

# Jason (Zhihang) Dong

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

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

## Summary

**Research Areas:** <1> Information Retrieval (Large-scale network, graph and time-series data) <2> Natural Language Processing (Natural language generation, Contextual NER and Sentiment Classification) <3> Applied Reinforcement Learning <4> Multimodal Machine Learning (Music Autocomposition, genre identification, and fake news detection);

**Software Engineering Experience:** (1) Business-Curated Interpretative Machine Learning Model (2) Deep Learning Optimization (Compression, Pruning, Scaling, Distillation) (3) End-to-End Pipelines for Recommendation, Ranking and Relevance with open-source frameworks (working experience with Airflow, Flink and Kafka and Zookeeper)

## Skills

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

<b>Programming Languages</b>	<b>Python, R, C++</b> , Scala, Java, Javascript, Typescript
<b>Machine Learning/CV</b>	<b>Spark, Torch, RLib, xgBoost, caffié, Tensorflow, OpenCV</b> , CTNK, EMR, Keras
<b>Databases</b>	<b>Neo4j, Cassandra, MongoDB</b> , postgre, RDS, mySQL, Hive, CouchDB, Lucene
<b>NLP</b>	<b>Transformer (XLM-ROBERTa), BERT, Tika, NTLK, SpaCy, TextBlob, gensim</b>
<b>Engineering</b>	<b>Airflow, Hadoop, H2O, ZooKeeper, Bash, ArcGIS, Flink, Kafka, tableau, AWS, Git</b>

## Work Experience

### JP Morgan Chase

*New York, NY*

SENIOR MACHINE LEARNING SCIENTIST (ASSOCIATE SR.), AI/ML GROUP

*Aug. 2020 - Present*

- Refined and Scaled the Transaction Flow Forecasting and Storytelling and the attrition risk models using deep ranking methods (4.\* times faster); Delivered the first owned project into production within **4 months** of joining the new team;
- Anomaly detection of large client-end financial data and Hosted the Sequence Model Reading Group on new research and engineering developments on time series data

### Microsoft Corporation

*Redmond, WA*

APPLIED SCIENTIST, INTERN

*Jun. 2019 - Sep. 2019*

- Research project on link prediction models for complex user communication networks using auxiliary network and deep graph network methods to boost user engagement with new product features;
- Built recommendation system on next user behavior forecasting with Transformer and produced visualization interface, increased accuracy by 2X%

### Amazon.com, Inc.

*Seattle, WA*

APPLIED SCIENTIST, INTERN

*Aug. 2018 - Dec. 2018*

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

### 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

During my Ph.D. years and early-career, I have also participated in a wide range of research lab activities that does not (or yet) yield a formal publication but a great grasp of the domain knowledge and research directives.

### ◇ Lab Rotations ◇

#### *Health Big Data*

- Yu Li Lab, Hong Kong SAR (Deep Sequence Model and Generative Model on Antimicrobial Peptides) (11/2020 - Now)
- Fred Hutchinson Cancer Research Center, Seattle, WA (Spatial Superresolution and Image Denoising) (09/2019 - 01/2020)

## Multimodal Learning

- Fenglong Ma Lab, Penn State University (Multi-Modal Learning on Fake News Detection (07/2020 - Now))

## Human-Trace Big Data

- Center for Statistics and Social Sciences, Univ of Wash (Spatial-Temporal Models on Human GPS-Trace Data)
- Center for Studies in Demography and Ecology, Univ of Wash. (Mapping and GIS)

## ◊ Service ◊

- Statistic Consultant, Department of Statistics, University of Washington (Spring 2019)
- The Landing Lab, an initiative to help underrepresented STEM Students Starting their First Data Science Project (2016 - 2018)

## Education

---

### University of Washington

M.S. STATISTICS (QUITTED PH.D. PROGRAM AFTER PASSING THESIS EXAM) / M.A. SOCIOLOGY

Seattle, WA

Aug. 2016 - Jun. 2020

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

### 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 *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2021)*.
- (S7) Zhihang Dong "Aspect-Aware Conversation Sentiment Models" Manuscript Submitted to *2021 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2021)*.
- (J6) 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].
- (W5) Dong, Zhihang, Birjal, A., Yan, X. "Solving Complex Network Link Prediction Problems using Auxiliary Graphs" (Microsoft Internship Project)
- (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*.