Changyao Chen

changyao.chen@gmail.com | www.linkedin.com/in/changyao-chen | www.github.com/changyaochen
917-438-8716, New York, NY

EXPERIENCE

Insight Data Science, New York, NY — May 2017 - Present
Data Science Fellow

- Processed, visualized, and analyzed 35 million Citibike trips from 2013 to 2016 using Python (matplotlib, bokeh) and PostgreSQL.
- Built a classification model to predict the destinations of CitiBike trips, by combining trip-level data and location features (*e.g.* US Census, Yelp search results), using Python (pandas, scikit-learn).
- Created a website (www.wisewheels.us) with flask and AWS, that makes interactive predictions based on user inputs.

Argonne National Laboratory, Lemont, IL — Aug. 2014 - May 2017
Distinguished Director's Postdoctoral Fellow

- Designed, tested, and patented new types of micro-mechanical clocks.
- Developed Python-based programs to improve the data (ASCII format) acquisition efficiency by 400%.
- Invited to speak at multiple international workshops and conferences.

Sendyne Corp., New York, NY — Feb. 2013 - July 2014
Modeling Engineer

- Designed experiments to measure voltage-current-capacity characteristics for commercial Lithium-ion batteries under different configurations.
- Developed computationally efficient runtime model to predict the remaining runtime and temperature of battery packs.
- Managed experimental data and simulation code (PSpice, C++) organizations.

EDUCATION

Ph.D., Mechanical Engineering — 2006 - 2013
Columbia University, New York, NY, *Chiang Chen Industrial Fellowship*

B.S., Precision Instrument and Mechanology — 2002 - 2006
Tsinghua University, Beijing, China

SKILLS

Machine Learning: general linear models, ensemble methods, deep learning.

- Collected and engineered college basketball team statistics from more than 75,000 regular season games (from 2003 to 2017).
- Built multiple models (random forest, logistic regression, neural network) to predict the final bracket for 2017 NCAA Men's Basketball Tournament. The overall performance exceeds benchmark.

Programming: Python (pandas, scikit-learn, xgboost, matplotlib, bokeh, flask, tensorflow), Matlab, R, C++, git.
Data Science: Hadoop, MapReduce, Pig, SQL, AWS products.
Statistics: A/B testing, hypothesis tests, continuous and discrete distributions.