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

SUMMARY **Computer Scientist** working in **Machine Learning** with 8+ years research experience.

INTERESTS Machine Learning, Deep Learning, Natural Language Processing

EDUCATION

Northwestern University, Evanston, Illinois USA
Ph.D., Computer Science **Sep '19**
Master of Science, Computer Science

Birla Institute of Technology & Science, Pilani, Rajasthan India
Master of Engineering, Software Systems **May '12**
Bachelor of Engineering, Chemical Engineering

PROGRAMMING SKILLS

Proficient: Python, Keras, Scikit-Learn, XGBoost, LightGBM, Shapley, NLTK, Gensim, Tensorflow, Pandas, Selenium
Familiar: OpenCV, Spacy, PySpark, PyTorch, Theano, R, MATLAB, C, C++, Java, LAMP, SQL, weka, HTML/CSS

PROFESSIONAL EXPERIENCE

Machine Learning Scientist, **American Family Insurance**, Madison, Wisconsin **Oct'19-**

- Developing deep learning models for predicting motor vehicle violation
- Designing insurance-based language models for predicting claims fraud

Data Science Intern, **Northwestern Mutual**, Milwaukee, Wisconsin **Jun - Aug '18**

- Developed distributed OCR algorithms for detecting responses from scanned questionnaires
- Designed a noise reduction algorithm to de-noise scanned and photocopied questionnaires

Research Intern, **Boeing Cybersecurity**, Sunnyvale, California **Jun - Sep '13**

- 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

RESEARCH PROJECTS

Research Assistant, **Northwestern University**, Evanston, Illinois **Sep '12** -

- Chemical Property Prediction using Molecular Fingerprints (Tensorflow, Scikit Learn)
  - ◇ 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 (CheMixNet architecture)
  - ◇ Designed Bagged Ensemble models for predicting power conversion efficiency of solar cells using chemical fingerprints, and achieved mean absolute percentage error between 1.5-2 %
  - ◇ Developed a transfer learning solution to predict solar cell properties with mean absolute percentage error below 1 % (SINet architecture)

- Developed Predictive Model for Additive Manufacturing (Tensorflow, Keras)
  - ◇ Created time series models for temporal analysis of temperature and heat flux data
  - ◇ Investigated Recurrent Neural Network models to predict point-wise temperature information for accelerating additive manufacturing simulations
  - ◇ Developed an iterative real-time predictive model using bagged decision trees

- Ensemble Learning-based Guided Optimization for Aircraft Design (MATLAB, Python)
  - ◇ 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

- Classification of Anonymous Posts using Recurrent Neural Networks (Keras, Scikit Learn)
  - ◇ Developed customized vector model using crowd-sourced (Urban Dictionary) & psycho-lingual (LIWC) dictionaries
  - ◇ Explored Word2vec, GloVe and FastText embedding schemes (Gensim)
  - ◇ Attained prediction accuracy of 79.8 % and 78.1 % using ensemble and LSTM models respectively

Select Publications (13 of 19)

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*

Z.Yang, D. Jha, **A. Paul**, W. Liao, A. Choudhary and A. Agrawal. "**Generative adversarial networks with mixture density networks for inverse modeling in materials microstructural design**", *20th SIAM International Conference on Data Mining (SDM) (under review)*

A. Paul, M.Mozaffar, Z. Yang, W. Liao, A. Choudhary, J.Cao and A. Agrawal."**A real-time iterative approach for temperature profile prediction in additive manufacturing processes**", *6th IEEE International Conference on Data Science and Advanced Analytics (DSAA)*

A. Paul, D.Jha, R. Al-Bahrani, W. Liao, A. Choudhary and A. Agrawal."**Transfer Learning Using Ensemble Neural Nets for Organic Solar Cell Screening**", *International Joint Conference on Neural Networks, 2019*

A. Paul, A. Furmanchuk, W. Liao, A. Choudhary and A. Agrawal. "**Property Prediction of Organic Donor Molecules for Photovoltaic Applications using Extremely Randomized Trees**", *Journal of Molecular Informatics, 2019*

A. Paul, W. Liao, A. Choudhary and A. Agrawal. "**Mining Anonymous Taboo Confessions using Psycho-lingual and Crowd-Sourced Dictionaries for Emotional Well-being**", *Journal of Health Informatics Research (under review)*

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, 2019*

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*

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 Sentiment Mining at 22nd Annual ACM Conference on Knowledge Discovery and Data Mining, 2016.*

Fellowships

McCormick Dean's Commendation Fellowship	**'18 Spring**
Predictive Science and Engineering Design Fellowship	**'16-'17**
Segal Design Fellowship	**'14-'15**
Walter P. Murphy Fellowship	**'12-'13**

Selected Teaching and Leadership

Teaching Assistant & Guest Lecturer, **Northwestern University** **Jan '14-**

- ◇ Prepared and delivered weekly lectures for 20-50 students
- ◇ Courses: Social Media Mining, Data Structures , Introduction to Programming (Python)

President/Vice-President/Treasurer, **Northwestern Toastmasters** **Sep '15 -May '18**

- ◇ Lead the Northwestern chapter of Toastmasters with over 30 graduate students, post doctoral fellows from 10 different departments
- ◇ Co-wrote proposal to The Graduate school and obtained 3000 USD to fund programming

Co-Facilitator, **Northwestern Dialogue Group** **Oct '16 - Sep '17**

- ◇ 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** **Jul '16**

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