Jin Xianwen

Data Engineer

February 28, 1995

Padua, Italy

+39 388 4953659

jinxw@live.it

Skills & Strengths

Technical Skills

- Good knowledge of Java and Flink Data Streaming Framework
- Knowledge of Hive, Hbase and Hadoop
- Knowledge of popular frameworks and APIs such as Spring/Spring Boot, JPA and Hibernate
- Knowledge of front-end technologies: HTML, CSS, Javascript
- Knowledge of C/C++ and Python

Soft Skills

creative problem solver | friendly | team player | flexible | like challenges | self motivated | hard-working | fast learner

Eager to deepen knowledge about...

- Big Data
- Machine Learning
- Algorithm Engineering
- Software Developer

Languages

Italian ●●●●●

Chinese ●●●●●○

English ●●●○○○

Interests & Activities

Ski | Anime | Travelling | Good Food | Technology

EXPERIENCE

2023 **Software Engineer** | **Huawei** | **Dongguan, China**

- Development and maintenance of CodeArts Pipeline based on Spring Framework
- Support the client for Application Maintenance

2021 - 2022 **Data Engineer** | **Technology Reply Financial Services** | **Padua, Italy**

- Design and implementation of offloading pipeline made of change data capture tool from mainframe, event streaming platform (Kafka + Apache Flink) and relational database for a main Italian banking group to address PSD2 regulation and open banking challenge
- Support the client for Application Maintenance
- Contact the vendor support for issue reporting/debugging

2017 - 2018 **Production Specialist** | **Adaptica Srl** | **Padua, Italy**

- Assembly of electronic optical instruments: VisioFit, 2Win and Kaleidos
- Provided temporary language (chinese) and technical supports to R&D office
- Participated to a brainstorming discussion about the company development

EDUCATION

2018 - 2020 **Master's Degree in Computer Engineering** | **Mark: 110/110**
Department of Information Engineering, University of Padua

Main Courses: Machine Learning, Big Data, Computer Vision and Web Application
Thesis: k-Center Clustering under Doubling Dimension

2014 - 2017 **Bachelor's Degree in Electronic Engineering** | **Mark: 105/110**
Department of Information Engineering, University of Padua

Main Courses: Fundamentals of Java, Digital Electronic, Laboratory of Industrial Automation
Thesis: The State of the Art of FPGA

PROJECTS

CodeArts Pipeline Back-end

- **Technology:** Spring Cloud, MyBatis, Redis, PostgreSQL, RabbitMQ, Python
- **Project Description:** The CodeArts Pipeline is essentially a visual automated task scheduling platform, which needs to be used in conjunction with the automated tasks of compilation and build, code check, cloud testing, deployment and other services in the software development cycle. According to the scenarios required by users, such as application deployment in the dev/test/production environment, these automated tasks could be customized and arranged. After a proper configuration, users can trigger the execution of the job with one-click avoiding frequent and inefficient manual operations.
- **Role:** Software developer and mainly responsible of collecting pipeline statistics. Also involved in customer support and software AM.
- **Management Tools:** IntelliJ IDEA, Maven, CodeHub, PyCharm, Insomnia, Datagrip.

OpenBanking Database

- **Technology:** IBM CDC, Kafka, Flink, OracleSQL, JDBC, SpringBoot, Hibernate
- **Project Description:** the project mainly realizes the real-time processing of the data on DB2 database and stores the result on Oracle database. The project consists of the following data pipeline: Change Data Capture (IBM CDC) to capture data manipulation operations, Apache Kakfa for temporary data storaging, Flink framework for real-time data processing and Oracle Database to store persistent data. Additionally a RESTful microservices based on Spring/Spring Boot, JPA, Hibernate frameworks are provided to retrieve data from Oracle Database.
  The core of the project is Flink DataStream Java API which provides a real-time processing based on event time of the data. The data processing aims to compute the core information of the customer account such as the balance and money transactions. As result the data on Oracle Database is much compact and faster to retrieve compared to DB2 Database. It has been observed an improvement of 30% in select query speed. Furthermore it provides additional information such as the history of account balance.
- **Role:** the main software developer with tasks including technical analysis, implementation of the software, intra-team coordinating and inter-team (such as DevOps) communication, customer support and software AM.
- **Management Tools:** IntelliJ IDEA, Maven, Git, Jenkins, Ververica, Kibana and Oracle SQL Developer.

k-Center Clustering with MapReduce

- Solving the problem using coreset approach and Farthest-first traversal algorithm
- Implementation using MapReduce with Java
- Optimized time and space complexity by using Doubling Dimension Notion

Traveler Salesman Problem Solver

- Implementation in C
- Use of IBM ILOG CPLEX Optimization Studio
- Linear integer programming
- Heuristic, Math-Heuristic and Meta-Heuristic algorithms such as 2-opt, Variable Neighbor Search and Local Branching

Teaching Platform

- Web application project
- Implementation by using HTML, CSS, Javascript for front-end
- Java, Java Servlet, JSP and Mysql for back-end