Chih-Feng Lin

120 York St APT 612, Jersey City, NJ 07302
(650)-526-8759 | chihfenl@alumni.cmu.edu | github.com/ChihFengLin | www.linkedin.com/in/chihfenl

Education

Carnegie Mellon University | Pittsburgh, PA
M.S. in Electrical and Computer Engineering | 12/23/2015
National Taiwan University | Taipei, Taiwan
M.S. in Applied Mechanics | 07/31/2011
National Taiwan University | Taipei, Taiwan
B.S. in Civil Engineering | 06/30/2009

Professional Experience

VP, Big Data Platform Architect | 01/15/2018 – Present
TD Securities | New York, NY

- Architect and develope big data backend system to support derivatives valuation service.

Data Engineer | 02/15/2016 – 12/31/2017
Bomoda (Acquired by Weber Shandwick) | New York, NY

- Designed ETL Process: Built a message queue system and API to digest raw data from vendor, integrated with cron jobs to automatically ship raw data to S3 data storage and a real-time error tracking system such as Sentry to improve the reliability of the process.
- Developed Python And Scala Based Backend System: Transformed the conventional single-machine data processing system into a distributed architecture that is deployed on Amazon AWS cluster using Spark framework, which significantly reduced the computation time from 10 hours to 1 hour and converted the data into clean and structured format.
- Bridged From Raw Data To Readily Accessible Information: Established the cron jobs that automatically ship cleaned data to the connected BigQuery database and designed all the schemas to normalize data, provided with an easy and queryable interface for data strategy team, which successfully doubled the company's productivity.
- Managed Product And QA Environment: Utilized Docker and Vagrant to build deployment-like server environment on the local machine. Programmed Ansible playbook to automate ETL process locally and deployed code on the product server after passing testing.

Software Engineering Intern | 06/21/2015 – 08/15/2015
Renault Innovation Silicon Valley | Sunnyvale, CA

- Prototyped In-Car Seat Driver's Heartbeat Detection System: Developed system including hardware and software. Embedded sensor into car seat to capture vibration signal from driver and communicated digital signal with computer via Arduino.
- Created Heartbeat Detection Algorithm: Analyzed the extracted siganl and designed a new algorithm based on time-frequency analysis to convert vibration siganl to human readable heartbeat number. Successfully achieved error rate within 10% as compared with ground truth.

Skills

Language:	Mandarin (Native), English (Fluent)
Programming:	Python, Scala, Java, SQL, Bash, C, C++, Matlab
BigData:	Hadoop, EMR, Spark, SparkSQL
DataAnalytics:	Machine Learning, Signal Processing
Database:	BigQuery, MySQL, MongoDB, HBase, S3
DevOP:	Amazon AWS, Ansible, Docker, Vagrant, Jenkins, NSQ, Kubernetes
WebServer:	Tornado, Flask, Tomcat
Others:	UnitTest, Git, Maven, Vim, Sentry, LaTeX

Professional Experience - Foreign

Software Engineer | 09/15/2012 – 09/30/2013
Innolux Corporation | Tainan, Taiwan

- Conceptualized and implemented new algorithm for Low Color Shift technique of oblique viewing angle of LCD using C and Python. Significantly reduced the color shift on oblique view by 30%

Research Development Engineer | 08/31/2010 – 09/01/2012
Shin Kong Memorial Hospital | Taipei, Taiwan

- Explored novel signal processing techniques applied in the sleep research field. Successfully created a new algorithm and python-based diagnosis model for Alzheimer's via sleep EEG. (http://goo.gl/k5Yu9i)
- Published a new feature of sleep EEG regarding non-dementia controls and Alzheimer's patients (Abstract has been issued on the journal of Sleep and Biological Rhythm).

Selected Projects

Predicting Aircraft Fuel Consumption Using Deep Learning - NASA Sponsored Project

- Established Stacked Autoencoder (SAE) model with 182 aircraft feed-in features and constructed the training workflow for multi-layer network architecture, which can be effectively applied to other prediction tasks. Succeeded in improving prediction rate by 40% on testing data as compared with pure linear regression model. (https://goo.gl/tcTZdR)

Cloud-Based Twitter Analytics Web Service

- Analyzed Twitter information and conducted the ETL process. Extracted one terabyte of Twitter data based on Python language. Transformed data into the correct format using the MapReduce architecture (Hadoop framework) and stored the data in the AWS S3 Bucket. Loaded parsed data into MySQL database on the Amazon EC2 instance (Linux environment).
- Designed MySQL table schema and specialized in performance adjusting, query optimization and index tuning. Successfully improved query response efficiency by horizontally scaling the MySQL database instance with Elastic Load Balancer (ELB).