Arindam Paul (440) 622-1087| arindam.paul@eecs.northwestern.edu| www.arindampaul.me|linkedin.com/in/arndmpaul/

SUMMARY — **Computer Scientist** working in **Data Mining** with 5+ years research experience seeking full-time opportunities

INTERESTS — Machine Learning, Deep Learning, Natural Language Processing, Materials Informatics, Social Media Analytics

EDUCATION

Northwestern University, Evanston, Illinois USA [GPA: 3.7 / 4.0] **Jun 2019**
Ph.D. Candidate, Computer Engineering (expected)
Master of Science, Computer Science **Sep 2014**

Birla Institute of Technology & Science, Pilani, Rajasthan India
Master of Engineering (with Honors), Software Systems, May 2012 **May 2012**
Bachelor of Engineering (with Honors), Chemical Engineering, Dec 2009 **Dec 2009**

PROGRAMMING SKILLS

Programming : Python, MATLAB, Java, C++, R, MySQL,HTML, CSS, JavaScript, PHP
Data Science: Keras, Tensorflow, Scikit-Learn, Torch, Gensim, NLTK, Pandas, Numpy, Matplotlib, Spacy

CONSULTING

Data Science Intern, **Northwestern Mutual**, Milwaukee, Wisconsin **Jun - Aug 2018**
- Developed distributed image to text conversion algorithms for detecting responses from scanned questionnaires
- Developed a noise reduction algorithm to denoise scanned and photocopied questionnaires

Data Science Consultant, **EDT** **June 2017 - Jan 2018**
- Designed models for profanity detection from company-wide email databases

Research Intern, **Boeing Cybersecurity (Narus)**, Sunnyvale, California **Jun - Sep 2013**
- Generated synthetic user profiles with different demographic and interest features for analyzing ads across profiles
- Developed a machine learning model for predicting user demographics and interests from ads

PROFESSIONAL & RESEARCH EXPERIENCE

Research Assistant, **Northwestern University**, Evanston, Illinois (2012 -)

- Developed Deep Learning-based Predictive Model for Additive Manufacturing (Tensorflow, Keras) **Nov 2016 -**
  - ◇ Created time series models for temporal analysis of heat flux data
  - ◇ Investigated Recurrent Neural Network models to predict point-wise temperature information for accelerating additive manufacturing simulations

- Chemical Property Prediction using Molecular Fingerprints (Tensorflow, Scikit Learn) **Mar 2016 -**
  - ◇ Developed a multi-input input neural network architecture by merging different molecular representations (SMILES and fingerprints) for predicting chemical properties and reduced the mean absolute error by half compared to state-of-the-art architectures
  - ◇ Designed Deep Neural Network and Random Forest models for predicting power conversion efficiency of solar cells using chemical fingerprints, and achieved mean square percentage error between 1.5-2 %

- Very Deep Neural Networks for Predicting Formation Stability (Tensorflow) **Mar 2016 - Sept 2017**
  - ◇ Constructed Neural Network Models with 18-25 layers to predict formation energy of a chemical compound
  - ◇ Attained 20 % higher accuracy than the state-of-the-art models using Random Forests that would allow domain scientists to explore millions of possible compounds

- Ensemble Learning-based Guided Optimization for Aircraft Design (MATLAB, Python) **Oct 2015 - Dec 2017**
  - ◇ Created intelligent sampling algorithms to explore the constrained search space for candidate microstructures
  - ◇ Developed Feature Ranking-based Technique for Search Space Reduction of Constrained Non-Convex Optimization
  - ◇ Achieved 100x more solutions compared to state-of-the-art methods that can accelerate the design-to-experiment life-cycle

- Convolutional Neural Nets for Thematic Image Classification in Pinterest (Torch) **Oct 2015 - Sep 2016**
  - ◇ Harnessed Association Rule Mining for thematic label curation
  - ◇ Developed ConvNet Models for hierarchical classification that led to automated image categorization based on themes

- Classification of Anonymous Posts using Recurrent Neural Networks (Tensorflow) **Jan 2015 - May 2016**
  - ◇ Generated vectorizer models using Word2vec trained on crowd-sourced (Urban Dictionary) & psycho-lingual (LIWC) dictionaries (Gensim)
  - ◇ Attained prediction accuracy of 79.8 % and 78.1 % using LSTMs and ensemble models respectively

Select Publications (9 of 13)

A. Paul, D.Jha, R. Al-Bahrani, W. Liao, A. Choudhary and A. Agrawal."**CheMixNet: Mixed DNN Architectures for Predicting Chemical Properties using Multiple Molecular Representations**", *NIPS Workshop on Machine Learning for Molecules and Materials, 2018*

A. Paul, P. Acar, W. Liao, A. Choudhary, V.Sundararaghavan and A. Agrawal. "**Microstructure Optimization with Constrained Design Objectives using Machine Learning-Based Feedback-Aware Data-Generation**", *Journal of Computational Materials Science, 2018 (in review)*

D.Jha, L.Ward, **A. Paul**, W. Liao, A. Agrawal, A. Choudhary and C. Wolverton."**ElemNet: Deep Learning the Chemistry of Materials From Only Elemental Composition**", *Nature Scientific Reports, 2018*

M.Mozaffar, **A. Paul**, R. Al-Bahrani, S. Wolff, A. Choudhary, A. Agrawal, K. Ehmann and J.Cao."**Data-Driven Prediction of the High-Dimensional Thermal History in Directed Energy Deposition Processes via Recurrent Neural Networks**", *Manufacturing Letters, 2018*

A. Paul, P. Acar, R.Liu, W. Liao, A. Choudhary,V.Sundararaghavan and A. Agrawal. "**Data Sampling Schemes for Microstructure Design with Vibrational Tuning Constraints**", *Journal of American Institute of Aeronautics and Astronautics, 2018*

J.Birnholtz, N.A.R. Merola, and **A. Paul**. "**Is it Weird to Still Be a Virgin?: Anonymous, Locally Targeted Questions on Facebook Confession Boards**", *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. ACM, 2015.*

R. Liu, D. Palsetia, **A. Paul**, R. Al-Bahrani, D. Jha, W. Liao, A. Agrawal and A. Choudhary. "**PinterNet: A Thematic Label Curation Tool for Large Image Datasets**", *Proceedings of the Workshop on Open Science in Big Data at IEEE Bigdata Conference, 2016.*

A. Paul, A. Agrawal, W. Liao and A. Choudhary. "**AnonyMine: Mining anonymous social media posts using psycho-lingual and crowd-sourced dictionaries**", *Proceedings of the Workshop on Issues of Sentiment Discovery and Opinion Mining at 22nd Annual ACM Conference on Knowledge Discovery and Data Mining, 2016.*

K Haribabu, **A.Paul**, and C. Hota "**Detecting Sybils in Peer-to-Peer Overlays using Psychometric Analysis Methods,**", *Proceedings of the 25th IEEE International Conference on Advanced Information Networking and Applications(AINA), Singapore, March 2011*

Selected Teaching and Leadership

President/Vice-President/Treasurer, **Northwestern Toastmasters** **Sep 2015 -**

- ◇ Lead the Northwestern chapter of Toastmasters; over 30 graduate students, post doctoral fellows from 10 different departments
- ◇ Organized 1.5 hour weekly meetings to improve student public speaking skills
- ◇ Co-wrote proposal to The Graduate school and obtained 3000 USD to fund programming
- ◇ Managed finances, prepared budgets for auditing and reconciled dues

Co-Facilitator, **Northwestern Dialogue Group** **Oct 2016 - September 2017**

- ◇ Facilitated dialogue in safe spaces for cultural exchange across international and domestic students
- ◇ Organized social events to enhance group cohesion

Organizer & Instructor, Machine Learning Workshop, **Northwestern University** **July 2016**

- ◇ Delivered and prepared talk attended by 70 graduate students and professors
- ◇ Designed coding assignments for the participants

Teaching Assistant & Guest Lecturer, **Northwestern University** **Jan 2014- June 2017**

- ◇ Prepared and delivered weekly lectures for multiple courses to 20-50 students
- ◇ Supervised course projects and provided subject matter expertise

Fellowships

Predictive Science and Engineering Design Fellowship	**2016-2017**
Segal Design Fellowship	**2014-2015**
Walter P. Murphy Fellowship	**2012-2013**