Yong Zhuang

Contact Information	yong.zhuang001@umb.edu 617-763-8919 https://yong-zhuang.github.io/

Research Interests

Artificial Intelligence, Machine Learning, Big Data Analysis, Feature Selection on Big Data, Spatio-temporal Data Analysis, Time Series Prediction

Education

University of Massachusetts Boston, Boston, MA

Ph.D., Computer Science, GPA: 3.906 — Sep. 2016 - Dec. 2021

M.S., Computer Science, GPA: 3.923 — Sep. 2014 - Aug. 2016

Harbin Engineering University, Harbin, China

B.E., Computer Science, GPA: 3.68 — Sep. 2001 - Jul. 2005

Highlighted Experience

Professional Services

- National Science Foundation(NSF) panel (2022).
- Program Committee Member in KDD, SDM, AAAI, CIKM.

Teaching Experience

- 4 semester as the instructor to counsel disadvantaged students to succeed in their compulsory courses such as Theory of Computation, Data Structures and Algorithms.
- 5 semesters as the teaching assistant in AI, Big Data Analysis, and Introduction to Computing
- 1 online tutoring material on time series forecasting

Industry Experience

- 5 years full-time working experience
- 2 semesters NSF Graduate Research Internship during Ph.D. study

Award

- Oracle Doctoral Research Fellowship in fall 2016, fall 2018
- The Randall Gates Malbone Fellowship in May. 2019

Selected Research Experiences

Feature Selection on Big Spatio-temporal Data

Given a large-scale Spatio-temporal database, effectively and efficiently identifying strongly related features and removing the irrelevant or less important features with respect to a target variable is a critical and challenging issue in many fields. In this work, I collaborated with scientists in climate science and water diplomacy at Tufts University to design a multi-Markov-blankets-based ensemble model "Galaxy" to identify precursors to heavy precipitation event clusters. "Galaxy" identified the cold surges along the coast of Asia as an essential precursor to the surface weather over the United States, which was confirmed by climate experts.

Deep Learning on Big Spatio-temporal Data

As the number, volume, and resolution of spatio-temporal datasets increase rapidly, Spatio-temporal dependencies of features become highly complex and hard to capture. With its strong hierarchical feature learning capabilities in both the spatial and temporal domains, deep learning has emerged as a promising method to address this challenge. In this study, I collaborated with scientists in the School of Criminology and Justice Studies at UMASS Lowell to integrate Convolutional Neural Networks and Recurrent Neural Networks to capture latent spatio-temporal features for predicting crime hotspots, resulting in a 21% improvement in F1-score.

Predicting the Long-Term Behavior of Chaotic Systems
Chaotic behavior is present in many nonlinear dynamical systems, including climate dynamics, weather prediction, and the Spatio-temporal dynamics of virus spread. To provide a reliable solution for these systems, it is necessary to handle their complex Spatio-temporal dynamics and their sensitive dependence on initial conditions. In this study, I collaborated with a research team at Tufts University to propose the Lyapunov Horizon loss, which measures how the error divergence of a forecasting sequence evolves in a chaotic system. We optimized this loss function using a new approach called "Horizon Forcing" on a recurrent "tower" architecture known as "Error Trajectory Tracing." This approach improves the predictive range of sequences in chaotic systems by over 20%.

Teaching Experience

Instructor
CS Seminar — Fall 2017, Spring, Summer, and Fall 2018

- Deliver a 60-minute presentation weekly, covering the materials in the computer science courses: Intermediate Computing with Data Structures (CS210), Advanced Data Structures and Algorithms (CS310), and Introduction to the Theory of Computation (CS420).
- Provide practical guidance and answer specific questions.
- Previewing upcoming course content, assignments and predicting future challenges.

Mentor
CS187: Science Gateway Seminar — Spring 2017

- Design a semester-long in-class project to expose undergraduate students to research methods.
- Closely advised students on this project, after which they presented their research posters at UMASS Boston's symposium.

Teaching Assistant
CS110: Introduction to Computing — Fall 2021
CS670: Artificial intelligence — Fall 2017, Spring 2019
CS697: Big data analysis — Fall 2016, Spring 2018

- Design semester-long in-class projects.
- Closely advised students on these projects.
- Grade students' homework and projects.
- Provide several lectures for each course.
- Hold Q&A sessions through office hours twice a week

Grant Writing Experience

Have experience writing two federal agency grants projects with my advisor,
Total Amount: $355,581

- Project 1 is about physics-guided deep model for flooding forecasting.
- Project 2 is about symbolic expression exploring using deep learning.

Industry Experience

Senior Machine Learning Engineer — Aug. 2022 - Apr. 2023
Constant Contact — *Waltham, MA*
Used Python, and AWS to build a machine learning system to estimate customer lifetime value (CLV) to help marketing teams develop marketing strategies.

- Conducted RFM analysis on eCommerce data to estimate future purchases and the average purchase value of customers..
- Developed and evaluated CLV using machine learning and statistical modeling techniques on RFM analysis results..
- Segment customers using CLV to tailor accurate marketing strategies.

Research Scientist, National Science Foundation Intern Jun. 2019 - Nov. 2019
Radial Analytics *Concord, MA*
Used Python, Tensorflow, and AWS to design a machine learning system to help hospital systems and physician networks provide patients with more effective care to meet their individual needs.

- Developed a machine learning pipeline that selects the most valuable features related to patients' health status using natural language processing (NLP) and causal-based feature selection.
- Built deep neural networks to identify patient candidates in different health levels with an average precision improvement of 17%.
- Determine the best predictive model for the given problem using feature selection and model selection.

Lead Software Engineer Mar. 2008 - Mar. 2013
Liaoning Triexcel Co., Ltd. *Anshan, China*

- Led front- and backend-focused teams to develop a GIS-based Geological Hazard Management System (GHMS) for data collection, synchronism, risk scoring, and investigation planning. Utilized: C# / ASP.Net / Oracle / MS SQLSERVER / LINQ / CSS / jQuery / AJAX / JSON and ArcGIS (Server, SDE, and Desktop).
- Led front- and backend-focused teams to develop an after-sales service management platform that streamlined the after-sale service process, improving efficiency and customer satisfaction. Utilized: C# / ASP.Net/ MVC / jQuery / AJAX and Bing Map.
- Spearheaded the development of a secure remote solution that enabled web-based remote control and system updates for Bank of Anshan terminals. Utilized: C# / ASP.Net / AJAX / MS SQLSERVER / CSS / JavaScript / video conversion and Bing Map.

Selected Peer Reviewed Publications

1. **Yong Zhuang**, Matthew Almeida, Wei Ding, Patrick Flynn, Shafiqul Islam, Zihan Li, and Ping Chen. "Horizon Forcing: Improving the Recurrent Forecasting of Chaotic Systems"
   Transactions on Intelligent Systems and Technology (Under Review)

2. **Yong Zhuang**, Matthew Almeida, Wei Ding, Patrick Flynn, Shafiqul Islam, and Ping Chen. "Widening the Time Horizon: Predicting the Long-Term Behavior of Chaotic Systems."
   The IEEE International Conference on Data Mining (ICDM), Orlando, Florida, Nov. 2022

3. **Yong Zhuang**, David Small, Xin Shu, Kui Yu, Shafiqul Islam, and Wei Ding. "Galaxy: Towards Scalable and Interpretable Explanation on High-dimensional and Spatio-Temporal Correlated Climate Data"
   IEEE International Conference on Big Knowledge (ICBK), Singapore, Nov. 2018

4. **Yong Zhuang**, Matthew Almeida, Melissa Morabito, and Wei Ding. "Crime Hot Spot Forecasting: A Recurrent model with Spatial and Temporal Information"
   IEEE International Conference on Big Knowledge (ICBK), Hefei, China, Aug. 2017

5. **Yong Zhuang**, Kui Yu, Dawei Wang, and Wei Ding. "An evaluation of big data analytics in feature selection for long-lead extreme floods forecasting"
   IEEE International Conference on Networking, Sensing, and Control (ICNSC), Mexico City, Mexico, Apr. 2016

6. **Yong Zhuang**, and Wei Ding. "Long-lead prediction of extreme precipitation cluster via a spatio-temporal convolutional neural network"
International Workshop on Climate Informatics(CI), Boulder, Colorado, Oct. 2016

7. Matthew Almeida, **Yong Zhuang**, Wei Ding, Scott Crouter, and Ping Chen. "Mitigating Class-Boundary Label Uncertainty to Reduce Both Model Bias and Variance"
ACM Transactions on Knowledge Discovery from Data(TKDD) 15.2 (2021): 1-18

Mentoring

Undergraduate and Pre-College Students

- Everest Yang, Pre-college student at UMASS Boston 2022
- Charlotte Yang, Pre-college student at UMASS Boston 2020

Professional Services

Grant Proposal/Fellowship Review

- National Science Foundation(NSF) panel 2022

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, 2023

Reviewer for Journal Manuscript Submissions

- IEEE Transactions on Knowledge and Data Engineering (TKDE), 2017
- ACM Transactions on Knowledge Discovery from Data (TKDD), 2018 - 2019
- Knowledge and Information Systems (KAIS), 2015 - 2017
- Applied Computing and Informatics (ACI), 2018
- International Journal of Information Technology and Decision Making (IJITDM), 2017
- Journal of Ambient Intelligence and Smart Environments (JAISE), 2017
- Spatial Algorithms and Systems (TKDE), 2017

Reviewer for Conference Manuscript Submissions

- International Conference on Machine Learning (ICML), 2015
- Knowledge Discovery and Data Mining (KDD), 2016 - 2021
- American Association for Artificial Intelligence (AAAI), 2016, 2021
- The IEEE International Conference on Data Mining series (ICDM), 2016 - 2018
- ACM International Conference on Information and Knowledge Management (CIKM), 2019 - 2020
- SIAM International Conference on Data Mining (SDM), 2021
- Pacific Asia Conference on Knowledge Discovery and Data Mining (PAKDD), 2016, 2018
- European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD), 2016 - 2017
- The ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL), 2016
- International Conference on Networking, Sensing and Control (ICNSC), 2016
- International Workshop on Climate Informatics (CI), 2016

University Services

The College of Science and Mathematics(CSM) Faculty Senate

Assistant UMASS Boston, Fall 2016

- Design a web application for CSM faculties to manage meeting agendas and minutes, proposals, and documents.

AI Association
Co-organizer UMASS Boston, Fall 2019
- Hold talks, seminars, workshops, and fun activities in AI.
- Web page: https://ai-umb.github.io/

Tech-writing Seminar
Organizer UMASS Boston, Fall 2021
- A seminar where students share good sentences from essays or articles, practice, and improve scientific writing skills.
- Web page: https://yong-zhuang.github.io/tech-writing

Activities

Microsoft's AI for Earth Summit
Member Redmond, WA, Oct 2016

The Eighth Annual "Science Engineering Technology in the CITY"
instructor Boston, MA, Apr 2016
- Give two demonstrations, "Image Printing" (a program that allows computers to copy paintings) and "Style Transfer" (a program that can convert photos into paintings)

Tech Savvy
Core Member & Instructor Boston, MA, Jun 2016, 2017
- Worked with Boston University, Harvard, MIT, etc. to organize a one-week Tech-Savvy camp to prompt STEM education among Boston intermediate school students.
- Organized interactive games and lectures to stimulate interest in machine learning among high school girls.

Technical Skills

- **Language:** Python, MATLAB, R, C#, Java, PHP, JavaScript, HTML, CSS, SQL, ASP.NET
- **Machine Learning Libraries:**, Tensorflow, Keras, Pytorch, Matplotlib, Pandas, Numpy, scikit-learn, seaborn, GPT
- **Database:** Oracle, Access, Microsoft SQL Server, MySQL
- **ArcGIS:** Map, Server, SDE and Desktop
- **Developer Tools:** Git, Docker, AWS, VS Code
- **Advanced Skills** Object-Oriented Programming, Design Patterns, MVC, Jquery, AJAX

Awards

Honors and Awards
- Oracle Doctoral Research Fellowship Award from the Collage of Science and Mathematics at UMass Boston, Jun. 2016, 2018
- The Randall Gates Malbone Fellowship in Mathematics and Computer Science, May 2019
- National Science Foundation (NSF) Graduate Research Internship Program, Jun. 2019
- Microsoft's AI for Earth summit Travel Grant, Redmond, WA, Sep., 2018
- Climate Informatics Workshop Travel Grant, Boulder, Colorado, Aug. 2016