Souradip Chakraborty |

Bangalore, India
(+91) 9038790361 • souradip24@gmail.com

Summary

Working at Fortune 1 company with significant impact affecting hundreds of millions of users.Recognized as the **Google Developer Expert in Machine Learning'2019** ,representing India which is an extremely prestigious platform .Batch Topper with record grades from Indian Statistical Institute.Co-authored several US patents and publications in Natural Language Processing,Representation Learning,Computer Vision and Machine Learning applications.Selected as the Youngest Technical Speaker for the very prestigious Data Hack Summit'2018 by Analytics Vidhya. Currently,I am also a Thesis Supervisor for students at upGrad's online Master's Program in Data Science as well as Machine Learning with Liverpool John Moores University (LJMU)

Research Interests

Representation Learning,Graphical Networks, Bayesian Inference and its subsequent applications in Multi-modal domains of Computational Linguistics (NLP) and Computer Vision..

Education

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**)
Master's Thesis: Developed a Hybrid Recommendation Engine deploying the social presence of the customers using NLP and Deep Learning Framework for e-commerce.
Advisor: Dr.B. S. Daya Sagar,SSIU,Indian Statistical Institute
Courses: Machine Learning,Statistical Learning Theory,Pattern Recognition,Mathematical Morphology, Multivariate Statistical Analysis,Bayesian Inference,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'2019

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. Representing India as one of the 13 GDEs in Machine Learning.

Work Experience

Walmart Labs
Research Engineer, Machine Learning *Feb 2018 – Present*

- **Automated Catalogue Management and Image Quality Assessment**
  MobileNet CNN Embedding and Structural Similarity Index with simulated noise was used in implementing no-reference human perceived quality assessment of catalogue images.
- **Scan from Kitchen to Cart:Vision based Recommender System**
  Developed an intelligent smart kitchen technology which will detect the ingredients present in kitchen if it falls beyond a threshold and recommend the same using Image processing and Convolution Neural Network and Time Series Forecasting.
- **Customer Intent based Recommender System**
  Developed a methodology for classifying and categorizing Store Item descriptions into Customer intent categories using a combination of Hierarchical clustering, Word embedding and Deep Neural Network classification. Built a Hybrid recommendation engine deploying the social presence of the customers using Gensim Word2vec embedding and cosine similarity.

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.

Patents

- Gregory Dixon,*Souradip Chakraborty*,Ojaswini Chhabra,Mallikharjuna Mv **"Reverse Engineering Food Ingredient Share estimation using Constrained Optimization"**, US Patent, Walmart Ref. 6031US01, Provisionally filed, 2019.
- *Souradip Chakraborty*,Mani Garlapati **"Retail Based Cost Reverse Engineering and Cost comparison within Item Similarity Clusters for Cost Negotiations"**, US Patent, Walmart Ref. 5928US01, Provisionally filed, 2019.
- *Souradip Chakraborty*,Mani Garlapati **"Systems and Methods for Identifying Negotiable Items (Cost Analytics)"**, US Patent, Walmart Ref. 5604US01, Provisionally filed, 2019.
- *Souradip Chakraborty*, Rajesh Shreedhar Bhat,Mani Garlapati , **"System and Method For Automated Electronic Catalogue Management and Image Quality Assessment"**, US Patent,US Patent, Walmart Ref.5118US01, Provisionally filed, 2018.
- *Souradip Chakraborty*, Rajesh Bhat, Mani Garlapati,Lakshmi Praneetha Kommuru, **"Generating Customized Alerts with Computer Vision and Machine Learning"**, US Patent, Walmart Ref. 5008US01, Provisionally filed, 2018.
- *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, Provisionally filed, 2018.
- *Souradip Chakraborty*,Ojaswini Chhabra, **"System and Method for Detecting Signature Forgeries"**, US Patent, Walmart Ref. 5603US01, Provisionally filed, 2019.

Publications

- Saswata Sahoo, *Souradip Chakraborty* **"Graph Spectral Feature Learning for Mixed Data of Categorical and Numerical Type"**, *submitted for IJCAI'19*
- Ojaswini Chhabra, *Souradip Chakraborty***"Siamese Triple Ranking Convolution Network in Signature Forgery Detection"**, *selected at NCMLAI'19,AICAAM'19,selected and submitted to Pertanika journal*.
- Mani Garlapati, *Souradip Chakraborty* **"Automated Catalog Management and Image Quality Assessment using Convolution Neural Networks and Transfer Learning"**, *selected at NCMLAI'19, AICAAM'19,selected and submitted to Pertanika journal*.
- Mani Garlapati, *Souradip Chakraborty* **"Graph Community Clustering and Retail Cost based Reverse Engineering"**, *accepted for NCMLAI'19, AICAAM'19,selected and submitted to Pertanika journal*.
- Mani Garlapati, *Souradip Chakraborty*,Rajesh Bhat,Lakshmi Praneetha Kommuru **"Customers consumption based Recommendation system"**, *accepted for "POSTER SESSION" at the Grace Hopper Celebration India (GHCI)'18 conference*.
- 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*

Blogs & Research Articles

- *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

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: elected 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