, and CI/CD for fraud detection models, ensuring seamless deployment and monitoring on the Azure platform.
* Automated repetitive banking tasks using RPA, Python, and **Linux Shell Scripts**, streamlining loan processing, customer onboarding, and backend operations.
* Deployed **Apache Kafka** for real-time event streaming, processing high-volume banking transactions and fraud alerts with low-latency data pipelines.
* Deployed **LangChain** with **Cosmos DB** and **Databricks**, enabling real-time fraud detection through dynamic transaction data retrieval. Integrated **LangChain’s** modular API with Kubernetes for rapid updates and model scaling in high-demand banking services.
* Automated **CI/CD** pipelines for Python microservices using **GitHub**, optimizing deployments for real-time banking applications with code quality enforced through **CodeQL** scans and custom build checks.
* Collaborated with cross-functional teams during **agile** ceremonies, including PI planning and sprint retrospectives, driving efficient delivery of real-time banking systems through code reviews and pair programming.

**Tools and Technologies:** Python,React (Hooks, Functional Components), Redux, Thunk, TypeScript, Bootstrap, Storybook, JEST, React Dev Tools, Python, Django, Asyncio, Rust, Clojure, gRPC, TensorFlow, Scikit-Learn, Kubernetes, Docker, Azure Services( Cosmos DB, Azure SQL Database, Blob Storage, Azure Data Factory, Azure DevOps, Azure Synapse, Azure App Services, Azure Kubernetes Service, Azure Text Classification), API, Langchain, Apache Kafka, Apache Spark, Terraform, Ansible, Postman, PyTest, PII, Selenium, Selenium Grid, OCR, Azure Machine Learning, MLOps, OpenCV, GenAI, RPA, Linux/Unix Shell Scripts.

**Client: UIPath. July 2020 – Aug 2022**

**Location: New York, USA.**

## Role: Data Scientist/ Python Developer.

* Worked on Python-based **ML** projects using **Pandas**, **NumPy**, and **TensorFlow** to analyze terabytes of transaction data for real-time customer segmentation and churn prediction.
* Optimized financial models with NumPy and implemented **ML** algorithms in **scikit-learn** for fraud detection and market trend analysis.
* Built real-time data dashboards using **Matplotlib**, **Seaborn**, and **Bokeh** to provide stakeholders with actionable banking insights.
* Automated large-scale workflows with **Dask** and **Scrapy**, integrating external datasets into financial models for comprehensive analysis.
* Developed **Flask**-based **RESTful APIs**, containerized with **Docker**, and deployed using **Kubernetes** for scalable, real-time ML predictions.
* Used **SQLAlchemy** for efficient database management and optimized queries, integrating relational databases with Python for real-time analytics.
* Automated ML model updates and workflows with **APScheduler** and **Celery**, adapting systems to real-time data changes without downtime.
* Implemented logging and monitoring using **Elasticsearch**, **Kibana**, and Python’s logging module to track real-time performance and troubleshoot proactively.
* Created interactive data dashboards using **Databricks** for real-time visualization, enhancing decision-making with up-to-the-minute insights.
* Created customer segmentation clusters with graph analysis (Breadth-First Search) using Python, **Pandas**, and network analysis libraries on Linux.
* Managed MySQL databases with **Linux automation scripts**, setting up cron jobs for indexing, backups, and real-time data updates.
* Managed **Docker**, **virtualenv**, and **Kubernetes** for consistent environments across development, testing, and production, ensuring smooth CI/CD.
* Handled version control with **Git**, following best practices like feature branching, and integrating **CI/CD** pipelines for automated testing and deployment.
* Worked in **Agile** teams, using **Jira** for sprint planning and task tracking, ensuring continuous delivery of Python-based ML models and data solutions.

**Tools and Technologies:** Python, Pandas, NumPy, TensorFlow, scikit-learn, Matplotlib, Seaborn, Bokeh, Dask, Scrapy, Flask, Docker, Kubernetes, SQLAlchemy, APScheduler, Celery, Elasticsearch, Kibana, Databricks, virtualenv, Git, Jira, Jenkins, GitLab

## Client: Soulpage IT Solutions Feb 2018 – July 2020

## Location: Hyderabad, India.

**Role: Python Developer.**

* Built and maintained **full-stack** applications using **Python**, **Django**, **MySQL**, and **PostgreSQL**, focusing on performance optimization and efficient data management.
* Automated real-time **ETL** processes using Python with **Databricks**, integrating **AWS Lambda** to handle backend automation, ensuring data was always up to date.
* Improved frontend functionality with **CSS** and **Bootstrap** while integrating **Redis** Cache with **Django** to speed up data retrieval in high-traffic applications.
* Designed and deployed **REST APIs** using **Python** and **Django**, hosted on **AWS EC2** and backed by **RDS** databases, ensuring high availability and scalability.
* Set up real-time data pipelines with **Apache Kafka** to handle large-scale data streaming, integrating with existing systems for advanced analytics.
* Managed complex **MySQL** and **Oracle DB** integrations, using Python for data extraction, and worked with Django to create user-facing dashboards for monitoring.
* Configured and managed **AWS EC2** and **RDS** instances with CloudFormation templates for automated infrastructure setup, improving scalability and reducing manual tasks.
* Adopted TDD for Python and Django development, writing test cases with **Pytest** to ensure stability and prevent issues in production.
* Streamlined **UNIX**-based job processing using Python’s OS module and **Bash scripting**, automating repetitive tasks to improve operational efficiency.
* Combined **Django** with **Redis** Cache to optimize performance and utilized **AWS Lambda** for various backend tasks, reducing downtime and improving response times.
* Worked extensively with **Pandas** for data analysis, developing **Python scripts** to manage and analyze large datasets stored in **MySQL**.
* Automated deployment and configuration tasks using **Ansible**, ensuring consistent environment setup across development, testing, and production.
* Created high-performance web applications with **Django** and **AWS**, leveraging **Kafka** for real-time data handling, and **MySQL** for reliable data storage.
* Utilized **Git** for version control, managing code across multiple branches, ensuring smooth collaboration, and maintaining a robust **CI/CD** pipeline for continuous integration.
* Actively participated in **Agile** development, contributing to sprint planning, daily stand-ups, and iterative feature development to align with project timelines.

**Tools and Technologies** Python, Django, MySQL, PostgreSQL, Databricks, AWS Lambda, CSS, Bootstrap, Redis Cache, Apache Kafka, Oracle DB, CloudFormation, AWS EC2, AWS RDS, Pytest, Bash, Ansible, Pandas, Git, CI/CD.

## Education:

## University: Stevens Institute of Technology

**Degree:** Master’s in Industrial/ Engineering Management

**University:** Jawaharlal Nehru Technological University Hyderabad

**Degree:** B.Tech in Computer Science Engineering.