HARSHVEER SINGH

8427337873 • Ludhiana, Punjab
harshveer321.code@gmail.com • Git: @llStringll • LinkedIn: @llharshveer-singhll

EDUCATION

Thapar Institute of Engineering and Technology (TIET) July'17 - Jul'21
Bachelor of Engineering(B.E.) in Electronics and Computer Engineering

EXPERIENCE AND TECHNICAL PROJECTS

NLP Engineer, AI Team, Mercer-Mettl Jun'21 - Present
Conducting research, developing POC baselines and deploying production ready pipelines for multiple projects.

NLP Engineer Intern, AI Team, Mercer-Mettl Jan'21 - Jun'21
Worked on improving NLI models, improved cohesion detection accuracy on English text by ~26%

Deep Learning Research Intern, CMS Experiment, CERN Jan'21 - May'21
Worked on building a transformer based model to isolate 'interesting' events from the background during the collision of subatomic particles.

Research Intern, Deptt. of Mathematics, IIT-Mandi Dec'19 - Jan'20
Provided an analytical study on the success of Batch Normalisation

Hackathons 2017 - 2018
SATURNALIA Hackathon '17 ranked 1st
PEC-FEST Hackathon '17 ranked 2nd
PEC-FEST Hackathon '18 ranked 1st

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 initialisation.

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.

SKILLS

Data structures and algorithms, C, C++, Python, PyTorch, NLTK, ScikitLearn, HuggingFace, NumPy, TensorFlow, Keras, Pandas, Matplotlib, Shell scripting.

TALKS AND PRESENTATIONS

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, PEC-FEST	Nov ‘17