

# 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 with a serverless **RAG** app using **Claude 3.5 Sonnet** on **AWS Bedrock**, transforming **200+** static BI dashboards into dynamic insights.
- Engineered **2 Lambda** functions handling query validation, row-level security, **Redshift** schema extraction, OpenSearch retrieval, prompt assembly, SQL generation, and LLM inference with VPC-secured Bedrock access, ensuring **99%** API uptime.
- Improved SQL generation accuracy by **28%** using **Cohere** embeddings with **OpenSearch** retrieval-based few-shot prompting.
- Fine-tuned **CodeLlama-7B-Instruct** using QLoRA and deployed it on SageMaker endpoints optimized with **NVIDIA TensorRT-LLM**, accelerating inference by **1.8x** and cutting memory usage by **50%**.
- Integrated **API Gateway** with a **Streamlit** frontend, delivering real-time tables and visualizations to **3000+** daily active users.

#### Data Science – Policy Grievance & Mis-selling Prediction

- Productionized a top-5 **LightGBM + CATBoost** ensemble to flag grievance/mis-selling risk across **809K+** policies, surfacing **93.8%** of risky cases in the top decile, empowering agents to preempt complaints.
- Developed a scalable ML pipeline on AWS **SageMaker**, containerized with **ECR**, and orchestrated automated weekly predictions reports with **Lambda** and **CloudWatch** cron jobs, reducing prediction latency by **3x**.
- Operationalized model outputs by tagging high-risk users as “Handle with Care” in internal systems and routing them to trained agents in real time, driving a **23.8% drop** in 3-month rolling average of complaints.

### Info Origin Inc. 🌐

Data Science Intern

May 2024 – August 2024

Remote

- Deployed a custom **spaCy NER** model with Streamlit UI to auto-extract key resume entities, cutting HR candidate screening time by **40%** and streamlining early-stage hiring workflows.
- Optimized hyperparameters for a custom **5-layer** neural network using Google’s NNLM embeddings and PyTorch **TensorDataset/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

- Automated ingestion of insurance quote data from Excel to **Oracle DB** using **Flask API**, and **SQLAlchemy**, cutting processing time by **97.8% (1.5 hours → 2 minutes)**, enabling instant policy generation for underwriters.
- Implemented a 4-factor validation pipeline (file format, encrypted token, template schema, metadata check) to detect obsolete or tampered Excel templates, reducing manual entry errors and rejections by **90%**.
- Designed a redundancy management system using an **isactive** flag for lead IDs, eliminating duplicate entries from reuploads.

## Projects

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

March 2025

- Architected a microservices-based car dealership app with a **React** frontend, **Django + SQLite** backend, review REST APIs with **Express.js + MongoDB**, sentiment analyzer on **IBM Cloud**, CI/CD with **GitHub Actions**, & deployment on **Kubernetes**.

### Biquadris 🌐 | *C++, XWindows, Git, Cygwin*

December 2024

- Built a two-player Tetris-style game in **C++** using modular **OOP**, applying **Factory Method** and **Observer** design patterns to support dynamic levels, special actions, and real-time text and graphical rendering via **XWindows**.

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

October 2022

- Developed an ML pipeline predicting Falcon 9 landing success with **87.5%** accuracy, using **GridSearchCV**-tuned **Decision Tree**; automated EDA of **90+** launches via **Db2**, SpaceX API, **BeautifulSoup**, and built **Plotly/Folium** dashboards.

## Technical Skills

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

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

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

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