



NAME	Dr. Zaiyong Zhang
POSITION	Data Scientist
KERNKOMPETENZEN	Data Science Big Data Artificial Intelligence
ABSCHLUSS	Doktor der Naturwissenschaften (Technische Universität München)
SPRACHEN	Englisch: verhandlungssicher; Deutsch: Grundkenntnisse; Chinesisch: Muttersprache

IT-KENNTNISSE

PROGRAMMIER – UND SKRIPTSPRACHEN	<ul style="list-style-type: none">• Ausgezeichnet: Python, SQL• Gut: R, Bash• Basiskenntnisse: C++, Scala
FRAMEWORKS	<ul style="list-style-type: none">• Pandas• Scikit-learn• PySpark• Matplotlib• Flask• Keras
SOFTWARE	<ul style="list-style-type: none">• Tableau• Git• Azure Databricks• Jupyter Notebook• Zeppelin Notebook• PyCharm• Docker• Docker Compose• Apache Kafka
DATENBANKEN	<ul style="list-style-type: none">• MySQL• PostgreSQL• Azure Data Lake• Oracle Database• AWS RedShift• Google BigQuery• MongoDB

METHODEN	<ul style="list-style-type: none">• Apache Hive• Apache Hbase
BETRIEBSSYSTEME	<ul style="list-style-type: none">• Agile Softwareentwicklung• Scrum Framework
SONSTIGES	<ul style="list-style-type: none">• Windows• Mac OS• Ubuntu
	<ul style="list-style-type: none">• Amazon Web Services• Microsoft Azure

ZERTIFIKATE / TRAININGS

	Microsoft Certified Azure Fundamentals Professional Scrum Master Oracle Certified Professional MySQL Developer Hadoop Platform and Application Framework
--	---

PROJEKTERFAHRUNG / BERUFSERFAHRUNG

01/2019 BIS HEUTE	<p>E.ON Digital Technology Rolle: External Data Scientist</p> <p>Aufgaben:</p> <ul style="list-style-type: none">• Data Scientist in strategischem Predictive Maintenance Projekt beim größten deutschen Energieunternehmen• Einsatz von Azure Databricks und PySpark (Anwendungsgebiete: Big Data Analyse, Visualisierung und Machine Learning)• Fokus des Projektes war die Untersuchung des Einflusses meteorologischer Daten auf die Haltbarkeit unterirdischer Mittelspannungskabel durch:<ol style="list-style-type: none">1. Exploration meteorologischer Daten2. Extrahierung von Merkmalen aus meteorologischen Daten3. Eliminierung redundanter und niedrig-relevanter meteorologischer Merkmale4. Integration meteorologischer Merkmale in bestehende Machine Learning-Modelle5. Anpassung und Optimierung der neuen Modelle sowie Präsentation der Ergebnisse vor den Stakeholdern• Verbesserung des Vorhersagemodells um 10% durch die Analyse meteorologischer Daten und die Anwendung von Machine Learning Algorithmen
-------------------	---

04/2017 BIS 12/2018

Technologien:

- Machine Learning, Python, Git, Microsoft Azure, Azure Databricks, Azure Data Lake, PySpark

AIDA Cruises

Rolle: Data Scientist

Aufgaben:

- Entwicklung eines Online-Empfehlungssystems basierend auf den historischen Buchungs- und Web-Trackinginformationen
 - Das Empfehlungssystem schlägt den Kunden u.a. personalisierte Kreuzfahrten, Restaurants, Flüge, Ausflüge sowie Aktivitäten an Bord und an Land vor
 - Integration des Systems in die interne Buchungssuchmaschine
- Technischer Projektleiter für Costa Asien
- Analyse des Kundenfeedbacks durch Natural Language Processing
- Handschrifterkennung mit Convolutional Neural Network
- Vorhersage des Kraftstoffverbrauchs von Kreuzfahrten
- Einrichtung und Wartung des Tableau Servers

Technologien:

- Machine Learning
- Maschinelle Verarbeitung natürlicher Sprache (Natural Language Processing)
- Business Intelligence

08/2014 BIS 04/2017

Avira Operations GmbH

Rolle: Virus/Data Researcher

Aufgaben:

- Verwaltung eines verteilten Dateisystems (100 TB+, 100 Millionen+ Datei, 150+ Windows/Linux Servers)
- ETL-Prozess für tägliche Batchdaten aus mehreren Quellen definieren
- Leitung der Werkstudenten
- Entwicklung und Umsetzung eines Google Play Crawlers, der alle Apps (>1.5 Mio.) von Google Play herunterladen kann

Technologien:

- ETL Pipeline
- verteiltes Dateisystem
- Hadoop

09/2011 BIS 07/2014

Technische Universität München

Rolle: Scientific Fellow

Aufgaben:

- Entwurf eines Algorithmus für die Proteinstrukturberechnung
- Open Source Software Entwicklung mit Python und C++

Technologien:

- Monte-Carlo-Simulation

09/2009 BIS 08/2011

University of Macau

Rolle: Junior Scientific Fellow

Aufgaben:

- Automatisierte Motor Diagnose durch Wavelet Packet Analysis und Support Vector Machines

Technologien:

- Wavelet Packet Analysis
- Support Vector Machines

08/2008 BIS 06/2009

Micro/Nano Electromechanical System Lab

Rolle: Werkstudent

Aufgaben:

- Automatisierte 3D Modellierung von Mikrosystem mit OpenGL und C++

Technologien:

- OpenGL