Yong Zhuang

+1 (617)-763-8919 | yong.zhuang001@umb.edu | yong-zhuang-b1847a114 | yong-zhuang.github.io

SUMMARY

Graduated Ph.D. in Computer Science with a focus in applied machine learning. Highly adept at feature selection, deep learning, and time series forecasting, but has a breadth of knowledge of data science techniques from work across various problem domains in industry and academia over the last seven years. During NSF Intern at Radial Analytics, Applied NLP and causal-based feature selection on clinical data to choose the most valuable features. Created deep neural networks to identify patient candidates across health levels with an average precision improvement of 17%. Have five years of software development experience leading a team of five before the Ph.D. study.

SKILLS

Languages : Python, Java, ASP.NET, C#, PHP, JavaScript, HTML/CSS, SQL
Machine Learning Packages : Tensorflow, Keras, Pytorch, Matplotlib, Pandas, Numpy, scikit-learn, seaborn
Database : MySQL, Microsoft SQL Server, Access, Oracle
ArcGIS : Map, Server, SDE and Desktop
Developer tools : Git, Docker, VS Code, Visual Studio, Eclipse
Advanced Skills : Design Patterns, MVC, Jquery, AJAX

EXPERIENCE

Senior Machine Learning Engineer August. 2022 - Present
Constant Contact Waltham, MA, U.S.

Ph.D. Candidate, Research Assistant Dec. 2014 - Dec. 2021
Knowledge Discovery Lab Boston, MA, U.S.
Worked as a Ph.D. Candidate / Researcher in applied machine learning at the Knowledge Discovery Lab in UMBs Computer Science department, focusing on modeling real-world datasets using feature selection and deep neural networks in Python. Major work including:

- Created multi-Markov-blankets-based methods to identify precursors to heavy precipitation event clusters. Designed multi-Markov-blankets-based methods to identify precursors to heavy precipitation event clusters. Our method identified the cold surges along the coast of Asia as an essential precursor to the surface weather over the United States, a finding which climate experts later corroborated.
- Created a new recurrent neural network architecture (Error Trajectory Tracing) to track the trajectory of the prediction error through the phase space, and a new training regime (Horizon Forcing) to expand the prediction horizon of time series by more than 20%.
- Created Spatio-temporal deep neural networks improving crime hot Spot forecast F1-score by 21%.
- Supervised undergraduate lab assistants and mentored junior Ph.D. students in data science methodology and programming best practice.

National Science Foundation Intern Jun. 2019 - Nov. 2019
Radial Analytics Concord, MA, U.S.

- Applied natural language processing(NLP) and causal-based feature selection to clinical data to select the most valuable features related to patients' health status.
- Created deep neural networks to identify patient candidates across health levels with an average precision improvement of 17%.
- Designed an auto-machine-learning platform that automatically generates the best predictive model for given needs using feature selection and model selection.

Software Engineer Mar. 2008 - Mar. 2013

Triexcel Co., Ltd. (2008 - 2011) & Huiyin Science and Technology (2011 - 2013) Anshan, Liaoning, China

- Led front-end and back-end development of a GIS-based geological hazard management platform for data collection, data synchronism, risk scoring, and investigation planning.
- Led front-end and back-end development of an after-sales service management platform to rationalize the after-sale service process and Increase efficiency and customer satisfaction.
- Spearheaded product development of a remote solution to provide secure web remote control and system update services for terminals of Bank of Anshan.

PUBLICATIONS

Zhuang, Y., Almeida, M., Ding, W., Flynn P., Islam, S., and Chen P., Widening the Time Horizon: Predicting the Long-Term Behavior of Chaotic Systems with Error-Trajectory Tracing and Horizon Forcing, Submitted to: The IEEE International Conference on Data Mining (ICDM)

Almeida, M., Zhuang, Y., Ding, W., Crouter, S. E., and Chen, P. (2021). Mitigating class-boundary label uncertainty to reduce both model bias and variance. ACM Transactions on Knowledge Discovery from Data (TKDD), 15(2), 1-18.

Zhuang, Y., Small, D. L., Shu, X., Yu, K., Islam, S., and Ding, W. (2018, November). Galaxy: Towards Scalable and Interpretable Explanation on High-Dimensional and Spatio-Temporal Correlated Climate Data. In 2018 IEEE International Conference on Big Knowledge (ICBK) (pp. 146-153). IEEE.

Zhuang, Y., Almeida, M., Morabito, M., and Ding, W. (2017, August). Crime hot spot forecasting: A recurrent model with spatial and temporal information. In 2017 IEEE International Conference on Big Knowledge (ICBK) (pp. 143-150). IEEE.

Zhuang, Y., Yu, K., Wang, D., and Ding, W. (2016, April). An evaluation of big data analytics in feature selection for long-lead extreme floods forecasting. In 2016 IEEE 13th International Conference on Networking, Sensing, and Control (ICNSC) (pp. 1-6). IEEE.

Zhuang, W. Y., and Ding, W. (2016, September). Long-lead prediction of extreme precipitation cluster via a spatiotemporal convolutional neural network. In Proceedings of the 6th International Workshop on Climate Informatics: CI.

EDUCATION

University of Massachusetts Boston Sep. 2016 - Dec. 2021

Ph.D, Applied Machine Learning, GPA 3.906 Boston, MA, U.S.

University of Massachusetts Boston Sep. 2014 - Jun. 2016

MS, Computer Science, GPA 3.923 Boston, MA, U.S.

Harbin Engineering University Sep. 2001 - Jul. 2005

BE, Computer Science, GPA 3.68 Harbin, Heilongjiang, China

HONORS and AWARDS

Oracle Doctoral Research Fellowship Award Jun. 2016, 2018

The Collage of Science and Mathematics

The Randall Gates Malbone Fellowship May. 2019

The Collage of Science and Mathematics

National Science Foundation (NSF) Graduate Research Internship Jun. 2019

National Science Foundation

PROGRAM COMMITTEE MEMBER

SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)	2022
The SIAM International Conference on Data Mining (SDM)	2022
AAAI Conference on Artificial Intelligenceg (AAAI)	2021, 2022
The ACM International Conference on Information and Knowledge Management (CIKM)	2019