Zhuoru (Simon) Lin

Address: Apt 2E 251 W109 St New York, NY 10025|Tel: (440)935-5423 | Email: Simoninoc@gmail.com | LinkedIn: http://www.linkedin.com/in/simonlzr | Github: http://github.com/ZhuoruLin

Skills

Deep Learning Library: Tensorflow, Torch

Programming languages: Python (Numpy, Pandas, Scikit-learn), R, SQL, MapReduce, Spark

Education

B.A IN PHYSICS AND MATHEMATICS (GPA:3.75) | SEP 2011- MAY 2015 | OBERLIN COLLEGE

- Related coursework: Statistical modeling, Bayesian computation, Statistical mechanics, Advance Laboratory

M.S IN DATA SCIENCE | SEP 2016-PRESENT| NEW YORK UNIVERSITY

- Related coursework: Natural Language Processing, Big Data, Machine Learning, Deep Learning

Experiences

RESEARCH AND DEVELOPMENT INTERN| SATEST EDUCATION INC| MAY 2016 – AUG 2016

- Design the database structure for the dictionary module in SAT test preparation iPad app
- Researched automated English part of speech (POS) tagging method

INTERN | GOMARKETING INC| AUG 2015 – FEB 2016

- Assisted senior developers in optimizing database structure of JBSHotels.com, a global accommodation wholesaler

RESEARCH ASSISTANT | OBERLIN COLLEGE DEPARTMENT OF PHYSICS | FEB 2015 – MAY 2015

- Worked as a team with Professor John Scofield to analyze building energy consumption data
- Wrote a R Shiny App to automate models' cross validation
- Accessed and found significant validation and data analysis problem in Energy Star's building benchmarking system

RESEARCH AND PRODUCTION INTERN | GOHMATH.COM | DEC 2014 – JAN 2015

- Cooperated with director to produce test preparation videos for Massachusetts Tests for Educator Licensure

QUANTITATIVE SKILL CENTER TUTOR |OBERLIN COLLEGE CLEAR OFFICE | FEB 2014 – MAY 2015

- Helped students build their quantitative skills including data analysis, modeling and problem solving

RESEARCH ASSISTANT | OBERLIN COLLEGE DEPARTMENT OF PHYSICS | DEC 2013- JAN 2015

- Worked independently with Professor Yumi Ijiri to analyze neutron scattering data from National Institute of Standard and Technology (NIST) to study magnetic properties of nanoparticles (Iron Oxide and Manganese Ferrite)

Data Science Projects

PREDICTIVE KEYBOARD

- Mined blogs, twitters and news corpus
- Built a Naïve bayes model with Kneser-Ney smoothing
- Deployed a R Shiny App online that prompts user input and makes prediction on the following words

ALLSTATE CLAIMS SEVERITY PREDICTION

- Built a feed forward neural network model in Tensorflow to predict claims severity based on 120 features

MNIST HAND-WRITTEN DIGIT CLASSIFICATION

- Developed own methods of images translation, rotation and skewing for artificial label-preserving data augmentation
- Trained a Denoising Auto-Encoder for pseudo-labeling
- Trained a convolutional neural network for semi-supervised image classification. Accuracy was more than 99%

WORD SENSES DISAMBIGUATION ALGORITHM FOR TEST PREPARATION (IN PROGRESS)

- Mine pretrained GloVe word embeddings and WordNet senses databases
- Build a language model to predict the correct sense given a selected word in a sentence
- Create a better sense matching dictionary to facilitate foreign language studying