**PAN HE**

ph2200@nyu.edu 718-877-4720

65 bay street, Jersey City, New Jersey, 07302

|  |  |
| --- | --- |
| **Education** | |
| 08/2020-06/2022 | **New York University** |
|  | **Major: Biostatistics**  *Courses: Deep Learning, Advanced Python, Machine Learning, Natural Language Processing, Data Structure, Algorithms, Java* |
| 09/2015-07/2020 | **Shanghai Jiao Tong University** |
|  | **Major: Preventive Medicine, Administration Management** |
| 12/2021- | **Extracurricular Studies(Udemy)**  *Full Stack programming, including HTML, CSS, Javascript, Node, React, MongoDB, Web3 and DApps* |
| **Projects** | |
| 01/2021-05/2022 | **Deep Learning and Machine Learning Project** |
| * Build Resnet18 and ViT from scratch with Pytorch * Used CKA (Centered Kernel Alignment) to analyze and compare the hidden layer representations within and across the neural networks * Build machine learning algorithm from scratch with Numpy, like Linear regression with Lasso, Logistic Regression, SVM , Decision Tree, Random Forest, Adaboost, K-Means. | |
| 03/2021-07/2021 | **Diabetes diagnosis based on machine learning** |
| * Preprocessed and balanced patient data such as multiple transaminase content, globulin content, cholesterol content, etc., and perform feature engineering. * Conducted a variety of regression prediction models, such as support vector machine, random forest, XGBoost, and explore the optimal parameter configuration * Model fusion for prediction of blood glucose, which is carried out through stacking, boosting and other methods * The importance of features is sorted, and several more important features are identified to provide additional reference for doctors' diagnosis. | |
| **Internship** | |
| 05/2021-08/2022 | **Shanghai heywhale tech***,* *Intern of data science group* |
| * Wrote online tutorial, including theorical and practical tutorials of cluster analysis, discriminant analysis and principal component analysis * Instructed Contestants to apply regression, cluster analysis and discriminant analysis methods to write data science projects, whose data set comes from Kaggle and other databases. * Analyzed the Contestants data of usage of heywhale website and products by SQL. | |
| 09/2020-03/2021 | ***Fusun Pharma****,* *Intern of statistical group* |
| * Wrote clinical research plan, designed statistical analysis plan, calculated sample size. * Conducted single factor analysis (t test, chi-square test, Kruskal-Wallis rank sum test) by SPSS and R to analyze the data from clinical trials, Performed exploratory factor analysis and built Logistics Regression model. * Medical industry research on COVID-19 vaccine and neutralizing antibodies. | |