Chenghao Ding

(217) 305-0997 ◊ frododch@gmail.com ◊ Web: chenghaoding90.github.io/ ◊ LinkedIn: linkedin.com/in/chenghaoding/

Education

University of Illinois at Urbana-Champaign
Ph.D. *in Nuclear Engineering* *Dec 2021*
Master of Science *in Applied Statistics* *May 2020*

Wuhan University
Master of Science *in Power Engineering* *Jun. 2015*

Hubei University of Science and Technology
Bachelor of Science in *Nuclear Engineering* *Jun. 2013*

Related Courses

Introduction to Data Science	Statistical Learning in Data Science	Advanced Data Science	Machine Learning
Applied Regression & Design	Statistics of Big Data & Clustering	Pattern Recognition	Statistical Learning

Experience

University of Illinois at Urbana-Champaign *Urbana, IL*
Graduate Research Assistant *May. 2018 - Present*

- Ph.D. Thesis: Global Heat Balance Model and Probability Distributions for Atmospheric Response
- Studied the global warming trend and effects of human activities on climate. Analyzed big climate data, and developed a mathematical model to fit the historical greenhouse gases emission data and did future climate projections.
- Applied different time series model to study the residuals and quantify the uncertainty. Provided suggestions to government policy makers.

Projects

CNN for Object Recognition in Images (case study on Fashion MNIST dataset) *Spring 2020*

- 60,000 images are loaded, One-hot coding is used for category labels, and a two-layer CNN model is built by Keras and 97.99% test accuracy is achieved
- Tuning hyperparameters with skopt, the best learning rate, filter size of the convolutional layer and the number of dense layer are found
- Test accuracy was improved to 98.36% and loss function converged quickly

Build a recommendation system on MovieLens 100K Dataset *Fall 2019*

- Singular Value decomposition (SVD) method is applied, 5 fold cross-validation for training and testing are computed
- Implemented item-based collaborative filtering algorithm through Hadoop MapReduce chaining jobs
- Made movie recommendations based on top-k user rating predictions, and missing rating pattern are inspected

CIFAR-100: object recognition *Fall 2018*

- Conducted data cleaning, exploratory analysis and feature engineering on CIFAR-10 dataset
- Implemented Pre-trained VGG16 Neural Networks and ResNet to train the CNN parameters
- Achieved 85% and 91% accuracy on testing dataset respectively for VGG16 model and ResNet model.

Face Detection with OpenCV *Spring 2018*

- OpenCV Cascade Classifiers was performed for face detection from the images in the CASIA-Webface dataset
- Minimized Triplet loss function to train embedding f(x) in CNN model, and both KNN and SVM are trained to classify the celebrities' faces
- Achieved 96% and 98% accuracy on test data respectively for KNN and SVM classifier, and embed data into 2D space to display identity clusters

Skills

Computer	Python, SQL, TensorFlow, Keras, R, MATLAB, Java, Hadoop, Amazon Web Service
Data Science	NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn, Plotly, XGBoost
Machine Learning	Predictive Modeling, Computer Vision, Recommender Systems