

# TEJODHAY BONAM

Email: [bonam.t@northeastern.edu](mailto:bonam.t@northeastern.edu) | Phone: (+1) 8573950779 | Portfolio: [tejodhayb.github.io](https://tejodhayb.github.io) | [LinkedIn](#)

## EDUCATION

**Northeastern University - Khoury College of Computer Sciences, MS Data Science** (May 2024) **Boston, MA**

Coursework: Supervised & Unsupervised Machine Learning, Data Engineering, DBMS, NLP, Algorithms, LLMs

- **Graduate Teaching Assistant (NLP, Data Science):** Supported 80+ students in key ML & NLP concepts - (data cleaning, model interpretability, optimization, hypothesis testing & inferential statistics) through OH & Python programming assistance)

**Sri Sathya Sai Institute of Higher Learning, BS Mathematics & Computer Science** (2021) **Bengaluru, KA**

- Headed the *Western Flute* section of the *Institute Brass Band*; Won *University Badminton Championship* (2 times)

## SKILLS & CERTIFICATIONS

**Programming Languages/ ML Frameworks/ Libraries:** Python, SQL, TensorFlow, PyTorch, Scikit-Learn, Pandas, NLTK

**Databases:** AWS S3, RDS, RedShift, Azure Synapse Analytics, Azure SQL Database, MySQL, MS SQL Server, MongoDB

**Software and Tools:** Apache Kafka, Spark, Airflow, Apache Flink, Docker, Terraform, MS Excel, Power BI, Tableau, Flask, DVC, CICD, GitHub Actions, MLflow, Agile, Scrum, Azure Data Factory, Azure Databricks, AWS SageMaker, AWS EMR

**Certifications:** AWS Data Engineer Associate (TBC), Microsoft Azure Data Engineer Associate, AWS Cloud Practitioner

## WORK EXPERIENCE

 **Accenture** **Data Engineering Associate** Mar 2022 – Sep 2022  
Hyderabad, IN

- As part of a top-performing 5-member Agile (Scrum) team, I actively collaborated with data engineering team to deliver 2 projects on time and within budget, applying data engineering best practices to develop robust and scalable data pipelines.
- Configured an E2E real-time data engineering solution to migrate data from on-prem SQL Server database to **ADLS Gen2** leveraging Azure Data Factory, **PySpark** transformations on **Azure Databricks** and loaded into **Azure Synapse Analytics**.
- Automated ETL pipelines orchestrated on **Azure Data Factory** enabling daily data refreshes to ingest OLAP data into Azure Synapse Analytics improving data processing efficiency by 35%, transforming data through bronze, silver & gold layers.
- Created cool data visualizations and delivered 15+ Power BI dashboards, reports to identify trends, patterns and anomalies.

**Arthashastra Intelligence** **Jr. Data Scientist** Jun 2021 - Mar 2022  
Hyderabad, IN

- **Star Projects:**
  - Client (Appliances Mfg., Consumer Goods)
    - User-engagement Centric Forum, integrated with Recommendation engine for Personalized Articles & Targeted Ads (**LDA, tf-idf, NER** - Content based, **Spectral, K-prototypes clustering** - Collaborative)
    - Social Media Analytics dashboard - **AWS Kinesis, Athena, SQL, Quicksight (30% boost in user engagement & CTR)**
  - Customer & Purchase analytics of an FMCG using Segmentation, Targeting, Positioning, Marketing Mix, Price Elasticity
    - Performed **K-means clustering** combined with **PCA** for customer segmentation. Analyzed purchase behavior through descriptive and predictive models for purchase incidence, brand choice, purchase qty.
    - Employed **A/B testing** to identify customer demographics for supervising marketing campaigns - (**110% sales growth**)

*Other Major Contributions:*

- Implemented regression & classification ML models, performed feature engineering in **Jupyter Notebooks** on 10+ client datasets and synthesized large amounts of data into succinct actionable insights to stakeholders (technical & non-technical).

## PROJECTS

 **Microsoft - Future Ready Talent** Feb 2022 – Mar 2022

[Facial Emotions based Music Recommendation](#) (OpenCV, Azure Face API, K-Means clustering)

- The web app gets live video feed from web cam, calls Face API that detects emotions. Accordingly, the app ranks recommended songs employing K-Means clustering. (Often used by my friends and family - **Satisfaction rate: 70%**)

[Transcribe 'n' Translate:](#) (Azure - Computer Vision, AI Translator, App Service, Streamlit UI)

- Extracts English text from uploaded img (PNG, JPEG) or docs (PDF) & translates to multiple languages (**Acc: 90%**)

### Portfolio Projects

[CMS Medicare Data Engineering Pipeline](#) (Python, AWS S3, Spark on EMR, Amazon RDS for MySQL, Power BI)

- Engineered an ETL pipeline sourcing 10GB data (Prescribers by Geography & Drug ~ 25M rows) from PostgreSQL & ingested into S3 Data lake, transformed using Spark on EMR, published to AWS RDS for MySQL, reports in Power BI

[Chicken Disease Classification with MLOps](#) (MLflow, DVC, Docker, AWS/Azure CICD, GitHub Actions, TensorFlow)

- Created an image classification deep learning model (CNN) following MLOps for CI/CD including logging, utils & exception handling, project workflows and deployed the pipeline in production ensuring optimal resource allocation.

**Northeastern University, LLM course projects**

[T-shirts Retail LLM web app](#) (Google PaLM, LangChain, MySQL, Hugging Face, Few-Shot learning)

- Developed Streamlit web app with user-friendly interface aimed at non-technical users which converts natural language questions into SQL queries and dynamically executes on MySQL database (**Generated output = Expected output: 98%**)

[PawPal:](#) Petcare Veterinary RAG chatbot (Llama2, LangChain, Multilingual E5 large, FAISS, AWS SageMaker)

- Built the app on open-source stack useful for Vet Doctors, Pet Lovers to provide rapid (**with response time of 0.4 sec**) and reliable veterinary advice (with **85% accuracy** of diagnoses and treatment recommendations verified by real Vets)

**P.S.** - Explore my [portfolio](#) for my other DE, DS, ML, DL, NLP and LLM projects!