

Bodhisattwa Prasad Majumder

La Jolla, San Diego, CA

(540) 467-6530

✉ bmajumde@eng.ucsd.edu

🌐 www.majumderb.com

Curriculum Vitae

Summary

Worked at Fortune 100 companies and researched at R&D giants with significant impact affecting hundreds of millions of users. Summa cum laude (Gold Medalist) from IIT Kharagpur. Co-authored over 10 publications at highly reputed venues, 4 US patents and 1 book in Natural Language Processing and Machine Learning applications.

Research Interests

Natural Language Processing, Machine Learning, Artificial Intelligence, Computational Linguistics

Education

2018 - Present **Ph.D., Computer Science (Artificial Intelligence Group)**, *University of California, San Diego*.

Advisor: Dr. Ndapa Nakashole, Artificial Intelligence Group, CSE

Broad direction: Few-shot learning (Meta learning, Multi-task learning, Transfer learning) for low-resource Natural Language Processing (NLP) tasks (Machine Reading such as QA, NER, RE); **Area:** NLP, Machine Learning

2015 – 2017 **Master's Degree**, *Indian Institute of Technology, Kharagpur*.

Major in Data Science & Machine Learning; Coursework at Indian Statistical Institute, Kolkata

GPA: 9.71/10; **Rank** — 1/51 (Gold medalist/ **Summa cum laude**)

Master's Thesis: Decoding Consumer Behavior in Retail via Statistical and Deep Learning Frameworks

Advisor: Dr. Animesh Mukherjee, CSE, Indian Institute of Technology Kharagpur

Courses: Machine Learning, Algorithms, Information Retrieval, Data Mining, Multivariate Analysis

2011 – 2015 **Bachelor of Engineering, Electronics & Telecomm. Engineering**, *Jadavpur University*.

GPA: 9.48/10; **Rank** — 5/70

Bachelor Thesis: An interactive GUI based Real-time pulse processing system - an on-line streaming data analysis framework (Variable Energy Cyclotron Center, Bhaba Atomic Research Center)

Work Experience

June 2017 – **Research Engineer, Machine Learning**, *Walmart Labs*.

- July 2018 ○ **Neuro-attribute tagging** using LSTM-CRF (with / without attention) to improve faceted product search
- **Conversational Agent** with entity extraction and resolution; Feedback based learning to improve chat experience
- **Attention based OCR** for text extraction and entity recognition from images to improve product catalogs

Selected Publications

- Amrith Krishna, *Bodhisattwa P. Majumder*, Rajesh Bhat, Pawan Goyal, “**Upcycle Your OCR: Reusing OCRs for Post-OCR Text Correction in Romanised Sanskrit**”, *CoNLL, 2018 (accepted for presentation)* [pdf]
- Amrith Krishna, *Bodhisattwa P. Majumder*, Pawan Goyal, “**An ‘Eklavya’ approach to learning Context Free Grammar rules for Sanskrit using Adaptor Grammar**”, *17th World Sanskrit Conference, 2018* [pdf]
- *Bodhisattwa P. Majumder**, Aditya Subramanian*, Abhinandan Krishnan, Shreyansh Gandhi, Ajinkya More, “**Deep Recurrent Neural Networks for Product Attribute Extraction in eCommerce**”, *Arxiv, 2018* [pdf]
- *Bodhisattwa P. Majumder*, Amrith Krishna, Unni Krishnan, Anil K Boga, Animesh Mukherjee, “**What’s in a ‘Meme’? Understanding the Dynamics of Image Macros in Social Media**”, *Preprint manuscript, 2018* [pdf]
- Denise M. Case, M. Nazif Faqiry, *Bodhisattwa P. Majumder*, Sanjoy Das, and Scott A. DeLoach, “**Implementation of a Two-tier Double Auction for On-line Power Purchasing in the Simulation of a Distributed Intelligent Cyber-Physical System**”, *Advances in Artificial Intelligence, Research in Computer Science, MICAI, 2014* [pdf]
- *Bodhisattwa P. Majumder*, M. Nazif Faqiry, Sanjoy Das, Anil Pahwa, “**An Efficient Iterative Double Auction for Energy Trading in Microgrids**”, *IEEE Symp. on Computational Intelligence Applications in Smart Grid, 2014* [pdf]
- Satrajit Mukherjee, Kunal Pal, *Bodhisattwa P. Majumder*, Chiranjib Saha, B. K. Panigrahi, “**Differential Evolution based Score Level Fusion for Multimodal Biometric Systems**”, *IEEE Symposium on CIBIM, 2014* [pdf]
- Satrajit Mukherjee, *Bodhisattwa P. Majumder*, Aritran Piplai, Swagatam Das, “**A Novel Fuzzy Non-homogeneity Measure based Kernelised Image Segmentation for Noisy Images**”, *FUZZ-IEEE, 2014* [pdf]

Book

- **Natural Language Processing: A Pragmatic Approach** [\[Website\]](#)
Anuj Gupta (IIT-Delhi), *Bodhisattwa P. Majumder*, Harshit Surana (CMU), Sowmya Vajjala (Microsoft Research)
In Progress. Expected to be published by the mid of 2019.

Selected Honors & Awards

- 2018-2019: **Department Fellowship**, PhD, Computer Science and Engineering, University of California, San Diego
- 2017: **Gold medal** and **Endowment** for the highest academic performance (Rank-1) in Master's, IIT Kharagpur
- 2016: **Finalist, Data Science Game '16**, Paris; Represented India (1 out of 3 teams), International Rank 14
- 2015: Endowed with **cash scholarship** for academic excellence (obtaining SGPA > 9.5), Indian Statistical Institute
- 2014: Officially entitled as **contributor** in **NSF-CPS** project (**CNS -1136040**) by PIs, Kansas State University
- 2011-2015: **4-year scholarship** for academic excellence, Ministry of Human Resource & Development, India

Patents

- Diptarka Saha, Debanjana Banerjee, *Bodhisattwa P. Majumder*, "**REDCLAN - Relative Density based Clustering and Anomaly Detection**", US Patent, 15/977167, Provisionally filed, 2018
- Sumanth Prabhu, *Bodhisattwa P. Majumder*, Rajesh Bhat, Vignesh S, Anirban Chatterjee, Gayatri Pal, "**Automated Extraction of Product Attributes from Images**", US Patent, Walmart Ref. 4909US01, Provisionally filed, 2018
- Aditya Subramanian, *Bodhisattwa P. Majumder*, Shreyansh Gandhi, Abhinandan Krishnan, Ajinkya More, "**System and Method for Product Attribute Extraction Using a Deep Recurrent System**", US Patent, Walmart Ref. 4661US01, Provisionally filed, 2017
- Subhasish Misra, Arunita Das, Amlan Das, *Bodhisattwa P. Majumder*, "**Analytical Determination of Competitive Interrelationship between Item Pairs**", US Patent, Walmart Ref. 4431US01, Provisionally filed, 2017

Selected Research Experiences

- July 2017 – **Computer Science Dept., University of California, San Diego.**
June 2018 **Advisor: Dr. Julian McAuley**
 - Distributed behavioral semantic representation of products from large-scale transaction data for product analogy
 - **Temporal Product Embeddings** for time-based recommendations, identifying substitutes and complements.
- July 2017 – **CNeRG, Computer Science Dept., Indian Institute of Technology Kharagpur.**
June 2018 **Advisor: Dr. Pawan Goyal**
 - **Grammar Induction from Morphologically-rich Languages**, Effectiveness of using **Adaptor Grammar** for various **supervised** tasks – compound type classification, identifying the source for derivational nouns in Sanskrit
 - **Structured Prediction & Syntactic Analysis** using PCFGs – *Poetry to Prose* conversion, Dependency parsing
- Jan 2016 – **CNeRG, Computer Science Dept., Indian Institute of Technology Kharagpur.**
June 2018 **Advisor: Dr. Animesh Mukherjee**
 - **Diachronic Study of Image Memes** via emerging conventions in a closed community in *Facebook*
 - **Motif identification** from temporal usage using Adaptor Grammar; Study of **linguistic properties** of memes
- Summer 2014 **Electrical and Computer Engineering, Kansas State University.**
Advisor: Dr. Sanjoy Das
 - Study of **Multi-Agent Systems (MAS)**, Computational **Mechanism Design**, Multi-tier Double Auctions
 - Extension of Kelly's Mechanism in an Iterative Double Auction framework to capture the asymmetric preferences

Competitions & Invited Talks

- **Finalist**, 2016, **Data Science Game** (Microsoft, Capgemini and AxaLabs), Paris – International Rank 14
- **Winner**, 2017, **GE Healthcare** – Built a deep learning based low-cost multitenanted Tele-ICU video monitoring solution
- **Invited Talk** at Walmart Labs, 2017, *on* Information Extraction from Images - Application in e-Commerce
- **Invited Talk** at Indian Statistical Institute, 2017, *on* Deep Neural Network: in light of Optimization and Regularization
- **Invited Talk** at Industry Conclave & Graduate Orientation, '18 *on* Data Science, Indian Inst of Management, Calcutta

Skillssets

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