

# ARYAN PATEL

Boston, MA | patel.aryana@northeastern.edu | Website | LinkedIn

## EDUCATION

**Masters of Science in Artificial Intelligence | Northeastern University**

Concentration: Machine Learning

Boston, MA

(Present)

**Bachelor of Technology in Computer Science & Engineering | Quantum University**

Specialization in Artificial Intelligence and Machine Learning

Dehradun

Aug. 2021 - Jun. 2025

## TECHNICAL SKILLS

**Programming & Tools:** Python, Git, Power BI, MS Office

**Frameworks & Libraries:** TensorFlow, PyTorch, Scikit-learn, Keras, XGBoost, Pandas, NumPy, Matplotlib, Plotly, MLflow, Seaborn, TensorBoard, Flask / FastAPI

**Machine Learning Techniques:** Regression Analysis, Predictive Modeling, Classification (Binary, Multiclass), Feature Engineering, Clustering (K-Means, DBSCAN, Hierarchical), Model Evaluation

**Advanced AI Methods:** Deep Learning, Neural Networks, Unsupervised Learning, Reinforcement Learning, Computer Vision, Natural Language Processing

**Statistics & Data Science:** Probability Theory, Hypothesis Testing, Statistical Inference, Descriptive & Inferential Statistics, Time Series Analysis

**Certifications:** Google Cloud Career Launchpad Data Analyst, Neural Networks and Deep Learning, Machine Learning with Python (Udemy), Deep Learning with TensorFlow

## PROFESSIONAL EXPERIENCE

**AI and Machine Learning Intern | YBI FOUNDATION**

Nov. 2023 - Oct. 2024

- Developed and deployed a Naïve Bayes classifier to categorize **10,000+** customer feedback samples, extracting actionable insights from unstructured data.
- Enhanced accuracy by **18%** through systematic preprocessing (tokenization, stopword removal, vectorization), advanced feature engineering, and iterative hyperparameter tuning.
- Conducted a thorough evaluation using metrics such as **accuracy (91%), precision (89%), recall (87%), and F1-score (88%)**, along with confusion matrix analysis to identify specific areas for performance enhancement.
- Documented the methodology, experimental design, and results in a structured report, enabling seamless knowledge transfer and serving as a blueprint for future machine learning projects within the organization.

**Machine Learning Intern | PRODIGY INFOTECH**

Jul. 2023 - Aug. 2023

- Built and validated a **regression model** to predict California housing prices with strong predictive accuracy.
- Conducted in-depth **exploratory data analysis (EDA)**, uncovering correlations between housing prices and **10+ demographic and economic indicators**, including income levels, population density, and proximity to the ocean.
- Streamlined the preprocessing pipeline by implementing procedures for missing value imputation, categorical variable encoding, and feature scaling, which reduced preparation time by **25%** while improving data quality.
- Collaborated with a 3-member intern team to compare regression approaches (linear, ridge, lasso), selecting the best-performing model for deployment.

## ACADEMIC PROJECTS

**Lung Damage Detection** (TensorFlow, TensorFlow Extended (TFX), Seaborn)

- Designed and trained a **CNN** for multi-class classification of lung diseases, including pneumonia, and lung cancer.
- Integrated the model with a **real-time ML pipeline** using TensorFlow Extended (TFX), ensuring smooth data ingestion, automated preprocessing, scalable training, and deployment capabilities in production-like environments.
- Designed the system to serve as a **diagnostic support tool**, potentially reducing manual diagnosis time by **30%** and improving early detection accuracy in clinical use cases.

**Multiclass Classification Using Neural Networks** (Google Colab, TensorFlow, Scikit-learn, Matplotlib)

- Developed and optimized a neural network on **50,000+ multi-class images**, achieving **92% test accuracy** through extensive hyperparameter tuning and architecture experimentation.
- Validated performance with confusion matrices and learning curves, demonstrating a **15% improvement over baseline models** while ensuring robustness and interpretability.
- Compiled results into a comprehensive report, including visualizations, error analyses, and comparisons against alternative approaches, establishing benchmarks for future research and scalable extensions of the project.