

# Yong Zhuang

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

## SUMMARY

---

Highly accomplished, skilled, and knowledgeable computer science PhD with 8+ years of software development experience and 7+ years applied machine learning experience. Proven success in building machine learning algorithms and predictive models for different industries. Highly adept at feature selection, time series forecasting, Spatio-temporal data analysis, and visualization. Passionate about solving real-world computational problems.

## 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

**Advanced Skills :** Git, Docker, Design Patterns, MVC, JQuery, AJAX

## EXPERIENCE

---

### Research Assistant

Dec. 2014 - Present

*Knowledge Discovery Lab*

Boston, MA, U.S.

- Worked as a Ph.D. Candidate / Researcher in applied machine learning at the knowledge discovery Laboratory of the University of Massachusetts, Boston. Focusing on developing applied machine learning algorithms to solve real-world computational problems.
- Research focuses on feature selection, time series forecasting, and Spatio-temporal data analysis.
- 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.

- Apply Machine Learning technologies in clinical data analysis to help hospital systems and physician networks deliver more efficient care to their patients, and to help patients better receive the necessary care, tailored to their individualized needs.
- Used feature selection, NLP, and deep learning.

### Software Engineer

Mar. 2008 - Mar. 2012

*Liaoning Triexcel Co., Ltd.*

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. [Demo](#)
- 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.

## RESEARCH PROJECTS

---

### Deep Symbolic Regression

- Recovering mathematical expressions from data via deep learning.

Long-term prediction of chaotic systems via machine learning

- Design, implement, train and evaluate deep sequence models for long-term prediction of chaotic systems.

### Interpretable Spatio-Temporal Modeling

- Identify the main sub-feature sets with predictive ability and interpretability from the original feature space via causal analysis.

### Crime Hotspot Forecasting

- Crime Hot Spot Forecasting via Spatio-temporal Deep Networks.

### Long-lead Heavy Precipitation Forecasting

- Using state-of-the-art online streaming feature selection methods to identify the strong relevant meteorological variables for heavy precipitation prediction, and using classification algorithm to forecast extreme floods.

## 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: IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) [Under Revision]

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

---

<b>University of Massachusetts Boston</b>	Sep. 2016 - PRESENT
<i>Ph.D, Applied Machine Learning, GPA 3.906</i>	Boston, MA, U.S.

<b>University of Massachusetts Boston</b>	Sep. 2014 - Jun. 2016
<i>MS, Computer Science, GPA 3.923</i>	Boston, MA, U.S.

<b>Harbin Engineering University</b>	Sep. 2001 - Jul. 2005
<i>BE, Computer Science, GPA 3.68</i>	Harbin, Heilongjiang, China

## HONORS and AWARDS

---

<b>Oracle Doctoral Research Fellowship Award</b>	Jun. 2016, 2018
<i>The Collage of Science and Mathematics</i>	

<b>The Randall Gates Malbone Fellowship</b>	Jun. 2016, 2018
<i>The Collage of Science and Mathematics</i>	

<b>National Science Foundation (NSF) Graduate Research Internship</b>	Jun. 2019
<i>National Science Foundation</i>	