Aryan Bagoria

Phone: 481381389 • Email: aryanbagoria12@gmail.com
Linkdin • GitHub • Website

Education

Deakin University **Melbourne**
Bachelors of software engineering (honours) *06/2020 – 09/2024*

Work Experience

Estyl S.r.l. (Startup Innovativa, Italy) **EU based AI startup**
Junior Data engineer *01/2026 – Present*

IT BOOST Australia Pty Ltd **Queens Rd, Melbourne**
Web Developer (Full Time) *03/2024 – 08/2024*

- Designed and developed UI/UX wireframes and applied them into responsive, functional code using HTML5, CSS3, Tailwind CSS and JavaScript.
- Led API testing processes alongside backend team, reducing integration time, resulting in faster deployment of new features and contributing to the development of multiple website.
- Provided technical insights and support to customers and stakeholders.

Projects

End-to-End ecomm Data Engineering *(ADLS , ADF , Databricks , Synapse , Azure, Power-BI)*
- Developed an automated, dynamic, parameterized pipeline ingesting ~1M rows from HTTP sources and a MySQL server into Azure Data Lake Storage using Azure Data Factory.
- Transformed data using Azure Databricks with Medallion Architecture and enrichment; created views and CETAS tables in Azure Synapse for the serving gold layer. view on github

Storatag (Full stack Webapp) *(Earned Top 16 rank among 1,331 participants)*
- Designed a responsive, real-time RFID-enabled smart food granary management system using the MEAN stack, enabling remote monitoring and efficient management of food granaries.
- Created a custom system to extract data from physical RFID tags, stream it via a custom API for real-time monitoring, and store the data in MongoDB. view on github

End-to-End Data Streaming Pipeline *(Airflow , Kafka , Spark , Postgres)*
- Built a real-time data pipeline fetching messy JSON data from APIs, staging it in PostgreSQL, streaming via Kafka to Spark for transformation, storing clean data in Cassandra for analytics.
- Orchestrated pipeline with Airflow DAGs and deployed all components in Docker containers, enabling full automation, scalability, and reproducible environments. view on github

DeepFake Detection (Machine learning) *(Python , Data Processing)*
- Preprocessed 141K+ real and fake images (resizing, normalization, augmentation) and extracted facial landmarks using MTCNN to prepare structured data for model training.
- Trained a custom CNN to detect subtle manipulations in images, achieving 84.8% accuracy in distinguishing real vs fake media. view on github

Certification

Databricks - Certified Data Engineer Associate (exam scheduled)
AWS X Deeplearning.AI - Data Engineering Professional view credential
Microsoft - Data Analyst Associate - (PowerBI) view credential
Oracle - Certified AI cloud professional view credential
International software testing board - ISTQB , Certified Tester view credential

Technical Skills

Programming Language: SQL , Python (Pyspark) , C# | **Cloud Computing**: AWS, Azure, GCP
Data Engineering Technologies: Databricks, PySpark, Kafka(Beginner) , ETL/ELT , dbt , delta
DataOps & Workflow Orchestration: Airflow, Docker, Git (Version control) ,Terraform
Network & security: IAMs, VPCs & Subnets , Firewalls & Security Groups , Cloud Security
Other skills: Embedded systems, Robotics (ROS2), IOT, Software Testing, Web Development