Arindam Paul

(440)622-1087| arindampaul.bits@gmail.com| www.arindampaul.me|linkedin.com/in/arndmpaul/

SUMMARY   **Computer Scientist** working in **Machine Learning** with 10+ years of 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

Experienced: Python, Tensorflow, H2O, Scikit-Learn, Dash/Plotly, Matplotlib/Seaborn, Pandas, XG-Boost, Selenium
Familiar: SQL, Shapley, NLTK, HuggingFace, Gensim, Spacy, CatBoost/LightGBM, PySpark, PyTorch/FastAI, HTML/CSS

PROFESSIONAL EXPERIENCE

Data & Applied Scientist, **American Family Insurance**, Greater Boston, Massachusetts **Oct'19-**

- Automatic Speech Analysis for Customer Calls :
  ◇ Harnessed and tuned Whisper.AI and Google's Chirp models for AI-driven transcription of customer calls
  ◇ Developed end-to-end system to process call transcripts into call summaries and extracted insurance-specific entities using PaLM, GPT-3 and GPT-4 models
- User-Based Insurance (in collaboration with major US automaker) :
  ◇ Developed generalized linear and additive models for usage-based auto insurance based on telematics features
- Claims-Channeling System:
  ◇ Co-Designed a multi-input, multi-label claims channeling system to route claims to relevant domain experts using the information (tabular + text) present in the claim which harnesses an insurance based language model using transfer learning to process the text data and thereby increase the accuracy of various downstream tasks
  ◇ Performed an ablation study based on different models, input/output type and day information to select the best models which get feed into a web based user interface
- Financial Forecasting
  ◇ Developed long and mid-term financial forecasting of KPIs using an ensemble ESRNN+SARIMA
- Motor Vehicle Violation:
  ◇ Developed an ML decision system for predicting motor vehicle violation risk
  ◇ Explored ordinal models using tree and neural networks including creating a custom ordinal loss function
- Leadership/Outreach:
  ◇ Host of the AmFam Fairness & Ethics in ML Seminar
  ◇ Collaborate with UW-Madison professors as part of Amfam Data Science Institute
  ◇ Mentored rotational associate data scientists

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

- Developed distributed OCR algorithms for detecting responses from scanned questionnaires

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

- Developed machine learning system using generated synthetic user profiles with different demographic and interest features for analyzing ads across profiles

RESEARCH PROJECTS

Research Assistant, **Northwestern University**, Evanston, Illinois **Sep '12 - Aug'19**

- 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

- 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 (12 OF 25)

Fa Li, Qing Zhu, Kunxiaojia Yuan, Fujiang Ji, **Arindam Paul**, Peng Lee, Volker C. Radeloff, Min Chen. "**Projecting large fires in the western US with a more trustworthy machine learning method**", *Earth's Future, 2024*

Y. Mao, M. Hasan, **A. Paul**, V. Gupta, K. Choudhary, F.M. Tavazza, W. Liao, A. Choudhary, P. Acar and A. Agrawal. "**An End-to-End AI-Driven Microstructure Optimization Framework for Elastic Properties of Titanium Beyond Cubic Crystal Systems**", *Nature Partner Journal on Computational Materials, 2023*

A.Dimri, **A.Paul**, D.Girish, P.Lee, S.Afra and A.Jakubowski. "**A Multi-input Multi-label Claims Channeling System Using Insurance-Based Language Models**", *Expert Systems With Applications, 2022*

Y.Mao, Z.Yang, D.Jha, **A. Paul**, W. Liao, A. Choudhary and A. Agrawal. "**Generative Adversarial Networks and Mixture Density Networks based Inverse Modeling for Microstructural Materials Design**", *Integrating Materials and Manufacturing Innovation Journal, 2022*

K.Ness, **A. Paul**, L. Sun and Z. Zhang. "**Towards a generic physics-based machine learning model for geometry invariant thermal history prediction in additive manufacturing**", *Journal of Materials Processing Technology, 2022 (Special Issue on AI in Advanced Manufacturing)*

R.Richards, and **A. Paul**. "**An Attention-driven LSTM Network for High Throughput Virtual Screening of Organic Photovoltaic Candidate Molecules**", *Solar Energy, 2021*

Z.Yang, D. Jha, **A. Paul**, W. Liao, A. Choudhary and A. Agrawal. "**A General Framework Combining Generative Adversarial Networks and Mixture Density Networks for Inverse Modeling in Microstructural Materials Design**", *NeurIPS Workshop on Machine Learning for Engineering Modeling, Simulation and Design, 2020*

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*

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), 2018*

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, D.Jha, R. Al-Bahrani, W. Liao, A. Choudhary and A. Agrawal."**CheMixNet: Mixed DNN Architectures for Predicting Chemical Properties using Multiple Molecular Representations**", *NeurIPS 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*

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.*

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- Jun'19**

- ◇ 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 and researchers
- ◇ Co-wrote proposal to The Graduate school and obtained 3000 USD to fund programming