Souradip Chakraborty |

College Park Maryland, USA
301-204-3898 • schakra3@umd.edu

Summary

Currently a 1st year Computer Science PhD student at the University of Maryland, College Park working in the theoretical and practical aspects of Deep Reinforcement Learning and Representation Learning.
In the past, worked at Fortune 1 company (Walmart Labs) with significant impact affecting hundreds of millions of users. Recognized as the **Google Developer Expert in Machine Learning'2019**, representing India. Batch Topper ***(summa cum laude)*** from Indian Statistical Institute. Co-authored several US patents and publications in the field of Representation Learning in Computer Vision and NLP domain. Selected as the Youngest Technical Speaker for the very prestigious Data Hack Summit'2018 & 2019 by Analytics Vidhya.

Research Interests

Deep Reinforcement Learning, Representation Learning, Probabilistic Machine & Deep Learning, Bayesian ML, Uncertainty, Robustness Generalisation

Education

University of Maryland, College Park
PhD, Computer science *Fall 2021 – Present*
Advisor: Dr.Furong Huang, Dr. Paratap Tokekar
Research: Theoretical and Practical aspects of Deep Reinforcement Learning, State Abstraction and Learning under Sparse settings

Indian Statistical Institute,Bangalore
Master's Degree *2016 – 2018*
MS with Major in Data Science & Machine Learning **GPA: 9.4**/10; **Rank — 1** (**Batch Topper**)
Advisor: Dr.B. S. Daya Sagar
Courses: Probability, Statistics, Inference, Statistical Machine Learning, Design of Experiments, Optimization and Reliability Theory.

Jadavpur University
Bachelor of Engineering, Electronics & Instrumentation Engineering *2010 – 2014*
GPA: 8.4/10; **Rank — 10**
Bachelor Thesis: Multivariate Anomaly Detection of Control Variable - an on-line multivariate anomaly detection framework of Pressure,Temperature,Flow,Level simultaneously using Mahalanobis distance and Hotelling T2 statistic.(Research Center, National Thermal Power Corporation Limited)

Google Developer Expert- Machine Learning

Recognized by Google as a Google Developer Expert in Machine Learning for my community contributions in Machine Learning research and Data Science and being a thought leader in sharing the most innovative ideas by speaking in various prestigious forums.

Work Experience

Walmart Labs
Research Data Scientist ||Statistical Machine Learning *Feb 2018 – Aug 2021*

- **Explainable Deep Neural Information Retrieval and Efficient De-Biasing of Sentence Vectors for enhancing the Search Mechanism||Information Handling Services**
- **Ingredient Share estimation using Constrained Non-Convex Optimization with Lagrange's' multiplier for Walmart Food Department**
- **No-Reference Image Quality Assessment & Disentangled Learning using Deep Autoencoders and SSIM**
- **Scan from Kitchen to Cart: Vision based Recommender System & Deep Matrix Factorization**
- **Customer Intent based Recommender System and Hierarchical Product Classification**

Amec Foster Wheeler Pvt Ltd.
Research Engineer, Control and Statistics *July 2014-2016*

- **Multivariate Statistical Quality Control in Oil and Gas field**
  Ensuring statistical control of process variable like Pressure, Level,Temperature,Flow through Statistical Control Charts and implementing Root Cause Analysis to identify deviation of process variables from target value.

Publications

- Amrit Singh Bedi, *Souradip Chakraborty*, Anjaly Parayil, Brian Sadler, Pratap Tokekar, Alec Koppel **"On the Hidden Biases of Policy Mirror Ascent in Continuous Action Spaces"**, *Under submission*(Link : *preprint*)
- *Souradip Chakraborty*, Ekaba Bisong, Shweta Bhatt, Thomas Wagner, Riley Elliott and Francesco Mosconi **"BioMed-BERT: A Pre-trained Biomedical Language Model for QA and IR"**, *28th International Conference on Computational Linguistics (COLING'2020)* (Website : *Link*)
- Saswata Sahoo, *Souradip Chakraborty* **"Graph Spectral Feature Learning for Mixed Data of Categorical and Numerical Type"**, *25th International Conference on Pattern Recognition (ICPR'2020)* (*Link*)
- *Souradip Chakraborty*,Sayak Paul,Aritra Roy Gosthipaty **"G-SimCLR : Self-Supervised Contrastive Learning with Guided Projection via Pseudo Labelling"**, *IEEE ICDM'2020, DLKT Workshop Proceedings* (*Link*)
- *Souradip Chakraborty*, Ekansh Verma, Saswata Sahoo, Jyotishka Datta **"FairMixRep : Self-supervised Robust Representation Learning for Heterogeneous Data with Fairness constraints"**, *IEEE ICDM'2020, DLC Workshop Proceedings* (*Link*)
- Ekansh Verma, *Souradip Chakraborty*, Vinodh Kumar **"Propaganda Fragment Detection using Diversified BERT Architectures based Ensemble Learning"**, *SemEval Workshop Proceedings, COLING'2020* (*Link*)
- Ekansh Verma, *Souradip Chakraborty*, Vinodh Kumar **"Deep Multi-level Fusion Learning Framework for Multi-modal Product Classification"**, *SIGIR ecom'2020 Workshop Proceedings* (*Link*)
- Ojaswini Chhabra, *Souradip Chakraborty* **"Siamese Triple Ranking Convolution Network in Signature Forgery Detection"**, *selected at AICAAM'19,selected and submitted to Pertanika journal* (*Link*)
- Soumya Dasgupta, Kaushik Halder, Shohan Banerjee *Souradip Chakraborty*,Amitava Gupta **"Stability anlysis and controller synthesis of networked control system (NCS) with arbitrary packet drop-outs"** *2nd International Conference on Electronics and Communication Systems (ICECS)'2015* (Link : *paper*)

Books & Live Projects

- *Souradip Chakraborty*, Sayak Paul **"Extractive Text Summarization of News Articles with NLP, Deep Learning, Python - An attention-based framework"**, ***Manning Books*** *Live Project* (Link : *Website*)

Patents

- Gregory Dixon,*Souradip Chakraborty*,Ojaswini Chhabra,Mallikharjuna Mv **"Reverse Engineering Food Ingredient Share estimation using Constrained Optimization"**, US Patent, Walmart Ref. 6031US01.
- *Souradip Chakraborty*,Mani Garlapati **"Retail Based Cost Reverse Engineering and Cost comparison within Item Similarity Clusters for Cost Negotiations"**, US Patent, Walmart Ref. 5928US01.
- *Souradip Chakraborty*,Mani Garlapati **"Systems and Methods for Identifying Negotiable Items (Cost Analytics)"**, US Patent, Walmart Ref. 5604US01.
- *Souradip Chakraborty*, Rajesh Shreedhar Bhat,Mani Garlapati , **"System and Method For Automated Electronic Catalogue Management and Image Quality Assessment"**, US Patent, Walmart Ref.5118US01.
- *Souradip Chakraborty*, Rajesh Bhat, Mani Garlapati,Lakshmi Praneetha Kommuru, **"Generating Customized Alerts with Computer Vision and Machine Learning"**, US Patent, Walmart Ref. 5008US01.
- *Souradip Chakraborty*,Mani Garlapati **"Architecturally-Distributed Apparatus and Method to Form and Leverage Clustered Content (Customer intent based recommendation system)"**, US Patent, Walmart Ref. 4970US01.
- *Souradip Chakraborty*,Ojaswini Chhabra, **"System and Method for Detecting Signature Forgeries"**, US Patent, Walmart Ref. 5603US01.

Blogs & Research Articles

- *Souradip Chakraborty* **"Detection of COVID-19 presence from Chest X-ray scans using CNN Class Activation Maps "**, *Towards Data Science,Medium'2020*
- *Souradip Chakraborty*,Rajesh Shreedhar Bhat **"Reducing the Carbon Foot Prints of CNNs at the cost of interactions-Depthwise Pointwise Convolution"**, *Towards Data Science,Medium'2020*

- *Souradip Chakraborty* **"Bayesian Thinking for Linear Regression @ Kaggle Days Meetup"**, *Towards Data Science,Medium'2020*
- *Souradip Chakraborty*,Amlan Jyoti Das & Sai Yashwanth **"Risks and Caution on applying PCA for Supervised Learning Problems"**, *Towards Data Science,Medium'2019*
- *Souradip Chakraborty*,Rajesh Shreedhar Bhat **"Why not Mean Squared Error(MSE) as a loss function for Logistic Regression?"**, *Towards Data Science,Medium'2019*.
- *Souradip Chakraborty***"Dimensionality Reduction in Supervised Framework and Partial Least Square Regression"**, *Analytics Vidhya,Medium'2019*.

Research Experiences

Google Developer Expert,Google
Data Science Research *Feb 2018 – July 2018*
Advisor Team: Google Research
- ***AI vs COVID-19 BioMedical Research*** Our goal is to make BioMedBERT a resource for biomedical researchers, doctors, and virologists, to augment their ability to sift through biomedical knowledge and existing research to extract novel insights and help them make new drug discoveries. (Submitted to COLING'2020)

Walmart Labs
Data Science Team *Feb 2018 – July 2018*
Advisor: Mallikharjuna Mv
- ***AutoElement*** Identifying the most suitable classifier and the classifier conditional hyper parameters using Meta Learning and Gaussian process for item description categorization task in retail stores.

Indian Statistical Institute Bangalore
Technical Report, Statistical Quality Control unit *Nov 2017 – Jan 2018*
Advisor: Dr. Boby John
- ***Development of an advanced Slopping Control Chart methodology for simultaneously monitoring multiple characteristics*** Developed a new methodology for simultaneous monitoring multiple output characteristics using Multivariate regression and Derringer function.

Australia and New Zealand Banking Group
Data Science Dept. *May 2017 – July 2017*
Advisor: Dr. Krishnendu Chandra
- ***State of cycle analysis***, Developed a Dynamic Index for dating Business cycles using Hodrick Prescott filter in the context of credit risk management for Mortgage portfolio in Australian economy. Cross-Validated the index for US-Economy.

Selected Honors & Awards

- 2019: Selected as the Technical speaker for the very prestigious **Data Hack Summit'2019 on Captioning & Attention Models by Analytics Vidhya**.
- 2019: Key-Note Lecturer to the Faculty of Presidency University,Bangalore in Statistical Learning Theory **Faculty Development Program**.
- 2019: Key-Note Lecturer to the students of Computer Science department of Coimbatore Institute of Technology **Machine Learning Workshop with Python**.
- 2019: Invited as a Technical Keynote speaker for **Target Talks AI Session-3** Bangalore'19.
- 2018: Selected as the Youngest Technical Keynote speaker for the very prestigious **Data Hack Summit'2018 by Analytics Vidhya**.
- 2018: **Batch Topper Certification** and **Endowment** for the highest academic performance (Rank-1) in Master's,Indian Statistical Institute.
- 2017: Selected at **Novartis Biocamp 2017** and represented ISI Bangalore in Novartis as a Data Scientist (top 50 nationwide).
- 2010-2014: **4-year scholarship** for academic excellence, Ministry of Human Resource & Development, India

Competitions

- **Runners up**,**Codeception 2019** Walmart Labs International Hackathon.
- **Rank-13**,**Crowd Analytix's Propensity to Fund Mortgages competition 2019** – Implemented LightGBM with error analysis to identify the curvature of the variables and interaction among the features in modelling the response variable.
- **Bronze Medal**,**Capillary Machine Learning Hackathon by Analytics Vidhya'2019** – Implemented Alternating Least Squares Method for Implicit recommendation.
- **Bronze Medal**,**WNS Analytics Wizard 2018 challenge** – An ensemble of Boosting and Deep Neural nets with synthetic minority oversampling was implemented to solve the classification problem with class imbalance.

Skillsets

- **Programming Languages:** Python, R, C/C++, MATLAB; **Web Development:** HTML
- **Deep Learning components:** *RNNs, *CNNs, *GANs, Attention, Capsule (* denotes variants)
- **Frameworks/Databases:** PyTorch, ,Keras,Tensorflow, PySpark, Teradata, MongoDB, Hive, SQL
- **Tools/Softwares:** NLTK, OpenCV, Octave, Docker, LaTeX 2_ε