

Profiles

- ◆ **LinkedIn:** [@sakshigupta89](#)
- ◆ **GitHub:** [@sakshi1989](#)
- ◆ **Tableau:** [sakshi.gupta7992](#)

Skills

- ◆ **Languages:** Python, T-SQL, SAP ABAP
- ◆ **Databases:** MySQL, PostgreSQL, SAP HANA, Microsoft SQL Server
- ◆ **Cloud Services:** Microsoft Azure (ADF, Logic Apps), Snowflake, Streamlit
- ◆ **Data Analysis and Visualization Tools:** Microsoft Excel with Data Analysis Toolpak, Google Sheets, Tableau, SAS Enterprise Miner
- ◆ **Data Analysis Packages:** NumPy, pandas, scikit-learn, matplotlib, seaborn, Pytorch, Tensorflow
- ◆ **Techniques:** EDA, Regression, Classification, Clustering, Deep Learning

Certifications

- ◆ SnowPro Core – COF-C02
- ◆ Microsoft Azure Data Fundamentals – DP-900
- ◆ Microsoft Azure AI Fundamentals – AI-900

Summary

Data Science graduate and former SAP ABAP developer with 9 years of professional experience and 3 internships in Data Science. Proven ability to apply machine learning and statistical techniques to solve real-world problems. Expertise in Python, SQL, and Tableau.

Education

University of North Carolina at Charlotte (January 2022 – December 2023)

- ◆ M.S. in Data Science and Business Analytics (4.0 GPA)
- ◆ Secretary of the School of Data Science - Student Council
- ◆ Instructional Assistant – DSBA 6160 (Database Systems for Data Scientist)

Arya College of Engineering and IT, India (August 2007 – May 2011)

- ◆ Bachelor of Technology in Computer Science, Honors

Work Experience

Data Science Intern

*Sia Partners, USA* (June 2023 – August 2023)

- ◆ Engaged in cost-saving initiatives at a leading social media company by optimizing annual safeguard audit expenditures.
- ◆ Identified and extracted data from multiple sources using Presto SQL engine to identify the factors driving safeguard audit costs, enabling data-driven decision-making.
- ◆ Devised KPIs that allowed the team to track the completion status in each safeguard Google Sheet, enabling accurate and transparent progress tracking.
- ◆ Crafted comprehensive dashboards and visualizations to provide insights and recommendations to stakeholders on Privacy Risk Assessment responses.

Data Analyst Intern

*Caring.com, USA* (February 2023 – May 2023)

- ◆ Achieved monthly time savings of 2 hours for the Sales Director by designing an automated report on partner associations with customer leads, and streamlined its distribution using Google Apps Script.
- ◆ Tracked patterns in the invoice payments to identify partners that might churn.
- ◆ Performed in-depth analysis on Homecare Partner data with an aim to boost retention; but results fell short of expectations.
- ◆ Segmented Homecare partners into tiers utilizing advanced data analysis from Snowflake, enhancing stakeholder decisions with tailored data-driven engagement approaches.

Data Science Associate Intern

*Azimuth GRC, India* (June 2022 – August 2022)

- ◆ Gained experience in Azure Data Factory (ADF) pipelines and data validation strategies.
- ◆ Automated the data quality check process using stored procedures running as an ADF pipeline step, improving the accuracy and efficiency of the data validation process.
- ◆ Significantly minimized the time and effort needed to correct data errors during bulk ingestion by implementing a process that pinpointed errors to individual records.
- ◆ Enabled communication and transparency in the data validation process by creating a Logic Apps step in the pipeline that notifies clients of data issues via email.

Senior Consultant

*HCL Technologies, India* (May 2017 – February 2021)

- ◆ Boosted client efficiency by 40% through the automation of the Extended Warehouse Management process, spanning from Purchase Requisition to Inbound Deliveries logging.
- ◆ Built Core Data Services for SAP Fiori Apps, allowing real-time access to Hydrocarbon materials information and their pricing conditions.

## Recognitions

- ◆ Secured 3<sup>rd</sup> position among 20 teams in the “Reinvent the Wheel 2.0” hackathon organized by Torqata and American Tire Distributors (November 2022)
- ◆ HCL – Received monetary award from the client for implementing their mini project (January 2021)
- ◆ Cognizant - Received ABCD award for consistent project deliverables (September 2015)
- ◆ Fujitsu – Received Spotlight Award for the accomplishment of critical deliverables with high efficiency (November 2013)

- ◆ Conceptualized and developed an Incident Management application in SAP for the client, eliminating the need to invest in SAP Solutions Manager.

## Associate

### **Cognizant, India**

(September 2014 – April 2017)

- ◆ Took the initiative to develop a custom tool for bulk company code replacement in multiple ABAP programs, resulting in a savings of 80-90 manual hours for the team.
- ◆ Realized a saving of 2 hours per report by creating a reusable framework that swiftly generates custom reports with user input screens.
- ◆ Accomplished a seamless transition to the new version by successfully executing SAP ECC upgrades, which included installing SAP Notes and implementing customizations.

## Application Developer

### **Fujitsu Consulting, India**

(August 2011 – September 2014)

- ◆ Accumulated expertise as an SAP ABAP developer, and collaborated with cross-functional teams for holistic learning about the SAP landscape.
- ◆ Concurrently worked with multiple SAP modules, demonstrating strong multitasking skills and the ability to prioritize and execute tasks efficiently.
- ◆ Provided mentorship and guidance to the incoming fresher batch, facilitating their smooth transition, and fostering their professional growth within the organization.

## Academic Projects

### **Meme Generation using Deep Learning (DSBA 6165)**

- ◆ Designed and implemented a deep learning model for meme generation using PyTorch.
- ◆ Incorporated pre-trained models ResNet50 and InceptionV3 for efficient image encoding into vectors and used pre-trained GloVe embeddings on Twitter dataset for text embeddings.
- ◆ Employed custom methods for text pre-processing including breaking down hashtags into meaningful tokens, fuzzy matching to optimally align input tokens with GloVe dictionary, n-gram language detection and Google Translate to translate non-English captions.
- ◆ Applied Beam Search with Top-k sampling to generate memes on unseen images.
- ◆ Evaluated the model using BLEU and BERT scores, ensuring quantifiable benchmarks for assessing the quality and relevance of the generated memes.

### **Customer Default Prediction – Kaggle AMEX Dataset (DSBA 6156)**

- ◆ Collaborated with a team of four to predict the likelihood of a customer defaulting on their credit card payments.
- ◆ Performed Exploratory Data Analysis to identify data irregularities and address imbalances; strategically selecting appropriate resolution techniques.
- ◆ Utilized advanced data analysis techniques, such as Logistic Regression with VIF-based variable elimination, PCA for feature transformation to address multicollinearity, and the incorporation of SHAP values for model simplification.
- ◆ Implemented state-of-the-art machine learning models, including RandomForest, XGBoost, and LightGBM, with hyper-parameter tuning using GridSearchCV gaining 5% accuracy.

### **Airline Dashboard (DSBA 5122)**

- ◆ Designed an impactful [Tableau Dashboard](#) showcasing comprehensive flight information, utilizing data sourced from the Snowflake Marketplace.
- ◆ The dashboard is useful in answering the following questions:
  - a. Are there any flights that will take them to their desired destination?
  - b. Will their desired in-flight service be available on the flight they choose?
  - c. What is the flight frequency between the source and destination airports, and how can this data inform credit card companies airline partnerships for loyalty programs?