YIXUE WANG

Data Analyst

yixue.wang@nyu.edu

201-208-9958

New York, NY

EDUCATION

M.S. Urban Informatics (GPA:3,7/4.0)

New York University

Aug 2014- Aug 2015

B.S. Telecommunications Engineering

Queen Mary, University of London

Sep 2010 - Jun 2014

HONORS

Academic Fellowship
New York University

First Class Honors
University of London

SKILLS

Programming Language

Python, R, JAVA, SQL, MATLAB, Javascript, HTML,

Data Analysis & Machine Learning

scikit-learn, Numpy, Pandas, R, MLlib, WEKA

Information Visualization

d3.js, Tableau, Matplotlib, ggplot, seaborn, bokeh

Geospatial Analysis

ArcGIS, Tableau, CartoDB, Basemap

Big Data Ecosystem

Hadoop, MapReduce, Hive, Pig, Spark

WORK EXPERIENCE

Data Analyst Jul 2015 - Present
@HaystaqDNA

- **The modeling and QA frameworks development and improvement**

  Improved and developed the data ingestion, pre-processing, modeling, evaluation, scoring and QA frameworks that the company uses as its core code set using scikit-learn, Pandas, Numpy, Matplotlib, LaTeX and MySQL

- **Exploratory data analysis**

  Manipulated, cleansed and processed raw data using Python and SQL. Analyzed processed data and drew conclusions for HaystaqDNA's clients.

- **Visualization tools development**

  Developed visualization tools for other analysts to visualize data using Pandas, Matplotlib, Basemap, Seaborn.

Graduate Research Assistant Oct 2014 - Jan 2015
@ New York University & University of Warwick

- **Emotion sensing analysis on social media data**

  Built web app for scripting participants' social media data in Python and created PostgreSQL database schemas.

RESEARCH EXPERIENCE

Student Sep 2014 - Aug 2015
@New York University

- **Machine Learning Application for Predicting News and Economic Crisis**

  Preprocessed and analyzed GDELT data using Hive to identify the most frequent events. Applied the frequent pattern growth algorithm with Spark MLlib and visualized time series data in Tableau and D3.js

  https://github.com/YixueWang/Yark

- **Fashion Analytics Research**

  Collected 2015 Spring and Summer fashion elements. Analyzed the elements trend from Twitter and Instagram API

- **Using Social Media to Predict Urban Transportation**

  Conducted correlation analysis between Instagram data and taxi data using Python, WEKA, R, Tableau, and CartoDB

- **Visualizing and Interacting with Large Decision Trees**

  Developed a D3.js visualization and interaction tool to visualize a decision tree built by sklearn.

  https://github.com/YixueWang/sp2015-group17