Name: ACHYUTHA PRANAVI
Role: Senior AI/ML Engineer | Data Analyst
Email address: achyuthapranavi268@gmail.com
Contact: +13479210136
LinkedIn: http://www.linkedin.com/in/achyutha-p-89a3b5216

Professional Summary

- Results-driven **Senior AI/ML Engineer** with **11 years of progressive experience** in data science, machine learning, and AI solutions across **finance, healthcare, retail, insurance, and telecom sectors**, delivering measurable business impact through advanced analytics and AI implementations.
- Expert in **cloud-native AI/ML solutions** on both **Microsoft Azure** (Azure ML, Databricks, Synapse Analytics, Cognitive Search, Azure OpenAI) and **AWS** (Redshift, S3, Lambda, Athena, Glue, QuickSight), with proven ability to architect and deploy scalable, production-ready systems.
- Advanced practitioner in **Large Language Models and Generative AI**, building **retrieval-augmented generation (RAG) pipelines** using **LangChain, Semantic Kernel, FAISS, and Azure Cognitive Search** to reduce hallucinations and enhance accuracy in knowledge-heavy financial and healthcare contexts.
- Developed **transformer-based NLP models** using **BERT, GPT, and Azure OpenAI** for financial document summarization, sentiment analysis, risk entity detection, and clinical note processing, enabling automated research and compliance workflows.
- Applied **time-series forecasting models** including **ARIMA, Prophet, LSTM, and Neural Prophet** for financial trend prediction, demand forecasting, and risk assessment in lending, trading, and retail operations.
- Built comprehensive **MLOps and CI/CD pipelines** using **MLflow, Azure DevOps, Docker, and Kubernetes**, ensuring reproducibility, monitoring, and governance in high-stakes ML deployments across regulated industries.
- Engineered **feature engineering and ETL pipelines** with **PySpark, pandas, Azure Data Factory, AWS Glue**, and **Delta Lake**, delivering enriched, high-quality datasets that significantly improved model performance and business outcomes.
- Implemented **responsible AI practices** using **SHAP, LIME, Azure AI Content Safety**, and fairness metrics to ensure compliance, explainability, and transparency across HIPAA-regulated healthcare and financial services environments.
- Designed **predictive models for fraud detection, claims management, and patient stratification** that achieved measurable results including **30% reduction in false claims**, **23% improvement in campaign targeting**, and **18% enhancement in fraud detection accuracy**.
- Applied **knowledge graph models** using **Neo4j and NetworkX** to map complex relationships among customers, transactions, and financial assets for advanced fraud detection, compliance analytics, and risk assessment.
- Leveraged **self-supervised learning and contrastive learning** techniques for pretraining on proprietary financial and healthcare text corpora, improving domain adaptation and embedding quality for downstream tasks.
- Built **synthetic data generation workflows** using **Gretel.ai** and internal libraries to augment imbalanced datasets while maintaining data privacy, security compliance, and regulatory adherence in sensitive industries.
- Created **real-time monitoring dashboards** using **Power BI, Tableau, Azure Monitor, Prometheus**, and **QuickSight** for tracking model drift, prediction quality, system anomalies, and business KPIs across 1,800+ retail locations.
- Led **end-to-end data pipeline development** processing high-volume transaction data, healthcare records, and telecom metrics, utilizing **distributed computing frameworks** and **cloud-native architectures** for scalable analytics.
- Delivered **cross-functional leadership** by mentoring junior engineers, conducting knowledge-sharing sessions on **MLOps, generative AI, Azure AI Studio**, and **foundation model fine-tuning**, while translating complex ML models into actionable business insights for stakeholders.

- Proven track record in **statistical modeling, A/B testing, cohort analysis**, and **demand forecasting** that directly contributed to revenue optimization, operational efficiency improvements, and strategic business decision-making across multiple industry verticals.

Technical Skills

Cloud Platforms	Microsoft Azure (Azure ML, Azure Databricks, Azure Synapse, Azure Cognitive Search, Azure OpenAI), AWS (Redshift, S3, Lambda, Athena, Glue, QuickSight)
AI/ML Frameworks	PyTorch, TensorFlow, scikit-learn, Keras, Hugging Face Transformers, LangChain, Semantic Kernel
MLOps & DevOps	MLflow, Azure DevOps, CI/CD Pipelines, Docker, Kubernetes, Prometheus, Azure Monitor
NLP & LLMs	BERT, GPT, Azure OpenAI, RAG (FAISS, Vector DBs), Text Summarization, Sentiment Analysis, Entity Recognition
Time-Series Forecasting	ARIMA, Prophet, LSTM, Neural Prophet
Big Data & ETL	PySpark, pandas, Azure Data Factory, AWS Glue, SQL, Delta Lake
Vector & Graph DBs	FAISS, Azure Cognitive Search, Neo4j, NetworkX
Responsible AI	SHAP, LIME, Fairness Metrics, Azure AI Content Safety
Data Visualization	Power BI, QuickSight, Matplotlib, Seaborn
Synthetic Data & Privacy	Gretel.ai, Differential Privacy, Data Augmentation
Programming	Python, SQL, PySpark, Bash
Knowledge Graphs	Neo4j, NetworkX, Graph Embeddings
Self-Supervised Learning	Contrastive Learning, Pretraining (Financial Text)

Educational Details

- Bachelors in computer science at Lovely Professional university — Aug 2007 to Jun 2011
- Masters in computer science at University of Central Missouri — Sept 2011 to Dec 2012

Certifications

- Microsoft Certified: Azure AI Engineer Associate - 2023
- AWS Certified Machine Learning – Specialty – 2021

Work Experience

Client: Jefferies Financial Group Inc, New York, NY — **May 2024 - Present**

Role: Senior AI/ML Engineer

Responsibilities:

- Designed and implemented **AI/ML solutions on Microsoft Azure**, leveraging services like **Azure Machine Learning**, **Azure Databricks**, and **Azure Synapse Analytics** for scalable, secure model training and deployment.
- Led the development of **end-to-end ML pipelines** using **Python**, **Azure ML SDK**, and **MLflow**, enabling reproducibility, CI/CD integration, and operational model governance.
- Developed and deployed advanced **NLP models** for financial document summarization, sentiment analysis, and risk entity detection using **transformer-based architectures** such as **BERT** and **Azure OpenAI Service**.
- Integrated **vector databases** like **Azure Cognitive Search** and **FAISS** to power **retrieval-augmented generation (RAG)** pipelines for LLM-based search over internal knowledge bases.
- Orchestrated **LangChain** and **Semantic Kernel** components to build LLM-powered agents for automating financial research and compliance workflows, improving analyst productivity.

- Built custom **RAG pipelines** with grounding and context injection strategies to reduce hallucinations and ensure factual accuracy in generative AI outputs.
- Applied **time-series forecasting models** including **ARIMA**, **Prophet**, and **LSTM**, to predict multi-asset financial trends and perform risk forecasting for lending and trading use cases.
- Enhanced LLM and NLP output quality using **prompt tuning**, **PEFT (parameter-efficient fine-tuning)**, and **reinforcement learning with human feedback (RLHF)** for domain alignment.
- Utilized **Azure Databricks Delta Lake** and **Unity Catalog** for unified governance, lineage tracking, and secure access to financial data lakes across teams and business units.
- Implemented **feature engineering pipelines** using **PySpark**, **pandas**, and **Azure Data Factory**, improving model performance through advanced feature transformations and temporal joins.
- Applied **knowledge graph models** using tools like **Neo4j** and **NetworkX** to map relationships across customers, transactions, and financial instruments for fraud detection and compliance analytics.
- Integrated **responsible AI tooling** including **Azure AI Content Safety**, **SHAP**, and **LIME** to ensure model transparency, fairness, and compliance with financial regulations.
- Designed and deployed **LLM-based systems** for financial document parsing, earnings summary generation, and real-time investor sentiment extraction via **Azure OpenAI**.
- Built **synthetic data generation workflows** using tools like **Gretel.ai** and internal libraries to augment imbalanced datasets and protect sensitive financial information.
- Developed **custom monitoring dashboards** using **Power BI**, **Azure Application Insights**, and **Prometheus**, to visualize model drift, prediction quality, and data anomalies in production.
- Led research into **self-supervised learning** and **contrastive learning** techniques for embedding and pretraining on proprietary financial text corpora.
- Implemented **data mesh principles** and domain-oriented ownership for scalable ML architecture across global financial data domains, fostering decentralized innovation.
- Mentored junior engineers and led internal sessions on **MLOps**, **generative AI**, **Azure AI Studio**, and **foundation model adaptation**, establishing best practices for production-scale AI in finance.

Client: HCA Healthcare Inc, Nashville, TN **Nov 2022 – April 2024**

Role: AI/ML Engineer

Responsibilities:

- Designed and deployed scalable **machine learning models** to automate risk prediction and patient stratification within healthcare plans, improving proactive care delivery.
- Collaborated with cross-functional healthcare and engineering teams to translate clinical goals into data-driven solutions using **Azure Machine Learning Studio** and **Azure Synapse Analytics**.
- Developed robust **data pipelines** and orchestrated ETL workflows using **Azure Data Factory** to ensure accurate ingestion and transformation of large-scale healthcare data.
- Engineered **predictive models** for claims fraud detection, leading to a 30% reduction in false claims through advanced **supervised learning** techniques.
- Built and maintained **CI/CD pipelines** for ML model deployment using **Azure DevOps**, improving delivery speed and ensuring reproducibility across staging and production environments.
- Utilized **PySpark**, **SQL**, and **Azure Databricks** for distributed data processing and real-time analytics in population health management systems.
- Created explainable AI solutions using **SHAP** and **LIME**, enabling compliance with healthcare regulations and enhancing model transparency for clinicians.
- Integrated external social determinant datasets with HCA data on **Azure Data Lake** to enhance patient outcome modelling through **feature engineering** and **data fusion**.
- Implemented **model monitoring dashboards** using **Power BI** and **Azure Monitor**, allowing continuous tracking of ML performance and drift in real-world settings.
- Conducted **hyperparameter tuning** and **model optimization** leveraging **Azure AutoML** for accelerated experimentation and model selection.
- Ensured all AI solutions complied with **HIPAA** standards, embedding privacy, security, and governance controls into the ML lifecycle.

- Supported the development of **NLP models** to analyze unstructured clinical notes, enabling structured data extraction and classification via **BERT** and **spaCy** frameworks.
- Led efforts in **data labelling strategies**, annotation workflows, and **active learning**, accelerating model training for medical image and document classification.
- Participated in design reviews and code audits to enforce **MLOps best practices**, including containerization using **Docker** and orchestration with **Kubernetes on Azure AKS**.
- Collaborated with stakeholders to define key metrics and success criteria for AI initiatives, aligning ML outputs with the goals of value-based healthcare.
- Applied **transfer learning** and **deep learning** models to diagnostic imaging use cases, improving anomaly detection accuracy across large datasets.
- Mentored junior engineers and data scientists on **Azure ML workflows**, code versioning using **Git**, and experiment tracking with **MLflow**.
- Delivered impactful presentations to both technical and non-technical audiences, demonstrating the business value and clinical implications of AI projects.

Client: Target Corp, Minneapolis, MN **Jan 2019 – Oct 2022**

Role: Senior Data Analyst

Responsibilities:

- Led the design and implementation of scalable **data pipelines on AWS**, utilizing **Amazon Redshift**, **S3**, and **AWS Glue**, to process and store high-volume retail transaction data.
- Delivered actionable insights by developing complex **SQL queries** and **stored procedures** to analyze customer behavior, product trends, and seasonal performance across Target's retail network of 1,800+ locations.
- Created dynamic **dashboard solutions using Tableau and Power BI**, enabling real-time executive reporting for inventory optimization and supply chain efficiency.
- Collaborated with cross-functional teams including marketing, merchandising, and e-commerce to align business goals with data strategies, improving campaign targeting by 23%.
- Executed deep-dive analysis on sales funnels, conversion rates, and clickstream data using **Python (pandas, NumPy)** to identify friction points in the customer journey.
- Managed the end-to-end lifecycle of data modeling projects, ensuring consistency across dimensions and metrics using **dbt** and **Snowflake**, enhancing data integrity across teams.
- Applied advanced **statistical methods and regression models** to predict demand fluctuations and optimize pricing strategies, contributing to a 12% revenue uplift in 2021.
- Automated recurring reporting processes with **Python scripting and AWS Lambda**, reducing manual workload by 40% and improving delivery accuracy.
- Utilized **Amazon QuickSight** for visual analytics on retail KPIs, integrating with **Athena** and **Glue Data Catalog** for seamless data querying and exploration.
- Led data governance initiatives, including **data quality checks**, lineage documentation, and access control using **AWS Lake Formation**, ensuring compliance with internal and regulatory standards.
- Integrated external market datasets (e.g., Nielsen, social media sentiment) to enrich internal datasets and support **competitive analysis** and localized merchandising efforts.
- Developed forecasting models using **Prophet and ARIMA** to assist in inventory planning and reduce overstock scenarios across warehouse hubs.
- Delivered training sessions and onboarding guides for junior analysts on **AWS tools**, data interpretation, and visualization best practices.
- Collaborated with DevOps to integrate data monitoring solutions using **CloudWatch**, improving system reliability and proactive issue detection.
- Conducted cohort and churn analysis using **SQL and Python**, identifying high-value customer segments and helping marketing focus on retention campaigns.
- Implemented best practices in version control, CI/CD pipelines, and analytics code management using **Git, AWS CodeCommit**, and **Jenkins**.
- Acted as a key liaison between business stakeholders and technical teams, translating analytical findings into business decisions that directly impacted revenue, customer engagement, and operational efficiency.

Client: Allstate, Northbrook, IL **Sept 2015 – Dec 2018**

Role: Data Analyst

Responsibilities:

- Conducted comprehensive data analysis on policyholder behavior, claim trends, and underwriting data to support strategic insurance pricing decisions and risk assessments.
- Utilized **AWS services (S3, Redshift, Athena, and Glue)** to manage and process large insurance data sets securely and efficiently within a cloud-based infrastructure.
- Developed and maintained **interactive dashboards using Tableau and Power BI** to track KPIs, claim volumes, and customer retention trends, increasing stakeholder visibility into performance metrics.
- Designed and implemented **ETL pipelines** leveraging **AWS Glue and Python** to automate data ingestion from multiple sources including internal CRM and third-party actuarial datasets.
- Collaborated with actuarial and underwriting teams to perform predictive modeling using **Python (pandas, scikit-learn)** for churn prediction and fraud detection in claims processing.
- Optimized SQL queries on **Amazon Redshift** to deliver near real-time reporting and improved system performance by 30% in monthly executive reporting.
- Participated in data governance initiatives to ensure **data quality, accuracy, and compliance with insurance regulations** including HIPAA and state-level privacy laws.
- Conducted detailed cohort analyses to identify customer behavior patterns and delivered actionable insights that directly improved cross-sell campaign effectiveness.
- Led efforts in integrating and standardizing structured and unstructured data across policy, claims, and customer interaction channels, using **Python and AWS Lambda** functions.
- Applied advanced statistical methods and **machine learning techniques** to support pricing optimization and dynamic segmentation of customers.
- Coordinated with business stakeholders to gather requirements and translate them into analytical models and visual solutions that directly influenced operational decisions.
- Created data dictionaries and maintained metadata repositories to ensure **data transparency and traceability** across functional teams and tools.
- Provided ad hoc reporting and strategic insights to support sales, marketing, and claims departments in aligning performance with business goals.
- Implemented data anomaly detection scripts in **Python and SQL** to flag outliers in claim submissions, improving fraud detection by 18%.
- Collaborated with DevOps and cloud engineering teams to ensure **secure and scalable deployment of analytics workflows using AWS IAM, EC2, and CloudWatch**.
- Regularly evaluated new market analytics tools and insurance industry trends to enhance analytical capabilities and align with evolving data science practices.

Client: Ooma Inc, Sunnyvale, CA **Feb 2013 – Aug 2015**

Role: Data Analyst

Responsibilities:

- Collaborated with cross-functional teams to **analyze large-scale telecom data**, improving decision-making processes and increasing operational efficiency by over 20%.
- Developed and maintained **automated dashboards and visual reports** using **Tableau and Power BI**, enabling senior leadership to monitor KPIs in real-time.
- Executed end-to-end data analysis projects by extracting data from **SQL databases**, transforming it through **ETL pipelines**, and performing statistical evaluation.
- Conducted deep analysis on customer behavior and churn patterns using **Python (pandas, NumPy)** and **R**, leading to actionable retention strategies.
- Worked closely with engineering and marketing teams to align data findings with **telecom industry trends**, enhancing product targeting and customer segmentation.
- Implemented **data validation and cleansing routines** to ensure data integrity and consistency across systems, improving report accuracy by 30%.

- Led efforts in **market segmentation analysis**, providing insight into user demographics and usage patterns which supported strategic pricing initiatives.
- Created predictive models using **machine learning techniques** to forecast call drop rates and optimize network resource allocation.
- Processed large datasets using **SQL Server** and **MySQL**, optimizing queries for analytical performance.
- Worked with **Oracle databases** and basic **ETL processes** to support telecom analytics.
- Assisted in migration projects by validating datasets post-transfer, ensuring smooth transition and compliance with industry standards.
- Supported monthly and quarterly business reviews with **trend analysis**, highlighting growth opportunities and operational inefficiencies in telecom services.
- Developed A/B testing frameworks to assess the impact of new telecom features, helping guide product development with data-driven insights.
- Provided stakeholder training on interpreting data reports and visualizations, promoting a **data-driven culture** across departments.
- Analyzed **call quality and VoIP metrics**, identifying bottlenecks and recommending improvements that enhanced service reliability and customer satisfaction.
- Maintained documentation for data models, processes, and business logic to ensure transparency and enable future scalability.