KUNAL SACHDEV

📞 437-663-6032 ✉ [Email](#) 🌐 [Portfolio](#) 🔗 [LinkedIn](#) 🐙 [GitHub](#) 🏆 [Credly](#)

EDUCATION

University of Waterloo Sep. 2023 – Present
Bachelor of Computer Science, Minors in Statistics and Economics *Waterloo, ON*

EXPERIENCE

Full Stack Software Developer - Cloud and ML Co-op January - April 2026
Nokia Canada *Kanata, ON*

- Incoming - Cloud and Network Services (CNS) team. AI/ML solutions for global networks.

Data Scientist Co-op Sept 2025 – Present
Ontario Lottery and Gaming Corporation *Toronto, ON*

- Deployed a **multimodal RAG** chatbot using **LangChain**, **Azure AI Search**, and **OpenAI** API (**GPT-4o**) to efficiently query insights from business documents, saving business teams **215+** hours per montth.
- Fine-tuned an **XGBoost** classifier using **Databricks**, **Hadoop**, and **PySpark** to segment **115M+** players into 6 categories, enabling targeted ad campaigns. Collaborated with stakeholders to design **A/B tests**, optimizing campaign engagement.
- Worked in an **Agile** environment, following **SDLC** best practices, creating **PRs**, and automating deployments using **CI/CD** pipelines (Jira, **GitHub**, TeamCity, **Docker**, Octopus Deploy).

GenAI Engineer Intern May 2025 – August 2025
Edelweiss Life Insurance *Mumbai, India*

GenAI-powered Business Insights with Visualization

- Built a serverless **RAG** app on **AWS Bedrock** enabling real-time conversational access to AWS Data Lake, replacing **200+** static BI dashboards and enabling **Rs. 300,000+** in estimated annual operational cost savings.
- Fine-tuned **CodeLlama**-7B-Instruct from **HuggingFace** using QLoRA on NVIDIA A10 GPUs; deployed model to AWS **SageMaker** with **TensorRT-LLM**, accelerating inference by **1.8x** and cutting memory usage by **50%**.
- Engineered **Lambda** functions for query validation, **Redshift** schema extraction, OpenSearch retrieval, and LLM inference; achieved **99%** API uptime and reduced **8,700+** annual emails per user via automation.
- Boosted SQL generation accuracy by **28%** using **Cohere** embeddings + **OpenSearch** retrieval for **few-shot prompting**. Integrated **API Gateway** + **Streamlit** UI, serving **1.5K+** daily active users with real-time tables/visuals.

Data Science – Policy Grievance & Mis-selling Prediction

- Productionized a **LightGBM** + **CATBoost** ensemble to flag grievance/mis-selling risk across **809K+** policies, surfacing **96.9%** of grievance cases in the top decile, empowering agents to preempt complaints.
- Developed a scalable ML pipeline on AWS **SageMaker**, containerized with **Docker** on AWS **ECR**, and automated weekly predictions with **EventBridge** Scheduler cron job, reducing prediction latency by **66.7%**.

Data Science Intern May 2024 – August 2024
Info Origin Inc. *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

PROJECTS

WLP4 Compiler | *C++, Bash, Linux, GNU, GCC, GDB*

- Built a compiler for WLP4, a subset of C, implementing Simplified Maximal Munch for **scanning**, LR(1) for **parsing**, context-sensitive analysis, and **MIPS** assembly generation, with **30%** improvement in performance using optimization techniques.

Car Dealership App | *[Architecture](#), Node.js, Express, MongoDB, Flask, Django, React, Kubernetes* | 🐙

- Architected a full-stack microservices app with a **React** frontend, and **3** services: **Django** + **SQLite** backend, reviews REST APIs with **Express.js** + **MongoDB**, and a sentiment analyzer on IBM Cloud; CI/CD with GitHub, & deployment on **Kubernetes**.

Biquadris | *[UML Diagram](#), C++, XWindows, Git, Cygwin, Bash, Linux, GNU, GCC, GDB* | 🐙

- Implemented a two-player Tetris-style game in **C++** using modular **OOP**, applying **Factory** and **Observer** design patterns to support **6+** blocks, **3+** levels, **3** special actions, and real-time text/graphical rendering via XWindows.

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

- 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 designed **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, Express, Node.js, Flask, Django, Dash, SQLAlchemy, MongoDB, MySQL, SQLite, Db2, Oracle
Technologies: Git, Docker, Kubernetes, Linux, bash, GNU, GCC, GDB