



# THARUN TEJ REDDY THODIMI

Experienced **Data Scientist** with nearly 6 years of expertise in developing solutions and pipelines to address complex business challenges. Proficient in GenAI, Statistical Analysis, Machine Learning, and NLP.

*Actively seeking opportunities in Machine Learning, NLP, GenAI, and Business Analytics to contribute to success and expand professional expertise.*

## CONTACT



tharuntejreddy112@gmail.com



+91- 8555-080-227



Telugu, Hindi, English

## CORE SKILLS

Data Science & Business Analytics

Mathematical & Statistical Concepts

GenAI

Data Mining & Visualization

Machine Learning Algorithms

Business Intelligence

NLP

Team Management

## IT SKILLS

Python

TIBCO Spotfire

Cloud Platform(AWS,Azure)

SQL,HTML,Streamlit

## DOMAIN SKILLS

Life Sciences

BFSI

## EXECUTIVE SUMMARY

- Working with **Accenture Solutions** as a **Data Science Consultant**
- Proven expertise in leveraging GenAI technology for content generation, content summarization and building QnA Chatbots
- Demonstrated success in developing Knowledge Graphs for scientific literature to support innovation in drug research and discovery.
- Showcased strong data science knowledge in various machine learning concepts; skilled in selecting features, building & optimizing classifiers using machine learning techniques
- Provided expertise in design and development of end-to-end pipelines.
- Directing full development cycle of planning, analysis, design, development, testing using data mining and data analysis tools and techniques
- Analyzed data insights and effectively communicated findings in a well-organized manner, fostering positive relationships with stakeholders and offering valuable guidance
- Recipient of multiple awards for consistent outcomes and value creation for clients
- Mentored junior data scientists, fostering their professional development and contributing to a collaborative and innovative team culture
- Capable to delve into the new Leading Technologies
- Good work ethics with excellent communication, presentation and interpersonal skills

## EDUCATION

- **PG Diploma (Data Science)**, from International Institute of Information Technology – Bangalore (IIIT-B) in 2021; secured 3.7/4.0 GPA.
- **B. Tech (ECE)**, from SRM Institute of Science & Technology, Chennai in 2018; secured 9.13 GPA.

## WORK EXPERIENCE

- **Accenture Solutions, Jun'21 – Till date**  
Data Science Consultant ( Oct'23 – Till date)  
Data Science Analyst (Jun'21 - Sep'23)
- **Tata Consultancy Services, Sep'18 – May'21**  
Junior Data Scientist

# PROJECTS

<b>Company &amp; Domain</b>	Accenture Solutions   Life sciences
<b>Project Name</b>	QnA Chatbot using GenAI
<b>Description:</b> A Leading Implant device manufacturing company wanted to develop a QnA Chatbot to expedite query resolution for call center agents assisting implant users.	
<b>Roles and Responsibilities:</b> <ul style="list-style-type: none"><li>– Designed the solution architecture to automate the manual process, ensuring timely and accurate responses by extracting relevant information from source documents.</li><li>– Leveraged the RAG architecture and GPT-3.5 models to build the chatbot, effectively reducing hallucinations in responses.</li><li>– Handled the integration of services on the Azure cloud platform.</li><li>– Trained and fine-tuned the GPT-3.5 models to improve the chatbot's performance and accuracy using <u>prompt engineering</u>.</li></ul>	
<b>Technology Stack:</b> Python, NLP, PostgreSQL, MongoDB Vector database, LangChain Framework, RAG Architecture, React JS	
<b>Cloud Platform:</b> Azure cloud platform (Azure App service, Blob storage, Databricks, Azure OpenAI)	

<b>Company &amp; Domain</b>	Accenture Solutions   Life sciences
<b>Project Name</b>	Knowledge Graph
<b>Description:</b> A Global Biopharmaceutical Company wants to build a Knowledge Graph based searchable repository of scientific literature to support innovation in drug research and discovery experiment pipelines. The main goal of this activity is to conduct a comprehensive analysis of the most pertinent and current scientific literature, focusing on identifying critical biomedical entities and relationships.	
<b>Roles and Responsibilities:</b> <ul style="list-style-type: none"><li>– Handled the development of a solution that identifies entities and tags relationships.</li><li>– Developed the end-to-end data pipeline for the solution in AWS cloud platform.</li><li>– Explored various models to identify state-of-the-art techniques for building a biomedical knowledge graph.</li><li>– Worked with the KAZU model for entity tagging, enhancing the accuracy of information extraction.</li><li>– Utilized the ChemProt model for relationship extraction, improving the comprehensiveness of the knowledge graph.</li><li>– Applied the latest advancements in the field that were most relevant and beneficial to the project's objectives.</li><li>– Played a key role in the project's success which garnered recognition and attracted the attention of numerous other clients.</li></ul>	
<b>Technology Stack:</b> Python, Neptune Graph DB, SQL and HTML	
<b>Cloud Platform:</b> AWS (Sagemaker, Lambda, EC2, API gateway)	

<b>Company &amp; Domain</b>	Accenture Solutions   Life sciences
<b>Project Name</b>	Content Generation using GenAI for GVD(Global Value Dossier)
<b>Description:</b> A prominent US Pharma Giant wanted to develop a content generation solution for Global Value Dossier (GVD) documents. Leveraged GenAI technology to process diverse information sources and deliver highly accurate content through prompt engineering.	
<b>Roles and Responsibilities:</b> <ul style="list-style-type: none"><li>– Designed solution architecture to read the user-defined configuration file which generates mappings internally allowing seamless integration of provided details into various sections.</li><li>– Developed the <u>mapping engine</u> to align input data with relevant content segments.</li><li>– Led <u>end-to-end workflow</u> design and implementation from configuration setup to content generation, including UI implementation.</li><li>– Focused on <u>prompt engineering</u> to ensure precise content generation.</li></ul>	
<b>Technology Stack:</b> Python, NLP, Streamlit, OpenSearch Vector database, LangChain Framework, RAG Architecture	
<b>Cloud Platform:</b> AWS (CodeBuild, ECR, Elastic Container Service, S3, API gateway, AWS Sagemaker)	

<b>Company &amp; Domain</b>	Accenture Solutions   Life sciences
<b>Project Name</b>	Spotfire dashboards (Drug development)
<b>Description:</b> Worked on building Spotfire dashboards for multiple use cases in R&ED domain, and few of the high- level project outline is as follows: RNA Sequence Visualization, MAPPs Reset, and Inventory Metric Analysis.	

**Roles and Responsibilities:**

- Acted as the primary point of contact for clients and scientists in the drug development team to gather dashboard requirements.
- Successfully developed and deployed various dashboards, including those for RNA sequence visualization, Inventory Management, and MAPPs Reset.
- Leveraged Iron Python for creating dynamic calculations in Spotfire, enhancing the dashboard's interactivity and functionality.
- Worked closely with the drug development team to ensure the dashboards met their needs and facilitated their work.
- Continuously improved the dashboards based on user feedback and evolving project requirements.

**Technology Stack:**

SQL, Tibco Spotfire, Python, Iron Python, Domino Model API, Domino Model APP, HTML.

**Company & Domain**

TCS | BFSI

**Project Name**

Customer Segmentation

**Description:**

Provide a strategy to bank by identifying the hidden patterns of customer behavior in opting to their price options. Price options with higher monthly average fees plays a key role in banks profit. Identify the hidden patterns and make changes to the price options to increases their profits by withstanding the customers. Increase, Decrease /Provide Negotiation on the Price options to provide benefits from the monthly revenue received from customers as monthly account maintenance fee.

**Roles and Responsibilities:**

- Taken all the necessary inputs from the dataset and did Exploratory data analysis to identify which Price range categories are bringing more income to the banks and identified the income range which bank is struggling to attract their customers and where bank can concentrate.

**Technology Stack:** Sklearn, Seaborn, Matplotlib, Numpy, Pandas.

**Algorithms Used:** Hierarchical clustering and K-Means Clustering.

## AWARDS AND ACHIEVEMENTS

- Achieved the “Super Squad Award” in Q2 2024 for significant contributions to the development of an end-to-end solution in GenAI from Accenture.
- Received “Beacon Award” for creating consistent outcomes and value creation for clients from Accenture
- Received “Distinguished Achiever” for the year 2022 from Accenture
- Received “Star Award” for the outstanding work delivered in the last 2 quarters from Accenture
- Received “Best Team Performer” for dedicated contribution as a team member from TCS

## CERTIFICATIONS

- AWS Certified Cloud Practitioner
- Prompt Engineering with GPT from ‘Coursera’
- ‘Generative AI for everyone’ Engineering with GPT from ‘Coursera’

## PERSONAL DETAILS

Date of Birth:

17th August 1997

Address:

Flat no: 502,Skanda heights, Apoorva Residency, Mitaganahalli, Bangalore- IN(562149).