Abhijit **D**eshmane

Email: abhijitdeshmane@gmail.com
Mob: +91 9158587706

Experience Summary:

AI/ML engineer with strong mathematical foundations and having hands on deep learning experience with LLM

ML Core Skill:

- Embedding
- Neural Network
- Backpropagation
- Optimization
- Fine tuning
- Data set

Tools & MLOps:

Python
PyTorch
Docker
Kubernetes
Git
Hugging face

Project:

Mini Transformer and GPT block (from scratch):

1. **Implemented a Transformer Encoder from Scratch (PyTorch / NumPy)**
   - Implemented Multi-Head Self-Attention manually: query/key/value projections, attention scores, softmax normalization, head concatenation, and output projection.
   - Built Positional Encoding (sinusoidal / learned) and integrated it with token embeddings.
   - Implemented **Layer Norm**, **residual connections**, and **feed-forward blocks (MLP)** identical to original Transformer architecture.
   - Trained the model on a custom dataset for tasks like classification, next-token prediction, or sequence reconstruction.
2. **Built a Trainable Mini-GPT Decoder Block**
   - Implemented causal (masked) self-attention used in GPT models to prevent information leakage from future tokens.
   - Implemented **GPT-style residual pathways**, pre-norm architecture, and GELU activation.
   - Created a scalable modular architecture so blocks can be stacked to form small GPT-like networks.
3. **End-to-End Training + Evaluation**
   - Trained the model on next-token-prediction using a small tiny Shakespeare dataset
   - Implemented cross-entropy loss, teacher forcing, gradient clipping, and custom training loop.
   - Achieved stable training and generated coherent text sequences.
4. **Engineering & Optimization**
   - Implemented batched attention for efficient GPU use.
   - Added weight tying between input and output embeddings.
   - Added support for attention masking, padding mask, and positional indices.
   - Visualized attention maps to interpret model behaviour.

Extractive Question Answering Model — SQuAD (HuggingFace Transformers)
Python, Transformers, PyTorch, Tokenizers, Datasets:

- Built an end-to-end QA system fine-tuning DistilBERT on the SQuAD dataset (2k training samples)
- Implemented custom preprocessing: sliding-window tokenization, overflow mapping, and character-to-token offset alignment for accurate span labeling.
- Implemented full post-processing pipeline with n-best decoding to generate final answer spans.
- Evaluated using official QA metrics (Exact Match & F1), achieving EM 57.8% / F1 64.9% on validation.
- Created reusable inference pipeline for answering arbitrary user questions.
- Experimented with multiple models (BERT, RoBERTa) and hyperparameters to analyze performance trade-offs.

Career Profile:

Working as **Associate Director** at UBS, Pune since 28 Aug 2023 to Till Date

Employment History:

Associate Director at UBS, Pune	Aug 2023 – Present
SDET Lead at Encora, Pune	Dec 2022 – Aug 2023
Sr. QA Automation Engineer at Happeo, Helsinki (Finland)	Mar 2022 – Aug 2022
Sr.Associate at Atyeti Inc. Pune	Jan 2021 – Feb 2022
Principal Engineer at Innominds, Pune	Oct 2018 – Jan 2021
Senior Software Engineer at HSBC, Pune	Aug 2016 – Jul 2018
Consultant at Capgemini, Pune	Jun 2013 - Aug 2016
Test Engineer at Syntel, Pune	Aug 2010 - Jun 2013

Education:

BE in Electronics from Walchand Institute of Technology in year 2009 with 67.19%

Personal Information:

Gender	: Male
Date of Birth	: 25 Nov 1987
Nationality	: Indian
Passport No.	: V6354961
Mailing Address	: Solitaire 906, Pune-Saswad road, Next to HP petrol pump, Landmark- HDFC Bank, Phursungi Phursungi, Pune 412308
E-mail ID	: abhijitdeshmane@gmail.com