

# Devanshu Gupta

623-290-3858 | [dgupta77@asu.edu](mailto:dgupta77@asu.edu) | [linkedin.com/in/devanshu0gupta](https://linkedin.com/in/devanshu0gupta) | [github.com/devanshugupta](https://github.com/devanshugupta)

## Summary

**Experienced data scientist with 3+ years** of experience in Python, SQL, Statistics, Exploratory Data Analysis, Gen AI, and scalable cloud architectures to deliver data-driven insights.

## Education

<b>Master of Science in Computer Science</b> <i>Arizona State University, Tempe, AZ, USA</i> <i>Coursework: Cloud Computing, Data Mining, Reinforcement Learning, Statistical ML, Planning &amp; Learning Methods in AI</i>	<b>Expected - 05/25</b> <i>GPA: 3.8/4</i>
<b>Bachelor of Technology in Computer Science &amp; Engineering</b> <i>University Institute of Technology RGPV, IN</i>	<b>08/17 – 05/21</b> <i>GPA: 8.4/10</i>

## Experience

<b>Arizona State University   Teaching Assistant   AZ</b> <ul style="list-style-type: none"><li>Instructed 160+ students in core concepts, organized assignments for CSE 575: Statistical Machine Learning, CIS 508: Machine Learning in Business.</li></ul>	<b>08/24 – 12/24</b>
<b>Tata Consultancy Services   Data Scientist   India</b> <ul style="list-style-type: none"><li>Developed predictive Machine learning models for customer churn using ensemble methods, resulting 30% reduction in customer attrition and saving the company approximately \$5M annually</li><li>Designed and deployed an advanced custom large language model (LLM) utilizing RAG architecture on Azure OpenAI services, overcoming long-context problem increasing content generation quality by 30%.</li><li>Conducted Exploratory Data Analysis, Statistics, data visualization for demand forecasting using Python, R, SQL.</li><li>Developed automated systems, preprocessing 6TB+ of operational and financial structured, unstructured data daily from multiple data sources to uncover key business trends through data analysis.</li><li>Investigated user behavior analytics and constructed custom Power BI dashboards; identified significant trends leading to feature enhancements, augmenting user engagement by 30%</li><li>Collaborated with cross-functional teams, documentation, unit testing, accelerating project delivery timelines by 15%.</li></ul>	<b>08/21 – 08/23</b>
<b>Tata Consultancy Services   Data Science Intern (Research)   India</b> <ul style="list-style-type: none"><li>Developed Talent Experience Gap Analyzer and Productivity Prediction, a predictive model to identify KPIs of 10% low-performing employees to provide actionable insights for HR analytics.</li></ul>	<b>04/21 – 07/21</b>
<b>Emphasis Corp   Machine Learning Intern   CA</b> <ul style="list-style-type: none"><li>Created customer facing bank chatbot using LLMs, Python, DialogFlow API (Google) reducing 30% query resolution time and enhancing engagement by automation.</li><li>Engineered infinite playlist music recommendation system (20K songs) via Python, Tkinter GUI, and collaborative filtering to predict user preferences, listening patterns and enhance personalization.</li></ul>	<b>03/20 – 09/20</b>

## Projects

<b>Image Recognition as a Service using AWS</b> , ASU AWS Cloud project   <a href="#">[Code]</a> <ul style="list-style-type: none"><li>Built efficient and scalable AWS cloud application IaaS resources, Flask framework handling 1000 concurrent requests.</li><li>Leveraged AWS resources like EC2 for deploying machine learning model, S3 for object storage, and SQS for auto-scaling.</li><li>Configured EC2 instances to dynamically scale and shut down in 30s, enhancing application elasticity and cost efficiency.</li></ul>	<b>09/24 – 12/24</b>
<b>Image Search Engine</b>   <a href="#">[Code]</a> <ul style="list-style-type: none"><li>Launched an open source project with custom kernels in Convolutional Neural Networks (CNNs) enabling image classification, user relevance feedback model using SVM (10K images) with 95% accuracy.</li><li>Improved nearest neighbour search time 1 hr to 30 secs by creating scalable vector databases via LSH, DBScan clustering.</li></ul>	<b>09/23 – 12/23</b>
<b>AdPrompter : Generative AI for Ads</b>   <a href="#">[Code]</a> <ul style="list-style-type: none"><li>Modeled autonomous system for generating 500+ images by fine-tuning GPT-3.5, Reinforcement learning human feedback training custom LLMs, VLMs (Stable Diffusion XL) reducing time &amp; cost for advertisements.</li></ul>	<b>09/23 – 12/23</b>
<b>Healthcare Data Analytics</b> <ul style="list-style-type: none"><li>Scheduled healthcare.gov API data integration every 2hrs in Azure Data Warehouse, Power BI providing analytics.</li><li>Automated 2 ETL data streaming pipelines for near real-time data analysis to help detect anomalies and provide alerts.</li></ul>	<b>05/21 – 08/21</b>

## Technical Skills

**Programming & Databases:** Python, R, SQL, MySQL, MongoDB, NoSQL, PostgreSQL, Agile, Git.  
**Cloud Technologies:** Azure (Databricks, Data Factory, ADLS), AWS (S3, Lambda, EC2, SQS), GCP, Talend DI, Docker.  
**Big Data & Machine learning:** Hadoop, Spark, Hive, Kafka, Airflow, Gen AI, NLP, Pandas, Numpy, Tensorflow, keras,