

Yijun Zhao

Associate Professor • Director of the MS in Data Science Program • CIS Department • Fordham University

EDUCATION

- PhD in Computer Science (2017), Tufts University
Research Interests: Machine/Deep Learning, Data Mining and Statistical Pattern Recognition
 - MS in Computer Science, University of Kansas
 - BS in Computer Science, Tianjin University
 - Certificate in Big Data Hadoop
 - Certificate in Apache Spark and Scala
 - Certificate in Quantitative Finance (<http://cqf.com>) – “Wilmott Award” winner
-

PROFESSIONAL EXPERIENCE

- 09/2018 - now, Assistant Professor and Director of the MS in Data Science Program
Computer and Information Sciences (CIS) Department, Fordham University
 - 09/2017 - 9/2018, Visiting Assistant Professor
Computer and Information Sciences Department, Fordham University
 - 04/2006 - 12/2010, Senior Quantitative Trader at Millennium Management, LLC; Ronin Capital, LLC.
 - 03/2004 - 03/2006, Vice President at Wells Fargo Advisors
 - 03/2002 - 04/2004, Assistant Vice President at HSBC Securities, Inc.
 - 08/1997 - 03/2002, Technical Specialist at Citigroup
-

RESEARCH COLLABORATIONS & ONGOING PROJECTS

- Joint work with Harvard Medical School and Brigham and Women’s Hospital
 - Predicting Disease Course of Multiple Sclerosis Patients
 - Predicting Secondary Progressive Multiple Sclerosis (SPMS) Using Retinal Images (in planning)
 - Joint work with NYU Grossman School of Medicine
 - Deep Learning for Detecting Abnormal Lesional Regions in Brain MRI Images
 - Deep Learning for Detecting and Reducing Motion Artifacts in 3D Brain MRI Scans
 - Deep Learning for Predicting Head Pose During Brain MRI Acquisition
 - Joint work with Harvard Medical School and Massachusetts General Hospital
 - Machine Learning for Monitoring and Predicting Lupus Disease Course
 - Machine Learning for Predicting Hydroxychloroquine Retinopathy
 - Educational Data Mining (Joint work with Fordham CIS faculty members)
 - Predicting MSDS Applicants’ Academic Performance Using Application Materials
 - Combat Systematic Racism, Gender Bias, and Cultural Bias in Academic Letters of Recommendation
 - Joint work with Psychology Department, Fordham University
 - Predicting College Students’ Coping Behavior and Stress Levels During the COVID-19 Pandemic
 - Joint work with Chemistry Department, Fordham University
 - Pedagogical Lab for Non-Specialists Students on Interpreting Spectroscopy Using Machine Learning
-

RESEARCH GRANTS & AWARDS

- 2022 - 2023 Fordham Faculty Challenge Grant on "Combating Systematic Racism and Gender Bias by Assessing Bias in Letters of Recommendation.," \$7,000
 - 2022 - 2023 Fordham Faculty Interdisciplinary Grant on "Development of a Machine Learning/Data Science Activity for the Physical Chemistry Laboratory," \$3,000
 - 2022 - 2023 Funding from GSAS Dean on "Address Gender, Racial, and Culture Biases in Academic Letters of Recommendation," \$11,888
 - 2022 - 2023 Funding from GSAS Dean on "Risk Stratification to Minimize Harms from SLE Treatments and Improve Outcomes," \$11,888
 - 2022 GSAS Co-Curricular Funds Award, \$300
 - 2021 - 2022 Fordham Interdisciplinary Research Award on "Psychology Meets Machine Learning: Predicting College Students' Coping Behavior and Stress Levels During the COVID-19 Pandemic," \$6,000
 - 2021 - 2022 Fordham Faculty Fellowship Award on "Predictive Algorithms for Multiple Sclerosis Disability Accrual"
 - 2021 - 2022 NIH K23 on "A Personalized, Patient-Centered Approach to Optimizing Lupus Care," \$19,220
 - 2021 - 2022 Funding from GSAS Dean on "Difference in Psychological, Cognitive, and Behavioral Patterns Between Graduate and Undergraduate Students During the COVID-19 Pandemic," \$15,146
 - 2021 - 2022 Funding from GSAS Dean on "Risk Stratification to Minimize Harms from SLE Treatments," \$15,146
 - 2020 - 2021 Fordham-NYU Research Fellow and Research Internship Award on "Deep Learning for Detecting and Reducing Motion Artifacts in Brain MRI Scans," \$9,000
 - 2020 - 2021 Funding from GSAS Dean on "Pedagogical Lab for Non-Specialists Students on Interpreting Spectroscopy Using Machine Learning," \$12,000
 - 2019 - 2020 Funding from GSAS Dean on "Machine Learning to Monitor and Predict Lupus Disease Course," \$12,000
 - 2019 - 2020 Funding from GSAS Dean on "Deep Learning for Reducing Motion Blurs in MRI Scans," \$12,000
 - 2016 Travel Award, The 22nd KDD Conference
 - 2015 Travel Award, The IEEE International Conference on Data Mining (ICDM)
 - 2013 - 2015 Henry Luce Foundation Fellowship
-

PEER-REVIEWED PUBLICATIONS

Journals

- J17. **Y. Zhao**, Y. Ding, H. Chekired, Y. Wu, and Q. Wang, "Ethnic Differences in Response to COVID-19: A Study of American Asian and Non-Asian College Students via Machine Learning Techniques," *Behavioral Sciences*, 2023. (IF 2.286)
- J16. **Y. Zhao**, D. Smith, and A. Jorge, "Comparing Two Machine Learning Approaches in Predicting Lupus Hospitalization Using Longitudinal Data," *Scientific Reports*, 2022. (IF 4.996)
- J15. **Y. Zhao**, Y. Ding, H. Chekerid, and Y. Wang, "Student Adaptation to College and Coping in Relation to Adjustment During COVID-19: A Machine Learning Approach," *PLOS One*, 2022. (IF 3.24)
- J14. **Y. Zhao**, Z. Qi, S. Do, J. Grossi, J. Kang, and G. Weiss, "Addressing Disparity in GRE-optional Admissions by Predicting GRE Performance Using Application Materials," Under Review.
- J13. **Y. Zhao**, S. Xu, J. Ossowski, "Deep Learning Meets Statistical Arbitrage: An Application of Long Short-Term Memory Networks to Algorithmic Trading," *The Journal of Financial Data Sciences*, 2022.
- J12. **Y. Zhao**, M. Qin, and A. Jorge, "A Calibrated Ensemble Algorithm to Address Data Heterogeneity in Machine Learning: An Application to Identify Severe SLE Flares in Lupus Patients," *IEEE Access*, 2022. (IF 3.476)
- J11. S. Li and **Y. Zhao**, "Addressing Motion Blurs in Brain MRI Scans Using Conditional Adversarial Networks and Simulated Curvilinear Motions," *Journal of Imaging*, 2022.

- J10. **Y. Zhao**, Y. Ding, Y. Shen, S. Failing, and J. Hwang, "Different Coping Patterns Among U.S. Graduate and Undergraduate Students During COVID-19 Pandemic: A Machine Learning Approach," *International Journal of Environmental Research and Public Health*, 2022. (IF 4.614)
- J9. **Y. Zhao**, Y. Ding, Y. Shen, and W. Liu, "Gender Difference in Psychological, Cognitive, and Behavioral Patterns Among University Students During COVID-19: A machine learning approach," *Frontiers in Psychology*, 2022. (IF 4.232)
- J8. A. Jorge, D. Smith, Z. Wu, T. Chowdhury, K. Costenbader, Y. Zhang, H. Choi, C. Feldan, **Y. Zhao**, "Exploration of Machine Learning Methods to Predict Systemic Lupus Erythematosus Hospitalizations," *Lupus*, 2022. (IF 2.858)
- J7. M. He, X. Wang, **Y. Zhao**, "A Calibrated Deep Learning Ensemble for Abnormality Detection in Musculoskeletal Radiographs," *Scientific Reports*, 2021. (IF 4.996)
- J6. H. Pardoe, S. Martin, **Y. Zhao**, A. George, H. Yuan, J. Zhou, W. Liu, O. Devinsky, "Estimation of In-Scanner Head Pose Changes During Structural MRI Using a Convolutional Neural Network Trained on Eye Tracker Video," *Magnetic Resonance Imaging*, 2021. (IF 5.119)
- J5. E. Thrall, S. Lee, J. Schrier, **Y. Zhao**, "Machine Learning for Functional Group Identification in Vibrational Spectroscopy: A Pedagogical Lab for Undergraduate Chemistry Students," *Journal of Chemical Education*, 2021. (IF 3.208)
- J4. S. Bai, **Y. Zhao**, "Startup Investment Decision Support: Application of Venture Capital Scorecards Using Machine Learning Approaches," *Systems*, 2021. (IF 2.895)
- J3. **Y. Zhao**, T. Wang, R. Bove, B. Cree, R. Henry, H. Lokhande, M. Polgar-Turcsanyi, M. Anderson, R. Bakshi, H. Weiner, T. Chitnis, SUMMIT Investigators, "Ensemble Learning Predicts Multiple Sclerosis Disease Course in the SUMMIT Study," *NPJ Digital Medicine*, 2020. (IF 15.357)
- J2. **Y. Zhao**, B. Healy, D. Rotstein, C. Guttmann, R. Bakshi, H. Weiner, C. Brodley, T. Chitnis, "Exploration of Machine Learning Techniques in Predicting Multiple Sclerosis Disease Course," *PLOS ONE*, 2017. (IF 3.24)
- J1. M. Kong, **Y. Zhao**, "Computing k-independent sets for regular bipartite graphs," *Congressus Numerantium Vol. 143(2000)*, pp. 65-80, 2000

Conferences

- C21. **Y. Zhao**, Z. Du, S. Xu, Y. Chen, C. Mu, and M. Ning, "Social Media, Market Sentiment and Meme Stocks," *IEEE COMPSAC*, 2023
- C20. **Y. Zhao** and T. Chitnis, "Dirichlet Mixture of Gaussian Processes with Split-kernel: An Application to Predicting Disease Course in Multiple Sclerosis Patients," *The International Joint Conference on Neural Networks (IJCNN)*, 2022
- C19. Q. Xu, M. Sun, B. Fu, **Y. Zhao**, "Deep Learning Based Parking Vacancy Detection for Smart Cities," *Hawaii International Conference on System Sciences (HICSS)*, 2022
- C18. D. Leeds, C. Chen, **Y. Zhao**, F. Metla, J. Guest, and G. Weiss, "Generalize Sequential Pattern Mining of Undergraduate Courses," *International Conference on Educational Data Mining (EDM)*, 2022
- C17. D. Y. Wang, Y. Wang, C. Zhong, and **Y. Zhao**, "US County-level Risk Factors Associated with COVID-19 Exacerbation During Vaccination Era," *IEEE COMPSAC*, 2022
- C16. W. Liu, J. Zhang, and **Y. Zhao**, "A Comparison of Deep Learning and Traditional Machine Learning Approaches in Detecting Cognitive Impairment Using MRI Scans," *IEEE COMPSAC*, 2022
- C15. Y. Wang, Y. Wang, C. Zhong, and **Y. Zhao**, "Assessing Deep Learning Approaches in Detecting Masked Facial Expressions," *IEEE COMPSAC*, 2022
- C14. **Y. Zhao**, J. Ossowski, X. Wang, S. Li, O. Devinsky, S. Martin, H. Pardoe, "Localized Motion Artifact Reduction on Brain MRI Using Deep Learning with Effective Data Augmentation Techniques," *The International Joint Conference on Neural Networks (IJCNN)*, 2021
- C13. Y. Xiao, **Y. Zhao**, "Preserving Gender and Identity in Face Age Progression of Infants and Toddlers," *International Joint Conference on Biometrics (IJCB)*, 2021

- C12. H. Yuan, W. Zheng, Y. Song, **Y. Zhao**. "Parallel Deep Neural Networks for Musical Genre Classification: A Case Study," *IEEE COMPSAC*, 2021
- C11. Y. Shi, Z. Wu, S. Zhang, H. Xiao, **Y. Zhao**, "Assessing Palliative Care Needs Using Machine Learning Approaches," *IEEE COMPSAC*, 2021
- C10. **Y. Zhao**, M. Berretta, T. Wang, T. Chitnis, "GRU-DF: A Temporal Model with Dynamic Imputation for Missing Target Values in Longitudinal Patient Data," *IEEE International Conference of Healthcare Informatics (ICHI)*, 2020
- C9. **Y. Zhao**, B. Lackaye, J. Dy, C. Brodley, "A Quantitative Machine Learning Approach to Master Students Admission for Professional Institutions," *International Conference on Educational Data Mining (EDM)*, 2020
- C8. **Y. Zhao**, Q. Xu, M. Chen, G. Weiss. "Predicting Student Performance in a Master's Program in Data Science using Admissions Data," *International Conference on Educational Data Mining (EDM)*, 2020
- C7. **Y. Zhao**, W. Wu, Y. Jin, S. Gu, H. Wu, J. Wang, X. Jiang, H. Xiao. "Predicting 30-Day Hospital Readmissions for Patients with Diabetes," *International Conference on Health Informatics*, 2019
- C6. **Y. Zhao**, T. Chitnis, T. Doan, "Ensemble Learning for Predicting Multiple Sclerosis Disease Course," *The 15th International Conference on Data Science*, 2019
- C5. **Y. Zhao**, S. Lebak, "Deep Convolutional Autoencoder for Recovering Defocused License Plates and Smudged Fingerprints," *The 15th International Conference on Data Science*, 2019
- C4. **Y. Zhao**, B. Ahmed, T. Thesen, K. E. Blackmon, J. Dy, C. Brodley, "A Non-parametric Approach to Detect Epileptogenic Lesions using Restricted Boltzmann Machines," *22nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, 2016
- C3. **Y. Zhao**, T. Chitnis, B. Healy, J. Dy, C. Brodley, "Domain Induced Dirichlet Mixture of Gaussian Processes: An Application to Predicting Disease Progression in Multiple Sclerosis Patients," *The IEEE International Conference on Data Mining Series (ICDM)*, 2015
- C2. **Y. Zhao**, C. Brodley, T. Chitnis, B. Healy, "Addressing Human Subjectivity via Transfer Learning: An Application to Predicting Disease Outcome in Multiple Sclerosis Patients," *SIAM International Conference on Data Mining (SDM)*, 2014
- C1. B. Ahmed, T. Thesen, K. Blackmon, **Y. Zhao**, O. Devinsky, R. Kuzniercky, C. Brodley, "Hierarchical Conditional Random Fields for Outlier Detection: An Application to Detecting Epileptogenic Cortical Malformations," *The 31st International Conference on Machine Learning (ICML)*, 2014

Abstract + Poster

- AP3. **Y. Zhao**, H. Yuan, S. Martin, H. Pardoe. "Deep Convolutional Neural Networks for Predicting Head Pose During Brain MRI Acquisition," *Visual Sciences Society (VSS) Annual Meeting*, 2020
- AP2. **Y. Zhao**, J. Ossowski, X. Wang, S. Li, S. Martin, H. Pardoe. "Deep Convolutional Autoencoder for Reducing Motion Artifacts in Structural Brain MRI Scans," *Conference of Organization for Human Brain Mapping (OHBM)*, 2020
- AP1. A. Jorge, Z. Wu, T. Chowdhury, **Y. Zhao**. "Exploration of Machine Learning Methods in Predicting Systemic Lupus Erythematosus Hospitalizations," *American College of Rheumatology (ACR) Convergence*, 2020

SERVICES

Department

- 2018-now, Director of the MS in Data Science (MSDS) Program
- 2018-now, Chair of the MSDS Curriculum and Admissions Committees
- 2018-now, Member of the CIS Executive Committee
- 2018-now, Member of the CIS Ph.D. Program Committee

University

- Fall'22 – Spring'23, Organizing Committee Member for Fordham Data Science Symposium
- 2020-now, Member of Fordham's Distinguished Fellowship Committee
- 2020-now, Member of Fordham's Core Curriculum Committee
- 2019 – present, Member of GSAS Council and DGS Forum
- 2020-now, Board Member of the McGannon Center for Communication Research

Academic Community

- Reviewer for various journals and conferences

<u>Journals</u>	<u>Conferences</u>
<i>Multiple Sclerosis Journal</i>	<i>AISTATS</i>
<i>Journal of Neuroimaging</i>	<i>KDD</i>
<i>IEEE Access</i>	<i>ICML</i>
<i>Nature Communications Journal</i>	<i>OHBM</i>
<i>Applied Soft Computing</i>	<i>ICDATA</i>
<i>IEEE Internet of Things Journal</i>	
<i>Sensors</i>	

- Session Chair for the 2019 International Conference on Cybersecurity (ICCS)
 - Session Chair for the 2019 International Conference on Health Informatics and Medical Systems (HIMS)
-

STUDENT ADVISING

Masters and Honors Theses Advisor

1. Zefan Du, MS thesis, 2023. *In Progress*
2. Man Qin, MS thesis, "A New Approach to Predict Lupus Flare Level Using Calibrated Ensemble," 2021
3. Tong Wang, MS thesis, "Deep Learning Approaches for Deblurring Motion-affected Brain MRI Scans," 2020
4. Hui Yuan, MS thesis, "Predicting Head Motions During Brain MRI Scans Using Deep Learning Models," 2020
5. Michael Tynes, MS thesis, "Tensor Factorization for Recommending Perovskite Crystallization Trials," 2020
6. Sarah Bai, Honors thesis, "Exploration of machine learning methods in venture capital decision support based on balanced scorecards," 2020
7. Matias Bartolome Berretta Magarinos, MS thesis, "GRU-DF: An RNN Model with Dynamic Imputation for Missing Values in Multivariate Time Series," 2019
8. Stephen Lebak, Honors thesis, "Deep Convolutional Autoencoder for Recovering Defocused License Plates and Smudged Fingerprints," 2019

Masters Theses Reader

1. Yiwen Chen, MS thesis, "A Resource-aware Container Management Scheme in Heterogeneous Cluster," 2021
2. Sam Stein, MS thesis, "Quantum Computing Aided Machine Learning Through Quantum State Fidelity," 2020
3. Yun Song, MS thesis, "Differentiate Containers Scheduling for Deep Learning Applications," 2020
4. Henry Gorelick, MS thesis, "Predicting Hearthstone Games with Combinatorial Fusion," 2020
5. William Charles, MS thesis, "Modeling Childhood Visual Development: The Benefits of Low Resolution Training Data on Category Learning," 2020

6. Tianyi Zhang, MS thesis, "Mining Course Groupings From Student Performance," 2020
 7. Tianyu Zhang, MS thesis, "Classifying Cognitive Tasks Based on Brain Activity," 2020
 8. Aaron Dharna, MS thesis, "Control and Creation: Coevolution of Game Levels and Gameplaying Agents," 2020
 9. Trevor Buteau, MS thesis, "Digging in the wrong place: why asking a computer which emotion a human is feeling is the wrong question, and what the right question might be," 2019
 10. Xiaoran Wang, MS thesis, "Improving Portfolio Performance Using Attribute Construction and Combinatorial Fusion," 2019
-

TEACHING EXPERIENCE

- **Fall 2017 - now, Assistant Professor, Computer and Information Sciences Department, Fordham University**

<u>Graduate courses</u>	<u>Undergraduate courses</u>
CISC 5790 Data Mining	CISC 4631 Data Mining
CISC 5800 Machine Learning	CISC 2000/2010 Computer Science II
CISC 6000 Deep Learning	CISC 1600/1610 Computer Science I
CISC 5950 Big Data Computing	
CISC 6080 Capstone Project	

- **Spring 2015 & Fall 2016, Adjunct Lecturer, Northeastern University, CCIS department**
CS 6220 Data Mining Techniques, a 4 credit graduate course. Full responsibility for the course.
- **01/2011 - 06/2012, Adjunct Lecturer, Computer and Information Sciences Department, Fordham University**
CISC 1100 Structure of Computer Science, a 4-credit core course for CS majors. Full responsibility for the course.