

NISHANT KUMAR

+91-9917760469 ♦ Haridwar, UTTARAKHAND

nishantsingh.talk@gmail.com ♦ www.linkedin.com/in/nishantksingh1 ♦ nishantksingh0.github.io

OBJECTIVE

Aspiring Data Scientist with hands-on experience in supervised and unsupervised learning, EDA, model evaluation, and deep learning techniques. Familiar with NLP, LLMs, vector databases, and deploying models using modern ML pipelines.

EXPERIENCE

Data Science Intern

Haridwar University

June 2025 – July 2025

Roorkee, Uttarakhand

- Applied ML Algorithms (Linear Regression, K-Means, ANN, CNN) to real-world datasets, using Numpy, Pandas, and Matplotlib for preprocessing and analysis.
- Utilized Scikit-learn and Tensorflow for end-to-end project workflows, enhancing model robustness, and evaluated models using metrics like F1 Scores, MAE, MSE and accuracy.

Data Science Intern

Onlei Technologies

Jan 2025 – Feb 2025

Remote

- Gained hands-on experience in building and evaluating regression-based machine learning models using real-world datasets.
- Applied data preprocessing techniques including data cleaning, merging, and handling missing/null values to improve model accuracy.
- Strengthened foundational knowledge of data analysis pipelines and practical model deployment workflows.

PROJECTS

Transformer Based Translation Model: Built core components of the Transformer using TensorFlow, including multi-head attention and positional encoding. Used the WMT dataset with SentencePiece, and tackled training issues like vanishing gradients and slow convergence due to limited compute.

Multi-Agent Debate Using LangGraph: Developed a multi-agent debate system using HuggingFace API, LangGraph, and Meta-Llama-3-8B-Instruct, enabling over 5 rounds of LLM-driven debates with 90% logical consistency as judged by an evaluation model.

RAG Generative Model: Built a Retrieval-Augmented Generation (RAG) system using GPT-2 for answer generation and Pinecone for vector-based document retrieval. Embedded textual data with SentenceTransformers to enable semantic search and integrated the pipeline to produce context-aware, relevant responses to user queries.

[\(22 more Open Source work\)](#)

SKILLS

Languages: Advanced-Python, C/C++, Java, JavaScript.

Libraries: NumPy, Pandas, Scikit-learn, OpenCV, Matplotlib, Keras, Seaborn, NLP, LoRA.

Frameworks: TensorFlow, LangChain, LangGraph, Flask, React.

Tools: Jupyter, Kaggle, Colab, VS-Code, Github.

Deployment: Git, Firebase, Fly.io, Vercel, Render, Gradio, StreamLit.

EDUCATION

Bachelor's in Computer Application, Haridwar University (Roorkee) Expected 2026
Relevant Coursework: Data Structures, DBMS, Operating Systems, Computer Organization Architectures, Data analysis algorithms, JAVA, C/C++, Python, Machine Learning

Primary/Secondary, Vidya Mandir Sec-5, BHEL (Haridwar), CBSE 2021 – 2023
Focused coursework in Physics, Chemistry, and Mathematics with a strong emphasis on problem-solving and mathematical reasoning.

CERTIFICATIONS

Reinforcement Learning (NPTEL) - Learned fundamentals of RL, bandit algorithms, TD and Q-learning, function approximation techniques, and advanced topics like DQN and Actor-Critic methods.

Caffeine fueled 24hr Hackathon (Haridwar University) - Handled team of 4 members and secured 3rd rank on Exam proctoring system project.

Building RAG Agents with LLM's (NVIDIA) - A hands-on, end-to-end implementation of a Retrieval-Augmented Generation (RAG) pipeline.

Data Science (British Airways) - Worked on airways data to perform data cleaning, exploratory analysis, and developed a highly-accurate Random Forest model for predictive insights.

GenAI Simulation (BCG X) - Analyzed 10-K filings from top companies like Microsoft, Apple, and Tesla, and built a rule-based chatbot to answer user queries based on financial report content.

GenAI Data Analytics (Tata) - Applied EDA and data structuring techniques for business intelligence use cases. ([30+ more certifications](#))

EXTRA-CURRICULAR ACTIVITIES

- Hands-on experience with building custom Transformer architectures and experimenting with LLMs.
- Preparing for advanced certifications like DP-100 to be a certified Data Scientist.
- Building a strong foundation in Azure ML deployment infrastructure.

LEADERSHIP

- Led team of 4 in University 24hr Hackathon and secured 3rd place with AI-based proctoring system project.
- Led a team for 3 full-stack live projects ([Resume Builder](#), [Proctoring System](#), [Betting App](#)).

Verification Links: bit.ly/4kySEQv