# Sitong Li

Edmonton, AB | sitongli1993@gmail.com | https://sitongli.github.io/ | +1 (226) 503 6105

## Skills & Others

**Programming**: Python, Java, JavaScript, XML, XSD, SQL, JSON, HTML, CSS

**Tools**: Apache Tomcat, Java Servlets, SPSS, Django, Android Studio, AWS, MongoDB, Kibana, scikit-learn, TensorFlow, PyTorch

**Courses**: Algorithm Design and Analysis, Internet Algorithmics, Readings in Algorithmics, Introduction to Data Science, Unstructured Data, AWS Essentials, Advanced Artificial Intelligence, Machine Learning, Probability, Linear Algebra, Calculus, Stochastic Process

**Interest**: Algorithm and problem-solving

## Education

**Western University**, London, ON — M.Sc. Computer Science (GPA: 4.00 / 4.00) — *Sep. 2018 – Nov. 2019*

**University of Science and Technology of China (USTC)**, Hefei, China — PhD neuroscience — *Sep. 2012 – Jun. 2018*

**University of Science and Technology of China (USTC)**, Hefei, China — B.Sc. Life Science — *Aug. 2008 – Jul. 2012*

## Project Experience

**AWS Based Job Recommendation Web Application Development** *Jan. 2020 – Feb. 2020*

- Created Java servlets with RESTful APIs to handle HTTP requests and responses.
- Built MySQL database on Amazon RDS to store position data from Github API.
- Designed algorithms (e.g., content-based recommendation) to improve job recommendation based on search history and favorite records)
- Deployed server to Amazon EC2 to handle 150 queries per second tested by Apache JMeter.

**Forum Platform Web Applications using Python Django Framework** *Dec. 2019 – Jan. 2020*

- Build a forum web application from scratch, including website wireframe, relational database design, implementation, optimization.
- The forum platform web application supports admin, user authentication, customized forms, various views, and automated tests.

**A Study on Deep Convolution Neural Networks for Salient Object Detection** **Supervisor: Charles Ling**

*View link to see the report* *May. 2019 – Jun. 2019*

- An introduction of visual saliency related topics with special emphasis on how SOD is related to other computer vision problems.
- A comprehensive review of how deep neural networks evolved for image classification and the development of convolutional neural networks for SOD problem.
- An empirical study of the most recent best-performing neural net for SOD- BASNet, including reproducing the experiments and an application of FOCAL loss in BASNet.

**Titanic: Machine Learning from Disaster**

*View this on Kaggle* *Feb. 2019 – Feb. 2019*

- Visualization to explore the features; imputation using median or mode value; classification of subjects by parsing their titles, then using the mean value of each class to impute the missing age values; discretization of continuous features and one-hot encoding.
- Tuning hyperparameters in models KNN, logistic regression, SVC, and random forest; Model ensemble using weighted voting.
- Top 7% out of all participants.

**A Review on Recent Advances in N-linked Glycoproteomics** **Supervisor: Kaizhong Zhang**

*View link to see the review* *Jul. 2019 – Aug. 2019*

- Background introduction of glycoproteomics and an overview of existed approaches and main challenges.
- An in-depth survey of recent advances in glycoproteomics, which are divided into different domains, including fragmentation strategies, glycopeptide-spectrum matching algorithms, false discovery rate (FDR) estimation, etc.
- A chapter introducing glycoproteomics research involving machine learning algorithms.

**Tree Alignment: Algorithms and Applications** **Supervisor: Kaizhong Zhang**

*View link to see the review* *Nov. 2018 – Dec. 2018*

- Introduction of widely used measures of similarities between trees, including tree edit distance, maximum agreement, tree alignment etc.
- An in-depth analysis of the representative tree alignment algorithm Jiang-Wang-Zhang's Algorithm and a derived algorithm for alignment between similar ordered tree. Time complexity of various measures is analyzed and compared.
- Introduction of applications of tree alignment algorithms, including comparison of RNA secondary structures and extraction of fields from HTML search results (This part is mostly done by my group partner Shabnam Shabni).

**Twitter Sentiment Analysis** *Dec. 2018 – Oct. 2018*

- Cleaning data using strategies including removing stop words, stemming, lemmmatization, etc. Feature engineering using classical natural language processing model N-gram and tf-idf.

- Various machine learning models, including maximum entropy, random forest, naive bayes. The one with highest cross validation accuracy was determined as our final model.
- Two twitter data sets on topics "Trump" and "feminism" were downloaded from web, and processed to assess the general emotional tendency towards these topics.

---

## Working Experience

**Volunteering Research Assistant, Computational Psychiatry** *Dec. 2019 – Feb. 2020*

- Preparing various medical data, including electrical health record cleaning and MRI data processing using FreeSurfer.
- In-depth investigation of the medical data to propose valuable scientific problems.
- Using brain surface areas and brain region volumes to predict whether a subject has schizophrenia and study gender-dependent classification.
- Using the survey data collected from Hebei to predict age of a subject. Using NHIS data to predict K6 score.
- Data visualization to explore feature value distribution and correlation between features. Remove pointless features or features with too many missing values. Using Bayes target encoding to missing values.
- Using RFE, XGBoost, Boruta, and Encoder-Decoder network to do feature selection or dimensionality reduction.
- Developing both classical and most recent machine learning algorithms, using Random Search to tune parameters and cross validation to select model.

**Software Engineering, HealthGauge** *Sep. 2019 – Oct. 2019*

- Wearable device application development and testing.
- Android BLE development.
- Algorithm development for ECG and PPG signal alignment.

**Teaching Assistant of Information Systems and Design, Western University** *Sep. 2018 – Apr. 2019*

- Giving interactive lectures on database management in class, assignment consulting and marking.

**Research Assistant in Neuroscience, USTC** *Sep. 2012 – Jun. 2018*

- Being involved in multiple research projects on visual system development, and play principle role in two of them.
- Participating in experimental design, operation, data analysis, thesis writing and reviewing, etc..
- Obtain Kwang-Hua Scholarship.