Xin Wang

Email: xinwang35314@gmail.com • Github: https://github.com/xin-wang-kr • Webpage: https://xin-wang-kr.github.io

RESEARCH INTERESTS

Computational Social Science, Natural Language Processing, Machine Learning, Network Science, GIS, Image Diagnosis, Artificial life, Human-centered Computing

APPOINTMENT

BIG DATA & AI Lab, IntelligentRabbit LLC. — Monmouth JCT, NJ
COVID-19 Research Team Leader — Jan 2020 – Present
Research Assistant — May 2019 – Jan 2020

AI HUB, GEC academy — Beijing, China
Instructor — Jan 2021 - Present

Tianjin Value Fair Technology Inc. — Tianjin, China
Machine Learning Engineer — Aug 2017 – Dec 2017

EDUCATION

Master of Science (M.S.) in Data Science (GPA: 4.0/4.0, Top 1) — Aug 2016 – Dec 2018
Data Science Institute, Saint Peter's University
- **Dissertation:** Industry Classification on Company Annual Report using Machine Learning
- Advisors: Dr. Finn Robert, Dr. Sylvain Jaume.

Bachelor of Science (B.S.) in Industrial Engineering — Sep 2010 – Jul 2014
Department of Aeronautical Engineering, Civil Aviation University of China
- **Dissertation:** Maintenance Plan Scheduling with Capability Constraints Based on Genetic Algorithm.
- Advisor: Dr. Weigang Zhang

PUBLICATIONS

[Book] Zhao P., **Wang X.**, & Chen X. (2021). *Applied Big Data Analytics and Its Role in COVID-19*. IGI Global. in Review.

[Journal] **Wang X.**, Chen X., Bolian L. & Zhao P. (2021). Constructing an Anti-Asian Hate Indicator for Pandemic-related Comments from Mainstream Media YouTube Channels. *International Journal of Society Systems Science (IJSSS)*. Publish in process.

[Journal] Zhao, P., Chen, X., & **Wang, X.** (2021). Classifying COVID-19-related hate Twitter users using deep neural networks with sentiment-based features and geopolitical factors. *International Journal of Society Systems Science*, *13*(2), 125-139. DOI: 10.1504/IJSSS.2021.116373

[Journal] **Wang, X.**, Zhao, P., & Chen, X. (2020). Fake news and misinformation detection on headlines of COVID-19 using deep learning algorithms. *International Journal of Data Science*, *5*(4), 316-332. DOI: 10.1504/IJDS.2020.115873

[Journal] **Wang, X.**, Chen, X., & Zhao, P. (2020). The relationship between Bitcoin and stock market. *International Journal of Operations Research and Information Systems (IJORIS)*, *11*(2), 22-35. DOI: 10.4018/IJORIS.2020040102 (Acceptance rate: 42%)

[Journal] Xiao H., **Wang X.**, & Zhao, P. (2019). Satellite Image Recognition for Smart Ships Using A Convolutional Neural Networks Algorithm. *International Journal of Decision Science (IJDS),10*(2), 85-91.

[Journal] Subramaniam V., Srungarapu G., Matijosaitiene I., Supe M., Agarwal A., Zhao P., **Wang X.**, Kwartler E. and Jaume S.(2016). Geospatial and Temporal Data Analysis on the New York City Taxi Trip. *International Journal of Data Analysis and Information Systems, 8*(2), 63-73.

[Journal] **Wang X.** (2015). Study on Scheduling Algorithm Based on Capacity Constraints of Maintenance Plan. *Industrial Engineering Practice*. ISSN 2304-5337.

RESEARCHES

Image Diagnosis on Skin Cancer — 2021
(Ensemble Learning, Image Recognition, Natural Language Processing, Python, PyQt5)
- Designed an interactive skin lesion diagnosis system through ensemble learning on multiple CNN models for image classification, including ResNet50, ResNeXt50, ResNeXt101, EfficientNetB4, MobileNetV2, MobileNetV3, and MnasNet.
- Designed user interface using PyQt5, which has the functions of uploading input image, updating model, and giving the diagnosis report.

COVID-19 Anti-Asian Hate in Social Media Research Project — 2020-2021
- **Twitter Research**
  (Geographic Information Analysis, Sentiment Analysis, Deep Learning, Python Geocoder, Twitter API)

- Led the team to collect over 10 million tweets by tracking pandemic-related Twitter users between January 15 and October 15, 2020, through setting up Twitter API on Amazon AWS EC2 cluster.
- Created new features using VADER sentiment analysis. Labeled each twitter user as constantly like, constantly dislike, swing, or does-not-care towards Trump, Biden, Republican, and Democratic.
- Visualized geopolitical difference on heatmaps with the combination of geolocation data and election information.
- Evaluated feature importance by build a random forest algorithm and analyzed the model performance by using only sentiment-based features in DNN, which was trained on AWS EC2 cluster.

- **YouTube Research**
  (Signal Processing, Sentiment Analysis, Time Series Analysis, Python, YouTube Data API, Facepager)
  - Collected 1,452,373 comments from COVID-19-related news from mainstream media YouTube channels using YouTube Data API and Facepager.
  - Extracted sample data based on Anti-Asian hate hashtags and keywords. Manually labeling the 3,756 YouTube comments to perform anti-Asian hate classification using SVM, Random Forest, LSTM, CNN models.
  - Established anti-Asian hate indicator to portray the tendency of hate over time through the generated daily hate signal.

COVID-19 Fake News Monitoring and Misinformation Detection System Project 2020
(Name Entity Recognition, Natural Language Processing, TensorFlow, Web Scraping, Flask, Python)
- Built a Python-based web scraper robot to collect 8,810 fake news during January, 2020 to August, 2020 from the CoronaVirusFacts/DatosCoronaVirus Alliance Database on Poynter.
- Collected and cleaned 1,673,354 real news by dealing with the semi-unstructured data from AYLIEN News API using Python programming.
- Performed LSTM, CNN, and DBN models on fake news detection and evaluated mainstream media credibility based on the an algorithm-based ranking methods.
- Built COVID-19 fake news detection application with interactive interface based on Flask framework.

Bitcoin Prediction Research Project 2019
(Time Sereis Analysis, Python Pandas, Yahoo Finance API)
- Collected historical data of S&P500, NASDAQ, and Dow Jones indices through Yahoo Finance API.
- Based on the predictive VAR model to analyze the impulse response between the parameters for stock and Bitcoin market.
- Utilized sliding window technique to optimize the prediction results.

Smart Ship Recognition on Satellite Imagery 2019
(*Image Recognition, Machine Learning, Python TensorFlow)*
- Extracted pixel data from JSON object to perform image classification using Python programming.
- Built Support Vector Machine, Random Forest, Logistic Regression and CNN models to recognize ships in the bay and sea areas from satellite images.

More researches and details are presented in Projects section at https://xin-wang-kr.github.io

HONORS & AWARDS

SCHOLARSHIPS

- **ALife 2020 Student Scholarship** 2020
  Awarded by ALife 2020 Organizing Committee
- **Second Prize Scholarship** 2012, 2013
  For academic excellence at Department of Industrial Engineering at CAUC (for top 10)
- **Third Prize Scholarship** 2011
  For academic excellence at Department of Industrial Engineering at CAUC (for top 15)

AWARDS AND COMPETITIONS

- **Outstanding Mobile App (Andriod mobile app - Java)** 2016
  Awarded by 2016 Google Developer StudyJams (100/400)
- **Third Prize of 2014 Cross-Strait Contest of Outstanding Bachelor's Degree Thesis** 2014
  Awarded by Chinese Institute of Industrial Engineering (10/15)
- **Outstanding Student Award** 2011, 2012, 2013
  Awarded as an excellent student at Civil Aviation University of China

MENTORING PROJECTS

BIG DATA & AI Lab, IntelligentRabbit LLC.

- A New Solution of the Social Distancing and Face Mask Monitor Using Deep Learning Algorithms
- Fake News Monitoring and Anti-rumor System Using DL & Blockchain
- AI Drives Music Self-creation with Deep Learning in Tensorflow

TEACHING

BIG DATA & AI Lab, IntelligentRabbit LLC.

Instructor Jan 2020 - Present

- Mobile App Development: Python Kivy (Fall 2020)
- Python Programming in AI (Summer 2020)

GEC Academy

Instructor, AI Hub Aug 2021 – Oct 2021

- Academic writing (Graduate level)

Teaching assistant Apr 2020 – Jan 2021

- Machine Learning in Biomedical Monitoring (by Prof Maarten De Vos at Katholieke Universiteit Leuven)
- Applied Machine Learning (by Prof Stephen Coggeshall at University of Southern California)
- Algorithmic Foundations of Learning (by Prof Patrick Rebeschini at Oxford University)
- Introduction of Machine Learning (by Prof Jovan Ilic at Carnegie Mellon University)

Data Science Institute, Saint Peter's University

Teaching assistant Jan 2017 – May 2018

- Data Mining (Graduate Level, Spring 2018)
- Statistical Programming: R and SAS (Graduate level, Spring 2017)

VOLUNTEERING

DataKind Global

Data Science Volunteer Jan 2019 – Present

- Google AI Impact Challenge
- data.org Challenge

Evaluated data reliability and human-centered ML rules for the challenge projects.

Marks JCH of Bensonhurst

Web Developer Volunteer Jun 2017 – Aug 2017

Developed class registration web application using Django, MySQL, HTML, CSS.

Dreamland Plus

Data Analysis Volunteer Nov 2015 – Aug 2016

Collected genealogy data from minority community groups. Generated data visualization dashboard and performed genealogy relationship analysis using Tableau.

CONFERENCES /WORKSHOP ATTENDED

- Virtual Complex Systems & Data Science Seminar Sereis, University of Vermont Nov 2021
- Samsung AI Forum, Samsung Nov 2021
- Virtual DataDive Event 2021, DataKind Global Sep 2021
- 2020 IDEAS Global AI Conference, International Data Engineering and Science Association Oct 2020
- ALife 2020, International Society for Artificial Life Jul 2020

TECHNICAL SKILLS

- **Programming:** Python, R, SAS, C++, Java
- **Mobile/web Application:** Django, Flask, Kivy, HTML
- **Big Data:** Hadoop, Spark, Amazon AWS, Hive
- **Database:** MySQL, PostgreSQL

LANGUAGES

- **English** (fluent)
- **Portuguese** (elementary)
- **Korean** (intermediate)
- **Chinese** (native)