

SONG WANG

sw299@njit.edu
(862) 763-2981
Basking ridge, New Jersey

Data Scientist

GitHub: <https://github.com/sw299>
LinkedIn: <https://www.linkedin.com/in/sw299>

Summary

Scientist with 3+ years of experience executing data-driven solutions to increase efficiency, accuracy, and utility of internal data processing. Experienced at creating data regression models, using predictive data modeling, and analyzing data mining algorithms to deliver insights and implement action-oriented solutions to complex technical and business problems.

Skills

Python/SQLite3/PostgreSQL	Machine learning	Natural language processing	Statistical analysis
Data mining/processing	Data visualization	Data cleaning	Experimental design

Recent Projects

20 News Groups | Github:

https://github.com/sw299/Thinkful_Capstone_Projects/blob/master/NLP_Capstone_Project.ipynb

- Applying NLP techniques to classify text documents into various news groups
- Tech: Language parsing supervised/unsupervised NLP Word2vec Text summarization/ extraction

Bank Marketing | Github:

https://github.com/sw299/Thinkful_Capstone_Projects/blob/master/capstone_supervised_learning.ipynb

- Utilized machine learning algorithms to predict if clients will subscribe deposit service and achieving 88% accuracy
- Tech: Knn/Random Forest/Gradient Boosting/SVM/Logistic Regression PCA Cross validation/Accuracy score

Credit Card Fraud Detection | Github:

https://github.com/sw299/Thinkful_Capstone_Projects/blob/master/capstone_unsupervised_learning.ipynb

- Conduct clustering algorithms and dimension reduction methods to identify fraud, and achieving a 90% accuracy
- Tech: KMean/Hierarchical/Gaussian Mixture Models PCA/TSNE/UMAP Rand Index/Confusion Matrix

Experience

ENCINA

Research Scientist

Basking Ridge, NJ
03/2018 – present

- Conducting statistical analysis, model estimation and lead research activities.
- Large-scale data cleaning, data interpretation, and cross-correlation analysis.
- Generate insights connecting simulation results with design and business problems and objectives.
- Work cross with vice president, and senior leaders in the engineering and research teams.

THINKFUL

Data Science

New York City
07/2019 – 11/2019

- Immersive program involving advanced statistics, data analysis, visualization, and machine learning techniques.
- Pursued more advanced topics such as Deep Learning with Keras and TensorFlow, CNN, RNN, NLP.
- Independently completed Capstone projects and articulate research findings with data scientists affiliated with Thinkful.
- GitHub URL: https://github.com/sw299/Thinkful_Capstone_Projects

Education

NEW JERSEY INSTITUTE OF TECHNOLOGY

Ph.D. Chemical Engineering

05/2014 – 12/2017

M.S. Chemical Engineering

01/2013 – 05/2014

B.S. Chemical Engineering

09/2008– 06/2012