**Xiangwen Liu**

**cmwenliu@gmail.com (903) 422-5139**

**Professional Summary:**

* Proficient in C++, Java, C#, .Net and Object-Oriented Design
* Experience in C++ with performance tuning and writing/debugging of parallel code
* Experiences of high performance parallel computation
* Ability to develop and optimize algorithms for efficiency and scalability
* Strong programming skills in Python, Java and Database SQL
* Solid understanding in Machine Learning algorithms, Statistics and Data Mining
* Experiences of Scientific tools including TensorFlow, Theano, Torch, Keras, NLTK
* Hands on experience in implementing LDA, Naive Bayes and skilled in Decision Trees, Linear and Logistic Regression, SVM, Clustering, Principle Component Analysis
* Experiences on Deep Learning models including CNN, RNN, LSTM, GAN, Autoencoder, RBM, DBM and DBN

**Education**

* Ph.D. May 2019 (Expected)

In Computer Science, University of Arkansas at Little Rock(UALR), AR, US

* Master. May 2014

In Computer Science, Texas A&M University, Commerce(TAMUC), TX, US

* Bachelor. June 2007

In Mechanical Engineering & Automation, Tongji University, Shanghai, China

**Work Experience:**

**Teaching Assistant TAMUC Aug 2012 – July 2016**

* Teaching graduate students on C++ and database in Lab classes
* Assisted professors to finish grading work of Computer Science courses

**Application Developer SIEMENS Jun 2007 – Dec 2011**

* tuned and wrote/debugged C++ parallel code
* Applied regression analysis to forecast sale of products and regional distribution
* Performed data profiling and data quality improvements in company Database
* Developed logical data models and physical data models using ER-Studio
* Created a nonlinear model to develop the motion detection of surveillance system

**Research Experience:**

**Graduate Research Assistant UALR Aug 2017 – Now**

* Implemented LSTM neural network model in combination with news data and tweets data conduct time series analysis dataset. Extracted new features and identified their importance
* Performed Latent Dirichlet Allocation (LDA) algorithm to conducted NLP topic modeling for BBC news dataset and find the most prevalent topic for each news.
* Predicted sentimental score using Support Vector Machines on text data using TFIDF embedding and word2vector (lightSVM-Multiclass)
* Analyzed unlabeled data using clustering and Silhouette score and using Convolutional neural network(CNN) and deCNN to transfer abstract features to new labeled data.

**Research Assistant Food and Drug Administration (FDA) Aug 2016 – July 2017**

* Improved the Detection Power for Ultralow-frequency Mutations of gene with deep neural network
* Performed data acquisition, data preprocessing, data engineering, features scaling, features engineering, statistical modeling (decision trees, regression models, neural networks, SVM, clustering) on FDA Label data
* Improved image data detection with dimensionality reduction using Principal Component Analysis and Autoencoder, avoid overfitting with K- fold cross validation.
* Created Complex Queries, Stored Procedures, Functions, Indexes, Packages and Materialized Views to access data from database using SQL Server2008