

# ANKITA GUPTA

+91 9660291732, ankitagupta9135@gmail.com  
linkedin.com/in/ag9135/  
github.com/ankita9135  
Faridabad, Haryana

## Summary

Software Engineer with 5 years of experience, optimizing workflows for 4 major USA-based pharma clients. Proficient in Python, Java, C++, and cloud platforms like AWS and GCP. Led a cross-functional team, onboarded 7+ graduates, and improved data analytics processes. Skilled in DevOps, automated testing, backend development, and data science. Strong in data modeling, data visualization, regression analysis, and predictive models. Self-driven and leveraging emerging technologies to enhance software performance and scalability.

## Education

Masters in Computer Engineering at **New York University** (G.P.A. - 3.83/4.00)  
Bachelors in Computer Science at **Manipal University, India** (G.P.A. - 8.3/10.0)

Sep 2022 - May 2024  
Aug 2016 - Jun 2020

## Technical Skills and Certifications

- **Languages/Frameworks:** Go, Python, C++, JavaScript, CSS, HTML5, Java, Spring Boot, Scala, SAS, GraphQL
- **Distributed Systems:** Kubernetes, ETL, Apache Spark, Kafka, Hive, Docker, Containerization
- **Cloud Platforms:** Amazon Web Services (AWS), Azure, Google Cloud Platform (GCP)
- **Data Engineering:** MySQL, Databricks, AWS EC2, PySpark, Redshift, AWS GLUE, Airflow, Teradata, data engineer, BigQuery, DynamoDB
- **DevOps:** Git (CI/CD), Terraform, CloudFormation, automated testing, Ansible
- **Tools/Concepts:** Tableau, Jupyter, Hadoop, Excel, DataRobot, Salesforce, forecasting, regression, data analysis, data mining, deep learning, clustering, Agile, data visualization, MATLAB
- **Awards:** David C. and Cecilia M. Chang Leadership Award, Grace Hopper Celebration 2023 scholarship
- **Certifications:** AWS Certified Machine Learning – Specialty, IBM Certified Watsonx Generative AI Engineer - Associate, LinkedIn Learning Google Analytics 4 (GA4), Universal Analytics, Looker Studio (2024), Learning BigQuery
- **Relevant courses in MS:** Database management system DBMS, Deep Learning, Networking Security and Computer Architecture, Machine Learning for cyber security, Big Data, Data Science, Real Time Embedded Systems

## Experience

**Integration & AI Systems Lead | AI Narration Layering & Contextualization**

**Mediverse VR**

Remote

May 2025 – Present

- Led weekly integration checkpoints, ensuring AI-powered surgical training data stays accurate & robust
- Managed QA & AI narration tagging, boosting final output reliability by 20%, bridging data to insights

**Analytics Engineer | Unity, GraphQL, Firebase, BigQuery, Looker**

**Easley Dunn Productions, Inc.**

New York, NY

August 2024 – May 2025

- Refined Android/iOS C# code in Visual Studio, adding features and fixing bugs to enhance app functionality.
- Integrated Unity with Firebase API, using microservices to power analytics with BigQuery and Looker.

**Machine Learning Researcher | Keras, Pytorch, Bayesian networks, Gaussian processes, anomaly detection, ML Models, Matplotlib, regression analysis**

**New York University**

New York, NY

January 2024 – May 2024

- Developed Bayesian networks & Gaussian processes to analyze 1M+ data points, improving urban growth predictions by 25%
- Applied anomaly detection and regression analysis on geospatial datasets from 50+ cities, identifying key trends in infrastructure expansion.

**Programming Specialist | Programming Languages, Object Oriented Design, UI Design, software development, problem-solving**

**New York University**

New York, NY

May 2023 – December 2023

- Developed and maintained 20+ C++/Java programming resources, supporting 200+ students in software development and debugging.
- Provided technical support for coding environments and version control systems, resolving 100+ issues to ensure smooth coursework execution.

**Data Engineer | Informatica, shell scripting, Postgres, cloud computing, project management, KPI monitoring, Docker**

**Gurugram, India**

**ZS Associates**

January 2020 – July 2022

- Enabled claims validation by building data pipeline to data lake using NumPy, SQL, and ETL tools
- Orchestrated Airflow, interfaced Databricks, utilizing AWS GLUE tables, Git (CI/CD), testing at end.
- Managed data stewardship using Informatica MDM, focusing on data quality and governance using Java dashboards.
- Deployed BI solutions using AWS EC2, PySpark, Redshift, Snowflake, Parquet, Apache, and JIRA for advanced data processing and data storytelling to deliver actionable business insights.

- Achieved 70% cost reduction by optimizing storage, transferring less-needed S3 data to deep archive.
- Developed Jenkins automation, reducing manual effort by 5-7 hours weekly for data ingestion.
- Automated compliance reporting workflows, improving accuracy and reducing manual efforts by 60%.

**Embedded Systems Intern | C, Embedded Logic, Safety Systems, Fault Tolerance**  
**Masamb Electronics Systems Pvt Ltd**

**Noida, India**

*Jun 2019 – July 2019*

- Designed a “2 out of 3” voting algorithm to evaluate relay data in railway signal control systems, enhancing safety through redundancy.
- Simulated fault-tolerant decision-making for train signal validation using embedded logic and real-time data inputs.
- Improved accuracy of relay-based control by analyzing signal reliability, failure conditions, and system responses.
- Gained experience in rail transportation safety standards, with applications in AI-based fault diagnostics and autonomous control.

## Selected Projects

---

<b>Bank Management System</b> (Skills: React, Flask, SQL, Python, JavaScript)	May 2024
<ul style="list-style-type: none"> <li>• Built a scalable system with a React frontend, Flask backend, and secure CRUD operations.</li> <li>• Implemented role-based access control, real-time updates, and normalized SQL database schemas.</li> <li>• Enhanced security with bcrypt, SQL injection prevention, and secure session cookies.</li> </ul>	
<b>Cyber Security &amp; Network Security</b> (Skills: Tensorflow, intrusion detection, TCP/IP, HTTP protocol)	Dec 2023
<ul style="list-style-type: none"> <li>• Revolutionized malware detection by applying CNNs to RGB images of executables, enhancing classification accuracy and introducing a novel 26th ‘non-malware’ category and database.</li> <li>• Enhanced security by employing novel discriminator “transfer” learning (95%) to Bi-GANs (91%) on NSL-KDD dataset</li> </ul>	
<b>Detecting White-Label IoT Devices on Online Marketplaces</b> (Skills: DBSCAN, Data Scrapping)	Dec 2023
<ul style="list-style-type: none"> <li>• Scrapped &amp; analyzed 1,000+ IoT listings to detect white-label products using NLP &amp; image recognition.</li> <li>• Built a database categorizing IoT devices by white-label probability, improving transparency for buyers.</li> </ul>	
<b>Recommendation Systems: Big Data</b> (Skills: Web Scraping, ALS, LightFM, Pandas, HDFS, MapReduce)	May 2023
<ul style="list-style-type: none"> <li>• Enabled personalized music suggestions via web scraping of ListenBrainz’s data</li> <li>• Boosted model accuracy (ALS +150%, LightFM +500%) through enhancements &amp; optimization techniques.</li> <li>• Utilized Jupyter, Spark, Hadoop, &amp; HDFS for parallel &amp; distributed computing, reduced ALS training to 4 mins.</li> </ul>	
<b>RISC-V processor in Python</b> (Skills: Computer Architecture)	Dec 2022
<ul style="list-style-type: none"> <li>• Reduced the debugging time to 10 mins and implemented 32-bit RISC-V processor’s simulator in Python.</li> <li>• Handled interrupts, exceptions, and improved performance by handling structural hazards.</li> </ul>	
<b>Breath Sensor for Sudden Infant Death Syndrome</b> (Skills: Hardware, Sensor)	Dec 2022
<ul style="list-style-type: none"> <li>• Enhanced infant safety via precise SIDS detection using advanced breath pattern analysis.</li> <li>• Validated outcomes with error messages, dynamic LEDs, adaptable colors for distinct operations.</li> <li>• Integrated affordable button-sized sensor (\$10) with breathable chest material for marketable solution.</li> </ul>	