Yingfa Xie

VERNON ROCKVILLE, CT 06066 | yingfa.xie@uconn.edu

EDUCATION

Expected 2024 **Ph.D. Student**: Statistics **University of Connecticut,** Storrs, CT

07/2020 **M.S.**: Applied Financial Mathematics **University of Connecticut,** Storrs, CT

12/2018 **M.S.**: Electrical Engineering **George Washington University,** Washington, DC

06/2016 **B.Eng.**: Microelectronics **Guangdong University of Technology,** Guangzhou, China

EXPERIENCE

08/2022 - Present

Research Assistant

University of Connecticut, Department of Statistics, Storrs, CT

- Performed data cleaning, visualization, and exploratory data analysis; handled missing value; conducted feature engineering in R and Python
- Predict future enrollment with logistic regression model, random forest model, and neural network

08/2021 - Present

Research Assistant

University of Connecticut, Department of Statistics, Storrs, CT

- Proposed to model recurrent events with the first hitting time (FHT) model of reflected Brownian motion
- Implemented efficient rejection sampling algorithm to generate random number from the FHT distribution
- Conducted inference with Bayesian framework using Markov Chain Monte Carlo
- Applied the FHT model to hypoglycemic events dataset and identified the risk factors of hypoglycemia

06/2019 - 08/2019

Data Scientist, Intern

JOYY Inc, Shanghai, China

- Contributed to the development of in-house facial recognition & verification system
- Developed an Optical Character Recognition (OCR) model for fraud detection that efficiently extracts identification information from images using Pytesseract and CV2 in Python; achieved more than 80% accuracy of information recognition
- Deployed the OCR model into credit assessment classification system used by risk management team

WORKING PAPERS

Xie, Y., Fu, H., Huang, Y., Pozdnyakov, V., and Yan, J. (2022): Recurrent events modeling based on a reflected Brownian motion with application to hypoglycemia.

PROGRAM LANGUAGES & SKILLS

Programming Languages: Python, R, SAS, C++, SQL

Framework & Tools: MySQL, SQL Server, R shiny, Ggplot2, Scikit-learn, TensorFlow