Yiqi Ma

ucbtmaj@ucl.ac.uk | 44-07774993209

E1 4GG, IQ East Court W811, 450 Mile End, London, UK

EDUCATION

University College London September 2024-September 2025(to be expected)

Master of Science in Ecology, Climate Change and Health

- Core Courses: Nature-Smart Challenge, Data Science for Ecology, Climate Change and Health, Human and Ecosystem Health in a Changing World, Climate Change and Health, etc.
- Online Courses: AI and Climate Change, GIS, Mapping, and Spatial Analysis, etc
- **GRE 337** (Quantitative Reasoning:170; Verbal Reasoning: 163; Analytical Writing:4.0)

University of Warwick October 2022-Febuary 2024

Master of Science in Statistics

- Core Courses: Data Mining, Statistical Learning and Big Data, Statistical Methods, Multivariate Statistics with Advanced Topics, Principles of Cognition, Bayesian Approaches in Behavioural Science, An Introduction to Statistical Practice, Risk Theory, etc.

University of Liverpool September 2020-July 2022

Bachelor of Science in Applied Mathematics

- Core Courses: Financial Mathematics, Statistical Theory and Method I & II, Methods of Operational Research, Complex Functions, Mathematical Models: Microeconomics and Population Dynamics, Geometry of Curves, Vector Calculus with applications in Fluid Mechanics, etc.

Xi'an Jiaotong-liverpool University September 2017-July 2020

Bachelor of Science in Applied Mathematics

- Core Courses: Introduction to Programming in Java, Advanced Linear Algebra, Dynamic Modelling, Abstract Algebra, Methods of Applied Mathematics, etc.

RESEARCH PROJECT

Effect of Stratif Index and Tree Rao on European Bird and Bat Species Richness Research Lead

Individual Project November-December 2024

- Objective: Investigated how vegetation structure (Stratif Index) and tree diversity (Tree Rao) impact biodiversity across 73 bird species and 27 bat species, using a dataset from six European countries. This research aimed to identify patterns that inform habitat restoration and biodiversity conservation.
- Data Analysis: Applied fixed- and mixed-effects Poisson GLMs to model species richness. Mixed-effects models included country-specific random effects, improving explanatory power (e.g., birds' AIC reduced from 937.58 to 901.68). Validated models through residual plots, Durbin-Watson tests, and comparison of AIC/BIC values.
- Key Findings:
  1. Birds showed a significant positive correlation with Tree Rao ($p = 0.012$), supporting the role of tree diversity in providing ecological niches.
  2. Bats exhibited a significant negative relationship with Stratif Index ($p = 0.0042$), indicating a preference for open habitats.
  3. Mixed-effects models highlighted country-level influences, such as differences in forest cover and land use, which contributed to species richness variability.
- Innovation: Enhanced biodiversity monitoring by integrating mixed-effects models, demonstrating how habitat heterogeneity influences species richness. These insights provide metrics for evaluating ecological restoration strategies, particularly for avian and bat populations.

CoNIC Colonic Nuclei Identification and Enumeration Challenges April-October 2023

- Aim: To utilise deep learning algorithms to accurately predict the cell counts of each type in a test image and the total number of cells in each image
- Innovation: the application of color space conversion technology (convert RGB images into HED color spaces) significantly improved the visualisation and recognition rate of nuclei
- Methods: Feature engineering: used pixel-level intensity analysis and morphological feature extraction to enhance the quality of input data; Employed PCA to reduce the dimensionality of the data and to analyse the correlation between features and cell counts; Modeling and forecasting: 1. Applied OLS and MLP regression models for preliminary prediction and analysed the difference in prediction accuracy between the two models; 2. Designed a customized CNN architecture and optimised for cell counting problems, including efficient feature extraction layer and accurate classification layer; Model evaluation and optimization: 1. Ensured the generalization ability and stability of the model through cross-validation and loss function optimization; 2. The evaluation includes RMSE, correlation coefficient analysis, and consideration of model interpretability and bias-variance tradeoffs

- Conclusion: The average prediction accuracy of the neural network on the test data reached 92%; A convergence plot showed a 40% reduction in losses after 10 training cycles; Compared with the real count, the average prediction error is reduced by 15%; The RMSE is reduced by 20%, with the Pearson correlation coefficient of 0.85, the Spearman correlation coefficient of 0.80, and the R^2value increased to 0.75

Developing a General-purpose Biomedical Data Analysis Tool based on Python Programming

Group Leader (4 members) September-December 2023

- Aim: To enable users to upload, process, explore, and visualize gene expression data, and also integrates basic differential expression analysis capabilities
- Aim: The primary goal of this project is to develop a versatile tool that enables users to upload, process, explore, and visualize gene expression data. This tool incorporates basic differential expression analysis capabilities, catering specifically to the biomedical field.
- Methodology and Tool Utilization:
  1. Overall Planning and Technology Selection:
     Conducted in-depth planning and analysis to ensure the project aligns with the unique needs of the biomedical sector.
  2. Data Processing and Exploratory Data Analysis Modules:
     - Utilized mathematical modeling and statistical analysis principles to develop robust data processing modules.
     - Integrated SciPy for advanced scientific computations, which enhanced the tool's capability to handle complex biomedical data structures and perform intricate mathematical operations efficiently.
     - Employed Statsmodels for its statistical modeling features, enabling the tool to carry out sophisticated statistical tests and data exploration. This was particularly crucial for understanding gene expression patterns and their statistical significance.
     - Incorporated Seaborn for its high-level visualization functionalities, allowing users to create informative and aesthetically pleasing plots. Seaborn's advanced graphical capabilities made it easier to interpret gene expression data visually.
  3. Web Crawlers and Database Interaction:
     - Developed web crawlers to automate the collection of relevant biomedical data.
     - Established interaction with databases for efficient data storage and retrieval, employing NumPy and Pandas for data manipulation and Matplotlib for preliminary data visualization.
  4. Differential Expression Analysis Module Design:
     Designed a highly configurable differential expression analysis module using advanced statistical methods. This adaptability ensured that the module could cater to various experimental designs in the biomedical field.
- Conclusion: The first version of this tool was successfully launched within the stipulated timeline and was well-received by its users.

PROFESSIONAL EXPERIENCE

SkyReach Environmental Technology Co., Ltd. March-June 2024

UAV Operations Intern

- Assisted in deploying UAVs for environmental monitoring, covering areas ranging from 5 to 50 hectares per session, including wetlands and urban green spaces.
- Conducted UAV setup and calibration, including IMU alignment and multispectral sensor adjustments, ensuring accurate data collection.
- Planned optimized flight paths with 70% forward and 60% lateral overlap, improving image precision for post-processing.
- Processed over 150 GB of aerial imagery using software such as DJI Terra to generate high-resolution orthomosaic maps, 3D models, and NDVI vegetation indices.
- Compiled operational reports evaluating UAV performance, including metrics like battery life and GPS accuracy, and provided actionable insights for ecological restoration strategies.
- Addressed technical challenges, such as GPS signal loss in mountainous regions, by incorporating ground control points (GCPs) to enhance positional accuracy.
- Collaborated with GIS teams to integrate UAV-derived DEM data for land-use change analysis and environmental restoration planning.

JD.Com, Inc. January-March 2023

Major Search User Behavior Analysis

- Analysed the behaviour of search users, understood the user characteristics, and knew the business iteration as Search is the traffic position second only to the home page in the golden process, and its user behaviour is significant to the overall conversion
- Business combing: familiar with the search module, understanding the key behavioural path of users, combing buried points and the logic of taking numbers

- Indicator system: used the funnel analysis model to build a large conversion funnel; selected GMV as the core indicator and process indicator disassembly; dimensional disassembly of key indicators from channels, geography and users, etc., to build the overall GMV improvement indicator analysis system
- Data analysis: started with the overall golden process conversion analysis to understand the general situation; conducted active analysis, retention analysis, purchase path analysis, analysis of popular products and RFM user value analysis for search users

Guizhou Transport Science Research Institute Shares Co., Ltd May-August 2022

Data Analyst

- Developed and optimised nano-biomimetic materials in the field of bone and cartilage repair and mastered the high-precision 3D bioprinting technology for multi-level biomimetic cartilage reconstruction
- Participated in the development of the 'An Yunbao' App to enable the system to automatically detect the dangerous behaviors of drivers and give corresponding instructions through deep learning
- Adopted the self-learning optimal control based on local data: iterative adaptive dynamic programming method, which finally made the application further optimized
- Successfully increased the accuracy by about 10 percents compared to the previous model
- Attended the First National Science Conference on Big Data and Artificial Intelligence (CSIAM-BDAI 2018) in Chongqing on behalf of the company, learning the optimization theory of deep learning, etc

EXTRACURRICULAR ACTIVITIES

Light Civil Unmanned Aircraft Safety Certification March 2024

- Obtained the Light Civil Unmanned Aircraft Safety Certification from the Civil Aviation Administration of China (CAAC) in September 2024.

Successful Participant in the American Mathematical Contest In Modeling (MCM) March 2020

- Built the STAR_Rating model and analyzed the distribution of each variable to establish the relationship between STAR_Rating and votes, consumption types, consumer sources
- Analysed main impact index of star rating (principal component analysis, hierarchical clustering, factor analysis, etc.), time-based reputation changes for each model (time series analysis) for each product and identified text-based metrics that give corresponding success or failure cues
- Established a comprehensive model based on Word2Vec and LSTM to analyze the sentiment tendency of the evaluation by using the sentiment analysis model

International Volunteer Programme in Sri Lanka June 2022

- Designed questionnaires and carried out in-depth interview of teenagers, set up community hub and organised activities to enrich their life
- Proposed advice regarding the employment opportunities and medical care of residents

SKILLS & INTERESTS

- **Languages**- Mandarin (native); English
- **UAV Operations**
- **Computer Skills**- Proficient in R, SQL, Python, Matlab, Excel, SPSS, Photoshop, Fluent in Word and PPT
- **Interests**- Avid Reader, Basketball, Guitar, Drum