

Jason (Zhihang) Dong

LOOKING FOR {APPLIED/RESEARCH/DATA/MACHINE LEARNING} {SCIENTIST/ENGINEER} POSITIONS

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

Summary

Research Areas: Natural Language Processing (Natural language generation and natural language understanding, NLP + Computer Vision (Image captioning & VQA)), Data Mining (Information Retrieval and Recommender System), Multimodal Machine Learning, AI + Music (Music Info Retrieval, Music Autocompletion, Music Autocomposition)

Industry Experience: 1.5 years (Internship included); Broad experience in prototyping Machine Learning and Deep Learning Methods to real, large and noisy data and communicating with non-technical audience

Skills

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

Programming Languages Python, R, Java, Scala, C++

Machine Learning Spark, Torch, xgBoost, **caffé**, Tensorflow, H2O, Hadoop, CTNK, EMR, Keras

Databases Neo4j, Cassandra, MongoDB, **postgre**, RDS, **mySQL**, **Hive**, CouchDB, Lucene

NLP Transformer, BERT, Tika, NTLK, SpaCy, TextBlob, **gensim**

Software & Misc. ArcGIS, Stata, **tableau**, GeoDa, SAS, Qt, AWS, Git

Experience

Microsoft Corporation

Redmond, WA

APPLIED SCIENTIST, INTERN

Jun. 2019 - Sep. 2019

- Complex Network Link Predictions on Dynamic Network User Communication Pattern using Auxiliary network;
- 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

- Multi-view Representation Learning on Conversational Sentiment Evaluation with Application on Customer Feedback Evaluation, f-1 score increased by 1X%
- Fraud Detection Models based on Merchant Feedback data

Deloitte Services LP.

Seattle, WA

SENIOR DATA SCIENTIST, TEMPORARY

Jun. 2018 - Aug. 2018

- End-to-End Resume Recommendation System (full-stack)
- Causal Inference Project on Human Resource Success (beat benchmark by 1X%)

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

Seattle, WA

RESEARCH ASSISTANT

Sep. 2018 - Dec. 2018

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

Center for Studies in Demography and Ecology

Seattle, WA

RESEARCH ASSISTANT

Jun. 2017 - Aug. 2017

- ArcGIS Project on Mapping Public Health Data

Project Experience

NLU & NLG

RESEARCH PROJECT

- NLU on Medical Crowdsourcing [J10]
- Contextual sentiment modeling in conversational setting [S5]
- Interpretable Data Science (Recommendation on Analytics) [S6]

Music@Work

RESEARCH PROJECT

- Classical music composer identification using multimodal ML [W11]
- Solo-piano music completion using Transformer

Graph Neural Networks on Large-Scale Network Data

RESEARCH PROJECT

- Link predictions using auxiliary network [W9] and community detections
- Causal inference on small data with perturbations [W7]

Recommender System and Ranking Methods

RESEARCH PROJECT

- Gathering Anchor Locations of Human Activities [J3]
- Recommender System using Density Ranking [W8]
- Spatial Trends of Demographic Data [P2]

Education

University of Washington

Seattle, WA

M.S. STATISTICS/ M.A. SOCIOLOGY

Aug. 2016 - PRESENT

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

Penn State University

State College, PA

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

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

- (W11) "Music At Work: A Multi-Modal Deep Learning Framework for Classical Music Auto-completion and Genre Identification" *Manuscript in Preparation*.
- (J10) Nora Kenworthy, Zhihang Dong, A. Montgomery, E. Fuller, L. Berliner "A cross-sectional study of social inequities in medical crowdfunding campaigns in the United States" [Preprint] *PLoS ONE*.
- (W9) Dong, Zhihang, Birjal, A., Yan, X. "Network Link Prediction using Auxiliary Network"
- (W8) Dong, Zhihang (2020). "Document Classification based on Aggregated Density Ranks".
- (W7) Zhihang Dong, W. Chen and X. Ge "Causal Inference on Bike Rentals using Anchor Regression"
- (S6) Zhihang Dong, S. Song and A. Wang "Toward a More Interpretable and Reusable Data Science: Analyses of Million Jupyter Notebooks" Submitted to *29TH ACM INT'L CONFERENCE ON INFORMATION AND KNOWLEDGE MANAGEMENT (CIKM 2020)*.
- (S5) Zhihang Dong "Aspect-Aware Conversation Sentiment Models" Submitted to *ACM Conferences on Knowledge Discovery and Data Mining (KDD 2020)*.
- (P4) T. Wu, Zhihang Dong, S. Song and M. Zhang "Interactive Attention Model Explorer for NLP Tasks with Unbalanced Data Sizes" Accepted: *The 13th IEEE Pacific Visualization Symposium (PacificVis 2020)* [Sample].
- (J3) Zhihang Dong, A. Dobra and Y. Chen "A statistical framework for measuring the temporal stability of human mobility patterns". [Preprint] *Journal of Applied Statistics*.
- (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*.