Zaiyong Zhang, Ph.D.

Data Science Consultant at Aequitas Software

(+49) 17684633900
zaiyong.zhang@yahoo.com
Hamburg, Germany

Profile

Full-stack Data Scientist with 5+ years of experience in Data Science, Data Engineering, Customer Analytics and data infrastructures.

Key Strengths

- Excellent knowledge of statistics, machine learning algorithms, and be able to implement them in Python.
- 5 years of experience of scientific/industrial data science solutions with Python. Quite good at data science related libraries: Numpy, Pandas, Scikit-Learn, Keras, Flask and Matplotlib.
- 2.5 years of experience of design and implement ETL data pipeline for real-time data stream.
- 1.5 year of experience of data visualization/Business Intelligence with Tableau.
- 1 year of experience of Hadoop ecosystem, including HDFS, Hive, Spark and so on.
- 3 years of experience of SQL and NoSQL databases, include MySQL, PostgreSQL, MongoDB, Redshift and so on.
- 1 year of experience of recommendation system for personalized offers.
- 2 years of experience of Amazon Web Services(AWS).
- 1 year of experience of Microsoft Azure.
- 2.5 years of experience of Agile/Scrum.
- 1 year of experience of technical lead of data science projects.

Working Experience

Data Science Consultant   *Hamburg, Germany 01/2019-Now*
Aequitas Software
- Work as an external Data Scientist at E.ON Digital and participate the Predictive Maintenance Project.
- Analyze the influence of weather on the outage of Power Grid.
- Optimize current Machine Learning Prediction models with new features.

Data Scientist   *Hamburg, Germany 04/2017-12/2018*
AIDA Cruises & Costa digital unit
- Develop and implement the **Recommendation system** to recommender personalized cruises, restaurants and onshore excursions to customers with **python** and **flask**.
- **Technical Project Lead** for Costa Asia. The projects including official website, WeChat MiniPrograms and CRM systems.
- Implement **NLP** for customer feedback analytics, including sentiment analysis, entity recognition with **python** and **spaCy**.
- Implement **CNN** for handwriting recognition with **python** and **Keras**.
- Implement **Time Series Algorithms** for cruise oil consumption prediction with **python**.
- Setup and maintain the **tableau server** on AWS for Costa Digital Unit.

Virus/Data Researcher   *Tettnang, Germany 08/2014-04/2017*
Avira Operations GmbH

- Manage a distributed dataset (100 TB+, 100 million+ entities) on 150+ Windows/Linux Servers.
- Design and implement the ETL process for daily real-time data stream from multiple resources.
- **Lead the Internships** and supervise their work.
- Develop and maintain a program that automatically monitor and download all Apps (>1.5 million) in Google Play in real time.
- Develop and maintain a program that automatically install installers (.EXE format) in Virtual Machine and collect all dumped files.

Scientific Fellow   *München, Germany 09/2011-07/2014*
Technische Universität München

- Scientific research work
  - Prediction of protein structures by analyzing chemical data with Monte Carlo method.
  - Design and implement a new algorithm form NMR NOESY assignment.
- Programming work
  - RosettaCommons developer: Code development within the well-known open-source software Rosetta in C++, in cooperation with worldwide partners.
    https://www.rosettacommons.org/
  - Code development within the biological software CS-Rosetta in Python.

Junior Scientific Fellow   **Macau 09/2009-08/2011**
University of Macau

- Ignition signal process and pattern classification for automotive engine Diagnosis with **Wavelet Packet Analysis** and **Support Vector Machines** and **MATLAB.**

Student Research Assistant   **Xi'an, China 08/2008-06/2009**
Micro/Nano Electromechanical System Lab

- Automatic 3D modeling of Micro electromechanical systems with **OpenGL and C++.**

Education

Technische Universität München (DE)   09/2011-07/2014
Doctor's Degree, with full fellowship
Computational Chemistry

University of Macau (MO)   09/2009-08/2011
Master of Science, with full scholarship
Electromechanical Engineering; GPA 3.8 of 4, top 10% of all students

Northwestern Polytechnic University (CN)   09/2005-08/2009
Bachelor's Degree
Mechanical Engineering; GPA 83.9 of 100, top 5% of all students

Publications

[1] G. Schot, **Z. Zhang**, R. Vernon, Y. Shen, W. F. Vranken, D. Baker, A. M. J. J. Bonvin, and O. F. Lange, "Improving 3D structure prediction from chemical shift data," *J Biomol NMR*, 2013.

[2] **Z. Zhang**, J. Porter, and O. F. Lange, "Robust and highly accurate automatic NOESY assignment and structure determination with Rosetta," *J Biomol NMR*, 2014.

[3] M. Akimoto, **Z. Zhang**, S. Boulton, R. Selvaratnam, B. VanSchouwen, M. Gloyd, E. Accili, O. Lange and G. Melacini, "A Mechanism for the Auto-Inhibition of HCN Channel Opening and its Relief by cAMP," submitted: *Journal of Biological Chemistry*, 2014

[4] W. Pak-kin, T. Lapmou, and **Z. Zhang**, "Ignition pattern analysis for automotive engine trouble diagnosis using wavelet packet transform and support vector machines," *Chinese Journal of Mechanical Engineering*, 2011.

[5] P. K. Wong, C. M. Vong, **Z. Zhang**, and Q. Xu, "Fault Diagnosis of Automotive Engines Using Fuzzy Relevance Vector Machine," in *Communications in Computer and Information Science*, 2011.

Certifications

Oracle Certified Professional, MySQL 5.6 Developer	Oracle
Machine Learning	Coursera
R Programming	Coursera
Hadoop Platform and Application Framework	Coursera
Neural Networks and Deep Learning	Coursera
Structuring Machine Learning Projects	Coursera
Apache Hadoop Essentials	MapR Technologies
Cluster Administration: Configure a MapR Cluster	MapR Technologies
Course Completion Confirmation, MongoDB for developers	MongoDB