Siinn Che

Address: 191 W. Woodruff Ave. Columbus, OH, 43210 • **Phone**: 510-967-5152 • **Email**: siinn.che@gmail.com
Blog: http://siinn.github.io • **Github**: github.com/siinn • **Linkedin**: linkedin.com/in/siinn-che

SUMMARY

Ph.D. student with extensive experience in **data analysis** with large experiment data at **CERN**. Proficient in data modeling, manipulation, and visualization with 5 years of experience in **Python** and C++. Strong background in **mathematics** and **statistics** from academic training. Enjoys working with others in a **collaborative** environment. Highly motivated to make a positive impact in **data science** with quantitative skills, knowledge and training.

RESEARCH EXPERIENCE

CERN (The European Organization for Nuclear Research) Geneva, Switzerland, 2013 – 2018
Graduate Researcher

- Data Analysis: Processed and analyzed **large data** (~200 TB) from **ATLAS** experiment with **reduction rate** < 5% in search for new physics.
- Data Mining: Developed a **data mining** algorithm to **improve** the efficiency of extracting rare physics data from 20% to over **90%**.
- Modeling: Designed Monte Carlo simulation of complex physics processes using random sampling and **statistical modeling**.
- Visualization: Developed and implemented a **visualization** tool for experiment data at the Ohio State using Python Web App

LBNL (Lawrence Berkeley National Laboratory) Berkeley, CA 2009 – 2012
Research Associate

- Modeling: Designed a simulation of complex molecular dynamics experiment using MATLAB

EDUCATION

Ph.D.	Experimental **Particle Physics**, The **Ohio State** University	2014 – 2017
	"Search for Long-lived Neutral Massive Particle with Displaced dilepton Resonance at the LHC"	
M.S.	Experimental Particle Physics, The Ohio State University	2012 – 2014
	"Higgs Boson as a Tool for Discovery at the LHC"	
B.A.	**Mathematics** & Physics, University of California, **Berkeley**	2007 – 2011
	Academic Achievement Scholarship Award, 2008, Linear Algebra (A+), Computational Physics (A)	

INDEPENDENT PROJECTS

End-to-End Project:	*NYC Rent Prediction*: Regression models are built to predict apartment rent in NYC using **RandomForest** and **GradientBoosting** regressors using data collected by **web scrapping**. Two regression models are compared using **learning curves** and **cross-validation** metrics.
Selected Kaggle Project	*Recognizing hand-written digits*: Analyzed and classified images of hand-written digits from MNIST dataset using Principal component analysis **(PCA)** and unsupervised learning (**K-Mean clustering)**.
In-course Projects	Natural language processing, Recommender system, SVM, Geographical Plotting.

PROFESSIONAL SKILLS

Data science	**Python** (pandas, numpy, scikit-learn, matplotlib, seaborn), **SQL**.
Machine Learning	**Scikit-learn** (model selection, regression, classification, clustering, PDA).
Research	C++, ROOT statistics package, Distributed computing.
Development Environment	UNIX (bash, ssh), svn, **Github**, LaTex, **Jupyter** notebook.