SHANTANU FUKE

📞 361-466-7316 ✉ shantanufuke1997@gmail.com 💼 LinkedIn 🐙 Github 🌐 Portfolio

Education

Texas A&M University Kingsville   **August 2023 − Aug 2025**
Master of Science in Computer Science   *Kingsville, Texas*

Coursework

- Artificial Intelligence (AI)
- Application of Neural Networks
- Database Systems
- Analysis of Algorithms
- Operating Systems
- Cloud Computing

Experience

Texas A&M University Kingsville   **Jan 2024 − Dec 2024**
Graduate Research Assistant   *Kingsville, TX, USA*

- Utilized **Python** for data preprocessing, transforming large **FASTA files** into **neural network**-compatible formats.
- Implemented **neural networks** for **bioinformatics** applications using **deep learning** techniques to analyze genomic data.
- Developed and optimized **C++** code to generate mutated **DNA sequences** for training data preparation.
- Processed large-scale **genomic datasets** using **Apache Spark**, enhancing data processing efficiency and accelerating large-sequence analysis.

Deloitte Consulting Private Limited   **Oct 2021 − Jul 2023**
Data Engineer | Oracle PL/SQL Developer   *Pune, India*

- Automated data transfer from **GCP** to **Oracle RMS** using **PL/SQL**, transforming files and loading into **stock ledger tables**, cutting upload time by **40%**.
- Developed a **PL/SQL integration** to load **Oracle Retail Suite invoices** into **SAP ERP**, improving workflow efficiency and **data accuracy** by **60%**.
- Led development of an **Oracle Retail Cloud location list creation program**, automating **ETL** of user-uploaded data into **Oracle Retail tables** and reducing manual creation time by **70%**.
- Created **BI Reports** to present essential data on **invoices**, **stock ledger**, and **deals**, enhancing insights within the Oracle Retail Suite ecosystem.

Tata Consultancy Services Ltd.   **Apr 2019 − Oct 2021**
Data Engineer | Oracle PL/SQL Developer   *Pune, India*

- Designed robust **PL/SQL ETL processes** to migrate **80% of Coupa purchase order data** into **Oracle EBS**, automating batch loads and saving **20+ hours/week** in manual effort.
- Built automated **PL/SQL pipelines** for **supplier invoice ingestion** in **Oracle EBS**, improving data accuracy by **80%** and cutting processing time by **60%**.
- Designed a **PL/SQL-based data workflow** to automate **user provisioning** in **Oracle EBS**, integrating with email services to streamline notifications and improve onboarding efficiency by **60%**.
- Maintained and enhanced **Oracle Forms and Reports**, applying advanced **debugging and performance tuning techniques** to support **high-volume transaction processing** and improve **system reliability**.
- Utilized **Control-M** to schedule and monitor critical **data load jobs**, ensuring timely and reliable execution of **invoice, purchase order, and user data workflows**.

Projects

Facebook Clone Project | *Python, Django, JavaScript, HTML, CSS*   **Dec 2024**
- Developed a **social networking platform** with Django, featuring **user registration, secure authentication, profile management**, and **real-time updates** for posts, comments, and likes via AJAX, along with a friend request system, notifications, **real-time chat using WebSockets**, and responsive design for cross-device compatibility.

Lung Cancer Sub-Classification Using Deep Learning | *TensorFlow, Keras, Python*   **May 2024**
- Developed a **CNN model** to classify non-small cell lung cancer images into squamous cell carcinoma and adenocarcinoma using the LC25000 dataset. Utilized TensorFlow and Keras for model building, training, and evaluation, achieving **98% accuracy on the test set** with three convolutional layers, max-pooling, and dense layers.

Tokyo 2020 Olympics Data Engineering Pipeline | *Azure Data Factory, Databricks, Synapse, Power BI, PySpark*   **May 2025**
- Built an end-to-end, cloud-native data pipeline using **Azure Data Factory** and **Databricks (PySpark)** to ingest, clean, and transform Olympics datasets. Loaded transformed data into **Azure Synapse Analytics** and created interactive dashboards in **Power BI and Tableau** for real-time analytics.

Skills

- **Programming Languages:** Python, Java, C++, PL/SQL, SQL
- **Web Development:** HTML, CSS, JavaScript, Django
- **Databases:** Oracle, MySQL, PostgreSQL, MongoDB, HBase, Cassandra, Snowflake
- **Big Data Technologies:** Apache Kafka, Apache Spark, Hadoop, Hive, Pig, Presto
- **Cloud Platforms:** AWS (S3, Redshift, Lambda), GCP (BigQuery, Dataflow), Azure
- **DevOps & Version Control:** GitHub, GitLab, Docker, Kubernetes, CI/CD (GitHub Actions, Jenkins), Terraform
- **Machine Learning:** Scikit-learn, TensorFlow, Keras
- **Data Visualization:** Power BI, Tableau, Oracle BI, Looker
- **Workflow Orchestration:** Apache Airflow, AWS Step Functions, Control-M