

# KUNAL SACHDEV

145 Columbia Street West, Waterloo, ON N2J 4C1

☎ 437-663-6032 ✉ Email 🌐 Portfolio 🔗 LinkedIn 🐙 GitHub

## Education

### University of Waterloo

Bachelor of Computer Science, Minors in Statistics and Economics

Sep. 2023 – Present

Waterloo, ON

## Experience

### Edelweiss Life Insurance

GenAI Engineer Intern

May 2025 – August 2025

Mumbai, India

#### GenAI-powered Business Insights and Visualization

- Enabled real-time conversational access to enterprise AWS Data Lake using a serverless **RAG** application integrated with **Claude 3.5 Sonnet** on AWS **Bedrock**, replacing **200+** static BI dashboards with dynamic insights.
- Engineered **2 Lambda** functions for query processing (input validation, **Redshift** schema extraction, OpenSearch retrieval, prompt assembly, LLM inference) with secure SageMaker/Bedrock access via AWS VPC and PrivateLink, ensuring **99%** API uptime.
- Used **Cohere** embeddings to index text-to-SQL example pairs in **OpenSearch**, retrieving top-5 similarity matches to support few-shot prompting, increasing SQL generation accuracy by **28%**.
- Fine-tuned **CodeLlama-7B-Instruct** using QLoRA and engineered production deployment via SageMaker endpoints leveraging DJL + LMI with **TensorRT-LLM**, resulting in **1.8x** faster inference and **50%** reduced memory footprint.
- Integrated **API Gateway** with a **Streamlit** frontend, returning real-time tables and plots to **3000+** daily active users seamlessly.

#### Data Science Project – Policy Mis-selling Prediction

- Designed, tuned hyperparameters, and deployed to production a **top-5 LightGBM ensemble** model to predict policy mis-selling with **85.5%** top-2 decile capture.
- Developed the entire ML pipeline to process **4.9M+** policy records on AWS **SageMaker**, containerized it with AWS **ECR**, and automated weekly email dispatch with AWS **Lambda** + **CloudWatch cron** jobs, achieving **3x faster** prediction cycles.
- Operationalized model output by tagging high-risk customers as “Handle with Care” in internal systems, reducing grievance escalation by routing flagged users to trained service specialists in real time.

### Info Origin Inc.

Data Science Intern

May 2024 – August 2024

Remote

- Trained and deployed a Named Entity Recognition model using **spaCy**, **Doccano**, and **Streamlit** to parse resumes, cutting candidate screening time for HR workflows by **40%**.
- Tuned hyperparameters for a custom **5-layer** neural network using Google’s NNLM embeddings, PyTorch **TensorDataset** and **DataLoader** for efficient batching/shuffling of **2,200+** articles, achieving **96.4%** accuracy in news classification.

### HDFC ERGO General Insurance

Software Developer Intern

April 2022 – June 2022

Mumbai, India

- Automate extraction, cleaning, and storage of insurance quote data from excel to **Oracle** db using a **Python Flask API**, Pandas, NumPy, and **SQLAlchemy**, significantly reducing policy generation time – from 1.5 hours to just 2 minutes (**97.8% faster**).

## Projects

### Winning the Space Race with Data Science 🚀 | BeautifulSoup, Db2, Matplotlib, Plotly, Scikit-learn

October 2022

- Developed a ML pipeline to predict Falcon 9 first-stage landing success, supporting **\$103M** per-launch cost savings.
- Automated collection and EDA of **90+** launch records with **Db2**, SpaceX API and web-scraping with **BeautifulSoup**.
- Designed web dashboards/maps with **Plotly** and **Folium** visualizing launch success patterns across launch factors.
- Tuned hyperparameters for **Decision Tree** with **GridSearchCV**, achieving **87.5%** cross-validation accuracy.

### Car Dealership Full Stack Application 🚗 | Node.js, Express, MongoDB, Flask, Django, React, Kubernetes

March 2025

- Architected a microservices-based car dealership platform with **React** frontend and **Django** + **SQLite** backend.
- Built a containerized backend dealership management service (**Express.js** + **MongoDB** + **Docker**) to serve dealer listings, and customer reviews via **RESTful APIs**.
- Hosted a microservice on **IBM Cloud Code Engine** to analyze customer sentiment and provide business insights.
- Implemented a **CI/CD** pipeline with **GitHub Actions**, and deployed the app stack to **Kubernetes**.

## Technical Skills

**Languages/Cloud:** AWS, IBM Cloud, Python, C, C++, R, SQL, JavaScript, HTML/CSS

**Machine Learning:** Hugging Face, Pandas, NumPy, Scikit-learn, PyTorch, TensorFlow, Transformers, Jupyter, BeautifulSoup

**Web Dev/Databases:** React, Node.js, Flask, Django, Dash, SQLAlchemy, MongoDB, MySQL, SQLite, Db2, Oracle

**Technologies:** Git, Docker, Kubernetes, Linux, bash