WANG LEI

E-Mail: lwang11@mtu.edu • Mobile: +906-231-3888
Term-time Address: DH Apt. 2015G, Woodmar Dr. Houghton, MI US

MAIN TECHNIQUE SKILLS INTRODUCTION

- Programming Language:

  - **expertised** in **Python(4 years)**. Open Source Package developer, full stack website developer. Open to open source. Familiar with python web server side development. **Have product experiences in Django, Flask and other python popular frameworks**. Familiar with RESTful api design. Focus on web safety development and server side database development. I use python mainly for business logics development.

  - C/C++ **Advanced** (6 years). Familiar with generic C development and Python C api. Familiar with Modern C++ (C++11/14 standard), C/C++ OOP. **Have hands on experience in C/C++ open source tool chains: Make/CMake, boost, glog/gflags, openGL, opencv, caffe, flex and other tools related to computing and graphics**. use C/C++ mainly for algorithms modules or modules demand high reliability and speed.

  - JavaScript(2 years experience). CJS packages developer. Familiar with ECMA5 and ECMA development tool chain. **Have product experience in working environment or Native Javascript development frameworks: node/webpack, React Native, html5/browser API, CSS3, WebGL, front tend template rendering engine**. I use javascript and HTML5 partially to support my UI development for better UX experience; partially as my first choice of human interaction interface to development product primitives.

- Dynamic Content Crawling & Website Development(Python & JS):
  - Website programming: experienced in ECMAScript5/6, node.js, phantom.js, XHR, HTML5 api, CSS and other front end tool chain.
  - Experienced in Python2/3 server side tool chain: urllib, Python2 selenium, twist api, scrapy, django, django-rest and other server side tool chain.
  - Proficient in RESTful api development; Familiar with HTTP 1.1 protocol(RFC 2616, 2617). Watching on HTTP 2.0 (RFC 7540).
  - Have absolute practical experience in WSGI(Apache Server >= 2.2) service deployment and release.

- Big Database Development Skills:
  - Familiar with CentOS 6.0 environment based tool chain: **Product experience in Hadoop, Hive, Spark, Jupyter-python, Linux txt processing tools and shell programming.** Understand Map Reduce based Hive HQL.
  - Familiar with product release procedure. Familiar with git, svn(some companies still host their projects using svn).
  - Have **work experience** in MySQL, Postgresql, Mongodb, Cassandra database (NoSQL)
  - Data scientist skills to support product operations work; discover and provide data intelligence for superiors.

- Machine Learning

- Hands on experience on various regression algorithms
  - Logistic Regression, Lasso Regression, Quantile Regression, CCA, PLSR (code),
  - Different artificial intelligence algorithms like 3 layers BP NN Algorithms, VGG net and its modification.
- Carried out research on Pls1 and PLSR and proposed a new method so sand feasible domain is expanded. Test results were positive to support my claim.
- Carried out research on various segmentation and aggregation algorithms in research institutes' programmes.
  - Otsu method, Random Walk, KMeans++, KNN (Learning to Hash), Level Set, differential geometry method for segmentation (PDE & differential geometry) and so on
- Carried out research using caffe on latest SSD & Fast-RCNN for autonomous vehicles multiple objects end to end recognition at multiple scales .
- Deep Study on C4.5, Adaboost, xgboost alorithms. Have half a year experience in xgboost algorithms in industry.
- Hands on experience in several industry ML frameworks.
  - Weka (java), sklearn (python), caffe、xgboost、tensorflow（c++）

- Mathematical Modelling
  - pattern recognition on direction of faces of people
  - Mathematics Analysis on Image Processing
  - Advanced Image segmentation Implementation Coursework
  - MCM/ICM: Modeling and Optimum of Management of the Park, CUMCM project

---

PHD RESEARCH INTREST

OCEAN DATA PROCESSING. MACHINE LEARNING.

EDUCATION

2015 - 2016 **Michigan Technological University**   **USA**

PhD Candidate in Mathematics, Full Scholarship

- Teaching assistant.

- recent works:
  - Finite Element Lagrange Method (code),
  - Finite Element Aggris Method (code),
  - Linear Algebra Survey

I decided to quit the program to purse Deep Learning relevant programme.

2013.8 – 2014.7 **Nanyang Technological University**   **Singapore**

Msc, in Digital Media Technology, early awarded, Computer Engineering

- Successfully finished Data Mining system in python: **Dissertation**, **codes** Provided

- Participating different programming contests: e.g. Regular Expression Parser in C++: **codes** provided
- Successfully finish and design Graphic Packages: <u>MeshViewer</u>, <u>Editable B-Spine Curve</u>, Animation IK algorithm, <u>A* Algorithm</u>

| 2009.9 – 2013.6 | **Chongqing University** | **Chongqing, China** |

- 4 yeas of study in mathematics and image processing

PROFESSIONAL EXPERIENCE

| 2016.11 - 2017.2 | *Autonomous Driving (L3, L4) Lab, Tsinghua University* | **Peking, China** |

Implement and apply end to end SSD algorithm on Tsinghua Mengshi Autonomous Vehicles.(Successful)

Reading & Learning Related Phd Papers, propose new information fusion algorithms to improve the reliability of the system.

| 2016.6 - 2016.10 | *DIDI Research Institute, <u>didichuxing</u>* | **Peking, China** |

Top Talents Internship Programme 2016 Member, R&D

- Kick off **Quick Car Intelligent Dynamic Price** project. Diving into DiDi production environment quickly, I delivered a new model combined with Linear Programming and Personal Price-Sensitivity Curve estimation. By drawing analogy between Computing Advertisement and Quick Car sharing pool model, I provide a detailed implementation solutions.

- Designing offline experiments and using xgboost algorithm to train data; established a ML codes base to leverage multi task problems. (<u>workshop</u>). A new HQL parser and unity test entry have been established within my work. This boosts my group productivity significantly.

- Using 2 days, finish a new project "Online Excel" using app flask so that operation members can edit configuration file in a much more friendly manner just liking using MS excel…

- Kick off **Expected Price Interval project** directed by Yu Jun. **We finished our KPI in a month**. In this project, **we first use Quantile Regression**, and construct a method to improve estimation qualities. This method can be applied in various similar projects.

| 2016.4 -
2016. 10 | **_Tiger Adsearch Member_** | **CA, USA** |

Principle algorithms designer, online/offline data designer

- Algorithms(inverse-index, tokens mathing, logistic Regression, Feature Engineering), CTR prediction pipline designer

- Data Updating Schedule (for locked dataset); hot/cold data exchange

- Index analyzer (Logloss, AUC, RMSE)

| 2016.2 -
2016-5 | **_WeiBoYi, Pte. Ltd_** | **Peking
, China** |

Data Mining Engineer (working logs)

- Kick off WeiBoYi OpenData Api project and responsible for DB design, api development, and standardisation.

- Lead Data Visualisation development and help on debugging the relevant, but not limited to data problems

- **Released** a Django-Restful based API (v 0.1.3); Customised a NestedRouter (Propose for PR in Django-Restful community, it is new design different what is existing in Django-Restful community); Build a customised Token and Signature based Auth(RFC 2616, 2617, i.e HTTP 1.1) base on the need that fast sharing in inner group. Hence I make it true to use url communicate date in RESTful style where the company use SQL to fetch data and transfer it in different machine; Support both Cross Origin and Jsonp visiting style.

- **Released** Visualisation package (CommonJS standalone js css files); Build a debug environment (Support google style searching and rendering) platform for my use and teach interns how to use it.

- Debug Support: Rlease deployment (RedHat CentOS), unity tests, documentation.

- Push the group to record everything they done in docs

The job was not a good match, and I am currently actively pursuing a good match.

| 2015.8 -
2016.3 | **_Guuud & ReadyObject_** | **Singapore** |

Senior Web Programmer

- Dynamic Rich Content Crawler Anonymously,
  has already passed experiment phases, now under development

2013.12

	• <u>ReadySales</u> developer. Sample work, <u>third party auth app for linkedin</u>(<u>code</u>)	2015.07
2014.8 – 2015.8	**Energy Research Institute @ NTU**	**Singapore**

Research Associate

- ICT Project: <u>Energy Analysis Dashboard(AJAX)</u>

2014.2 – 2014.8	**Agency of Science, Technology and Research (A*STAR)**	**One North Singapore**

Research Intern, Algorithm designer, Human Speech Lab

- Improve and implement Partial Least Regression Algorithm based on my mathematical deduction and experiment: **<u>Report provided</u>**

2012.7 – 2013.4	**Neusoft**	**ShenYang, China**

Entry Level Software Engineer Trainee

MORE ABOUT PROJECTS

| <u>Signautre Authenti-caiton Pack-age</u> | I freshly publish this package on Github (16 Oct. 2016). Signature Authentication proposal-5 proposed in around 2015 (the quality of the previous is not so good). Few projects can support it, and they mainly use digest authentication. After reading some coders work, I find that they are not what the proposal tells (e.g. signature package supported by django-rest framework). Hence I decided to made my own.

I also provide you a demo project (RESTful api based) so you can navigate through it more quickly. You can refer to my project based on GNU License. |
|---|---|

Ajax Crawler for SNS	①Wrote web data AJAX (dynamic) crawling tool (webkit engine based) and analysis (machine learning and data mining).
	②One approach I did is using selenium to communicate with webkit env host. I hence submit a pull request to scrapy(extension to http11 module of scrapy). Then I quickly realise selenium is not efficient for communication using wireless protocol for crawling.
	③I found that I can mimic method used by "ghost driver" (core of selenium to support phantom) but use another version of protocol for communication…. Then I did a lot of experiments on Phantomjs which is a headless browser. I use both http and inter-process methods to do my work. It was successful.
	①WeKnow (**team leader** download our presentation for VC): I am the people who build the team and fully immersive the product life circle from UI designing, usage scenario to technical implementation: **Demo** Provided. Our vision is "type less, find more from others!"
Mass Data Reader	①**Mass data reader** provided that data cannot be loaded into memory, using double hash strategy, programming in C++
	②Firstly I spliced a very large file into small files and read them into memory buck (1024) and sort them. Each file will be mapped by a hashtable I wrote. This is the first level hash map. For each hash key I used second level hash table to indicate each individual item.
	③The very important things are "popup search method" and random variables design.
	①I found ORM is extremely hard for complex sql execution. I hence made my own **SQL lexcial parser** based on the technique I picked up when I was developing a lexical parser for regular expression.
PyMatrix	①**Author** of **matrix_array**, which is complement of numpy, published in PyPI. N-Demension data container, brand new design to overcome the following problems in numpy (2014):
	②Numpy str is not elegant, well designed. Numpy didn't support object container concept for universal programming purpose (Latest version has supported object container by default). Numeric representation is not elegant, for example dot and element wise dot. In a mater of fact both of them.
ERIAN Dashboard	①**Creator** of NTU ERIAN lab, Energy Analysis Dashboard(AJAX). Big Data Enterprise Architecture Design and Project Deployment: Three Ties Framework: DAO design(New framework based on Finite SQL Template States Machine for universe purpose based on oracle mysql connector implementation) -->Database Horizontal and Vertical Sharding --> Statistic Assumption (Filter) --> Data Preprocessing --> Data Mining (KMeans++, Market Basket Analysis) --> Machine Learning (Energy consumption prediction based empirical data).

EXTRA-CURRICULAR ACTIVITIES

2012.7 - 2013.4	Design Pattern, Advanced C++ training, Final Big Project	**Neusoft**
2014.8- present	NTU Toast Master Club Member	
2014.3- 2014.5	Image and video processing: From Mars to Hollywood with a stop at the hospital, certificate: outstanding, <u>Licence</u> 9EB534GEY9	**Duke University**
2011.2 - 2012-12	Image processing, Student Research Training Program (SRTP) funded by Chongqing University (CQU), team leader certificate : successful Licence 120434	
2010, 2011	Mathematic Olympics: awarded third price in Chongqiqng by China Society for Industrial and Applied Mathematics (CSIAM)	
2011.7	Mathematical Modeling, Shenzhen Cup, problem A, top three in Chongqing University.(7, 2011)	
2010.9 - 2011.9	Appointed by CQU Student Union (CQUSU) as President, Debating Team of Mathematics and Applied Mathematics (9, 2010 - 9, 2011)	
2010.3 - 2011.3	Vice President, College Student Mathematical Modeling Association	

<u>CONTACT LEI ON LINKEDIN</u>