**Jinwei (Edward) Kang**

[edwardkang3636@gmail.con](mailto:edwardkang3636@gmail.con) (626)-695-1218 <https://github.com/edwardkang1004>

Work Authorization: U.S. Citizen

**TECHNICAL SKILLS**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Data Manipulation | **Data Analysis** | **Predictive Model** | **Data Visualization** | **Programming** |
| Excel | A/B Testing | Gensim | Matplotlib | MATLAB |
| Pandas/ Numpy | Bayesian Inference | Keras | PowerPoint | Python |
| NLTK | Boostrap Inference | Scikit-Learn | Seaborn | R Programming |
| SQL | Time Series Analysis | TensorFlow | Tableau | Spark |

**PROJECTS**

**Kaggle Natural Language Processing on Home Depot Product Search Relevance Model (09/2019 – 11/2019)**

* Text pre-processing by correcting spelling mistakes, removing tags, accented characters, special characters, stopwords, stemming, and lemmatization.
* Features engineering by using TF-IDF, Word2Vec, Doc2Vec, and FastText.
* Used ridge regression, bagging, random forest, XGBoost, LightGBM, and stacking to build a predictive regression model to predict the relevance score of the Home Depot product search.
* Achieved top 7% ranking (ranked #134 among 2124) with RMSE 0.46185.

**UFO Sightings Predictive Model (07/2019 - 09/2019)**

* Data wrangling and exploratory data analysis of 50,000 UFO sightings reports.
* Used logistic regression, KNN, SVM, naive Bayes, AdaBoost, XGBoost, and random forest to build a predictive classification model to predict the month of UFO appearance.
* Built a predictive model to predict the month of UFO appearance base on weather, geography, and population information with a 65.47% accuracy score and 70.09% AUC.

**WORK EXPERIENCE**

**Data Analyst (02/2019 - Present)**

Tak Shing Hong Inc. *Industry, California*

* Conducting researches on sales and products of both retail and e-commerce by collecting data from the data warehouse and Google Analytics using SQL.
* Conducting statistical analysis and predictive models on sales, products, and customers by applying methods such as bootstrap Inference, time series analysis, clustering, regression, and classification using Microsoft Excel, Python, and R.
* Conducting and presenting data visualization and reports using Tableau and Microsoft PowerPoint.
* Successfully identified high spending and repeat customers, and helped to develop new e-mail marketing strategies to increase 15.58% of the sales and 7.33% of conversion rate compared to last year in six months.

**Investment Consultant Internship (06/2016 - 09/2016)**

Shenzhen Zhen Jing Yi Investment Management co., Ltd. *Guangdong, China*

* Successfully helped a client make about 10% profit in two weeks.
* Successfully collaborated and contributed to a six people investment team to achieve the monthly team goal of getting over a hundred thousand investment funds.
* Successfully achieved the trainee goal of convincing four clients to open an investment account in about three months.

**EDUCATION**

**Bachelor of Science in Computational Statistics (09/2015 - 09/2018)**

University of California, Davis