Gurpreet Singh

B. Tech, Department of Computer Science and Engineering

(+91) 90055-28086 | guggu@iitk.ac.in | fat-fighter.github.io | github.com/fat-fighter

Education

Year	Degree	Institute	CPI / Score
2015 - 2019	B. Tech, Computer Science and Engineering	Indian Institute of Technology, Kanpur	9.3 / 10.0
2015	Class XII (CBSE)	Delhi Public School, Kalyanpur	94.4%
2013	Class X (CBSE)	Delhi Public School, Kalyanpur	9.6 / 10.0

Scholastic Achievements

Got Academic Excellence Award, IITK for the year 2016-2017 — 2018

Got selected for ACM ICPC Regionals, 2017 at two sites, Chennai and Amritapuri (Overall Rank - 70) — 2017

Secured All India Rank 174 in the JEE Advanced examination amongst 1.2 lakhs shortlisted students — 2015

Kishore Vaigyanik Protsahan Yojana (KVPY) Scholar with All India Rank 464 — 2015

Projects

DISCRETE VARIATIONAL AUTOENCODERS AND STOCHASTIC BLOCK MODELS

Fall 2018 (Ongoing)

Prof. Piyush Rai — *Undergraduate Project*

- Surveyed continuous relaxations to discrete latent variables such as Gumbel-Softmax, Spike-and-Exp, Overlapping, Gumbolt, etc.
- Implemented GumBolt relaxation for binary latent variables with RBM prior using tensorflow and performed analysis on MNIST dataset
- Augmented GVAEs with binary latent embeddings to offer interpretable latent representations, imitating mixed membership models
- Tested the resultant model for link prediction on graph datasets such as Citeseer and Cora and achieved superior results to baseline models

MIXTURE OF EXPERTS USING DISCRETE VAE

Fall 2018

Prof. Arnab Bhattacharya — *CS685: Data Mining*

- Proposed a novel model using the VAE framework for clustering in latent space, extending the ideas of the VaDE model
- Modeled the cluster assignment using a deep neural network, and facilitated fast inference by using Gumbel-Softmax relaxation
- The proposed model worked better than VaDE on clustering tasks. Analysis was performed on MNIST and Spiral datasets
- Further extended the proposed model to work as a gating function for Mixture of Experts tasks and achieved performance exceeding that of some naive baseline models

INCREMENTAL NEURAL NETWORKS TRAINING

Spring 2018

Prof. Purushottam Kar — *CS777: Statistical and Algorithmic Learning Theory*

- Two layer NNs can be represented as an ensemble of multiple single node hidden layer networks, which can be individually trained using generic boosting methods (gradient boosting), which also afford definite theoretical convergence guarantees
- Applied gradient boosting to train two layer networks incrementally and studied the convergence analysis under various constraints
- Implemented incremental NN training in python using sklearn, and applied for Softmax Regression on the MNIST Dataset
- Applied incremental training as pre-training, along with backpropagation for fine-tuning and observed remarkably better convergence

SURVEY ON METHODS FOR CONVEX OPTIMIZATION

Spring 2018

Prof. Purushottam Kar — *CS777: Statistical and Algorithmic Learning Theory*

- Surveyed prominent Gradient Descent based techniques (SGD, AdaGrad, etc.) for optimization and perused the convergence bounds of each
- Reviewed and paraphrased a paper which disproves guaranteed convergence of Adam for even convex objectives using a counterexample
- Identified inconsistencies within the convergence proof for Adam as an attempt to explain its incorrectness

CLUSTERING AND MOE FOR ARBITRARY SHAPED CLUSTERS

Spring 2018

Prof. Piyush Rai — *CS698X: Bayesian Modelling and Inference*

- Studied VAEs and surveyed clustering models (iWMM, SVAE, VaDE, etc.) for data existing in non-Gaussian shaped clusters
- Implemented Variational Deep Embeddings (VaDE) in Tensorflow to experiment on MNIST and spiral dataset to learn arbitrary shaped clusters
- Proposed gating functions based on VaDE and Stick Breaking-VAE for mixture of experts models

JAVA TO X86 ASSEMBLY COMPILER

Spring 2018

Prof. Subhajit Roy — *CS335: Compiler Design*

- Developed an end-to-end compiler in node.js for a subset of Java language to compile into x86 Assembly using jison (for parsing)
- Implemented advanced features such as classes and type casting, along with support for floats. Adjudged one of the best projects

MACHINE COMPREHENSION USING MATCH-LSTM

Spring 2018

Prof. Harish Karnick — *CS671: Natural Language Processing*

- Surveyed various models for Machine Comprehension (FastQA, R-Net, Match-LSTM, etc.) and implemented Match-LSTM using Tensorflow
- Experimented with SQuAD and combated inefficiency of Match-LSTM to apply separate attention mechanisms for different question types
- Additionally, introduced simple changes to loss function to improve the EM score on SQuAD by a total of over 5%

SCALING RECOMMENDATION SYSTEMS USING K-MEANS CLUSTERING *Fall 2017*

Prof. Purushottam Kar *CS771: Introduction to Machine Learning*

- Used K-Means clustering to divide users into cliques, and applied Collaborative Filtering independently within each clique
- Clustered songs based on MFCC features using K-Means and quantified user features based on song clusters from the user's learning history
- Applied the model on MSD. Also proposed simple exploration strategy based on song clusters to allow variations in suggestions provided

NACHOS OPERATING SYSTEM *Fall 2017*

Prof. Mainak Chaudhuri *CS330: Operating Systems*

- Implemented basic operating system functions (Fork, Join, etc.) on a truncated NachOS code (provided) in C++ programming language
- Implemented and evaluated performance of algorithms for various scheduling processes and various page replacement strategies
- Implemented Shared Memory Allocation, Demand Paging and various Page Replacement Algorithms

** Code and reports for all projects are available at https://github.com/fat-fighter*

Work Experience

GOLDMAN SACHS | Summer Analyst *Bangalore, May'18 - Jul'18*

Objectives

(i) Introduce changes in existing models for asset liability gap management for deposits and clearing house initial margin, and
(ii) Build a lite calculator for customer margin allocation for proper internalization

- Understood working of financial firms, along with basic concepts of Asset-Liability Gap (AL Gap) Management
- Identified bugs in pre-written code, and augmented features for proper AL Gap Management
- Built a waterfall logic for customer margin allocation, taking various parameters into consideration
- Built a greedy strategy for margin allocation per stock, improving the run time of the allocation logic

INMOBI | Data Science Intern *Bangalore, May'17 - Jul'17*

- Extracted Features from Ad creative images using OpenCV (in python) and Google Cloud Vision API
- Analysed Pearson Correlation with the Click Through Rate (CTR) and used variable selection (Weka) to detect explainable features
- Created a python server to handle feature extraction and prediction for building suggestions for possible Ad enhancements based on CTR

EXXAMM.COM | Web Development Intern *Delhi + Remote, Jan'16 - Jul'16*

- Lead architect of the core content engine and front end web interface
- Designed and developed a dashboard to add and edit questions using PHP and MySQL

INMOBI | Software Engineering Intern *Bangalore, Dec'15*

- Worked in Strategic Advertising for better in-app product discovery and user experience
- Developed a Curator Tool which scraped and rendered information from the web using python

Technical Skills

Programming/Scripting	C/C++, Python, Bash, Octave/MATLAB, R, LaTeX
Assembly Languages	MIPS, Verilog
Web Development	PHP, Javascript, JQuery, MySQL, CSS/HTML, node.js
Utilities and Tools	Git, Linux Shell Utilities, Tensorflow

Coursework

Computer Science and Engineering

Data Mining (I)	Networks (I)	Bayesian Modelling and Inference (A*)
Statistical Learning Theory (A)	Natural Language Processing (A)	Introduction to Machine Learning (A)
Compiler Design (A)	Operating System (A)	Algorithms - II (A)
Computing Laboratory - II (A)	Computing Laboratory - I (A)	Data Structures and Algorithms (A)
Computer Organization (A)	Fundamentals of Computing (A*)	

Miscellaneous

Probability and Statistics (A)	Linear Algebra	Numerical Methods in Engineering

** (I - Ongoing Course) (A* - Excellent Performance) (A - Very Good Performance)*

Positions of Responsibility

ACADEMIC MENTOR | ESC101: Fundamentals of Computing *Counselling Service, IITK, '16 - '17*

- Took multiple doubt clearing sessions, both hall level and institute level
- Took multiple institute level classes and also gave one-on-one tutoring to a few students