Duong Van Chien

Nam Tu Liem, Ha Noi, Viet Nam ● chiendv.sthn@gmail.com ● (+84) 962145390

PROFILE SUMMARY

Senior Data Engineer with over 7 years of proven success in delivering end-to-end data solutions across diverse business domains. Demonstrated ability to independently lead full-cycle data projects from gathering business requirements and collaborating with stakeholders, to architecting scalable data platforms, building ELT pipelines, and deploying reporting systems. In addition to technical execution, consistently took ownership of project planning, progress tracking, and post-deployment maintenance, ensuring timely delivery and long-term sustainability of data solutions.

Proficient in SQL,PLSQL, Python, Java, Spark, Airflow and AWS, with deep expertise in data modeling, data warehouse design, and workflow orchestration. Known for bridging the gap between business and technology, ensuring data systems are aligned with strategic goals. Strong ownership mindset, capable of managing project timelines, resolving technical challenges, and delivering high-quality outcomes.

EDUCATION

Bachelor's Degree in Computer Science   Sep 2010 – May 2014
Ha Noi University of Industry

EXPERIENCES

1. United Insurance Company of Viet Nam – *Technical*   Aug 2022 – Present

Project : Building data pipeline on Cloud

- **Requirement & Planning:**
  Collaborated with cross-functional teams, including Business Analysts and internal users, to gather, clarify, and document business requirements originating from the data team of the parent company in Japan. This involved understanding data needs, integration points, and expected outcomes to ensure alignment with business objectives and technical feasibility.

- **Data Storage & Modeling:**
  Designed and implemented a cloud-based data architecture using Amazon Web Services (AWS), with a centralized data lake hosted on Amazon S3 and structured data stored in AWS Glue Data Catalog.

- **Data Pipeline & Orchestration:**
  AWS Glue for data extraction, transformation, and loading (ETL) .
  Amazon S3 for raw and processed data storage.
  Amazon Athena for serverless querying and reporting.
  Integrated data ingestion from on-premises data warehouses, performed necessary transformations, and stored the results in S3 for the data team from mother company.

- **Deployment & Maintenance:**
  Implemented Infrastructure as Code (IaC) using AWS CDK
  Built and maintained CI/CD pipelines with GitLab, automating deployment of ETL jobs, Glue scripts.

Project : End-to-End Development of UIC Report System

- **Requirement & Planning:**
  Collaborated with cross-functional teams (Finance, Sales, UW, Claim Actuarial…) to gather and clarify business requirements, and design a scalable architecture for reporting and analytics. In addition, worked independently with individual departments to build tailored reports and provide reliable data solutions, ensuring that each team's operational and analytical needs were met..

- **Data Warehouse & Modeling:**
  Designed and implemented a dedicated on-premise Data Warehouse decoupled from the core insurance system, ensuring reporting workloads do not impact transactional DB and application performance. Applied star schema design, partitioning, and indexing strategies to optimize query performance and scalability, accelerating complex reports multiple times and enabling advanced actuarial, finance, and sales analytics**.**

- **Data Pipeline & Orchestration:**
  Developed batch ETL pipelines using Python (Pandas), Spark and Airflow to ingest data from Oracle, ensuring data consistency and incremental loading.

- **Reporting Optimization:**
  Optimized complex SQL queries and dashboard queries (e.g., aggregations, joins) by rewriting inefficient statements and applying database tuning techniques such as indexing, partitioning, and temporary tables... These improvements significantly accelerated report generation and enhanced overall system responsiveness, enabling Finance, Sales, UW, Actuarial… teams to access insights faster and improve performance of business operation

- **Analytics & Dashboarding:**
  Developed and Implemented reporting and dashboarding solution with Metabase, strategically chosen over custom development to reduce delivery time and maintenance overhead. Enabled self-service analytics for business teams with dynamic filtering, drill-downs, and export features, supporting > 80 active users while maintaining a 1-person data team.

- **Deployment & Maintenance:**
  Containerized the application using Docker and managed CI/CD pipelines via GitLab.

2. Fubon Life Insurance – *Data Engineer*   Oct 2017 – Jul 2022

Project : Development Data Pipeline and Design Report for Agent Portal

- **Requirement & Planning:**
  Collaborated with cross-functional teams to gather business requirements, and support design a scalable architecture for reporting and analytics.

- **Data Warehouse & Modeling:**
   Implemented data partitioning and indexing strategies to improve query performance.

- **Data Pipeline & Orchestration:**
  Developed batch Celery pipelines using Python (Pandas), to ingest data from Oracle, ensuring data consistency and incremental loading.

- **Reporting Optimization:**
  Optimized complex SQL queries and dashboard queries (e.g., aggregations, joins) to reduce report generation time.

- **Web Portal & Dashboarding:**
  Designed and built a self-service Django-based web portal with interactive dashboards using Plotly, JavaScript, and HTML/CSS, enabling business users to dynamically filter, and export reports.

Project: Development and Maintainance of Data Pipeline, Features, and Reports for Sale & Agent Module

- **Requirement & Planning:**
  Collaborated with cross-functional teams to gather business requirements for the Sale & Agent module, leading to the design of the architecture and development plan for new data pipelines and the reporting system..

- **Data Pipeline Development & Maintenance:**
  Developed new batch data pipelines using Spring Batch and Java Spring, PLSQL procedure. Concurrently, maintained and enhanced existing legacy batch jobs to ensure data consistency, improve performance.

- **Report Development & Data Export:**
  Designed and built new business reports using Crystal Reports, Excel and maintained existing ones. Also implemented and maintained functionality for exporting raw data for business requirement

- **Query & System Optimization:**
  Wrote and optimized complex SQL/PL-SQL procedures, functions, and queries for new developments. Also refactored and tuned existing queries and database objects to minimize processing time and report generation.

Project: Migration of Agent Performance & Sales Management Systems to Core Insurance Platform

- **Requirement & Planning:**
  Implement migration of Agent Performance System and Sales Management System into the company's Core Life Insurance Platform, ensuring data consistency and business continuity.

- **Data Migration & Transformation:**
  Analyzed legacy system schemas and mapped to Core System's data model, rebuilding ETL batches using Java, Oracle for compatibility.
  Implemented audit log tables to track data lineage and migration integrity

- **Data Pipeline & Orchestration:**
  Design Java Batch jobs to align with Core System's standards, optimizing performance.
  Refactored Oracle Packages/Procedures for critical workflows (e.g., agent commission payout, sales performance aggregation) with error handling and logging.

- **Collaboration:**
  Worked with sale, ops, actuarial and finance teams to validate business logic in batches/reports.

OTHER EXPERIENCES

- Fullstack developer(2014–2023): Experienced in building scalable web applications using Java (Spring), Python (Django), React, jQuery, HTML, CSS across both frontend and backend layers.

TECHNICAL SKILLS

- **Programming/Scripting**: Python (Pandas, PySpark, Airflow), Java, SQL (Advanced Query Optimization)
- **Big Data and Data Engineering**: Apache Spark (Spark SQL), Apache Kafka, Workflow Orchestration(Apache Airflow)
- **Database Management**: MySQL, Oracle, Microsoft SQL Server, Data Warehouse.
- **Cloud and Devops**: AWS (S3, Glue, Athena…), Docker , Gitlab(CI/CD)
- **Development Workflow**: Agile, Jira, Confluence, Git Version Control

LANGUAGE

English – Fluent Communication Skills: Fluent in English for technical communication and cross-functional collaboration in speaking and writing.