

# YuKai Wu

email: [ywu048@fiu.edu](mailto:ywu048@fiu.edu)

call: 7866083426

## EDUCATION

- 2009.09—2013.07 AnQing Normal University (bachelor degree), major Information and Computing Science
- 2018.05—2019.12 Florida International University (master degree), major computer science

## PROFESSIONAL EXPERIENCE

- **2015.9 – 2018.1 HangZhou ShuYun Information Science Corporation (Algorithm Engineer)**
  - **Design and Development of Recommendation System for commodities**

The company started to develop a framework for Recommendation System which can train model from users features data and generates a list of recommendation commodities. In this project, My duty is to develop a back-end part that can get data from several recommendation lists and merge them by specific algorithm, finally push them to user end. I'm also in charge with developing some recommendation algorithms for some particular scenarios. This recommendation framework mainly serves the advertising business.
  - **Design and Development of TaoBao, TMall User Portrait System**

Depending on user shopping records, I establish a label system, representing user feature space, for all of them, such as Preference for Clothes, Preference for Cosmetics, Preference for Food etc. It is advantageous to divide group users into several categories and improve the precise of recommendation system.
  - **Analysis user comment data with NLP**

The company needs to extract short labels from user comments to help our customers fully understand what the user feel about their product. I use RNN with attention model to extract the key information in the short comment, analyse these data and display the results.
- **2015.6 – 2015.8 ShangHai YiChen Information Science Corporation (Big Data Development Engineer)**
  - **Design and Development of Recommendation System for short videos**

The company started to establish recommendation system. what I do is storing user's history data in Hive, training the Recommendation Models with Spark and storing those models in

HDFS. I develop a framework that automatically deploys the trained models that had been loaded in HDFS. Recommending short videos for users and putting the results into Hbase are my next work. The frame has RESTful API for responding the request from other front-end programme.

- **2014.5 – 2015.5 ANHUI XiangXing Information Science Corporation (Data Mining Engineer)**
  - **Development of Data Mining System Based on Spark Framework**

The company wants to create a cloud computing platform for the enterprise that is beneficial for them to analyse huge amount of data. Therefore, my task is to integrate Spark into this System and extend the function and algorithm of MLlib.
  - **Data Modeling of Customer Loss with Anhui Telecom User Data**

Help our customer to build a ML model to predict whether the user would lost or not in the next month. The Telecom company gives me lots of user behavior data without sensitive part. The label is binary, meaning whether the user would lost in next month. I use previous four-month data to extract the user's feature, and then input the features into model to predict the user's situation in next month.

## RESEARCH / PROJECT EXPERIENCE

- **2019 Reinforcement Learning Project: ([Git-Repo](#)) :**
  - **Navigation:** Training an agent to learn how to get maximum bananas in one episode, using the Deep Q-Network(DQN) Algorithm.
  - **Reacher:** Training a robot arm to quickly catch a target object. In this project, I use the Deep Deterministic Policy Gradient(DDPG) with continuous action space to train 20 virtual agents simultaneously to reduce the training time.
  - **Tennis:** Two players compete with each other. The project is training two agents to play the tennis game. I use the DDPG to train multi-agents in competitive environment.
  - **Half Field Offense in [Robocup 2D Soccer](#):** Training an agent to kick the ball to the goal. This environment action space is parameterized action space which combines the discrete value and the continuous value.
- **2018 NLP Project ([Git-Repo](#), [Paper](#)) :**

We use four kinds of algorithms to complete Name Entity Recognition(NER) task and compare them.
- **2019 Image Process ([Git-Repo](#), [Report](#)) :**

For this project, My target is to display the intermediate processes of image-style transformation, how to transform a photo from a original photo to a style photo and what the processes look like. I use VGG19 and transform-net (Johnson, J. et al. 2016) to train model for image-style transformation. In each layer of transform-net, I use DeconvNet to get visible image from intermediate data to display the changing process.

- **2019 Distributed Framework:**

It is a distributed system framework and has a master and various plumbable components. The plumbable components can be designed to divers function components. The master manages all components. The system has a pub-sub system to support communication between all components. The component can publish and subscribe topics. This project is based on AKKA concurrency framework with Scala programm language. Here is a Distributed Web Crawler based on this distributed framework, "**crawlnet**" ([Git-Repo](#)).

- **2015 Recommendation System** ([Git-Repo](#)) :

The project target is to design and develop a recommendation system to recommend commodity for users. The architecture of the project includes four parts. I use Hive as my data warehouse that stores all orders data, user and commodities data. I use Spark to generate models used to recommend commodities and store the result into Hbase database. And then we design restful API to receive requests and return the recommendation commodities for users. The algorithm I use in this project is user-item interactions matrix, Collaborative filtering, and TF-IDF etc.

## PROFESSIONAL SKILLS

- **Deep learning** development skills:

- **Tensorflow, pyTorch** and relevant **Python lib**, Such as **Numpy, Pandas** etc.

- **Reinforcement learning:**

- Get the Nanodegree in Udacity online course.

- **MDP, DQN, DDPG, PPO, A2C, A3C**, .

- **NLP**

- **word2vec, RNN, LSTM, seq2seq**, and **attention model**.

- **Image Process**

- **CNN, VGG, ResNet, DeconvNet, Image Style\_Transfer**

- **Big Data** development skills:

- **Spark** distributed computing system, experience for dealing with **large-scale data sets** and the **spark-mllib** machine learning algorithm library.

- **Recommendation System:**

- **Framework** design and development

- Design **Recommendation Algorithm**

- Design **User Portrait**

- **AKKA concurrency framework**, and **Real-time stream processing** with Spark.

- **Big Data** related systems, such as: **Hadoop, Spark, Hive, Hbase, Kafka, Elasticsearch** etc.

- Relational database **Mysql** and non-relational database, such as : **HBase, Mongoddb** and **Redis**

- **Development Skills:**

- Programming language **Scala, Python, Java, C/C++, SQL**, etc.

- Tools: **git**, **vscode**, **vs**, **JetBrains**, **sbt**, **Maven**, **CMake**, **Amazon Web Services(AWS)**

## **CERTIFICATE**

[Udacity Nanodegree:](#)

## **WEBSITE**

[Github](#), [Linkedin](#), [Personal Website](#)