**ZHAOHUI (MARVIN) MAN**

[**manzhaohui1975@gmail.com**](mailto:manzhaohui1975@gmail.com) **|** (**984)2157418 |** Linkedin: <https://www.linkedin.com/in/zhaohui-marvin-man-2832142b/>

124 St. Andrews Ln., Chapel Hill, N.C.

**EDUCATION AND CREDENTIALS**

**University of North Carolina at Chapel Hill (Expect to graduate in December 2019)**

Candidate of MSIS (Master of Science Information Science)

**CHIP (Carolina Health Informatics Program)**

Clinical Information Science Certificate Program

**PLA （People’s Liberation Army）University of Foreign Language, China**

MA in English Language and Literature

BA in ETS (English for Science and Technologies)

**SKILLS AND CERTIFICATS**

**Computer Skills**: Python (Numpy,Pandas, Matplotlib, Scipy, Scikit-learn, Tensorflow, etc.), R, SAS, Tableau, SQL, VISIO

OpenRefine, LucidChart, LightSIDE, Apache Spark, MySQL, Machine Learning toolkits, Access, Qualtric

**Web Skills**: HTML, CSS

**Data Analysis Skills**: Manage EHR(Electronic Health Record) , Build dashboard, Define business metrics, Machine Learning and

Deep learning algorithms such as Linear Regression, Logistic Regression, Random Forest, SVM, KNN,

CNN, RNN, and TensorFlow; NLP (Natural Language Processing); Sentiment Analysis

**Certificates:** TESOL (Teaching English to Speakers of Other Languages) Advanced Certification, TESOL Trainer Certification

**Language Skills:** Mandarin (native), English (proficient)

**PROJECTS**

***Suspended Stock Sentiment Analysis (Product Control of Credit Suisse)***

• Use stock suspension summarization as input

• Apply different sentiment analysis models using various machine learning algorithms to predict the sentiments of news articles

***Data Analysis,*** May 2019-Dec. 2019

• ETL(Extract, Transform, Load) Perinatal data (baby, mom, fetus data) from the Epic Electronic Medical Records

• Integrate that data into a dataset using SQL and Python

• Use Tableau to create dashboards to find data insights for patient care, patient safety, quality/process improvement

• Use machine learning algorithms to make better predictions of the risk of PPH (Post Partum Hemorrhage)

***Optimal Workstation Design for Radiation Oncologists in*** ***UNC Cancer Hospital,*** May 2019- Dec. 2019

• Do literature review of ergonomic work space design

• Analyze video and text data from experiment

• Use machine learning algorithms to process data and validate the quantification of metrics involved in the ergonomic analysis

such as cognitive workload, situation awareness and performance

***Data Analysis*** ***for*** ***Carolina Applied Informatics Research (CAIR) Group (http://thecairgroup.web.unc.edu) at UNC-Chapel Hill,***

January 2019-May 2019

• Evaluate virtual urgent care on four domains- access to care, effectiveness, experience and economic impact

• Compare the UNC virtual urgent care with in-person urgent care centers on these four domains, and develop broad framework for

evaluation of virtual urgent care centers

• Use Python libraries ‘googlemaps’, ‘uszipcodes’, ‘pandas’, and ‘numpy’ for feature engineering

• Use Tableau to visualize the result of data analysis

***Crowdsourcing and Image Analysis Project in UNC Kenan Flagler Business School*,** Jan. 2019, ongoing now

**Data Collection**

• Develop and implement code to scrape image data from Design Hill website, a specialized logo competition platform

• Design local database for data storage and normalize the database using SQL skills

• Execute website scrape, data cleansing, and data pre-processing for future analysis

**Data Analysis**

• Use OpenCv to quantify the similarity between images

• Use K Means Clustering to calculate the optimal number of clusters of images from each logo contest and submitted by each

designer as an indicator of the diversity of the works from contests and designers

• Research new methodologies and machine learning algorithms that could produce better, more accurate, measures of similarity,

originality, learning, etc.

***Sentiment analysis of Yelp restaurant reviews,***  October, 2018

• Use python to scrape customers’ review data from Yelp

• Sample and label the dataset, randomly split them into training and test sets

• Select feature representation using LightSIDE and Python coding

• Compare the accuracy of three classifiers, namely, Naïve Bayes, SVM, Logistic Regression and Random Forest

• Use the classifier with the best accuracy to predict the sentiment of customers and recommend restaurant based on the prediction

***Information Gathering Plan for Reducing Pharmacy Inventory at Local Pharmacy*** May，2018

• Use system analysis models to analyze the flow of medicine in one Harris Teeter pharmacy

• Come up with a proposal as to augment the efficiency of pharmacy inventory operations and to reduce the workload of pharmacists

**PROFESSIONAL EXPERIENCE**

**Business Analyst Intern:** Credit Suisse, Raleigh, N.C., U.S.A.

**Data Analyst Intern:** UNC Center for Maternal and Infant Health (UNC CMIH), N.C., U.S.A.

**Research Assistant:** UNC Department Of Radiation Oncology, UNC-CH, N.C., U.S.A.

**Carolina Applied Informatics Research Group Member:** School of Nursing, UNC-CH, N.C., U.S.A.

**Graduate Research Assistant:** Kenan Institute of Private Enterprise, Chapel Hill, N.C., U.S.A.

**Writer and Translator:** Craft Translation (New York)

*Realize the glocalization of international brands including UPS, Gulfstream, Cisco, Dove and Mars*

**Lecturer:** SIAS International University, China (U.S. Sponsored)

**Lecturer**: PLA University of Foreign Language, China

**PUBLICATIONS**

**Paper: *Computer-assisted Language Testing: Validity Analysis*,** MA dissertation

***Constitution, Attributes and Assessment of Mobile English Teaching****,* published on *Distance Education, China*