

LONG WANG

Email: lwang010@gmail.com | Tel: 412-880-9374
Github: [longw010](https://github.com/longw010) | Blog: <https://longw010.github.io/>

EDUCATION

Carnegie Mellon University (CMU) Pittsburgh, PA, US	May 2018
Master of Science in Computational Biology, GPA: 3.63	
Huazhong Agricultural University (HZAU) Wuhan, China	Jul 2015
Bachelor of Science in Bioinformatics, Valedictorian, GPA: 3.68	

WORK EXPERIENCE

Subtle Medical. Melon Park, CA.	Aug 2018 - Present
<i>Machine Learning engineer</i>	
<ul style="list-style-type: none">Developed and validated prototype to improve MRI images with various acceleration factors in plane, and resolve the quality issuesRan benchmark testing on hardware and different inference platforms	
SemanticMD, Inc. Pittsburgh, PA.	May 2017–Aug 2017
<i>Deep Learning R&D Intern</i>	
<ul style="list-style-type: none">Took the lead on 3 deep learning pipelines developments, including sperm cells detection and recognition, diabetic retinopathy patient identification, and a 2-chambers vs. 4-chambers ultrasound videos classification.	
Jianbing Yan's Lab <i>National Key Laboratory of Crop Genetic Improvement.</i> Wuhan, China.	Aug 2014–Jul 2015
<i>Research Assistant (published one paper)</i>	
<ul style="list-style-type: none">Developed a tree-based machine learning method to detect SNP associations with rare variants in maize	
Lingling Chen's Lab <i>College of Informatics, HZAU.</i> Wuhan, China.	Jul 2013–Jun 2016
<i>Research Assistant (published three papers)</i>	
<ul style="list-style-type: none">Developed an ensemble model to infer transcriptional regulation using RNA-seq in Prokaryotes.Developed a statistical method to predict protein-protein interaction network and performed network analysis.	

SELECTED COURSE PROJECTS

N-Gram probabilistic model (Java, HBase Independent)	
<ul style="list-style-type: none">Ran ETL on raw data and implemented a N-Gram language model using MapReduce to predict the next wordRead plain text input file from HDFS, wrote statistical result to HBase, and connected the client interface with database to run phrases auto-completion.	
Diabetic Retinopathy Classification with Imbalanced, Low-resolution retina images (Python Independent)	
<ul style="list-style-type: none">Designed and implemented a pipeline to do the classification (applied contrast limited adaptive histogram equalization (CLAHE) to enhance the contrast of the images, constructed a GAN to run data augmentation and applied transfer learning to avoid overfitting; embedding resolution-aware CNN with VGG16)Improved the accuracy from 92.9%(baseline) to 96.8%.	
Machine learning in Large datasets (Python Independent)	
<ul style="list-style-type: none">Designed an efficient way to run a scalable regularized logistic regression on stochastic gradient descent (10x faster than the classical LR model in the 0.2x memory)Implemented an automatic reverse-mode differentiation platform to train the CNN models (such as LSTM)Implemented a memory-optimized label propagation algorithm in Spark	
3D Structure Reconstruction of Human Chromosomes from Hi-C data (Python Team)	
<ul style="list-style-type: none">Generated a polynomial curve fit between the interaction frequency (Hi-C data) and relative distance among nucleosome beads using the least squares method.Estimated the 3D coordinates for each nucleosome bead by posing the distance information to a constraint convex problem (find the minimal from the objective function using gradient descent)	

SKILLS

Programming: Python, Java, Scala, Bash, R, Matlab, Golang, Perl, C++.
Tools: Hadoop, Spark, Tensorflow, Pytorch, OpenCV, Hbase, MySQL, HDFS, MapReduce, Kafka, Scikit-learn.

ACADEMIC HONORS AND AWARDS

MSCB Merit Fellowship , <i>Computational Biology Department. CMU, USA.</i> (twice)	Aug 2017
Honorable Mention, Interdisciplinary Contest in Modeling , <i>Comap Inc, USA.</i>	Apr 2014
First Prize, Chinese Undergraduate Mathematical Contest in Modeling , <i>Ministry of Education, CN.</i>	Nov 2013
National Scholarship , <i>Ministry of Education, China. (for top 0.5% students)</i>	May 2012