**HONGYUAN CUI**

|  |  |  |  |
| --- | --- | --- | --- |
| 636 Summerlea St. Apt.8 | Pittsburgh, PA, 15232 | feanor090@gmail.com | (412)652-8319 |

|  |  |
| --- | --- |
| **EDUCATION** | **University of Pittsburgh** *Pittsburgh, PA* Expected 4/2017  Master of Information Science GPA: 3.90/4.0  Core Courses: Algorithm Design, Database Management, Cloud Computing, Information Storage & Retrieval, Data Mining, Machine Learning, Information Visualization  **Guangdong Medical College** *Dongguan, China* 6/2015  Bachelor of Information Management and Systems GPA: 3.32/4.0  Core Courses: Operation System, Website Design, Computer Network And Data Communication, C++ Programming Language, Data Structure, Software Engineering, Information Security Technology |
| **PROFESSIONAL EXPERIENCE** | **.NET Engineer Intern- SVC Health Science IT Department of University of Pittsburgh** 5/2016-8/2016   * Developed student enrollment websites using **ASP.net** and **SQL Server**. * Modified and improved application process logic using state machine and managed user roles with different actions for system robustness. * Built Website Architecture with **MVC** (Model View and Controller)and **Three Layers** **Architectures**. * Designed and modified websites with Javascript, JQuery, CSS and AJAX. * Created and modified SQL Server Reporting Services (SSRS) reports to support application decision of enrollment committee and interviewers.   **Intern- Shenzhen Annet Information System Co,. Ltd** 11/2014-2/2015   * Managed data backup and server cluster in Oracle and SQL Server Database. * Integrated data from various sources using web service or database view to provide overview of data. |
| **TECHNICAL**  **SKILLS** | **Programming**: Java, R, MATLAB, Python, SQL, ASP.NET, Node.js, C#, C++, JavaScript, scala  **System:**Windows, Android, AWS EC2 Linux, Hadoop  **Tools**: SQL Server, MySQL, Oracle, Visual Studio, AWS, Slack, Elastic Search, Kibana, D3.js, SSRS |
| **ACADEMIC**  **PROJECTS** | **Yahoo Music Recommendation System in Hadoop** 1/2016-5/2016   * Stored data in Hadoop Distributed File System and preprocessed data using Map-Reduce and Scala. * Clustered music using Mahout (Data Mining Library in Hadoop) to group music of similar users’ reviews. * Implemented music Recommendation system using Item-Based Collaborative Filtering.   **Prediction and Detection of Crisis Intensity: Graph-Based Analysis** 1/2016-12/2016   * Used **2.3 GB** events dataset from ICEWS (Integrated Crisis Early Warning System) and crisis dataset from GTDS (Ground Truth Data Set). Then cleaned and transformed events data into graph model. * Visualized events dataset and crisis dataset on a world map using D3.JS, and implement advanced interaction for users to explore trends of data, view the entirety and detail of countries’ events, filter data, and compare data among countries. In order to improve visualization performance, data is transformed by aggregation. * Used above visualization to explore relation between crisis and events, and attested it in following steps * Calculated topological value (e.g. betweenness, page rank, hub score, authority) as prediction features * Made prediction for the number of crisis in the next month with SVM, Decision Tree, Naive Bayesian, KNN, Logistic Regression.   **House Spy- A chat robot based search engine to find desired house** 8/2016-12/2016   * Crawled house information with **Scrapy** (a python web crawler framework) from various house websites such as craigslist, rent.com, Zillow, etc. * Used Logstash (an ETL tool for Elasticsearch) to clean and load the crawled data into **Elasticsearch** and index them for later search. * Built a **Slack** chat robot to let user search their desire house by chatting with robot. Used **Botkit** to build the chat logic and do the NLP processing. * Deployed the system on **AWS** (Elasticsearch on EC2 Linux, back-end system on Beanstalk)   **Kaggle Titanic Passenger Data Analytics** 8/2015-12/2016   * Built Neural Network model (backpropagation and optimized with **Pruning**) and Naive Bayesian Network models using Matlab. * Classified and Improved Prediction Accuracy using bagging (combine various classifiers). |