Raajitha Muthyala

(317) 500-7147 raajithamuthyala@gmail.com linkedin.com/in/raajithamuthyala github.com/RaajithaMuthyala

Professional Summary: Health Informatics professional with a pharmacy background specializing in healthcare data analytics, machine learning, and clinical decision support systems. Experienced in extracting actionable insights from medical data using advanced NLP techniques, predictive modeling, and FHIR-based EHR integration. Skilled in translating complex healthcare data into meaningful visualizations and evidence-based recommendations.

Education

Indiana University Indianapolis **Aug 2023 - May 2025**
Master of Science in Health Informatics *Indianapolis, Indiana*

Jawaharlal Nehru Technological University **July 2018 - July 2022**
Bachelor of Pharmacy *Hyderabad, Telangana, India*

Experience

Indiana University Indianapolis **Aug 2024 - May 2025**
Data Analyst at the Purkayastha Lab for Health Innovation *Indianapolis, Indiana*

- Lead the design and implementation of machine learning algorithms for clinical data analysis and predictive modeling.
- Develop and maintain NLP systems to process unstructured healthcare data using biomedical ontologies.
- Create scalable clinical decision support frameworks for large-scale patient datasets.
- Collaborate with clinical researchers to translate complex healthcare requirements into technical specifications.
- Present analytical findings to stakeholders and recommend evidence-based interventions.

Indiana University Indianapolis **Aug 2023 - Aug 2024**
Research Analyst at IU BioHealth Informatics Research Center at Indianapolis *Indianapolis, Indiana*

- Gather and organize healthcare policy data from multiple international sources, including the Czech Republic, Montenegro, and Australia, to support international benchmarking efforts.
- Perform data preprocessing, cleaning, and validation according to research protocols.
- Conduct preliminary statistical analysis to identify significant trends and patterns.
- Support research teams in data-driven comparative policy assessment.

Viatris **Jan 2023 - Jul 2023**
Data Analyst *Hyderabad, India*

- Execute analytical testing protocols using HPLC-MS instrumentation.
- Manage data collection, processing, and reporting using Empower software.
- Ensure accuracy and consistency in pharmaceutical test data.
- Support quality control procedures for drug development processes.
- Contribute to data-driven workflow optimization initiatives.

MedBound **Jan 2022 - Dec 2022**
Data Analyst *Hyderabad, India*

- Process and interpret clinical data using standardized medical coding systems (ICD-10, CPT).
- Design and implement dashboards to track key healthcare performance indicators.
- Generate regular reports on HEDIS metrics and care coordination outcomes.
- Collaborate with clinical teams to identify quality improvement opportunities.
- Ensure data accuracy and compliance with documentation standards.

Skills

Languages: Python, R, SQL, XML, JSON

Frameworks/Libraries: NLTK, SpaCy, Hugging Face, transformers, Tensorflow, PyTorch, Keras, Numpy, Pandas, Scikit-Learn, Scipy, Matplotlib, seaborn, Pyplot, Plotly, ggplot2, R packages

Tools: Microsoft Excel, Outlook, Jupyter, RStudio, Oxygen XML, DHIS2, Tableau, Postman, Github, PowerBI, GitLab, Airflow

Technologies/Platform: EDA, Data Visualization, REST, APIs, Machine Learning, Natural Language Processing, Clustering, Regression, Classification, Sentiment Analysis, Topic modeling, Time Series analysis, Statistical Analysis

Data Standards: SNOMED, ICD, LOINC, CPT, RxNorm, NDC, HL7, FHIR

Soft Skills: Decision making, Problem solving, Active listening, Empathy

Projects

Capstone Project - EviLivBot | *Streamlit, Transformers* **Jan 2025 - Present**

- Fine-tuning the LlaMA 3.1 8B model on domain-specific datasets related to liver disease and Social Determinants of Health (SDOH) for biomedical entity recognition and explainable AI.
- Developing a Streamlit-based interface to visualize extracted biomedical entities and deliver evidence-backed insights from the fine-tuned model.
- Designing a pipeline to integrate large language model outputs with structured biomedical evidence, enhancing model transparency and domain-specific reasoning.

Evaluating LLMs for Medical Coding&Readmission Risk| *Prompt Engineering, Biomedical AI* **Sep 2024 - Jan 2025**

- Conducted a comparative evaluation of LLMs (ChatGPT-4, LLaMA-3.1, Gemini, DeepSeek-R1, OpenAI-O3) on diagnostic and prognostic tasks using the MIMIC-IV dataset.
- Designed zero-shot prompting strategies and created structured clinical input from discharge summaries to assess primary diagnosis generation, ICD-9 coding, and readmission risk prediction.
- Implemented rationale extraction and semantic evaluation using SciBERT and cosine similarity to quantify model alignment with ground truth data.

Biomedical NER and QA System using PubMed Data | *Named Entity Recognition, Streamlit* **Aug 2024 - Dec 2024**

- Extracted and preprocessed PubMed abstracts for entity recognition and question-answering tasks using biomedical ontologies like UMLS.
- Applied Named Entity Recognition and fine-tuned a T5-small model to generate and answer clinical questions from extracted entities.
- Deployed the QA system via a Streamlit interface for interactive exploration of biomedical information.

Biomedical Data Analytics | *R, DHIS2, XML, NLP* **Aug 2024 - Dec 2024**

- Visualized reproductive health trends using DHIS2 and developed time-series forecasting models (ARIMA, polynomial regression) to predict malnutrition and birth rates. Conducted geographic mapping to analyze ANC2 coverage by chiefdoms, focusing on underserved areas near forest reserves.
- Performed sentiment analysis based on the team members' interpretations of the data using Natural Language Processing.
- Created a DHIS2 dashboard and provided actionable recommendations for resource planning and targeted healthcare interventions, addressing policy implications for improved healthcare delivery in Sierra Leone.

Healthcare Data Integration and Interoperability | *XML, APIs, JSON, Postman* **Jan 2024 - May 2024**

- Designed and implemented workflows using HL7 CDA, FHIR, XSLT, and REST APIs to enable seamless data exchange between EHR (OpenMRS) and labpratory systems (OpenELIS).
- Mapped clinical data to ICD-10-CM, SNOMED-CT, RxNorm, and LOINC to ensure standardized and accurate communication in healthcare use cases.
- Designed UML/BPMN models for e-prescription workflows; validated pipelines using Oxygen XML Editor and Postman, ensuring seamless interoperability.

Statistical Analysis of Blood Metrics & CKD Risk | *R, Statistics* **Jan 2024 - May 2024**

- Designed a predictive model for CKD using logistic regression in R, identifying significant predictors such as hypertension, bacteria, blood urea and blood glucose levels, achieving actionable insights.
- Conducted comprehensive data preprocessing, visualization, and hypothesis testing to ensure robust data quality and reliable outcomes.

CVD Insights: A Data-Driven Cross-Sectional Analysis | *SQL, Python, ML* **Aug 2023 - Dec 2023**

- Performed comprehensive Cardio Vascular Disease (CVD) analysis using the correlations between various risk factors (age, gender, height, weight, blood pressure, cholesterol,glucose, smoking status, alcohol consumption, and physical activity).
- Developed and evaluated predictive models (Random Forest, Logistic Regression, SVM), leveraging Python tools (Pandas, Matplotlib, scikit-learn) to generate actionable healthcare insights.

Formulation of Maraviroc Floating Tablets | *DDS, CRF, Pharmaceutics* **Sep 2021 - Jun 2022**

- Designed and developed effervescent floating tablets of Maraviroc using low-density polymers to improve gastric retention and drug bioavailability.
- Utilized direct compression techniques with agents like HPMC K100M, Carbopol 940, and Xanthan Gum for controlled drug release.
- Evaluated physical parameters and in-vitro performance to optimize therapeutic efficacy for enhanced patient outcomes.

Relevant Publications

∗ Evaluating Reasoning Capabilities of Large Language Models for Medical Coding and Hospital Readmission Risk Stratification with Zero Shot Prompting (Preprint)- **JMIR 2025**