

**Dokyun Lee**  
Jon M. Huntsman Hall 5th Floor  
3730 Walnut Street  
Philadelphia, PA 19104

917.750.9963 (phone)  
Leedokyun@gmail.com  
www.leedokyun.com

## Education

- **Wharton, University of Pennsylvania** Philadelphia, PA  
*Ph.D in Information Systems and Economics* Aug. 2010 - May. 2015
- **Yale University** New Haven, CT  
*Masters in Statistics* Sep. 2009 - May. 2010
- **Columbia University** New York, NY  
*Bachelors in Comp Sci -Machine Learning & AI track* Sep. 2005 - May. 2009
- **Stuyvesant High School** New York, NY  
*High School* Sep. 2001 - June. 2005

## Research Interests

Information Systems and Economics, Causal Inference, Network Science, E-Commerce.

## Work Experience

- **Thompson Reuters** New York, NY  
*Machine Learning (Natural Language Processing) Engineer* May 2010 - July 2010
  - Building a Classifier/Automated Information Extractor Natural Language Processing Engine for any unstructured market data news feeds such as BusinessWire by using SVM/Proprietary Algorithms/Heterogeneous Ensemble Learning with an extreme low latency (few ms) from scratch
- **Snac Inc.** New York, NY  
*J2ME Mobile Widget Developer* October 2007 - June 2008
  - Use of Jscript and proprietary language to develop widgets that run on java enabled cellphone
  - Used ERD, Created database schema, backend php, info scraping php scripts
- **Blackrock** New York, NY  
*Electronics Communications Network/ Trading Application Developer* June 2007 - August 2007
  - Wrote part of ALADDIN (Blackrock's Proprietary software for trading/portfolio/market analysis)
  - OOP PERL, Commons Chain Apache, ALADDIN serves 6000 clients, 200+ financial companies
  - Developed a new web application that reports external trade and internal trade matching error
- **Ourtaste** New York, NY  
*Ruby on Rails Developer Intern* January 2007 - June 2007
  - Added features to the website using ruby on rails
- **Center for Computational Learning System at Columbia University** New York, NY  
*Research Assistant for Machine learning algorithm development team* June 2006 - August 2006
  - Wrote variety of customized Machine Learning Software parts using Python

## Leadership Activities

- **Statistical Consulting** New Haven, CT  
*Statistical Consulting in a Class* 2010 - current
  - People from variety of different disciplines come in with data set and we are paired up to advise them with techniques that can be used to discern anything interesting from the data
- **Columbia ACM (Association for Computing Machinery)**  
*Vice President/ Start-up Corporate Affairs* 2007-2009

- Planning events, Social events, Tech talks, Startup talk, Programming contest, Hacking contest
- Invited: Steve Wozniak, Larry Walls, Co-Creator of VMWare
- Events with: Google, Microsoft, Facebook, Amazon, VMWare, IBM, Many Start-ups

- **Korean Student Association**

- *Webmaster*

2007

- Learned *Project managing/ defining user needs*. Fast Iteration.

- **Youth Community Project Team**

- *The Captain/Leader*

2001-2005

- Community service team entirely planned/organized by students. ( 50 students, +\$4000 raised, 6 fundraising events, 4 seminars, 3 screen plays, 2 educational videos produced. Members (including myself) received presidents award for community service). Led the entire team through discussion/planning : 1/wk for the entire team and 1/wk for the officers.

## Teaching & TA Experience

### Wharton School of Business, UPenn

2011

*Teaching Assistant (Management Information Systems OPIM 210)*

- Taught by professor Eric K Clemons: <http://opimweb.wharton.upenn.edu/people/faculty.cfm?id=26>

### Wharton School of Business, UPenn

2011

*Teaching Assistant (Development of Web-Based Products and Services OPIM 654)*

- <http://karlulrich.pbworks.com/w/page/33435276/2011-San-Francisco-Web-Workshop>

### Columbia University

2008

*Teaching Assistant (Intro to Computer Science in Java CS1004)*

- Hold review sessions, Grade Homeworks, Hold office hours

### ESL(English as Second Language) Class teacher

*Teacher/Organizer*

2002

- Planned and taught senior center ESL class. 15 students

## Original Research & Presentation

”Propensity Score Method in Network Data:A Simulation Study of Bias” Wharton, 2011 Summer

”Johnson-Lindenstrauss Lemma and its effect on Supervised Learning” 30 minute Presentation for Multivariate Statistics, Yale University, 2010

”The Effect of Random Projection on Learning Methods” Masters Research, Yale University 2009-2010

”Electricity Distribution Network Visualization” Undergraduate Research, Columbia University, Summer 2006.

## Honors

**2010-2014** Wharton Doctoral Fellowship

**2009** The Andrew P. Kosoresow Memorial Award for Excellence in Teaching and Service (Columbia University Computer Science Department)

**2005-2009** Columbia College Global Scholarship / Global Scholar

**2005-2009** Scholarships: Olegario Lorenzo Fund/ Oak Fund/ Starr Fund/ Hank Hannah Choi Fund/ etc

**2005-2007** Dean’s List

**2005** Dr. Richard Hampton Award for Excellence in Physics

**2005** Dr. Beller Memorial Award for Excellence in Physics

**2005** The Highest: Physics Achievement Award

**2005** The Highest: Physics Bowl Score

**2004** Citibank and Korean Air-Line Scholarship (Project Proposal essay winner for translation service)

## Skills

**Languages:** Java Perl Python Matlab R Lisp PHP C/C++

**Exposure:** Liblinear, Ensemble Method, ANTLR, WEKA, Information Retrieval, Large Scale Data Mining, Text Clustering/Classification, Personalization, J2ME, OLAP, ORDBMS, Oracle 10g, Data Warehousing, Web Scraping, Recommendation System, Expert System(built : <http://projects.seas.columbia.edu/save/>)

**Language:** Fluent/Write/Read in Korean