Contact

www.linkedin.com/in/hu-tiger-cao-412883b6 (LinkedIn)

Top Skills

Successfully completed three acturial exams (p, FE, C)

Acturial Analysis and Financial Risk Analysis

Statistical Analysis and Modeling

Languages

English (Professional Working)

Chinese (Native or Bilingual)

Certifications

SAS Base Programing for SAS 9

SAS Advanced Programing for SAS 9

Honors-Awards

Yongling Liu's Award for Outstanding Dissertation

Knobble Scholarship

Hu (Tiger) Cao
Senior Data Scientist at Ephesoft, Inc.
Cypress, California

Summary

An intelligent and experienced Data Scientist with a great passion for solving complex problems using fully equipped expertise in data mining, statistical modeling, machine learning/Deep Learning, and Natural Language Processing (NLP), plus strong work ethics, self-motivated and quick learning capability, to leverage a heavy dose of programming skills and applied statistics.

Experience

Ephesoft, Inc.
2 years 5 months

Senior Data Scientist
November 2019 - Present (5 months)
Irvine, CA

Data Scientist
November 2017 - October 2019 (2 years)
Irvine

Named-Entity Recognition using Natural Language Processing (CoreNLP, SpaCy, NER, Glove, Word2Vec, LSA, Document Classification, Topic Modeling) ;
Invoice Key-Value Extraction (PDF-TIF-OCR, Image Pre-processing, Tesseract OCR, AWS Textract, Text Mining, Feature Extraction, XGBoost) ;
Invoice Address Extraction (CoreNLP, Hierarchical/K-means Clustering, Document Layout Analysis)
Personal Identifiable Information (PII) Extraction (Correlation, Clustering, and Classification, RegEx, CoreNLP, Association Rule, Named-Entity-Recognition, Graph Data Structure) ;
Deep Learning in Image Text Recognition, Table Detection, and Classification (TensorFlow, Keras, TFLearn, CNN, LSTM);
Cloud Services (AWS S3|EC2|SageMaker|Comprehend|Machine Learning| Macie|Textract|Lambda Function, Google Vision|NLP|DLP)

Bank of Hope

Senior Model Validator

June 2017 - November 2017 (6 months)

Performed comprehensive validation of bank-wide models (DFAST, PD/LGD/ EAD, PPNR) ;
Developed Probability of Default(PD) and Loss Given Default (LGD) models (Fractional Response Model and Beta Regression Framework (Beta, Beta mixture and zero-one-Inflated Beta)

QuickBridgeFunding
Data Scientist

June 2015 - May 2017 (2 years)

Irvine, CA

Credit Risk Modeling; Predictive Analytics; Unbalanced Sampling; AdaBoost Algorithm; Machine Learning; Extract, Transform and Load (SQL, JSON, XML, MongoDB), Renewal Forecasting, Response Modeling, A/B-Testing and Evaluation of Direct Marketing Campaigns

California State University, Los Angeles
Staff

March 2013 - August 2013 (6 months)

Los Angeles

Teaching Undergraduates about Physics 101 and Lab Experiments

University of California, Riverside
Postdoctoral Schoolar

May 2011 - March 2013 (1 year 11 months)

Riverside

Physical Properties and Structures of Superconductor

Indiana University Bloomington
Research Scientist

January 2009 - April 2011 (2 years 4 months)

Bloomington, Indiana

Simulation, Instrumentation, and Neutron Scattering

Education

University of California, Riverside

Master's degree, Statistics · (2013 - 2015)

Virginia Polytechnic Institute and State University

Doctor of Philosophy (Ph.D.), Materials Scientist · (2003 - 2008)

Shanghai Institute of Ceramics, Chinese Academy o f Sciences

Master's degree, Materials Science · (1999 - 2002)

Nanjing University of Technology

Bachelor's degree, Materials Science · (1995 - 1999)