

Report of the IIT Bombay MInDS workshop, 18 Jan 2020
Prof. Ankur A. Kulkarni

The Dean R&D, Prof. Milind Atrey initiated the speech thanking all the dignitaries, speakers, and participants. As a whole, 140 faculty members and around 250 students (MTech and PhD level) registered for the event. He emphasized that the institute has growing interests in the areas concerning AI/DS not only in core departments like CS or EE, but the majority of the other departments. He mentioned that a number of companies in Japan have been approached to potentially fund the MiNDs department. He concluded the speech by thanking the organizers again, in particular Profs. Sudarshan, Ankur, and Manjesh.

Next, Prof. Subhasis Chaudhuri (Director, IITB) mentioned about the recent trends in the areas of Artificial Intelligence and how different institutes over the globe initiated several endeavors in this line. However, he was very clear about the need of academia-industry interaction in these areas. IIT Bombay is planning to start AI related courses at different levels. In particular, he mentioned that another committee at IIT Bombay has been working on revamping the current academic curriculum and that committee also recommends the need of AI degree at the UG level. He concluded with a note of thanks and also mentioned that such an event (the workshop) would help connect people from different spheres: academic, industry, and external agencies.

The schedule and list of talks in the workshop are given in the Annexure.

Session 1

The session was chaired by Prof. Amit Sethi.

Prof. Ganesh Ramakrishnan:

Prof. Ramakrishnan (CSE) summarized the works done by Profs. Subhasis Chaudhuri (EE), Rajbabu V (EE), Biplab Banerjee (CSRE), and him in the areas of computer vision and video analytics. He showcased several video analytics products developed by his group for video summarization, person search, text based video retrieval, text recognition from videos, number-plate recognition, and OCR, respectively. Besides, he mentioned about the work on crowd-sourcing noisy labels to deal with the problem due to label scarcity. For the recent works from the VIP lab, EE, he highlighted the works on image co-segmentation, zero-shot learning, person re-identification, and abnormal activity recognition.

Prof. Amit Sethi:

Prof. Amit Sethi presented on behalf of Profs. Ajit Rajawade (CSE), Suyash Awate (CSE), Sharat Chandran (CSE), and himself.

Prof. Rajawade currently focuses on medical imaging and under-water imaging. He is very much interested in the image acquisition part and works on the compressive sensing based techniques. The major focus of Prof. Awate has been onto (a) medical image computing where he looks into problems including statistical shape analysis, CAD, image quality enhancement detection and delineation, deep learning for segmentation, and (b) neuro image analysis where several problems related to MRI imaging are tackled. Prof. Sethi also presented the works of Prof. Sharat in the area of visual data captured by quadcop which can be used for inference even in no GPS areas. Prof. Sethi himself works in the domain of medical imaging and proposed several deep learning based models for regions of interest detection, segmentation, and modeling the deep learning systems for small sized datasets.

Prof. Biplab Banerjee:

Prof. Biplab Banerjee (CSRE) presented the work of his group in the area of satellite image analysis using deep learning. His main focus has been in developing deep learning models for improved overall scene understanding. He mentioned four major research directions: multi-label aerial scene recognition, multi-task learning for joint semantic segmentation and depth perception from monocular images, multi-modal learning in the context of remote sensing, and unsupervised domain adaptation.

Prof. Preeti Jyothi:

Prof. Preeti Jyothi (CSE) presented the works carried out by Prof. Preeti Rao (EE), Prof. Rajbabu (EE), and herself in the domain of speech processing. She started with the challenges of Automated Speech Recognition (ASR) for regional languages and then mentioned about their recent work in accent aware ASR. Code switching is another research area of interest to her group where she mentioned about their recent works using recurrent networks. They are currently focusing heavily on Indian language ASR with support from several industries. On the other hand, Prof. Rao and Rajbabu focus on speech enhancement, joint audio-visual analysis.

Prof. Puspak Bhattacharya:

Prof. Puspak Bhattacharya (CSE) presented the research works of Prof. Ganesh Ramakrishnan (CSE), Prof. Soumen Chakraborti (CSE), and himself. The emphasis on NLP at IIT Bombay started almost 20 years back and currently the research group is well-known for their endeavors in diverse topics including computational politeness, question generation from texts, knowledge graph based approaches for reasoning, complex question-answer modeling, to name a few.

Prof. Sunita Sarawagi:

Prof. Sarawagi (CSE) presented on behalf of Prof. Vivek Borkar (EE), Prof. Manoj Gopalkrishnan (EE), Prof. Shivaram Kalyanakrishnan (CSE), and herself. Prof. Shivram currently works in the domain of reinforcement learning for agent based modeling. Specifically, he deals with multi-arm bandit, evolutionary computing, and AI for societal good. Prof. Borkar has pioneered in the areas of distributed optimization, safe reinforcement learning, and reinforcement learning for structured policies. Prof. Sarawagi has been working on domain adaptation, sequence to sequence learning using graphical models, and the usage of high-level supervised which goes beyond the usual exemplar based approaches. Prof. Manoj has deep interests in molecular computing and deals with machine learning techniques to reason about the chemical reaction networks.

Keynote address:

Dr. Mayur Datar, chief data scientist from Flipkart India delivered the keynote and mentioned about how Flipkart extensively relies on machine learning and statistical approaches for building their recommendation systems. He browsed through some of the optimization techniques they follow for online learning and to deal with noisy data, he also spoke about the customization of the Flipkart platform for regional customer support. He expressed the willingness to collaborate with the faculties at IITB in areas concerning NLP, Speech analysis, and large-scale intelligent system development.

Session 2.A

The session was chaired by Prof. S. Sudarshan and the session had two talks.

Prof. Soumen Chakrabarti:

The first talk was by Prof. Soumen Chakrabarti (CSE) about social media and network analysis. He spoke about the scourge of fake news, and an architecture "Kauwa Kaate" which has been developed at IIT Bombay to battle this scourge. He also spoke about influence propagation in social networks, aging-related effects in social networks, privacy-preserving friend suggestions, and designing graph algorithms for long-term welfare.

Prof. Sanjeev Sabnis:

The second talk was by Prof. Sanjeev Sabnis (Maths). He spoke about activities in the statistics group at IIT Bombay related to big data. These include techniques like Sure Independence Screening, Zero Inflated Binomial model, and curve registration.

Session 2.B

This session had two talks, both delivered by students in the robotics domain. First was on the use of computer vision for Autonomous Underwater Vehicles (team Matsya). The second was on self driving cars on using guidance and computer vision.

Panel Discussion on academia / industry interaction in AI/ML

The panelists were:

Moderator: Prof. Sunita Sarawagi from IITB (**SS**)

Dr. Mayur Datar from Flipkart (**MD**)

Mr. Abhishek Shekhar from JPMC (**AS**)

Dr. Sanjay Bhat from TCS (**SB**)

Mr. Jayesh Pillai from L&T-NxT (**JP**)

SS began the discussion by noting that there is considerable interest in AI/ML research in academia as well as the industry. Given this, there are complementarities between the two that can be exploited fruitfully. Specifically, industry offers early access to practical problems, while academia offers depth, perspective (and bright and ambitious students).

SS moderated the discussion by posing the following questions.

Q1: List AI / ML problems in your organization where there is scope for academia interaction

AS: Many classical models in finance ignore details and practical considerations like transaction costs. ML tools can be used to address these issues.

Training machines to hedge against the risk associated with financial contracts, training networks to price contracts, high frequency trading, etc. all represent interesting problems, and also come as interns.

JP: L&T works on data streaming from machines in factories. Machine failures can be very expensive => need predictive maintenance using ML tools.

L&T is also interested in smart city projects, including smart traffic management, facial recognition, etc. Also navigation solutions for army (autonomous navigation/combat) manned / unmanned vehicles.

SB: Predicting when a certain feature in time series will recur, for example, predicting the peak load time in the grid.

Sometimes, datasets come from clients, where the data is not necessarily collected in a sound manner. How to 'clean' the data? More importantly, how does one figure out what can be done using a given dataset?

MD: Adapting research in Indian context. Mixed language queries.

Can you predict the success/failure of a script with actors, etc.? For content creation.

Optimization of hyper-local delivery systems.

FinTech: Can you deduce the credit worthiness of a customer given publicly available data.

Q2: Sharing of data: How can data be shared with academia for research purposes?

MD: Tell us how to anonymize data so that we can share the data with you.

AS: Need a platform for industry-academia interaction for data sharing.

JP: Might need tripartite arrangement, since L&T does not own the data itself.

SB: As a society, we don't collect a lot of data. IIT faculty who are on board/advisories should use their clout to encourage data collection and make it public.

MD: Some data, including product reviews on Flipkart, are already public.

Public Q: What are some best examples of academia-industry interactions, and bad examples.

MD: Good papers have come out of academia-industry interactions. Bad experiences: sometimes, priorities shift quickly in the industry, leaving projects in limbo.

AS: Understand the domain first, before jumping the ML gun.

Session 3

This session was concerning Decision Making and was chaired by Prof. Ankur A. Kulkarni and contained three talks -- on logistics, on game theoretic decision-making, and on online learning.

Prof. Jayendran Venkateswaran:

Prof. Jayendran (IEOR) discussed several projects that IEOR faculty are working on:

- Computer vision for autonomous cars. For example, why did the vehicle brake?
- Data analytics for electric vehicles. Route planning, charging, scheduling
- Predicting the health of a vehicle
- RL in random environments: change point detection
- Robust supply chain management – minimize the impact of an epidemic
- Performance of assembly centres (spread across India) – deduce quality of a centre based on operations data
- Railway operations: timetable management
- Container logistics: fleet sizing, repositioning decisions...
- Data driven methods for after-sales market for car manufacturers

Prof. Ankur A. Kulkarni:

The projects described by Prof. Kulkarni (SysCon) include:

- How learning influences behaviour: A model for recommendation systems
- Control over social networks: Control of opinion dynamics.
- Routing and scheduling for logistics

- Minotaur toolkit for integer programming

Agent-based modelling: can overcome data inadequacies, can design agents who are also learning. Examples of applications of high frequency trading and incentive design for insurance agents.

Prof. Jayakrishnan Nair:

Prof. Nair (EE) discussed several projects in the domain of online learning:

- Dynamic pricing on a crowd sourcing platform
- Resource allocation in wireless networks
- Network utility maximization where the utility functions have to themselves be learnt
- Learning distribution properties via oracle queries
- Dynamic provisioning and caching in edge computing
- Environment oblivious, risk aware bandit algorithms

Session 4

The session was chaired by Prof. Manjesh Hanawal. There were 5 talks in this session, almost all of which were oriented towards bringing in domain knowledge more strongly into ML based approaches.

The first talk was presented jointly by Prof. Usha Anantharaman (SOM) and Prof. Rajendra Sonar (SOM) on the aspects of **AI in management practices**. After an initial overview of areas such as applied statistics, statistical learning and data mining, Prof. Anantharaman focused on the important problem of **making decisions under unbalanced data** (imbalanced data), where each of the class/ behavior is not of the same size, and the overall problem could be overwhelmed by the class behaviour which is the largest. She dwelt on feature selection and modeling aspects of such problems in the presence of between class and within class imbalances, and also on class overlaps. Subsequently she also described the delightful problem related to **statistical analysis of paintings**, and dwelt on *how does one analyze art*. She touched upon aspects such as classification based on colors, style, school genre and presented an example of the use of CNN to analyze style and content of paintings.

Prof. Sonar then talked about different aspects of knowledge based expert systems and case based reasoning methods. He made a strong case for how data based representations do not take care of semantics and presented a Web based RAD platform that was developed by him.

The second talk was by Prof. Anupama Kowli (EE) and was on the use of AI/ML based approaches in the **area of power systems control**. She presented many problems associated with synchrophasor analytics such as oscillation monitoring, short term load forecasting, prediction of peak load, and also problems associated with updating the grid model, calibration of instrument and bias detection. She next talked about the problem of how data analytics could help in the supply-demand balance, in the context of inertia control of wind farms. Some aspects of demand side management such as flexibility in adjusting the demand balance (loads) were highlighted. Next IoT approaches for end-use control as well as problems associated with sensor fault detection, sensor placement, analytics for load control were also highlighted.

The third talk was by Prof. Asim Tewari from ME department, and was focused on various aspects of the use of **AI in manufacturing**. He mentioned initiatives by various countries in this area, and talked about the Samartha Udyog Bharat 4.0 initiative. He then talked about several problems / case studies such as machine effectiveness monitoring, data analytics in the tool condition monitoring, breakout detection in continuous casting, development of cyber twin for 3-

axis CMC, and also presented some of the work that was initiated at IITB under the **connected campus initiative**.

The fourth talk was by Prof. Sanjeeva Shrivastava (BSBE) and was in the area of **proteogenomics and machine learning for health care**. He talked about the use of AI tools for analyzing several problems such as proteomics based investigation of brain tumors, ANN based localization of meningioma. He dwelt on the problem of analyzing why some tumors are more aggressive than others. He then presented approaches to identify biomarkers using support vector machines (SVM), application of ML In malarial proteomics, and different aspects associated with building human brain proteomic map.

The final talk of the session was by Prof. Mani Bhushan (ChE), and he presented about the various research activities being carried out in the Chemical Engineering Department. He highlighted several developments related to **soft sensing, inferential estimation** for difficult to measure variables in chemical engineering. He also mentioned about projects being taken for safety critical installations such as nuclear reactors (**condition based monitoring**) as well as fighter aircrafts (**control loop performance monitoring**). He also highlighted several other industrial projects directed towards achieving **operational excellence** in manufacturing using data based and model based analytics. He also described many ongoing and completed works at the intersection of chemical process control engineering and AI/ML.

Annexure: Schedule and list of talks

	Time		Agenda	Speaker (Profs.)
	8:30 am - 9.00 am		Registration	
	9.00 am - 9.05 am		Welcome Address	Dean R&D - Milind Atrey
Session 1 9.05 am -10.35 am	9.05 am - 9.20 am		Vision / Video Analytics	Ganesh Ramakrishnan Rajbabu Subhasis Chaudhuri
	9.20 am - 9.35 am		Image Processing and Analytics	Amit Sethi Suyash Awate Sharat Chandran Ajit Rajwade
	9.35 am - 9.50 am		Satellite Image Processing	Biplab Banerjee B Krishna Mohan
	9.50 am - 10.05 am		Speech	Preethi Jyothi Preeti Rao
	10.05 am - 10.20 am		Natural Language (Text/language/sentiment)	Pushpak Bhattacharyya Ganesh Ramakrishnan Soumen Chakrabarti
	10.20 am - 10.35 am		ML / DL / RL Foundations	Sunita Sarawagi Shivaram Kalyanakrishnan Vivek Borkar Manoj Gopalkrishnan
	10.35 am - 11.05 am		Tea Break & Poster session	Networking
Keynote	11.05 am - 11:45 am	11.05 am - 11.45 am	Data Science Applications in E-commerce	Dr. Mayur Datar (Chief Data Scientist, Flipkart)
Session 2.A	11.45 am - 12.15 pm	11.45 am - 12.00 pm	Social Networking	Soumen Chakrabarti Abir De S. Sudarshan Kameswari Chebrolu
		12.00 pm - 12.15 pm	Math / Statistical Foundations	Sanjeev Sabnis Radhendushka Srivastava
Session 2.B	12.15 pm - 12:35 pm	12.15 pm - 12.25 pm	Autonomous Underwater Vehicle	Mohan Abhyas + AUV team Leena Vachhani Hemendra Arya
		12.25 pm - 12.35 pm	Self-driving cars (SeDriCa)	Hemant Kumawat Atharva Jaipurkar S. N. Merchant

	12.35 pm - 2.00 pm		Lunch	Networking Posters/Demos +
Panel	2.00 pm - 2.45 pm	2.00 pm - 2.45 pm	Panel discussion	Sunita Sarawagi Dr. Mayur Datar (Flipkart) Mr. Abhishek Shekhar (JPMC) Dr. Sanjay Bhat (TCS) Mr. Jayesh Pillai (L&T-NxT)
Session 3	2.45 pm - 3.30 pm	2.45 pm - 3.00 pm	Logistics / Supply-chain / IE	Jayendran V Narayan Rangaraj N. Hemachandra
		3.00 pm - 3.15 pm	Decision Making 1 (optimization, game theory, control)	Ankur Kulkarni Mallikarjun Rao D. Manjunath Ashutosh Mahajan
		3.15 pm - 3.30 pm	Decision Making 2	Jaykrishnan Nair Manjesh Hanawal Sharayu Moharir Nikhil Karamachandani
	3.30 pm - 4.00 pm		Tea Break & Poster session	
Session 4	4.00 pm - 5.15 pm	4.00 pm - 4.15 pm	AI in Management	Rajendra Sonar Usha Anantakumar Piyush Pandey
		4.15 pm - 4.30 pm	Power Systems Control/IoT	Anupama Kowli S. A. Soman Mukul Chandorkar B.G. Fernandes
		4.30 pm - 4.45 pm	Manufacturing	Asim Tewari
		4.45 pm - 5.00 pm	Proteomics/Genomics	Sanjeeva Srivastava Santosh Noronha Swati Patankar
		5.00 pm - 5.15 pm	Chemistry/Chemical Engg	Ravi Gudi Mani Bhushan Sharad Bhartiya R. Sunoj
	5.15 pm - 5.30 pm		Future Plans for MInDS Centre	AI&DS Committee Members (S. Sudarshan)
	5.30 pm - 5.40 pm		Vote of Thanks	Associate Dean R&D
	5.40 pm - 6.00 pm		High Tea	