

# Tarun Kumhari

<mailto:tarun.95532@gmail.com> | 9553273998 | [www.linkedin.com/in/tarunkumhari](https://www.linkedin.com/in/tarunkumhari) | [www.github.com/t95532](https://www.github.com/t95532)

## Objective

Detail Oriented Data Science Associate with experience in data science and analytics. Proficient in translating business problems into actionable insights, developing reports, and contributing to project management activities. Seeking a challenging role where I can leverage my analytical skills and contribute to data driven decision making process.

## Professional Experience

### Junior Data Scientist | Voziq AI Hyderabad | July 2024 – Present

- Created standard customer profiles by segmenting customers based on behavioural and demographic attributes.
- Analysed large datasets to identify trends, patterns, and actionable insights, presenting them through effective reports.
- Developed, tested, and optimized predictive models to improve customer behaviour analysis and retention strategies.
- Leveraged statistical models and machine learning algorithms to predict customer churn and enhance client retention.
- Refined machine learning models, improving accuracy and delivering impactful client outcomes.
- Designed and implemented survival curves to analyse customer retention trends and forecast churn probabilities.
- Developed production-ready decks and presentations to showcase key insights and recommendations to stakeholders.

### Advanced ML and Gen-AI Intern | Innomatics Research Labs Hyderabad | Nov 2023 – May 2024

- Developed Natural Language Processing (NLP) techniques, including text preprocessing, feature engineering, and sentiment analysis.
- Applied advanced NLP algorithms (word embeddings, sequence-to-sequence models, attention mechanisms) to real-world datasets.
- Explored Generative AI (LLMs like OpenAI and Google Gemini) and learned prompt engineering for task-specific optimization.
- Contributed to multiple projects involving model development, evaluation, and deployment.

### Data Scientist Trainee | TuringMinds.AI Hyderabad | March 2022 – April 2023

- Worked on data and business research projects, secondary research testing, and results-oriented intelligence, supporting informed decision-making.
- Participated in projects by assisting systems support group with design, programming, research, testing, and implementation aspects.
- Assisted in troubleshooting and problem-solving with respect to production or issuance issues, ensuring smooth operations.
- Translated business problems and data needs into low-complexity analyses and recommendations, contributing to improved decision-making processes.
- Contributed to the development of descriptive models using key performance KPIs to identify systematic issues and perform root cause analysis.
- Developed standard and customized reports providing valuable insights to stakeholders.

# Education

## TKR College of Engineering and Technology

Bachelor of Technology | Electronics and communication Engineering | 7.28 CGPA

July 2018 – June 2022

# Software Skills

Programming Languages	: Python, SQL
Frameworks and Libraries	: Pandas, NumPy, Matplotlib, Seaborn, Scikit-Learn, Keras, TensorFlow, Flask
Databases	: MySQL, DB Browser, Microsoft SQL Server, Chroma DB
Machine Learning	: Algorithms, Data Collection, Data Cleaning, Preprocessing, Feature Engineering, Data Visualization, Model Building, ML-Ops
Deep Learning	: Artificial Neural Networks, CNNs, RNNs
Natural Language Processing	: NLTK, Spacy, Gensim, Text Embeddings, Language Modelling, Transformers
Cloud Computing	: AWS (S3, EC2, CloudWatch, IAM, RDS)
Generative AI	: Large Language Models (LLMs), Prompt Engineering, OpenAI, Google Gemini, Langchain
Version Control	: Git, GitHub
DevOps & CI/CD	: Basic knowledge of containerization (Docker), understanding of CI/CD pipelines.
Tools	: Tableau, Power BI, Excel, PowerPoint

# Projects

## Retrieval-Augmented Generation (RAG) Based Search Engine

- Developed a Retrieval-Augmented Generation (RAG) search engine integrating Google’s generative AI for context-based document retrieval and AI-generated responses.
- Connected to Chroma DB for storing and retrieving document embeddings, ensuring fast and relevant search results.
- Built a pipeline to retrieve the top 5 relevant documents and enhance AI-generated responses based on context.
- Designed an interactive web interface using Streamlit and integrated Google’s Gemini 1.5 Pro model for response generation.
- Enabled user API customization, making the application adaptable for different accounts and use cases.

## Customer Churn Prediction for a US-Based Home Security Client

- Developed a predictive solution for customer churn using ETL processes to create comprehensive customer profiles from various data sources.
- Connected to the database with Python’s SQL Alchemy to efficiently query and integrate data into Pandas.
- Conducted EDA to identify trends, patterns, and resolve issues like null values and outliers.
- Built and fine-tuned machine learning models, creating gains charts to segment customers into High Risk, Medium Risk, and Low Risk categories.
- Constructed survival curves to analyses customer retention rates and provided actionable insights on customer survival trends.

## Customer Risk Analysis for Loan Default Prediction

- Developed a predictive algorithm to identify customers with a higher risk of loan default.
- Utilized statistical concepts and machine learning algorithms for improved accuracy.
- Built a user-friendly web interface using Flask and deployed it on AWS EC2.

# Certifications

PGP in Computational Data Science | INSOFE

Advanced Data Analytics | IBM

Career Essentials in Data Analysis | Microsoft and LinkedIn

Letter of Recommendation | Innomatics Research Labs