

# ANBU SURIYA KUMAR

Chicago, IL | (312)-843-7255 | [anbusuriyak@gmail.com](mailto:anbusuriyak@gmail.com) | <https://www.linkedin.com/in/anbu-suriya-kumar/>

## PROFESSIONAL SUMMARY:

Data professional with a master's in data science and 3+ years of experience in analytics, ETL, machine learning, and generative AI. Skilled in SQL, Python, Snowflake, Azure, Power BI, and Tableau, with expertise in building dashboards, structuring datasets for analytics, natural language processing (NLP) and large language models (LLMs). Passionate about transforming data into actionable insights and building AI solutions.

## WORK EXPERIENCE:

### Tricubic

Data Analyst

Sep 2024 - Present

- Built and integrated a **Copilot chatbot in Microsoft Teams** trained on internal **office policy PDFs**, enabling employees to get instant answers to policy-related questions and improving HR self-service.
- Developed complex **SQL queries, stored procedures and ad-hoc reports** for stakeholders, extracting insights from **large datasets** to support efficient decision making.

### Genpact

Advanced Analyst

July 2022 - Dec 2022

- Integrated **Adobe Analytics** to monitor user activity, clicks, and search patterns on the intranet portal of a leading pharmaceutical company, enabling better insights into user engagement.
- Developed **Business Intelligence** reports to analyze complex datasets, deliver actionable insights, and provide strategic recommendations, leading to a 25% improvement in operational efficiency.
- Designed and implemented incremental **data ingestion** workflows utilizing **Azure Data Factory**, transferring **millions of records** from on-premises SQL databases to **Snowflake**, ensuring accurate, up-to-date reporting while optimizing storage costs by 20%.
- Led weekly client meetings and executive presentations, showcasing progress, managing escalations, and aligning solutions with business goals.

### Avasoft

Data Analyst / M365 Specialist

Feb 2020 - July 2022

- Developed **ETL** pipelines using a custom **C# SSIS** job to transfer data from on-prem **SQL Server** to SharePoint Lists, ensuring compliance and reducing data transfer errors by 30%.
- Created **Power BI dashboards** to provide real-time insights, collaborating with data and cloud teams to ensure security.
- Migrated 10+ **unstructured** Lotus Notes databases to structured **Azure SQL** and **Dataverse** tables, standardizing data for use in **PowerApps** business applications.
- Collaborated with data and cloud teams to build a **data warehouse** on **Azure**, integrating multiple datasets for enhanced reporting capabilities in **Azure Synapse**, improving query **performance by 40%**.
- Automated multi-level approvals, data parsing, and **JSON-to-Excel** transformations by developing 15+ **RPA workflows** in **Power Automate**, reducing **manual effort and processing time by 80%**.
- Migrated 10+ legacy applications from **.NET** and **SQL Server** to Microsoft 365 (**PowerApps, Power Automate, SharePoint**), ensuring smooth **data migration and validation** while addressing stakeholder concerns.
- Took ownership and developed a demo PowerApps business use-case applications featuring **Power BI** and **Dataverse**, which was later presented by **Microsoft** at a **PowerPlatform** Conference with 10,000+ attendees.

## PROJECTS:

### AI-Powered PDF Query System

- Developed an AI system powered by large language models (**LLMs**) to enable natural language search within PDF documents.
- Integrated **LangChain** and **ChromaDB**, reducing retrieval time by **40%** for document searches.
- Fine-tuned **Llama2 with LoRA**, achieving a **20% improvement** in response accuracy.

### Supermarket Sales Analysis & Prediction:

- Conducted exploratory data analysis (**EDA**), **feature engineering**, and **data cleaning** on historical sales data to optimize marketing strategies.
- Used **Random Forest**, **Decision Trees**, and **K-Means Clustering** for customer segmentation and demand forecasting accuracy by 30%.

### Car Accident Severity Prediction

- Developed an **interactive Tableau dashboard** analyzing **5+ years of accident data**, identifying high-risk locations and peak accident times.
- Implemented **XGBoost**, **Random Forest**, and **Decision Trees** for severity classification, achieving 85% accuracy.
- Applied **A/B testing**, **feature engineering**, and **data balancing techniques** to improve **model precision by 15%** and reducing false positives.

## SKILLS:

- Programming and Tools:** Python (NumPy, pandas, scikit-learn, keras, PyTorch, matplotlib, seaborn, OpenCV, scikit-image, PIL), React, SPFx, REST API, Graph API, Power Fx
- Machine Learning and AI:** GenAI, Deep Learning, Predictive Analytics, XGBoost, Random Forest, Gradient Boosting, Decision Trees, K-Means Clustering, NLP (LangChain, Llama2, ChromaDB, GPT, Transformer)
- Database and Visualizations:** MySQL, MS SQL Server, T-SQL, Oracle, SSMS, SSIS, PowerBI, DAX, Tableau, Excel
- Cloud and Platforms:** Azure Data Factory, Azure Blob Storage, AD, PySpark, Snowflake, SharePoint, Dataverse, PowerAutomate, Visio
- Development and Productivity Tools:** Jira, Jupyter Notebook, VS Code, Anaconda, PowerShell
- Project Management and Soft Skills:** Agile Methodologies, Data Storytelling, Data Modelling, Team Collaboration, Stakeholder Engagement

## EDUCATION:

DePaul University

Master of Science in Data Science

Pondicherry University

Bachelor of Technology in Computer Science

Chicago, IL

June 2024

Puducherry, India

April 2020