CHRIS LI

Data Scientist

Eager to make a difference in data world

CONTACT INFORMATION

Email Address: junjie.li@rhsmith.umd.edu
Phone Number: 2404136563
Linkedin: www.linkedin.com/in/ds-chris-li
Github: https://github.com/kkkris7

ACADEMIC BACKGROUND

UNIVERSITY OF MARYLAND
COLLEGE PARK, MD   **DEC 2019**
MS in Business Analytics, GPA 3.7

DONGBEI UNIVERSITY OF FINANCE AND ECONOMICS
DALIAN, CHINA   **JULY 2017**
BS in Applied Mathematics, GPA 3.6

SKILLS

- **Programming Languages:** Python (Numpy, Pandas, Tensorflow, Keras, Opencv, Scikit-learn, BeautifulSoup, Xpath), R, Javascript, Object-Oriented Programming
- **Deep Learning:** TensorFlow, Keras, Torch, Skilled in Computer Vison Modeling
- **Machine Learning Algorithm:** Linear/Logistic Regression, Random Forest, XG-boost, Ensemble Method, K-means
- **Big Data:** Spark, Hadoop, HIVE, Pig, EMR
- **Web Development:** JavaScript, HTML, CSS, LXML, Django
- **AWS:** S3, EC2, Kinesis, Athena, Lambda, EMR, Glue, Sagemaker
- **Database Management:** SQL, MySQL, DynamoDB, RDS
- **Agile Development Method:** Scrum, JIRA

LEADERSHIP

LEAD OF AMAZON WHOLEFOODS CASE COMPETITION (WINNER)

- Delivered a new price strategy about the free delivery fee service to client and enhanced the estimated profit by 200%
- Led the team and processed thorough understanding of case objectives to ensure the team operated as a cohesive unit

PROFESSIONAL CAREER

Data Science Engineer Intern
Hitachi Vantara | May 2019 - Present
- Developed an image processing and computer vision pipeline to detect and localize vehicle and license plate from multi-sourced real-time video stream using OpenCV, YOLO V3, deep ANPR, Keras and PyTorch
- Speeded up the algorithm by 10X through GPU processing with AWS EC2
- Performed data labeling, data exploration, data cleaning and hyperparameter tuning to optimize the loss and accuracy in the training procedure and achieved 95% accuracy
- Built a multi-classification model for food recognition using Wide-Resnet, TensorFlow 2.0, Google Colab and TPU processor

Data Research Analyst
Morningstar, Inc | Aug 2017- May 2018
- Developed a web application to automatically extract financial product data with Python and reduce 80% manual hours
- Imported US fund data to database with 98% accuracy using Microsoft SQL Server and VBA

Research Assistant
Dongbei University of Finance and Economics | June 2016 - June 2017
- Collected, processed and cleaned 20000+ data from National Bureau of Statistics database
- Classified and grouped indicators by using principal component analysis, factor analysis and rotated component matrix with R

PROJECT EXPERIENCE

Product Recommendation System
Big data, AWS| June 2019
- Dumped the GB level server logs data into Amazon S3 buckets by using Kinesis Data Firehose and EC2 instance
- Generated a recommendation model based on the order data in S3 by running ML lib on Apache Spark and EMR cluster

NBA League Championship 2019 Winner Prediction
Machine Learning, Web Scraping| April 2019
- Scraped thousands of data from internet by using object-oriented programming with Python and stored the data into MySQL DB
- Built a logistic regression model and Monte Carlo simulation to predict play-off teams and champion and achieved 93% accuracy
- Developed a user-friendly front-end web interface with Django, CSS and HTML to interpret prediction results

Hospital Patients Return Rate Prediction
Machine Learning, R| June 2019
- Collected, cleaned and re-structured the raw data to fit for different requirements for various models
- Implemented an ensemble ML model with multiple algorithms including XG-boost, Random Forest and LDA in R to predict the patients return rate of 5 hospitals by learning their historical digital analytics data and increase the accuracy to 80%

Restaurant Database Management
SQL, Database Management| June 2019
- Led a team of four to write user stories, design ER diagram, and build a database in Microsoft Server SQL
- Analyzed data in SQL, visualized the results with Tableau to provide comprehensive information