Sai Sravan Chintala

AI/ML Engineer | NLP, LLM Workflows & Intelligent Automation

Harrisonville, MO | +1 913-378-6834 | saisravan@belvionus.com | LinkedIn

PROFESSIONAL SUMMARY

AI/ML Engineer with 5+ years of experience building machine learning, NLP, Generative AI, and intelligent automation solutions across enterprise data environments. Experienced in developing Python-based ML and data pipelines using Azure Data Factory, Azure Databricks, PySpark, and SQL, with hands-on expertise in feature engineering, model development, evaluation, and data validation. Skilled in LLM application development using GPT-4, Llama 3, LangChain, Hugging Face, RAG, prompt engineering, and fine-tuning techniques. Experienced in transforming structured and unstructured data into production-ready AI workflows through APIs and scalable data architectures.

TECHNICAL SKILLS

Programming & Data: Python, SQL, Pandas, NumPy, Data Cleaning, Data Transformation, Data Validation, Exploratory Data Analysis, Statistical Analysis

Machine Learning: Supervised Learning, Unsupervised Learning, Classification, Regression, Predictive Modeling, Feature Engineering, Model Selection, Hyperparameter Tuning, Cross-Validation, Anomaly Detection, Clustering

AI, NLP & Generative AI: Natural Language Processing, Generative AI, LLM Workflows, Llama 3, GPT-4, LangChain, Hugging Face, Retrieval-Augmented Generation (RAG), Fine-Tuning, Prompt Engineering, Text Preprocessing, Tokenization, Feature Extraction, Text Classification, Semantic Analysis, Intelligent Automation, OCR

ML Libraries & Development: Scikit-learn, XGBoost, Imbalanced-Learn, OpenCV, Tesseract OCR, Scikit-Image, Streamlit, REST APIs, ML Pipelines, Model Evaluation

Data Engineering & Cloud: ETL, Distributed Processing, Hadoop MapReduce, mrjob, Apache Cassandra, Microsoft Azure, Azure SQL Database, Databricks

Analytics & Databases: KPI Development, Advanced Data Visualization, Altair, Microsoft Power Apps, Relational Databases, NoSQL Databases

PROFESSIONAL EXPERIENCE

DICEUS | AI/ML Engineer **Sep 2026 - Present**

- Developed Python-based machine learning workflows to identify unusual transaction patterns across 150K+ transaction records, supporting anomaly detection and investigation of potentially suspicious activity across transaction datasets.
- Applied NLP techniques to extract relevant entities and transaction indicators from unstructured text, preparing structured information for downstream analytics, feature preparation, and model-based classification workflows.
- Integrated GPT-4 and LLM capabilities into financial knowledge workflows, evaluating responses against 40+ test queries to assess relevance, response consistency, and alignment with information retrieved from source content.
- Built data transformation logic using PySpark and Delta Live Tables (DLT) to organize incoming datasets into Bronze, Silver, and Gold layers, improving data readiness for downstream analytical and machine learning workloads.
- Evaluated machine learning models using validation datasets and classification metrics, reviewing false positives and misclassified records to identify performance gaps and opportunities for model refinement.
- Developed customer segmentation workflows using behavioral and account-related features, applying clustering techniques to identify distinct patterns across 5 customer behavior groups and support customer activity analysis.
- Created reusable Python components connecting data preparation, NLP processing, and model inference, reducing duplicated implementation effort by 15% across experimental AI/ML workflows and supporting consistent development practices.

Mphasis | AI/ML Engineer **Jul 2022 - Aug 2024**

- Processed 80K+ borrower records through Python-based credit analysis workflows, combining credit histories, repayment behavior, and income details to prepare structured inputs for creditworthiness assessment and risk classification.
- Implemented credit data extraction and preprocessing workflows to standardize borrower information, calculate debt-to-income ratios, and prepare 12+ credit attributes for consistent credit-risk evaluation and predictive modeling.
- Built machine learning classification models using Scikit-learn to categorize borrower risk profiles based on repayment behavior, credit utilization, outstanding liabilities, and application attributes for credit assessment workflows.
- Engineered borrower-level credit-risk indicators using Pandas and NumPy, preparing 10+ derived features to support predictive analysis, identify relevant borrower characteristics, and distinguish patterns across credit-risk categories.
- Evaluated model performance through cross-validation, precision, recall, F1-score, and confusion matrices, comparing 3 classification approaches to identify performance trade-offs and assess credit-risk prediction reliability.
- Developed Python and SQL validation routines to reconcile borrower records, detect missing credit attributes, and investigate data inconsistencies across multiple input datasets used in credit assessment and model development.
- Integrated credit-risk predictions and assessment outputs into application workflows through REST APIs, validating 5 key response fields to support reliable data exchange between connected software components and consuming applications.

Cybage Software | Machine Learning Engineer **Jun 2019 - Jun 2022**

- Contributed to the development of a predictive analytics module for business software, using Python, Pandas, and NumPy to transform historical business data into model-ready datasets for predictive applications.
- Implemented classification models with Scikit-learn to predict categorical business outcomes, integrating model training and prediction logic into the analytics workflow to support software-based business predictions.
- Developed regression models to estimate numerical business measures from historical records, evaluating prediction errors to assess the reliability of generated outputs and compare model performance across evaluation datasets.
- Built data preprocessing components to handle missing values, outliers, duplicates, and inconsistent formats, improving dataset completeness by 11% before model integration and supporting more consistent training inputs.
- Engineered and selected predictive features using statistical analysis and correlation testing, improving model performance by 12% during evaluation and strengthening the quality of inputs used in predictive modeling.
- Used SQL to retrieve, join, aggregate, and validate source data, ensuring consistent inputs for the software platform's predictive analytics components and downstream model development workflows.
- Created reusable Python modules for data preparation, model training, and performance evaluation, supporting repeatable testing and integration of predictive capabilities into the application across business prediction use cases.

SELECTED PROJECTS

Financial Fraud Detection & Risk Scoring System

Python, Scikit-Learn, Pandas, Imbalanced-Learn, Streamlit

- Engineered a fraud-classification pipeline across 6.3M+ financial records, achieving 95% recall on highly imbalanced transaction data through feature engineering and threshold tuning.
- Deployed a Streamlit inference application delivering sub-second predictions and risk assessments, demonstrating an end-to-end workflow from preprocessing and modeling to application-level inference.

Automated Currency Authentication & OCR Verification

Python, OpenCV, Tesseract OCR, Scikit-Image

- Developed a dual-layer currency verification solution combining micro-text analysis and multilingual OCR to automate denomination recognition and currency authentication.
- Implemented LBP texture descriptors and Tesseract OCR, reducing validation latency from hours to seconds on standard CPU hardware.

Public Health Insurance Demographic & Trend Dashboard

Python, Streamlit, Altair, Pandas

- Developed an interactive dashboard analyzing 5 years of U.S. Census Bureau SAHIE county-level health insurance data across demographic and income metrics.
- Automated data ingestion, cleaning, standardization, and visualization workflows through dynamic KPIs, trends, and county-level rankings.

EDUCATION

University of Central Missouri - Lee's Summit, Missouri **Aug 2024 - May 2026**

Master's in Data Science and Artificial Intelligence | GPA: 3.8 / 4.0

CMR Engineering College - Hyderabad, Telangana **Aug 2015 - May 2019**

Bachelor of Technology (B.Tech.)

CERTIFICATIONS

- **Databricks Certified Data Engineer Associate - Jun 2024**
- **Microsoft Certified: Power BI Data Analyst Associate (PL-300) - Feb 2024**
- **Microsoft Certified: Azure Data Engineer Associate (DP-203) - Dec 2023**
- **Microsoft Certified: Azure Data Fundamentals (DP-900) - Nov 2023**
- **Microsoft Certified: Azure Fundamentals (AZ-900) - Mar 2022**