



# LEON LEE

Data Scientist | Machine Learning Engineer

leon.lee@berkeley.edu  
https://www.linkedin.com/in/leon1110  
https://github.com/leon-lee-jl

## EDUCATION

- University of California, Berkeley** Master of Information, Focus on Machine Learning 2016.08 - 2018.05
- Xi'an Communication Institute** Master of Communication and Information System 2007.09 - 2010.06
  - **Second Prize** for Military Science and Technology Progress 2009
  - **First Prize (top 4%)** in National Postgraduate Mathematical Contest in Modeling 2008

## SKILLS

- Languages** - Python, Java, C#, SQL, JavaScript
- Machine Learning** - Tensorflow, Scikit-learn
- Big Data** - Hadoop, Spark, HDFS
- Natural Language Processing** - NLTK, Gensim
- Database** - MySQL, SQLServer, HBase
- Tools** - Node.js, Docker, Vagrant, Unity, D3.js

## EXPERIENCE

- AutoCAD Data Scientist / Machine Learning Engineer Intern** Autodesk Inc. @San Francisco, CA 2017.06 - 2017.08
  - Worked as software engineer and data scientist to incubate future AutoCAD features by leveraging compound data.
  - Trained CNN model with transfer learning method to recognize CAD drawings, with 80% test accuracy.
  - Developed a search engine and web app to reorganize users' CAD library and help users retrieve content for reuse.
- Virtual Reality Developer / Researcher with Google Daydream Team** UC Berkeley @Berkeley, CA 2017.08 - Present
  - Developing a game to support children's making and storytelling in VR, a research project collaborating with Google.
  - Implementing real time data streaming from Kinect to Unity application with Oculus Rift and Google Cardboard.
- Full Stack Data Engineer** International Computer Science Institute @Berkeley, CA 2017.02 - 2017.05
  - Designed and implemented system infrastructure of AppCensus to examine mobile apps violations of Children's privacy regulations of COPPA, by monitoring and analyzing apps' network traffic data.
  - Led development of web application to visualize data collected by mobile apps, and analyze regulations compliance of censored apps.
- System Engineering Manager** Tianao Internet Technology Limited @Guangzhou, China 2014.03 - 2016.07
  - Executed log data mining to design an evolving travelling dictionary to analyze users' interests, which were used to dynamically adjust website layout and report monthly and quarterly hotspots to marketing and sales departments to adjust strategy, contributing to a 20% increase in business.
  - Led research in massive network packages analysis in Hadoop cluster to detect and visualize potential threat.
- Senior Software Engineer** 41st Group Army, PLA @Liuzhou, China 2010.07 - 2014.02
  - Managed network deployment and security protection of three-layer B/S structured logistics business system and solved data integration and real time sharing among different logistics departments.
  - Achieved data fusion and accessing of heterogeneous data in Integrated Commanding Platform and enhanced comprehensive data query and retrieval efficiency by more than 30%.

## PROJECTS

- Course Recommendation by Semantic Similarity Analysis (Natural Language Processing)** 2016.09 - 2016.12
  - Implemented UC Berkeley's course recommendation system for 537 courses by semantic similarity analysis.
  - Designed D2V+LSI algorithm to build similarity index, which multiplied LSI vector with D2V vector to solve discreteness of LSI similarity and closeness of D2V similarity; reached 90% students' satisfaction through A/B testing.
- iRaccoon Reading Management Application (goo.gl/ZZeT7q)** 2016.09 - 2016.12
  - Developed a personal reading application, which helps users scrape online articles, save and organize personal reading data, such as reading annotations and comments, and helps explore visualizations of reading habits.
  - Led the project and did full-stack development from database (MySQL), web server (Flask), to front-end (Bootstrap).
- Integrated Network Safety Mechanism Research (National High-Tech R&D Program)** 2008.08 - 2010.04
  - Researched integrated network key management, authentication protocol, simulation and realization.
  - Simulated designed protocol in OTcl and C++ by modifying agent on top of CMU Wireless Extension of NS2.
  - Realized combined public key generation and distribution system with Mathlib in VC; programmed in C++ with MFC.
  - Published four papers in Core Journals and a patent (goo.gl/b3Nkym).
- National Defense Data Dictionary Gateway** 2009.02 - 2009.08
  - Developed an application layer gateway for early detection and prevention of SQL injection and XSS attacks.
  - Implemented SSL authentication module and integrated with inner PKI system; programmed in Java with JSSE.
- Data Access and Authentication System** 2007.11 - 2008.04
  - Developed a system that initiated an encrypted explorer in U-disk to meet needs of removable device security, and fulfilled core data access control for different users by authenticating with certificate embedded in the explorer.
  - Led software module and database design, programmed in C++ with MFC and tested performance with LoadRunner.