**NISHANT KUMAR**

[Portfolio](https://nishantksingh0.github.io) | [Mail](mailto:nishantsingh.talk@gmail.com) | [Github](https://github.com/NishantkSingh0) | [Linkedin](https://www.linkedin.com/in/nishantksingh1) | [Kaggle](https://www.kaggle.com/nishantsingh96)

**Summary**

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.



**Internships**

* Haridwar University – Data Science Intern June 2025 - July 2025
* 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.
* Onlei Technologies – Data Science Intern Jan 2025 - Feb 2025
* 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.



**Key Projects**

* Transformer Based Translation Model (From Scratch)
* TechStack: Tensorflow, Transformer Architecture, NumPy, WMT dataset, Sentencepiece.
* Implemented core components of the Transformer architecture including multi-head attention, positional encodings and deep stacked layers.
* Faced and analyzed training challenges due to limited compute and data, deepening understanding of vanishing gradients and model convergence issues.
* Multi-Agent Debate Using LangGraph
* TechStack: HuggingFaceAPI, LangGraph, LangChain-Core, Meta-Llama-3-8B-Instruct.
* Created a multi-agent debate system using LLMs that generated 5+ debate rounds with 90% success rate in logical consistency as evaluated by the judge model.

[**16+ More projects with proper readme and workflow**](https://github.com/NishantkSingh0?tab=repositories)



**Tools & Technologies**

* **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, Azure, Vercel, Render, Gradio, StreamLit.



**Education**

* Bachelor’s in Computer Application 2023 – 2026

Haridwar University

GPA 7.9

* Primary/Secondary (PCM) 2021 – 2023

Vidya mandir sec-5 BHEL (Haridwar) CBSE

GPA 7.3



**Certification**

* Reinforcement Learning – NPTEL
* Learned key reinforcement techniques like UCB, PAC, Q-Learning, LSTD/LSTDQ etc..
* Caffeine flued 24hr Hackathon – Haridwar University
* Handled team of 4 members and secured 3rd rank on Exam proctoring system project.
* GenAI Data Analytics – Tata
* Applied EDA and data structuring techniques for business intelligence use cases.

[**18+ more certifications**](https://drive.google.com/drive/u/2/folders/1FEHD7C-eIT7S2FtjRGr9v2KXnOVP4trG)



**Key Achievements**

* Led team of 4 in Hackathon and secured 3rd place with AI-based proctoring system project.
* Hands-on experience 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.
* Led a team for 3 full-stack live projects (Resume builder, Proctoring System, Betting App).



**Verification links:** <https://bit.ly/4kySEQv>