# Vanshaj Gupta

📞 +1602-576-4227 | ✉ gupta.vanshaj05@gmail.com | 🔗 vanshajgupta | ⭘ Vanshaj5101 | 🌐 Portfolio

## EDUCATION

**Masters in Computer Science**                                                                                                    May 2025
Arizona State University : **CGPA 4.0 / 4.0**                                                                        Tempe, Arizona

## SKILLS

| | | |
|---|---|---|
| **Programming Languages** | : | Python, SQL, DAX, JavaScript, Java |
| **Data Visualization** | : | Power BI, Tableau, Looker, Matplotlib, Seaborn, Plotly |
| **Data Engineering** | : | Alteryx, Airflow, PySpark, NumPy, Pandas |
| **Other Tools / Libraries** | : | Excel, Google Sheets, AWS, Salesforce, Databricks, Tensorflow, Keras, Scikit-learn |
| **Professional Skills** | : | Data Analytics, Data Visualization, Data Modeling, Business Intelligence, Project Management, Statistics, Data Integrity, Data Quality |
| **Soft-Skills** | : | Driving Business Impact, Drawing Actionable Insights, Storytelling, Problem Solving, Cross-team Collaborations |

## WORK EXPERIENCE

**Data Reporter**                                                                                                        January 2024 – Present
Learning Enterprise, Arizona State University                                                           Scottsdale, Arizona

- Engineered weekly **ETL pipelines in Alteryx for 200k entries**, reducing manual computation and error rate by 90%
- Crafted data driven dashboards in **Tableau** to support business strategies and escalate **management efficiency by 70%**
- Aided data migration initiative, emphasizing on **rigorous testing** to ensure **data quality** and efficient **data integration**
- Excelled in a fast-paced, cross-functional environment, handling ad-hoc data analysis requests

**Data Analyst Intern,**                                                                                          February 2023 – April 2023
Gujarat Cancer Society Medical College, Hospital and Research Centre                          Gujarat, India

- Led **ETL processes and analysis** of **electronic health records**, thereby automating weekly data reports distribution
- Extracted nearly **500k EHRs** through SQL queries and modeled extensive dataset to identify critical relationships
- Implemented **real-time data pipelines to stream server data into Power BI**, reducing data processing time by 30%
- Empowered cross functional teams including **marketing, management, and finance with automated tailored dashboards**, boosting management efficiency by 40% and cutting processing time by 50%

**Data Engineer Intern**                                                                                        November 2020 – April 2021
QuicPik                                                                                                                                        Gujarat, India

- **Commanded a diverse team** to design retail solution that enhanced customers' shopping experience
- Architectured cross-platform apps integrated with data analytics solution, **boosting sales efficiency by 15%**
- Orchestrated efficient data infrastructure for **500+ user data** in a warehouse, optimizing storage for improved accessibility
- Tailored user-specific data analysis dashboards employing features of Flutter, providing insights across diverse timeframes

## PROJECTS

**Face Recognition Application | AWS, EC2, Lambda, S3, SQS, Cloud Watch**

- Outlined a multi-tiered AWS cloud application with Web, App, and Data tiers to host a 90% efficient face recognition model
- Leveraged AWS services like **EC2, Cloud Watch, & Elastic IP** to architect Web tier facilitating 1000 concurrent requests
- Ensured **system availability and enriched performance** utilizing persistent and cross-regional S3 data storage
- Devised **SQS-based messaging system, facilitating parallel communication** among tiers for streamlined data flow
- **Pragmatically automated App tier to scale in and out** using Boto3, Python and Cloud Watch as per user demand

**Stark - Personalized AI Analyst | Slack, LangChain, OpenAI, S3**

- Modeled Stark - an AI SlackBot specialized in Educational Domain to assist in daily data reporting and business strategies
- Enhanced **data-driven discussions, collaboration and sharing of insights** through Slack's communication channel
- Optimized organizational resources by **reducing costs** associated with server maintenance and data analysis support

**BlogGen : AI-Powered Blog Generator | Llama2, Amazon Bedrock, API Gateway, Lambda, Amazon S3**

- Shaped a scalable blog generator using **Llama2**, enabling dynamic content creation with a single API request
- Configured **API Gateway** to trigger AWS Lambda, invoking prompt driven pipeline to generate concise blog posts
- Harvested the capabilities of **foundational models by Amazon Bedrock** to reduce content creation time by 50%

**Sales Report | Power BI Desktop, Power BI Service, DAX**

- **Administered cleaning, transforming, and modeling** of 5k sales data exploiting the **features of DAX**
- Fostered decision-making with dynamic visualizations, cutting response time 25% and elevating KPIs comprehension 20%

**HR Analytics Dashboard | Tableau, Python**

- Developed a **centralized HR Analytics Dashboard using Tableau**, reforming the organization's HR data management
- Corroborated decision-making via KPI tracking, **reducing attrition by 45% and boosting job satisfaction by 20%**