Harshveer Singh
https://llstringll.github.io/
harshveer321.code@gmail.com • Git: @llStringll

SKILL MATRIX:

Skill	Experience	Companies
C/C++, Python, Shell scripting, Pytorch	5 years	College/IIT Mandi/ Vaultedge/ChartChat/CERN
Statistics, Probability, Linear Algebra, Bayesian Inference	5 years	Research work/College/Vaultedge/Merce-Mettl/Repodox/IIT Mandi/CERN
NLP, NLI, Huggingface,	5 years	College/Research work/Vaultedge/Merce-Mettl/Repodox/ChartChat
VertexAI, Docker, SageMaker	2 years	Vaultedge/Mercer-Mettl

SUMMARY:

- **Data Scientist & NLP Engineer** with 5+ years of experience in machine learning, AI, and NLP, specializing in Python, C/C++, and PyTorch.
- Proven track record in **enhancing model performance** by 10-30% and developing **CI/CD pipelines** for efficient model deployment.
- Expertise in **Natural Language Processing (NLP)** techniques, **Bayesian inference**, and **cloud platforms** like VertexAI and AWS SageMaker.
- Research experience at **CERN** and **IIT Mandi**, focusing on innovative AI solutions and deep learning models.
- **Hackathon achievements** include top rankings in national competitions (1st in SATURNALIA, 2nd in PEC-FEST, and 3rd in IITB-TECHFEST).
- Delivered technical presentations on advanced AI topics like **variational inference** and **inductive bias in machine learning**.

TECHNICAL SKILLS:

Programming
- C, C++, Python, Shell scripting
- PyTorch, TensorFlow, Keras
- ScikitLearn, HuggingFace, NLTK
- Pandas, NumPy, Matplotlib

Data Analytics
- Statistics, Probability
- Data visualization with Matplotlib
- Data manipulation with Pandas
- Bayesian Inference

Machine Learning
- Theoretical machine learning and deep learning
- Optimisation algorithms
- Loss Landscape Analysis
- PyTorch, TensorFlow, Keras
- ScikitLearn
- Bayesian Inference

Natural Language Processing (NLP)
- NLP techniques and models
- Natural Language Inference (NLI)
- Conversational Agents
- HuggingFace, NLTK

Cloud Management
- VertexAI
- AWS Sagemaker
- Docker

TECHNICAL PROJECTS:

ChartChatAI visit here

- An in-house multi modal large model (finetuned Llama) with basic RAG layer to provide candlestick chart analysis in various formats. Have constent growth in user-base with healthy and suggestive feedback as well.
- Production stack nodejs, CSS/JS, weaviate vector db, pytorch, llammacpp
- Also launched on Product Hunt here

Repodox visit here

- An attempt at very large scale and efficient RAG pipeline over public, personalized GitHub repos, to create a system to easily fetch and discuss and generate specifics from a code repo
- Production stack nodejs, CSS/JS, weaviate vector db, pytorch, llammacpp
- Work in progress.

PROFESSIONAL EXPERIENCE:

Applied AI, Vaultedge Pvt. Ltd. **Oct'22- Present**
Data Scientist

- Enhanced production model metrics by ~11%, and without regressions.
- Implemented model interpretability techniques
- Developed and integrated CI/CD pipelines for seamless model deployment, reducing the need for manual human intervention and accelerating the model update cycle.
- End-2-end implemented robust data preprocessing pipelines to clean and improve noisy real-world data for language models in production.
- Trained, optimized, and deployed multi-lingual language models in production

Mercer-Mettl **Jun'21 - Sep'22**
NLP Engineer, AI Team

- Improved inter-sentence and intra-sentence cohesion measuring pipeline precision by 25%, involved feature understanding, reducing the output feature space of BeRT by putting a posterior on latent space
- Implemented production ready email formality checking pipeline for business environments, involved
- topic modeling of a raw real-life email set that had noisy lexical structure, through improvised LDA
- Analyzed on-prod spell-check pipeline and suggested and upgrade with specific fine-tuning, that increased recall and precision of the model by ~13% and ~32%.

Mercer-Mettl **Jan'21 - Jun'21**
NLP Engineer Intern, AI Team

- Worked on improving NLI models, improved cohesion detection accuracy on English text by ~16%

CMS Experiment, CERN **Jan'21 -May'21**
Deep Learning Research Intern

- Worked on building a quasi-linear attention model to isolate 'interesting' events from the background
- during the collision of protons with Low-Z targets.

IIT-Mandi **Dec'19 - Jan'20**
Research Intern, Department of Mathematics

- Provided an analytical study on the success of Batch Normalization [Blog]

RESEARCH EXPERIENCE:

Cross-layer residual connection transformer **Oct'20 - Nov'20**

- Developed a novel architecture, which has a recursively "smooth" loss surface, allowing the possibility

- of reaching more generalized minima, even in the absence of good parameter initialization.

Adversarial Training for Facebook's Blender **June'20 - Aug'20**

- Created a self-play regime for conversational agents, and extending that to a competitive conversation where an agent discriminates the output distribution of the other agent against human dialogue distribution.

Poly encoder regime for fine-tuning decoder-only model (GPT-2) **May' 20**

- Showed that a decoder model fine-tuned like such on language modeling apparently is more robust to inductive bias than encoder model even though encoder reached better recall@k/C score

Analytical study of the success of Batch Norm **Nov'19 - Dec'19**

- Showed that batch normalization smooths the loss surface and how it brings that effect, through the study of eigenvalues of the hessian of weight matrix. [Blog]

Beta2 variation regime for Adam Optimizer **May'18 - Jul'18**

- Developed a regime for varying beta2 hyper-parameter of Adam, preventing Adam from getting stuck
- in sub-optimal minima.
- A similar result was also shown in a subsection of Sashank J. Reddi et al.

HACKATHONS AND TECHNICAL PRESENTATIONS:

- **Hackathons, 2017 2018**
  - SATURNALIA Hackathon '17 ranked 1st PEC-FEST Hackathon '17 ranked 2nd
  - IITB-TECHFEST Hackathon'18 ranked3rd
- Causality and its importance in variational inference and EM, TIET Jan '20
- Inductive bias in machine learning models, TIET Oct '19
- Effect of constraining the posterior to Gaussian in VAEs, IITB-TECHFEST Nov '17