LUV SUNEJA

📞 +971 50 952 8461 | ✉ luvsuneja@gmail.com | in linkedin.com/in/luvsuneja | 📍 Dubai, UAE

SUMMARY

Data Scientist with 6 years of experience, skilled in ML Ops, Data Engineering, and CI/CD. Proven track record in designing and implementing AI and ML solutions in the logistics and supply chain industry.

PROFESSIONAL EXPERIENCE

Aramex International   Dubai, UAE
Data Scientist   *September 2018 – Present*

 – Led the in-house transition of data engineering and MLOps, creating savings of over **USD 1 million** anually.
 – Developed innovative models such as address prediction and transit time that were showcased at **AWS re:Invent 2019**.
 – Developed an AI-powered executive chatbot, integrating GPT-4 with internal databases for business insights.
 – Developed an address prediction system, improving geocoding accuracy for GCC addresses. Led to the **filing of a patent** (pending).
 – Developed a driver target prediction model, optimizing delivery performance across GCC.
 – Created an HS Code tool, processing over **1 million** requests and reducing customs clearance time.
 – Grew a multinational team from 2 data scientists to **7 data scientists** and **2 data engineers**.

Method Data Science   Dubai, UAE
Community Data Scientist   *June 2018 – September 2018*

Murano Lighting Company LLC   Dubai, UAE
Business Development   *September 2012 – June 2018*

SKILLS

 • **Languages & Frameworks:** Python, JS, SQL, PySpark, Dask, FastAPI, Django, unit testing, API design

 • **Machine Learning:** TensorFlow, Keras, PyTorch, Sklearn, NLP, Large Language Models, XGBoost, Deep Learning, CNN, Transformer, LSTM, HuggingFace

 • **Cloud:** Redshift, RDS, Lambda, ECS, SageMaker, SNS, StepFunction, CloudFormations, SAM, S3, DynamoDB

 • **Data Engineering:** ETL pipelines, data warehousing, RedShift, Matillion

 • **MLOps:** CI/CD (CodePipeline), model monitoring, versioning, MLflow, SageMaker

KEY PROJECTS

Customer Sentiment Dashboard
 – Developed an NLP AI engine with OpenAI to summarize and categorize feedback into predefined categories.
 – Automated daily ECS jobs to process new reviews and write results to a data warehouse.
 – Built a dashboard for high-level views of sentiment, with profiling by categories, subcategories & countries.
 – *Tech Stack: OpenAI, S3, ECS, EventBridge, Redshift, MSSQL, PowerBI, CloudWatch Alarms, SNS, PowerBI*

AI-Powered BI Assistant
 – Created NLP-to-SQL chat using RAG and OpenAI models, transforming user chat into actionable information.
 – Boosted productivity by integrating the assistant into chat platforms such as Teams via Power Automate
 – Enhanced user experience with context memory & implemented an RBAC policy to ensure data security.
 – *Tech Stack: React, Python, Langchain, Weaviate, Lambda, PostgreSQL, RAG, OpenAI, FastAPI, Docker, EC2*

Data Engineering In-house Transition
 – Transitioned from vendor to in-house data engineering & MLOps team, saving **USD 877,670 annually**.

- Spearheaded end-to-end data engineering & MLOps: monitoring, API maintenance, ETL pipelines, infrastructure management, software engineering & maintenance, cost optimization.
- Achieved **USD 219,000 yearly savings** through cost optimization of cloud services.
- Enhanced project delivery efficiency by **40%** through successful insourcing from external vendors.

Harmonized System Code Classification
- Built deep learning model to predict HS codes for incoming shipments, enabling customs duty calculation.
- Implemented end-to-end solution from data extraction, cleaning to modeling & deployment.
- API processes over **1 million** requests annually, significantly reducing customs processing time.
- *Tech Stack: Python, keras, convolutional neural networks (CNN), SageMaker, Lambda, API Gateway.*

Transit Time Prediction
- Reduced inbound call volume by **15%** by developing a predictive model for shipment transit times.
- Used a variety of features such as origin, destination, weight, value, historical transit times.
- *Tech Stack: Keras, Neural networks, Xgboost, Lambda, API Gateway, SageMaker, Dynamodb*

Geocoding and Sequencing System
- Developed a CNN based impersonal geocoder and address completion API using hierarchical clustering.
- Developed end-to-end event-driven micro service architecture for address geocoding and sequencing.
- The geocoder forms a core service supporting use cases such as sorting, sequencing and route optimization.
- *Tech Stack: API Gateway, RDS, Redshift, ECS Fargate, EventBridge, Lambda, Docker, SAM, CloudFormation, CNN, Keras, CodePipeline*

Driver Targets Prediction
- Created ML model to set fair delivery targets, considering traffic, inter-stop distance, and cash-on-delivery.
- Engineered accurate stem time calculations with live geolocation data, addressing edge cases and data gaps.
- Deployed model for target-setting across **GCC**, processing **160,000** driver trips monthly.
- *Tech Stack: Linear Regression, API Gateway, ECS, Matillion, Docker, DynamoDB.*

Delivery Risk Prediction
- Developed models to predict and mitigate shipment risks, enhancing delivery success rates.
- Implemented tree-based classification algorithms to identify risky shipments and failure causes.
- Engineered aggregate features for consignees, shippers, and locations using Uber's H3 spatial indexing.
- Utilized Redshift window functions and PySpark for large-scale data profiling.
- *Tech Stack: Cloud Development Kit, Glue, Step Functions, RDS, PySpark, LightGBM, SageMaker*

Shop and Ship Unknown Customer
- Developed an algorithm to identify customer addresses for Aramex's Shop and Ship service, resolving ambiguities when customers omit their unique mailbox number.
- Applied weighted-Levenshtein distance, recency-frequency scores, and custom rules based on UAT feedback.
- API processes approx. **5,049** requests/mo, reducing customer service interactions; improving SNS lifecycle.
- *Tech Stack: AWS EventBridge, Redis, SageMaker, DynamoDB, API Gateway*

Strategic Data Platform Initiative
- Supported requirements gathering and stakeholder interviews for a global strategic data platform.
- Developed project charter and RFP for an integrated BI and AI solution spanning multiple regions.

EDUCATION

Georgia Institute of Technology	Atlanta, USA
MS in Electrical and Computer Engineering (3.84/4.00; Full GRA Scholarship)	*2010 – 2012*
BITS Pilani-Dubai	Dubai, UAE
BE(Hons) in Electrical and Electronics Engineering (9.8/10; Rank 2)	*2005 – 2009*

CERTIFICATIONS

- AWS Certified Developer – Associate
- AWS Certified Machine Learning – Specialty