

WENJIA ZHAI

(420) 890-4973 | zhai.we@husky.neu.edu

EDUCATION

Northeastern University

M.S. in Bioinformatics, GPA: 4.0

Boston, MA

May 2020

Certification in Data Analytics

Coursework: Statistics, Data Mining, Machine Learning, Data Visualization

Massachusetts Institute of Technology

Certification in Data Science and Big Data Analytics

Boston, MA

Jan 2019 – Mar 2019

Coursework: Data preparation and transformation, Regression, Classification and Clustering, Neural Network, Recommendation Systems

WORK & ACADEMIC EXPERIENCE

Project: Extraction and Analysis of Opinions in News in Python

Team Leader

Virtual

July 2019

- Collected ~100 articles from main stream media websites (CNN, Fox, New York Times, Washington Post)
- Extracted opinions from articles, analyzed opinion texts by using coreNLP, clustered articles with similar opinions
- Constructed a website using Flask and Bootstrap to visualize the analysis

Project: Petfinder.com Adoption Prediction in R

Team Leader

Boston, MA

Mar 2019 – May 2019

- Utilized R to perform data collection, data loading on multiple sources of data and data transformation
- Developed statistical inferences, multi-dimensional and interactive charts to analyze adoption rate using ggplot2 and shiny
- Created multiple linear regression model to quantify the adoption speed by analyzing all possible consequences related to pet demographic info

Project: Pet Segmentation Clustering in Python

Team Leader

Boston, MA

Mar 2019 – May 2019

- Explored the insight of data by comparing distribution of features in train and test data and conducting word cloud
- Performed various types of feature engineering to pets' age, breed, color and some other categorical parameters for model preparation and clean parameter distribution
- Drafted KNN method to group 8 marked group from 15 thousand dogs and cats and evaluate the performance via AUC value check and parameter tuning using Python
- Ran each pet profile's description through Google's Natural Language API, providing analysis on sentiment and key entities and utilize this supplementary information for pet description analysis

OTHER EXPERIENCE

Co-author

July 2019

- Author of chapter 2: *Smart Search Strategy* in book: *Artificial Intelligence and Natural Language Processing*
- Publisher: China Machine Press (In Progress)

TECHNICAL SKILLS

Data Science: Data Mining, Feature Engineering, Data Modelling, Data Visualization, Regression Model, Classification Model, Clustering Model, Neural Network, Natural Language Processing, Time Series Analysis

Programming Language: R, Python, Perl

Programming Tools: Regular Expression, SQL (MySQL, SQLite), MongoDB, Shell, Git, Google Cloud Platform

Web Development: HTML, CSS, JavaScript, Flask, Bootstrap

Operating System: Windows, Linux/Unix (MacOS, Ubuntu, Red Hat)

Report Writing: Microsoft Office (Word, Excel, PowerPoint), LaTeX, Markdown