Ankita Pichuka

ankitapichuka@gmail.com | Mobile: 857 565 6050 | LinkedIn | GitHub | Portfolio

Summary

Master's candidate in Analytics (Applied Machine Intelligence) at Northeastern University (GPA: 3.83/4.0) with 2.5+ years building ML models, automating data pipelines, and delivering analytics across business and technical teams. Proficient in Python, SQL, R, Scikit-learn, TensorFlow, PyTorch, and HuggingFace Transformers, with hands-on experience in supervised learning, NLP, generative AI, and data visualization using Power BI and Tableau. Deploys production-ready solutions and communicates findings clearly to both technical and non-technical stakeholders.

Technical Skills

ML & AI: Python | Scikit-learn | TensorFlow | PyTorch | XGBoost | Logistic Regression | Random Forests | Gradient Boosting | Classification | Regression
Generative AI & NLP: HuggingFace Transformers | LangChain | spaCy | NLTK | NLP | LLMs | RAG Pipelines | Prompt Engineering | OpenAI API
Data Analysis: Pandas | NumPy | SQL | R | PySpark | Statistical Analysis | Data Wrangling | Exploratory Data Analysis | A/B Testing
Visualization & Reporting: Power BI | Tableau | Matplotlib | Seaborn | Plotly | Excel | Dashboards | KPI Reporting | Stakeholder Communication
Platforms & Tools: Snowflake | Databricks | PostgreSQL | MySQL | MongoDB | AWS | Azure | Docker | Git | Jupyter | MS Teams
Additional: R | Java | JavaScript | C# | HTML/CSS | Agile | MLFlow | TensorBoard | Hyperparameter Tuning | Model Evaluation

Work Experience

Data Analyst @ *OpenIntervue – Bengaluru, India* **Mar 2024 – Jan 2025**
- Built and evaluated NLP and ML models using HuggingFace Transformers and Scikit-learn, applying supervised learning techniques to enterprise text classification that improved screening accuracy by 30%.
- Designed automated Python data pipelines using Pandas and NumPy that cut manual processing effort by 40%, delivering weekly analytical reports with consistent data quality checks to leadership.
- Presented model performance and findings to technical and non-technical stakeholders, participating in code reviews and contributing to software development across the full production lifecycle.

Operations Analyst @ *Esycommerce Pvt Ltd – Mumbai, India* **Jul 2022 – Feb 2024**
- Applied machine learning and statistical modeling using Python and SQL to analyze large-scale datasets, building 10+ dashboards that reduced operational cycle time by 30% and improved data accuracy by 28%.
- Implemented generative AI and NLP techniques for unstructured text analysis, collaborating across engineering, product, and operations teams to deliver insights through data visualizations and performance reports.
- Automated financial forecasting and reporting workflows, documenting processes and contributing to system improvements that scaled operations across multiple client engagements.

Business Analyst Intern @ *Esycommerce Pvt Ltd – Mumbai, India* **Feb 2022 – Apr 2022**
- Analyzed 5+ datasets using Python and Excel, applying classification and regression techniques to produce 10+ analytical reports that surfaced trends and supported management decisions.

Education

M.P.S. in Analytics (Applied Machine Intelligence) | Northeastern University Boston, MA
Relevant Coursework: Machine Learning, Deep Learning, NLP, Generative AI, Data Science, Statistics, Python, SQL **GPA: 3.83/4.0** | Expected May 2027

Bachelor of Commerce | Andhra University Visakhapatnam, India
Relevant Coursework: Financial Accounting, Business Statistics, Quantitative Methods, Mathematics, Economics **GPA: 3.36/4.0** | 2019 – 2022

Projects

NLP Resume Screening Engine | *Python, Scikit-learn, spaCy, NLTK, HuggingFace* | Live App
- Fine-tuned HuggingFace Transformer models for enterprise NLP classification, processing 100+ records at production scale with automated evaluation pipelines — deployed live at OpenIntervue.

Customer Churn Risk Analyzer | *Python, SQL, Pandas, Scikit-learn, XGBoost, Power BI* | Live App
- Trained and evaluated supervised ML models including XGBoost, Random Forests, and Logistic Regression with hyperparameter tuning, achieving 87% prediction accuracy in a fully deployed Streamlit application.

ImmiQ – US Immigration Trends Tracker | *SQL, PySpark, Snowflake, R, Tableau* | GitHub | Live App
- Built data pipelines across a 7+ entity relational database using PySpark and SQL, reducing preprocessing time by 80% and delivering Tableau dashboards that surfaced trends for stakeholder decision-making.

Certifications

Google AI Essentials Specialization | Google Mar 2026
Google Data Analytics Professional Certificate | Google 2024
AWS Cloud Practitioner | Amazon Web Services 2024