Yan Zhang

Worcester County, Massachusetts, United States

✉ zhangyanallen@163.com   📞 7742536194

in linkedin.com/in/yan-zhang-8960ab231

Summary

I have 6 years of experience in data mining and data science with Python and Java, massive data processing, and ETL development. (Hadoop, Teradata, Oracle). In addition, I am also equipped with website development skills with JAVA, AWS. Now I am studying at WPI for higher education in Master of Science in Computer Science.

Experience

Full Stack Engineer
Merck

Nov 2019 - Aug 2021 (1 year 10 months)
Refined an application to automatically calculate the sales bonus using a machine learning model with python(pandas, NumPy, Sklearn), autotest pages with Java(React and Java). In addition, I also enabled the ETL process with information to ensure the daily task run smoothly.

Created a model for the audit department to find false product flows using Python and established a website with DJANGO for team members to view the daily reports.

Worked with colleagues to refine the existing model from Python 2.7 to Python 3.6. Moreover, we added new features for the model to make it fitter to the current business requirements.

Cooperated with Accenture for the database migration projects, transferring data from Teradata to AWS. In addition, redesigned the database with the star model and snowflakes model. Moreover, wrote SHELL to ensure data integrity.

Software Engineering Specialist
Mooton

Nov 2017 - May 2019 (1 year 7 months)
Built RFM model with SQL and HIVE for the game planning team and used Tableau to build a reporting system for the boss.

Founded a recommendation system by FM model with Python and Wrote a Java GUI for the business colleagues to use. Improving the in-game purchasing rate by 30% and 2% DAU. Providing the company of profit by almost 100K $ per day.

Senior Data Analyst
Roche

May 2015 - Oct 2017 (2 years 6 months)
Worked at data department, coaching the lab colleagues and database members to learn basic machine learning knowledge(Regression, SVM, Boosting/Bagging, and so on).

Cooperated with the R&D department to complete phase 1 research for the oncology medicines. Deploying the PCA to complete feature filtering and finding outlier attributes.

Worked on the cardiovascular products at phase 2. Using python with logistic regression, SVM, and neural models to do prediction, picking up the best model from R square and accuracy with ROC.

Participated in the medical tumor CT image recognization project. Deployed the HAAR and Adboost with PYTHON.

Senior Database Administrator
Standard Chartered Bank
Mar 2013 - Mar 2015 (2 years 1 month)
Created daily backup schedule with disks and RMAN catalog. Deployed SEHLL for backup in different timezone and situations.

Monitored the AWR reports and looked for the weakness of the database. Find issued SQL with high resources usage and optimize them with DBMS package or check system views and SQL execution plans to find the bottleneck.

Established the RAC and DG to ensure the database's high performance and high availability.

Education

Worcester Polytechnic Institute
Master's degree, Computer Science
2021 - 2023

Sanda University
Bachelor's degree, Computer Science
2008 - 2012

Licenses & Certifications

Oracle Database 11g Administrator Certified Master - Oracle
C14363-02

Skills

database administration • Java • Python (Programming Language) • Data Analysis • Data Mining • Amazon Web Services (AWS) • D3.js • Tableau • React.js • Oracle Database