

Pratheeksha Nair

I am in the final year of my 5-year integrated Masters (Bachelors plus Masters of Engineering) in Computer Science at International Institute of Information Technology-Bangalore (IIIT-B). I am expected to graduate in July, 2019. My areas of research interests are **Machine Learning** and **Artificial Intelligence**. I have been working on projects related to these fields over the past year. I am particularly interested in the applications of AI in the medical field. Significant part of my work was focused on Natural Language Processing and to a small extent Computer Vision.

I am planning to do a one-year thesis as part of my Master's degree. The topic of the thesis is being finalized, however I will likely be working with generative models like GANs in the medical field. Currently I am in the process of understanding GANs and exploring varied application areas of this advanced technology. I am very excited to work in areas that potentially integrate AI and Computational Biology, particularly Computational Neuroscience.

Academics

At the end of 8 semesters in IIIT-B, I have obtained a Cumulative GPA of 3.56/4.0 and am among the top 5/62 students of my class. Some of the courses that I thoroughly enjoyed taking include Machine Learning, Machine Perception, Advanced Machine Perception, Math for ML, Digital Signal Processing, Python programming and Technology and Society.

During my courses, I picked up various skills including proficiency in **Python** scripting, hands-on experience with deep learning libraries like **Pytorch**. I am quite comfortable with using tools like **LaTeX** and **Matlab**. I have used **MySQL** databases extensively.

Research experience

- **Method Summarization from Code - Internship at IBM Research AI Lab, India**

(Guide: **Rahul A R**, May'18 - Aug'18)

The aim of the project was to generate short 2-3 word descriptions of small code snippets (like method bodies). These short summaries will denote what the method accomplishes. We modelled this as a simple summarization + translation task of NLP because we wish to translate a long piece of text in one language (any programming language - we worked with Java) to a short sequence in another language (English). We used a novel method of code representation. As opposed to existing methods of treating the entire code as an English paragraph, we used a unique and succinct representation as input to a neural machine translation model and this resulted in our beating the state of the art results. We made use of the open-source **OpenNMT** library.

- **A Scalable Clustering Algorithm for Serendipity in Recommender Systems.**

(Under review at **SAREC 2018** (workshop at **ICDM**). Guide: **Prof. Shrisha Rao**, Jan'18 - Aug 18)

One area of Machine Learning that caught my interest was that of Recommender Systems. I worked on a modified version of Collaborative Filtering. SC-CF is a scalable clustering algorithm that efficiently tackles the problem of high sparsity and overspecialization in Recommender Systems. The goal was to effectuate serendipity in movie recommender systems with a simple clustering algorithm, that considers diversity, unexpectedness and relevance while making recommendations.

- **A Faster Sampling Algorithm for Spherical K-Means**

(Under review at **ACML 2018**, Guide: **Dr R Pratap**, Jan'18 - May'18)

This work was on statistical Machine Learning where we developed a Markov Chain (MCMC) based sampling algorithm for Spherical K-Means clustering of large document corpuses. This algorithm takes only one pass over data and gives 'k' cluster centers as opposed to vanilla Spherical K-Means and Spherical K-Means++. We find that our algorithm reduces the computational complexity of the ++ algorithm and obtains significant speed-up while retaining the clustering quality.

- **Detection of star clusters using Pattern Analysis - Indian Institute of Space Science and Technology (IIST)**

(Guide: **Prof. Sarita Vig**, May'17 - July'17)

Worked in the Earth and Space Science Department for a duration of 7 weeks on a direct application of pattern analysis techniques for a highly relevant problem statement in astrophysics – star cluster detection from large astronomical datasets. For this, an algorithm that can automatically detect potential star clusters from given star position data is needed. My work was concerned with using the K-Nearest Neighbor algorithm for this. We used KNN for both detection of star clusters and density estimation. We also carried out exhaustive comparative analysis between this method and previously used methods for cluster detection. Finally we estimated a background density level on real sky data sets which will serve as a standard for cluster detection. The most challenging part of this work was figuring out how to work with astronomical data and understanding the relevance and significance of its various attributes.

- **Content-Based Image Retrieval using Neuromorphic Computing.**

(Guide: **Prof. Shrisha Rao**, Aug'17 - Dec'17)

Content-Based Image Retrieval (CBIR) is a technique of identifying matching images based on the image contents like intensity, shape, etc. The project looks at cognitive methods of achieving CBIR using neuromorphic

architecture which is widely gaining popularity. We made use of a Pulse Coupled Neural Network (PCNN) model to achieve image matching and PCNN-Icons for feature extraction. PCNNs act as simulations of the brain and mimic the neuron firings of the brain. We were faced with computational power limitations and hence had to work with small datasets and could not exploit the full potential of this architecture. This was a challenging part in the project.

Achievements

Driven by an earlier interest in coding, I participated in a series of coding competitions conducted by [HackerRank](#). These contests test coding skills and knowledge in Data Structures and Algorithms in a highly competitive environment. Some of my performances are summarized as follows.

- **World Rank 7** ([HackerRank Women's Cup 2015](#))
7th rank out of 1000+ participants in the world and 3rd in the country. I was also featured in a [YourStory](#) article.
- **World Rank 23** ([Adobe CODHERS Codesprint 2016](#))
23rd rank out of 1500+ participants all over the world and 21st place in the country.
- **World Rank 49** ([Women's CodeSprint 2016](#))
49th rank out of 2500+ participants all over the world and 19th place in the country.

The following are my test scores :

- GRE – 158 Verbal, 169 Quant (Total 327)
- TOEFL – yet to be given

Non-Academics

- I am a **TA** for the **Machine Learning - 101** course taken at IIIT-B at an introductory level (August - December, 2018). I am assigned with the tasks of taking weekly tutorial sessions, setting questions for quizzes, planning assignment questions and projects. Along with for 4 other TAs, we handle a class of 240 students.
- As a part of the **Student Mentoring Program** at IIIT-B, I served as a mentor for 3 first-year students during the academic year of 2017-2018. Mentorship was in terms of academics as well as their cultural and social lives within the college. It was an enriching and fun experience.
- I was a curator at the TEDx event held at IIITB in 2018. My responsibilities included managing and assisting the speakers with their speeches. The theme for the event was "Demystify"
- When I'm not busy working or coding, I'm often reading, writing or dancing. I am an avid reader and have been one since I was a child. I particularly enjoy works by Dan Brown, Agatha Christie and Khaled Houssini. I have two blogs that I try to keep updated - a [personal](#) one and a [technical](#) one. I enjoy dancing and am an active member of the Dance Club at IIIT-B. I play the [veena](#), an Indian classical music instrument, and have received formal training for the same.

References

1. [Prof. Dinesh Baby Jayagopi](#) (jdinesh@iiitb.ac.in)
2. [Prof. G Srinivasaraghavan](#) (gsr@iiitb.ac.in)
3. [Prof. Shrisha Rao](#) (srao@iiitb.ac.in)
4. [Prof. Sujit Kumar](#) (sujitkc@iiitb.ac.in)

Contact

Email : 96.pratheek@gmail.com

Ph. No : +91 9535953152

[LinkedIn profile](#)

[Github Handle](#)