Xi (Bill) Chen | Resume

+1 857 209 1002 • +1 309 455 4260 • billchenxi@gmail.com
https://billchenxi.github.io • https://www.linkedin.com/in/billchenxi/

Graduate bioinformatician with Statistic MA certification completing the final year of a Ph.D degree. Passionate about Big Data, Machine Learning and AI research, with strong technical and interpersonal skills, adept at working in teams and successfully delivering projects. I was awarded **Microsoft Azure Research Funding** in 2016 and gained a **Grow with Google Challenge Scholarship** in 2018.I am a certified **SAS** programmer and certified **Deep Learning** and **Cuda** instructor.

Education

- **University of Kentucky** — **Kentucky**
  Ph.D. in Biochemistry (Bioinformatics), — *2013–2019*
  Dissertation: Automatic 13C Chemical Shift Reference Correction of Protein NMR Spectral Data Using Data Mining and Bayesian Statistical Modeling (https://doi.org/10.13023/etd.2019.057)
- **University of Kentucky** — **Kentucky**
  Master (Cert) in Statistics — *2016–2019*
- **Udacity**
  Self-Driving Nano-degress (Expected) — *2018–2019*

Technical Skills

- **Programming Languages:** CUDA, C++, R, Python3, SAS and Matlab (OK). Please refer to:
  - **Hackerrank:** `https://www.hackerrank.com/billchenxi`
  - **NVidia:** `https://nvidia.qwiklab.com/public_profiles/5521a192-c2e9-4899-9750-500959646159`
- **AI/DL/RL/ML Knowledge:** TensorFlow, PyTorch. Please refer to:
  - **Courser:** `https://www.coursera.org/user/902cbf6a236e768db1c185c96778971b`
- **Cloud Computing Skills:** AWS, Google Cloud (GCP), Shiny. Please refer to:
  - **GCP/AWS**: `https://qwiklabs.com/public_profiles/032b735c-3942-4f65-96e2-46bc821a884a`

Publication & Software & Packages & Talks

- Papers:
  - Automatic 13C Chemical Shift Reference Correction for Unassigned Protein NMR Spectra, *Journal of Biomolecular NMR, Aug 2018* DOI: 10.1007/s10858-018-0202-5 (link)
  - Parallelized Interactive Machine Learning on Autonomous Vehicles, *NAECON Dec 2018* DOI: 10.1109/NAECON.2018.8556776 (link)
  - Deep Learning by Doing: The Nvidia Deep Learning Institute, *Journal of Computational Science Education, Dec 2018* DOI: 10.22369/issn.2153-4136/10/1/16 (link)
  - Pan-Cancer Epigenetic Biomarker Selection from Blood Sample Using SAS®, *MWSUG, Sep 2018* (link)
  - Finite Mixture-of-Gamma Distributions: Estimation, Inference, and Model-Based Clustering, *Advances in Data Analysis and Classification, Jan 2019*
- Package
  - BaMORC *CRAN, 2018* CRAN GitHUB
  - BMRBr *CRAN, 2017* CRAN GitHUB
- Workshop
  - CUDA Programing Workshop *(UK ACM 18)*

- Deep Learning for Computer Vision Workshop *(UK ACM 19)*

Notable Projects

- **(Deep Learning) Deep Learning for Cancer Classification with Gene Expression Data:** *Deep learning approach to classifying cancer types. Data from NIH Cancer Data public repo, with 32 different cancer types. The model accuracy is more than 97% with 0.2% false positive/negative rates.*
- **(Reinforcement Learning) Parallelized Interactive Machine Learning on Autonomous Vehicles:** *Development of an interactive reinforcement learning pipeline for autonomous driving from a game simulation. The simulator is from Microsoft AirSim and driving environment is simulated from Unreal Engine.*
- **(App Prototyping/Dev) CatsHub App Entrepreneurship:** *Development of student-oriented phone application that simplifies student's life by providing a one-stop shopping service hub. The original idea was from the University of Kentucky student organization. We pitched and created a prototype, and won venture reward.*

Previous Employment

- **Deep Learning Institute** **NVIDIA**
  University Ambassadors / Deep Learning Institute (DLI) Certified Instructor *2018–Present*

  Certified to teach Fundamentals of Computer Vision (CV), Fundamentals of Deep Learning for Multiple Data Types (MDT), and Fundamentals of Deep Learning for Natural Language Processing (NLP). Contents include: Cuda, Deep Learning, Neural Network, Image Classification, Object Detection, Image Segmentation, Word Generation, Image and Video Captioning, Text Classification, Text Translation, and etc..

  Technologies: TensorFlow, Cuda, PyTorch, Caffe2, Keras, etc.

- **Moseley Bioinformatics Lab** **Lexington KY**
  Graduate Research Assistant *2013–Present*

  Worked on construction of Protein NMR Reference Correction and Protein NMR Deuteration Level Detection frameworks. Designed and implemented the corresponding software package (BaMORC) and web applications systems. Created R package 'BMRBr' for easy download and processing of BMRB files.

  Technologies: R, R Package, RStudio, Python, Shiny, Docker.

- **Prof. Derek S. Young, Dept. of Statistics, University of Kentucky** **Lexington KY**
  Research Collaborator *2017–Present*

  Worked on simulation for the Mix-Gamma Model with R. Jobs included generation of a large amount of simulated data from gamma distributions, implementation of the testing code, unit test, and preparation and design of submission shell code for HPC (High-performance Cluster).

  Technologies: SAS, R, R Package, HPC Management, Bash

Reward and Certification Highlights

- SIAM International Conference on Data Mining Dissertation Travel Award (2018)
- MWSUG Conference Paper Scholarship (2018)
- Grow with Google Challenge Scholarship (2017/18)
- SC18 Paper Scholarship (2018)
- SAS Certified Clinical Trials Programmer, Statistical Business Analyst, Advanced Programmer
- Deeplearn.ai Certification Series (Coursera)