

## Anjali YellapuntulaVenketa

anjali.yv01@gmail.com | linkedin.com/in/anjaliyv | www.anjaliyv.com

### EDUCATION

---

**University of North Carolina, Chapel Hill, NC, United States**

**Aug 2023 – Apr 2025**

*Master of Science in Information Science*

RELEVANT COURSEWORK: Information Retrieval, Introduction to Deep Learning, Visual Analytics, User Interface Design, Usability Testing and Evaluation, Consumer Health Information, Modern Web Programming

**Veerмата Jijabai Technological Institute (VJTI), Mumbai, India**

**Aug 2019 – May 2023**

*Bachelor of Technology in Electronics Engineering*

### WORK EXPERIENCE

---

**Data Scientist (Graduate Assistant)**

*UNC University Libraries, Chapel Hill, NC*

**Apr 2024 – Dec 2024**

- Collaborated with subject matter experts (SMEs) to engineer searchable archives of 100+ South Carolina laws by creating structured metadata using OCR and NLP (SpaCy, regex), while maintaining version control with Git for the “On The Books: Algorithms of Resistance” project.
- Deployed ML models to tag 3,000+ Jim Crow laws, revealing discriminatory language patterns and supporting transparency in historical legal documentation.

**Research Contributor**

*UNC Gillings School of Global Public Health, Chapel Hill, NC*

**Jan 2024 – Dec 2024**

- Achieved 96% accuracy in Post-Traumatic Stress Disorder (PTSD) prediction with Support Vector Machine (SVM) model out of 4 models by applying feature selection and supervised ML models to survey data from 129 healthcare professionals, collected using validated psychometric instruments.
- Conducted statistical modeling (hypothesis testing, regression, SHAP) on 130+ features to identify key predictors of clinician depression and PTSD.
- Lead author of a full paper accepted for presentation at MedInfo 2025, presenting machine learning-driven insights into clinician mental health and contributing empirical evidence to global health policy.

**Data Science Intern**

*Exponentia.ai, Mumbai, India*

**Jun 2024 – Aug 2024**

- Built scalable ML pipelines in Databricks using LLMs for unstructured data analysis, reducing manual effort by 30% and delivering clean, parameterized notebooks for repeatable reporting.
- Enhanced personalized search on a web platform by applying prompt engineering and Retrieval-Augmented Generation (RAG) to identify user behavior patterns and generate tailored insights.

**Data Analyst Intern**

*PricewaterhouseCoopers (PwC), Kolkata, India*

**May 2022 – Jul 2022**

- Accelerated audit reporting by 50% by building a voice-enabled Business Intelligence (BI) assistant using Streamlit and Google Cloud Platform (GCP), allowing audit teams to query KPIs and generate reports in real time.
- Designed and integrated an NLP-powered chatbot with speech recognition to automate metric retrieval, eliminating manual report generation steps.
- Delivered a fully functional web app that enabled auditors to access key metrics in under 1 minute, improving decision-making speed and operational efficiency.

### PUBLICATIONS

---

- Anjali Yellapuntula VENKETA, Lukasz MAZUR, Karthik ADAPA, “Modeling Workplace Predictors of Depression and PTSD in Healthcare Professionals Using Machine Learning,” *Proceedings of MedInfo 2025*, Taipei, Taiwan, August 2025 (Accepted for full paper presentation, with publication expected November 2025).

### TECHNICAL SKILLS

---

- **Languages & Tools:** Python, SQL, Git, Databricks, Google Cloud Platform (GCP), Tableau, OpenAI APIs
- **Core Libraries:** Scikit-learn, Pandas, NumPy, TensorFlow, PyTorch, Hugging Face Transformers, SpaCy
- **Techniques:** Supervised & Unsupervised Learning, Predictive Modeling, Feature Engineering, Fine-tuning, Transfer Learning, Exploratory Data Analysis (EDA), Text Classification, Neural Networks
- **Specialization:** Machine Learning, Natural Language Processing (NLP), Large Language Models (LLMs), Prompt Engineering, SHAP, Optical Character Recognition (OCR)