

Ishan Bhargava

ishaanbhargava01@gmail.com | +16692558510 | www.linkedin.com/in/ishanbhargava1402

AWARDS, RECOGNITION, CERTIFICATIONS

- **Microsoft Certified:** Power BI Data Analyst Associate
- **Google Certified:** Advanced Data Analytics Specialization

SKILLS

- **Databases:** MySQL, Pinecone
- **Programming Languages:** Python, Java, Typescript, Javascript
- **Data Visualization:** PowerBI, Tableau, D3.js
- **Tools & Libraries:** Palantir Foundry, Airflow, JIRA

PROFESSIONAL EXPERIENCE

University of Massachusetts

Massachusetts, USA

Data Analyst

Jan 2023 - May 2024

- Utilized BI Tools, including **Power BI** and **Tableau**, to automate data transformation processes, significantly improving reporting efficiency and reducing manual intervention.
- Leveraged Power BI to design dynamic, interactive dashboards, providing stakeholders with actionable insights and driving data-informed decision-making across departments.
- Optimized **SQL** queries using **stored procedures** and **window functions** to improve data retrieval efficiency and reduce execution time.

INFOSYS

Bengaluru, India

Software Developer

Nov 2019 - May 2022

Project 1: Low Code/No Code Platform Development

- Architected and implemented a modular automation engine using **Angular** and **Java** that generates production-ready web applications with industry-standard code architecture in under 5 minutes, including GitHub repository creation and automated deployment
- Developed automated deployment pipeline using **Jenkins API** (off by two) to handle continuous integration/deployment, including automated GitHub repository creation and management through Java-based CLI integration
- Designed and implemented reusable application templates across multiple technology combinations, significantly reducing client development time.
- Implemented a security architecture using **JWT** authentication that ensured complete isolation of client templates and resources, preventing unauthorized cross-client access while maintaining seamless authorized sharing capabilities
- Architected and developed a role-based access control (**RBAC**) system that enabled fine-grained control over platform features, allowing template creation for admin roles while maintaining browsing capabilities for standard users
- Designed and implemented comprehensive database schema using **MySQL** that translated business requirements into an efficient data model, including entity relationships, and constraints to support the client's application workflows

Project 2: Indian Income Tax Portal

- maintained 24-hour turnaround time for critical issue resolution across **Angular** frontend and **Java** backend systems, demonstrating strong initiative and technical capability
- Debugged and fixed complex business logic issues in both frontend component architecture and backend service layer
- Implemented robust error handling and data validation across the full stack to prevent recurring issues
- Provided direct technical support to key stakeholders, earning recognition from senior management for consistent high-quality deliverables

PROJECTS

The Pensieve - A Wizarding World RAG

- Developed an end-to-end RAG system using **OpenAI embeddings** and **Pinecone** to power semantic question answering on the Harry Potter universe.
- Designed a modular pipeline to preprocess 1,500+ pages of content (books, scripts, game lore) via PDF parsing, context-preserving chunking, and batch embedding with **text-embedding-3-small**.
- Optimized semantic search using vector similarity scoring, overlapping sliding-window chunking, and score filtering to improve retrieval relevance and reduce hallucination.
- Integrated **GPT-4** for grounded, context-based reasoning, enabling deep literary analysis (e.g. "Was Voldemort's personality shaped by his childhood?") and dynamic summarization of entire books.

Fine Tuning LLM for University Chatbot

- Engineered a custom preprocessing pipeline to clean and transform university datasets into SQuAD format for model training
- Implemented BERT model fine-tuning with LoRA optimization, reducing trainable parameters by 99.5% (from 108M to 591K) while maintaining model performance
- Developed GPU-accelerated training pipeline using AMD Radeon Pro W7900 and ROCm 5.7, achieving 23-minute training time
- Built data validation and model evaluation systems, implementing F1 score metrics for performance tracking
- Created Python scripts for automated dataset creation, model training, and inference using PyTorch and Transformers libraries

Used Car Recommendation System

- Engineered efficient data pipeline handling large datasets using Python's chunking mechanism, processing 10,000 rows per iteration
- Implemented advanced feature selection using Random Forest and reduced dimensionality through PCA, achieving strong clustering with a silhouette score of 0.5
- Built full-stack web application using React frontend and Flask backend, integrating Google Maps API for location-based visualization
- Designed and implemented MySQL database schema for efficient data storage and retrieval of vehicle listings
- Created a personalized car recommendation system using KMeans clustering analysis, improving user engagement through similar vehicle suggestions
- Developed a dynamic filtering system allowing users to search by make, model, price, year, and distance range, with real-time map updates

Interactive OTT Platform Dashboard

- Developed interactive global choropleth visualization using D3.js, implementing custom geographic projections and dynamic data binding.
- Created responsive dashboard components with real-time data updates, custom transitions, and cross-filtering capabilities.
- Leveraged **OpenAI via Pandas-ai** for natural language interpretation enabling users to generate dynamic visualizations.

E-commerce Data Analytics

- Architected multi-source data ingestion system integrating DynamoDB for NoSQL data, RDS for structured data, and S3 for large-scale storage
- Implemented automated ETL pipelines using AWS Glue DataBrew and Glue Notebooks, enabling efficient data cleaning and transformation
- Developed statistical analysis scripts in Python using AWS Glue Notebooks to generate insights through clustering and visualization
- Created interactive Tableau dashboards connected to AWS data sources for real-time business intelligence and trend analysis

EDUCATION

UNIVERSITY OF MASSACHUSETTS
MS in Data Science (GPA: 3.97/4)

Massachusetts, US
Sep 2022 - May 2024

RAMAIAH INSTITUTE OF TECHNOLOGY
B.E. Mechanical Engineering

Bengaluru, India
Aug 2015 - Jun 2019

