

# Jason (Zhihang) Dong

LOOKING FOR MACHINE LEARNING ENGINEERING/DATA SCIENCE/ RESEARCH SCIENTIST ROLE

☎ (814-548-6383) | ✉ zdong@uw.edu | 🏠 www.zdong.org | 📱 zdong1 | 🌐 zhihangdong

## Summary

---

**Research Areas:** Information Retrieval (Large-scale network, graph and time-series data), Natural Language Processing (Natural language generation, video question answering, zero-shot learning) Multimodal Machine Learning (Music Auto-composition, genre identification, and fake news detection)

**Software Engineering Experience:** Near-daily use of Tensorflow and frequent user of PyTorch. Experience on ML product pipelines (recommendation, fraud detection, forecasting and relevance/ranking). Industry experience in using Flink, Airflow, NodeJS and Neo4j

## Skills

---

NOTES: Skills with **bold** mean work/research-level experience & ranked with familiarity

**Programming Languages** Python, R, C++, Scala, Java, Javascript, Typescript

**Machine Learning/CV** Spark, Torch, xgBoost, **café**, Tensorflow, **OpenCV**, CTNK, EMR, Keras

**Databases** Neo4j, Cassandra, MongoDB, postgres, RDS, MySQL, Hive, CouchDB, Lucene

**NLP** Transformer (XLM-ROBERTa), BERT, Tika, NTLK, SpaCy, TextBlob, gensim

**Engineering** Airflow, Hadoop, H2O, ZooKeeper, NodeJS, Bash, ArcGIS, Flink, Kafka, **tableau**, AWS, Git, GeoDa, SAS

## Experience

---

### JP Morgan Chase

*New York, NY*

SENIOR MACHINE LEARNING ASSOCIATE, AI/ML GROUP

*Aug. 2020 - Present*

- Transaction flow forecasting and storytelling, revamping the risk evaluation models in production;
- Anomaly detection of large client-end financial data, scaling the time-series prediction model to accounts of various sizes

### Microsoft Corporation

*Redmond, WA*

APPLIED SCIENTIST, INTERN

*Jun. 2019 - Sep. 2019*

- Developed link prediction models for complex user communication networks using auxiliary network and deep generative methods to boost user engagement with new product features;
- Recommendation system on next user event and user behavior forecasting, increase accuracy by 2X% (anonymous for privacy issues)

### Amazon.com, Inc.

*Seattle, WA*

APPLIED SCIENTIST, INTERN

*Aug. 2018 - Dec. 2018*

- Developed a multi-level representation learning method on sentiment evaluation tasks with conversation texts and applied the model on Customer Feedback Evaluation data, f-1 score increased by 1X%
- Analyzed large-scale merchandise data and designed fraud detection models based on merchant feedback data, improved accuracy by 1.X%

### Deloitte Services LP.

*Seattle, WA*

DATA SCIENTIST, TEMPORARY

*Jun. 2018 - Aug. 2018*

- Designed end-to-end resume recommendation System, Taught a social network analysis workshop, and developed a causal inference project on human resource success metrics (beat benchmark by 1X%)

## Project Experience

---

### Fenglong Ma Lab

*University Park, PA*

RESEARCH INTERN

*Jun. 2020 - Present*

- Working with Prof. Ma on an individual project focusing on multi-modal learning models for adversarial generation/detection of fake news with multiple media (e.g. text and images)

### Fred Hutchinson Cancer Research Center

*Seattle, WA*

RESEARCH ASSISTANT

*Sep. 2019 - March 2020*

- Image Denoising and Spatial Superresolution Project on Single-Cell Data

## Center for Statistics and Social Sciences

RESEARCH ASSISTANT

- Novel Methods to validate GPS Data Information Coverage (Improves Efficiency by 3X%)

Seattle, WA

Sep. 2018 - Dec. 2018

## Center for Studies in Demography and Ecology

RESEARCH ASSISTANT

- ArcGIS Project on Mapping Public Health Data

Seattle, WA

Jun. 2017 - Aug. 2017

## Department of Statistics, UW

CONSULTANT

- Working with clients from different departments leading to the right model choice for their data-science related questions
- Receive media coverage of one applied work with public health faculty that led to a paper published to PLOS One

Seattle, WA

## Music Learning Lab

LAB LEADER

- Working with several undergraduate and graduate students on music auto-composition and genre identification projects
- More than 25 months of project management and team working experience in a research collaboration setting

Seattle, WA

## Education

---

### University of Washington

M.S. STATISTICS/ M.A. SOCIOLOGY

- Clarence and Elissa M. Schrag Endowed Fellowship (only 2 per class)

Seattle, WA

Aug. 2016 - Jun. 2020

### Penn State University

B.S./B.A. QUINTUPLE MAJORS AND TRIPLE MINORS IN STATISTICS, SOCIAL SCIENCE (MULTIPLE MAJORS & MINORS), PHILOSOPHY AND GEOGRAPHY

State College, PA

Aug. 2012 - May. 2016

## Courses

---

2019-2020 **Fairness in ML**, Data Science for Human Well-Being, Reinforcement Learning

2019 **Online and Adaptive Learning, Optimization**, Computer Vision, Alg. via Geometric Lens

2018-2019 **Machine Learning and Big Data**, Advanced Database Management, Advanced NLP Methods

2017-2018 **Statistical Learning**, Design & Analysis of Algorithms, Representation Learning, Bayesian Statistics

## Publications

---

[J] JOURNAL [P] CONFERENCE PROCEEDINGS [W] WORKING PAPER [S] UNDER REVIEW

(S8) "A Multi-Modal Learning Framework for Classical Music Auto-completion and Genre Identification" Manuscript Submitted to *Thirty-fourth Conference on Neural Information Processing Systems (NeurIPS 2020)*.

(J7) Nora Kenworthy, Zhihang Dong, A. Montgomery, E. Fuller, L. Berliner (2020) "A cross-sectional study of social inequities in medical crowdfunding campaigns in the United States" Accepted to: *PLoS ONE* [Preprint].

(W6) Dong, Zhihang, Birjal, A., Yan, X. "Solving Complex Network Link Prediction using Auxiliary Graphs" (Microsoft Internship Project)

(S5) Zhihang Dong "Aspect-Aware Conversation Sentiment Models" Manuscript Submitted to *The 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP 2020)*.

(P4) T. Wu, Zhihang Dong, S. Song and M. Zhang (2020) "Interactive Attention Model Explorer for NLP Tasks with Unbalanced Data Sizes" Accepted: *The 13th IEEE Pacific Visualization Symposium (PacificVis 2020)* [Full Paper].

(J3) Zhihang Dong, A. Dobra and Y. Chen (2019) "A statistical framework for measuring the temporal stability of human mobility patterns". Accepted to: *Journal of Applied Statistics* [Full Paper].

(P2) — (2017). "Estimation and Extrapolation of Spatial Trends in Mortality Data using Bayesian APC Modeling". Accepted to *International Conference on Population Geography*.

(P1) — (2016). "Theorizing Urban Neighborhoods: Mapping the Interneighborhood and Intra-neighborhood Networks and Criminogenic Factors on Street Crime Victimization". Accepted to *American Society of Criminology Annual Meeting*.