Md Enayat Ullah, Senior Undergraduate
Department of Mathematics and Statistics, IIT Kanpur
enayat@iitk.ac.in | (+91)8953440996 | home.iitk.ac.in/∼enayat

---

Education

Indian Institute of Technology, Kanpur *(2012 - present)*
Bachelor of Science, Mathematics and Scientific Computing
Cumulative Performance Index: *8.13*

All India Senior School Certificate Examination *(2011)*
SS Vidya Mandir, Madhupur; Aggregate : *90.6%*

Indian Certificate of Secondary Education Examination *(2009)*
Carmel School, Madhupur; Aggregate : *95.2%*

Research Interests

- Machine Learning, Natural Language Processing
- Statistical Learning Theory, Optimization

Projects and Internships

Cross-lingual Plagiarism Detection [poster] [report] *(Ongoing)*
Prof. Amitabh Muherjee, Indian Institute of Technology, Kanpur

- Performed joint learning of word vectors in unified multilingual distributional space from document aligned comparable corpora.
- To estimate the robustness of the multilingual word space, two tasks: Bilingual Lexicon Extraction and Suggested Word Translation in Context are performed.
- Trained a Deep Recursive Autoencoder with dynamic pooling to generate phrase representations which are fed to an SVM for paraphrase detection.

Domain Invariant Transfer Kernel Learning [slides] [report] *(Fall'15)*
Prof. Harish Karnick, Indian Institute of Technology, Kanpur

- Based on the work of **Long et al** [link], proposed to implement a learning model which generalizes across training and testing data with different distributions.
- Designed a family of spectral kernels by extrapolating target eigensystem on source samples to reduce the Nystrom Approximation error in the RKH Space.
- Plugged the obtained domain-invariant Kernel matrix to an SVM which outperformed the traditional SVM on benchmark text and image datasets.

N-body Simulation using Sampling [poster] *(Summer'15)*
Prof. Geoffrey Fox, Indiana University, Bloomington, USA

- Worked on optimising the n-body formulation of Deterministic Annealing, used to perform Multi-dimensional scaling of a high dimensional data.
- Studied methods of solving n-body problems such as Hierarchical partitioning data structures, treecodes, fast multi-pole methods and Barnes-hut simulation.
- Designed a topological spherical embedding to distribute point into near and far groups, wherein the contribution of the near points are to be approximated.
- Used Importance sampling to heuristically to sample from a distribution different from the inherent distribution.

Aspect based Sentiment Analysis [poster] [report] *(Spring'15)*
Prof. Amitabh Mukerjee, Indian Institute of Technology, Kanpur

- **Sem-Eval 2015 task** [link]: Identification of every entity E and attribute A pair E#A in a text towards which an opinion is expressed, and adjudging it's polarity.
- Constructed features based on word-vectors, n-grams, parse trees,POS tag and out-of-domain, publicly available sentiment lexica (wordnet, sentiwordnet).
- Trained a Conditional Random Field(CRF) for sequential learning of aspect term, and a Maximum Entropy Classifier to adjudge the polarity.

Forest Cover-type Classification Problem [report] *(Spring'14)*
Prof. Amit Mitra, Indian Institute of Technology, Kanpur

- Attempted the **Kaggle challenge** [link] of classifying forest cover type by building classification models based on the dataset.
- Employed various classification techniques such as Neural Networks, SVM, Logistic Regression, Naive-Bayes classifier, CART and Random Forests.
- Classification by Random Forests on a 10-fold cross validation training set was the most accurate with an accuracy of 87%.

Random Graph models of Social Networks [slides] *(Spring'15)*
Prof. A K Lal, Indian Institute of Technology, Kanpur

- Studied Random Graphs and their properties, degree distributions, scale-free graphs and small-world networks.
- Studied about phase transitions and random graph models such as Erdos-Renyi model, Configuration model and preferential attachment model.

Phonotactic Constraints in McGurk Fusion [poster] [report] *(Fall'14)*
Prof. Amitabh Mukerjee, Indian Institute of Technology, Kanpur

- On cognitive audio-visual speech perception establishing the role of Phonotactic constraints towards producing a bias in McGurk Effect.
- Conducted experiments to conclude that Phonotactic constraints can diminish McGurk fusion rate when phonetic licensing biases against the fusion expected.

Causal Relationships Between Econometric Parameters [report] *(Fall'14)*
Prof. Amit Mitra, Indian Institute of Technology, Kanpur

- Used Time Series Econometric modelling to analyze the data of the policy macroeconomic variables using Augmented Dicky-Fuller and Granger Causality test.
- Established that both FDI inflows and Exports have a direct causal linkage with the GDP of India but there is no reciprocal causality between them.

Summer Intern at Aurus Network Infotech Pvt. Ltd. *(Summer'14)*

- Worked with a team of developers towards building a novel e-commerce educational platform, based on the PHP framework Yii.
- Developed the Relevance Algorithm module to sort courses based on an aggregated scoring system parametrized on sale, recency and rating-reviews.
- Implemented SMS-based phone-number verification using SMS service APIs.

Philosophical Problems from the Standpoint of AI [term paper] *(Spring'15)*
Prof. A.V. Ravishankar Sharma, Indian Institute of Technology, Kanpur

- Compiled a brief summary of the paper "Some philosophical problems from the standpoint of artificial intelligence" by John McCarthy and Patrick J. Hayes
- Attempted to put forth a concise version of this 50 page paper while keeping the main ideas intact.

Scholastic Achievements

- Ranked in **Top 0.5%** (amongst 0.5 million students) in IIT-JEE 2012.
- Ranked in **Top 0.3%** (amongst 1.1 million students) in AIEEE 2012.
- Recipient of **Inspire** Scholarship awarded by Department of Science and Technology, Government of India.
- Secured **99.2 percentile** in National Cyber Olympiad 2009.
- Awarded the 3rd best position for manufacturing a working model of an engineering design(rope-making machine) amongst over 400 students in the course.

Relevant Coursework

Machine Leaning:

- Artificial Intelligence Programming
- Learning with Kernels
- Probabilistic Machine Learning
- Online Learning and Optimization
- Natural Language Processing
- Stats. and AI Tech. in Data Mining
- Time Series Analysis

Mathematics:

- Linear & Abstract Algebra
- Probability and Statistics
- Topics in Topology
- Real & Complex Analysis
- Graph Theory
- Partial Differential Equations

Other Relevant Courses:

- Introduction to Programming
- Introduction to Electronics
- Introduction to Cognitive Science
- Data Structures and Algorithms
- Theory of Computation

Online Courses:

- Machine Learning (Andrew Ng)
- NLP (Dan Jurafsky)
- Deep Learning (Nando de Freitas)
- Cryptography (Dan Boneh)
- Data Scientists Toolbox (Jeff Leek)

Technical Skills

Programming: C, C++, Java, Python, R
Web Development: HTML, PHP, JavaScript, Yii, Node.js
Other Tools: Bash, Matlab, Git, LaTeX, Android SDK, Visual Studio, Adobe Photoshop
Operating Systems: Windows, Linux(Ubuntu)

Extra-Curricular Activities

- An active member of Quiz Club, participated in various Intra and Inter-college Quizzes.
- An Active Member of Special Interest Group on Machine Learning(SIGML).
- Worked with Pulkit Aggarwal in Winter Hackathon'14 to develop Infexious, spatially local Social Networks, working on Android devices employing Bluetooh-LE.
- Among the top 10 best coded applications in the country for Hitch-a-ride, a taxi-pooling Windows phone app in Microsoft Code.fun.do
- Co-ordinator, Crypto (Techkriti): Formulated questions for the online cryptographic treasure hunt, which witnessed participation from more than 1000 people.
- Secured 1st position in App development competition(Takneek) with SOS, an Android application which facilitates Disaster Relief efforts to aid Uttrakhand flood victims.

References

Available on request.