Pranav R

Bengaluru, KA, India | 7411257234 | pranavramesh84@gmail.com | LinkedIn | GitHub | Portfolio

Data Scientist with 3+ years of industry experience expertise in predictive modeling, deep learning, and data visualization. Proficient in delivering data-driven solutions, optimizing workflows, and uncovering actionable insights. Demonstrates exceptional problem-solving, communication, and time-management skills. Based in Bengaluru, India.

Technical Skills

Programming Languages: Python, SQL, JavaScript, HTML/CSS
ML Frameworks & Libraries: Pandas, NumPy, Scikit-learn, TensorFlow, PyTorch, Generative AI, LLMs
Databases: MySQL, Microsoft SQL Server, SharePoint
Automation & BI Tools: Tableau, Power BI, Power Apps, Power Automate, Alteryx, Excel
Big Data & Cloud Technologies: Hadoop, Azure DevOps

Education

Master of Science in Data Analytics Dec 2024
San Jose State University, San Jose, California, USA | GPA: 3.55/4

Bachelor of Engineering in Computer Science Aug 2020
Dr. Ambedkar Institute of Technology, Bengaluru, India | CGPA: 9.15/10

Professional Experience

Senior Data Scientist | *Decision Point Analytics, Bengaluru, KA, India* May 2025 – Present
- Scraped retail pages using BeautifulSoup and used a Groq-hosted LLaMA 3 model via LangChain to extract key product data into structured JSON, reducing manual prep by 35%.
- Applied ARIMA and SARIMA models on retail sales data to explore capacity forecasting and seasonality patterns, building reusable forecasting components for client-ready use cases.
- Selected optimal ARIMA/SARIMA parameters by applying normal and seasonal differencing to achieve stationarity, and using ADF test, PACF/ACF plots to determine p, d, q values for improved model fit.

Software Engineer - AI/ML | *NeuroLeap Corp, Costa Mesa, CA, USA* Feb 2025 – May 2025
- Reduced inference latency by 35% in speech-to-text systems using Whisper (tiny, small) and Sphinx through model tuning and audio preprocessing.
- Improved noisy-environment transcription accuracy by 20% via data augmentation and WER-driven fine-tuning.
- Visualized model performance using Matplotlib and Seaborn, creating bar charts for WER and latency comparison across audio samples, and grouped bar charts for transcription accuracy under varying noise conditions.

Data Scientist Intern | *MarketMakerCRE, Pensacola, FL, USA* Jul 2024 – Aug 2024
- Optimized data collection by 35% through Python-based web scraping with BeautifulSoup and Selenium, ensuring seamless integration with REIT datasets.
- Enhanced real estate market predictive analytics using XGBoost & Random Forest, reducing MSE from 2.5 to 1.2 (52% improvement).
- Automated data validation with Azure DevOps CI/CD, improving real-time insights by 30%.
- Developed 2+ Tableau dashboards that reduced decision-making time by 25%, supporting 20+ real estate leaders.
- Led A/B testing on the MarketMakerCRE website, increasing user engagement by 15% and optimizing conversion rates by 10% through data-driven insights.

Software Engineer 1 | *Juniper Networks, Bengaluru, KA, India* Jul 2020 – Jan 2023
- Built end-to-end data pipelines in SQL and Alteryx, optimizing processing time by 25% for financial and operational analytics.
- Automated client notifications with Power Automate, ensuring timely contract reminders via JSON attachments, boosting PM team efficiency by 30%.
- Integrated PS/AS Credits dashboard with Power Apps, driving a 23% increase in PM team productivity.
- Built a global Power Apps job rotation tool with automated multi-stage approvals and notifications via Power Automate, streamlining internal job transfers across AMER, EMEA, CALA, and APAC.
- Analyzed 1,700+ customer engagements, achieving a 20% boost in customer satisfaction.
- Created 7+ financial Tableau dashboards, tracking KPIs, monitoring performance metrics, facilitating data-driven decisions for over 100+ stakeholders, improving decision-making efficiency by 40%.
- Collaborated with cross-functional teams to develop 15+ Power Apps tools, enhancing workflows for 80+ daily users and improving operational efficiency by 33%.

Intern 3 Professional Services | *Juniper Networks, Bengaluru, India* Jan 2020 – Jun 2020
- Automated SQL-based reporting processes using Power Automate, reducing manual effort by 10+ hours weekly, improving overall team productivity.
- Streamlined data extraction with advanced SQL queries and complex joins, cutting report generation time by 18%.

Projects

TailorHire: AI-Powered Job Outreach Tool | Python, LangChain, Groq (LLaMa 3), ChromaDB, Streamlit (GitHub)

- Built an AI-powered job application assistant that scrapes job listings, applies HTML parsing and noise filtering to clean text, extracts roles and skills using LLaMA 3 via LangChain, and generates personalized cold emails.
- Integrated ChromaDB vector search to match extracted skills with candidate GitHub projects, enhancing email relevance by showcasing actual experience.
- Deployed the solution with Streamlit, enabling users to input job URLs and instantly generate tailored, professional outreach messages.

Fake Job Detection using NLP and Machine Learning | Python, BeautifulSoup, NLP, TF-IDF (GitHub)

- Scraped job listing data from the EMSCAD website using Python's BeautifulSoup and Requests libraries to build a clean dataset for analysis.
- Applied NLP techniques such as TF-IDF vectorization and NLTK for text preprocessing and feature extraction from job descriptions.
- Built a binary classifier to detect fake job listings using logistic regression, SVM, and decision tree models.
- Improved recall to 0.91 in the decision tree model through tuning and dataset balancing for enhanced fraud detection accuracy.

Soil Health Prediction – Fresno County | Python, Scikit-learn, TensorFlow, GeoPandas, Streamlit (GitHub)

- Integrated multi-source data (satellite, sensors, weather, soil, crop) spanning 1987–2022 for Fresno County to assess soil health and environmental factors.
- Applied data preprocessing and exploratory analysis using transformations and visualizations to reveal seasonal and long-term trends in soil and weather variables.
- Built and compared 8 models—including ARIMA, Random Forest, XGBoost, LSTM, and hybrid LSTM+RF—to predict key soil parameters like temperature, moisture, pH, carbon, and nitrogen.
- Developed an interactive GUI with geospatial maps, time-series graphs, and filters, enabling users to explore soil and crop data for informed, sustainable agriculture decisions.