

# TEJ GANGUPANTULA

+1 (209)-600-5060 | [tejgangupantula@gmail.com](mailto:tejgangupantula@gmail.com) | [Portfolio](#) | [LinkedIn](#) | [GitHub](#)

## EDUCATION

**University of California, Santa Barbara**

September 2023 – December 2026

*B.S. Statistics and Data Science*

*Santa Barbara, CA*

- **GPA:** 3.6
- **Relevant Courses:** Data Science with R/SQL, Probability and Statistics, Data Science Concepts and Analysis, Regression Analysis, Statistical Machine Learning, Stochastic Models & Applied Probability, Calculus I-III, Linear Algebra, Differential Equations

## SKILLS

**Languages:** Python, R, SQL, Java, Javascript, HTML

**Libraries:** Scikit-learn, TensorFlow, pandas, NumPy, Matplotlib, Seaborn, Streamlit, OpenAI API, spaCy, Langchain

**Tools/Frameworks:** Git, Jupyter, Google Cloud Platform (GCP), FastMCP, Django, Tableau

## PROFESSIONAL EXPERIENCE

**AI Engineer Intern**

June 2025 – September 2025

*Agentman*

*Berkeley, CA*

- Built and developed agentic AI systems for healthcare operations, using HIPAA APIs and GCP, reducing manual processing by ~70% and saving practices \$20K+ per user annually.
- Designed operational agents and cloud pipelines that connected APIs, and EHR data streamlining workflows and saving 3+ staff-hours daily while boosting reliability through improved QA and schema validation.
- Owned full initiatives end-to-end, collaborating across product and engineering teams, and gained expertise in cloud architecture, API integrations, and healthcare compliance workflows.

**Machine Learning Intern**

July 2024 – September 2024

*FusionCare*

*Davis, CA*

- Built a Python-based clinical text processing pipeline to ingest, clean, and structure de-identified patient notes, applying NLP-based PII masking (spaCy/medspaCy) to safely prepare data for downstream analysis.
- Integrated OpenAI GPT models with structured prompting and output validation using LangChain to convert unstructured clinical notes into consistent, JSON-formatted visit summaries and outcome signals, improving care plan evaluations by 30% and saving clinicians 4–5 hours weekly.
- Designed and validated Random Forest models in scikit-learn to predict patient outcomes from structured visit data, evaluating performance with patient-level splits and cross-validation to support exploratory risk stratification and inform future workflow integration.

## PROJECTS

**Brain Tumor Classification** | *Python, TensorFlow, Scikit-learn, NumPy, pandas*

January 2025 – April 2025

- Developed a Convolutional Neural Network model using TensorFlow to classify brain tumors (glioma, meningioma, pituitary tumor, no tumor) with 95% accuracy on test data, leveraging a dataset of 7,000+ MRI scans.
- Implemented data preprocessing and augmentation techniques (rotation, flipping, brightness adjustment) to improve model performance and ensure robust predictions; visualized results using matplotlib, scikit-learn and seaborn.
- Demonstrated potential to revolutionize diagnostics by enabling early, accurate, and accessible AI-powered brain tumor classification, supporting personalized care and improved patient outcomes.

**NBA Finals Outcome Prediction Analysis** | *Python, Scikit-learn, NumPy, pandas*

April 2025 – June 2025

- Developed a Random Forest classification model in Python to simulate and predict 2025 NBA Finals outcomes, achieving 62% test accuracy using team-level regular season and playoff metrics.
- Engineered a Monte Carlo simulation to estimate championship probabilities over a best-of-seven series, predicting a 75.5% chance of the Oklahoma City Thunder winning against the Indiana Pacers (24.5%).
- Cleaned, processed, and analyzed historical NBA data with pandas, numpy, and scikit-learn, enabling interpretable model outputs and supporting data-driven sports analytics.