



謝 曉鋒 シャ ギョウホウ

DATA ENGINEER

HRT（謙虚、尊敬、信頼）という原則を大切にしています。

- ・謙虚 (Humility) とは、自己を過信せず、むしろ失敗から学び、常に成長し続ける姿勢を持つことです。私は自分にはまだ知らないことがたくさんあると認識し、日々の経験から学ぶことを重視しています。
- ・尊敬 (Respect) とは、チームメンバーや同僚と協力する際、お互いを尊重し、個々の能力や貢献を高く評価することです。私は他の人々の視点やスキルに価値を見出し、共同で目標を達成するために協力することを大切にしています。
- ・信頼 (Trust) : チームの一員として、他のメンバーに対して信頼を寄せ、仕事を任せることです。私はチームメンバーが有能であり、正しいことを行うと信じ、互いにサポートし合う信頼関係を築くことを重視しています。

✉ xiexiaofeng0226@gmail.com

🌐 <https://github.com/xiaofengxie0226>

スキル・資格

- Google Professional Data Engineer
- Google Professional Cloud Architect
- AWS Certified Solutions Architect Associate
- Tableau Desktop Certified Associate
- 基本情報技術者
- 統計検定2級
- TOEIC : 825点
- G検定

技術スタック

- プログラミング言語 : Python, Scala, SQL, Golang
- ビッグデータ技術: Apache Spark
- データパイプライン: Apache Airflow
- ETLツール : Apache Beam
- クラウドプラットフォーム : GCP, AWS
- データ可視化ツール: Looker, Tableau
- コンテナ: Docker, Kubernetes
- バージョン管理: Github

言語

- 日本語 : ビジネス会話
- 英語: 日常会話
- 中国語: ネイティブ

学歴

- 名古屋大学大学院経済学研究科卒業

職歴

2021/12/16 ~ 現在

ディップ株式会社

1. データ基盤パイプライン構築・運用保守

- 【内容】 WebSiteDB・CRMDBからBigqueryに集めるパイプラインを開発・運用保守
- 【期間】 2021.12月~2022.5月
- 【職業】 データエンジニア (6人体制)
- 【開発言語】 Scala, Python, SQL
- 【DWH】 Bigquery
- 【DB】 WebSiteDB: Oracle (S3)、CRMDB: MySQL
- 【ETL】 Composer (Apache Airflow)、dataproc (Apache Spark)
- 【バージョン管理】 Github
- 【役割】 データパイプライン一部設計と実装及びリファクタリング

2. Looker開発・移行案件

- 【内容】 全社のBIツールの入り替えとしてLookerの導入からユーザ受け入れまで
- 【期間】 2022.3月~2022.8月
- 【職業】 データエンジニア (6人体制)
- 【開発言語】 SQL, LookML
- 【DWH】 Bigquery
- 【BI】 Looker
- 【役割】 一部マートのLookML開発担当、データ検証、LookerSSOの設定

3. 外部サービスとのデータ連携

- 【内容】 GoogleAnalytics、Adobe Analytics、ADEBIなどの外部サービスにより取得したデータとBigqueryの連携
- 【期間】 2022.5月~現在
- 【職業】 データエンジニア (ほぼ1人対応)
- 【開発言語】 Python, SQL
- 【GCPサービス】 BigqueryのAPI、Cloudstorage、Dataflow、Cloudfunctions、Cloudscheduler
- 【バージョン管理】 Github
- 【役割】 要件定義、データ品質の担保、メンテ

4. 分析者環境の構築・保守運用

- 【内容】 社内のデータサイエンティストの要望から、スケーラブルなGPUが利用できる分析環境、または機械学習モデルデプロイ環境の構築
- 【期間】 2022.9月~現在
- 【職業】 PM・開発者 (2人体制)
- 【スタック】 docker、GCP (GKE、GAE)、python
- 【役割】 プロジェクト計画、要件定義、スケジュール&予算管理、一部の開発

5. GCPログ統合管理の開発

- 【内容】 GCP上にあるデータ基盤・関連するプロダクト・予測モデルのログに対して、利用者の要望により必要な情報だけBigqueryに出力し、Slackまで通知する
- 【期間】 2023.5月~現在
- 【職業】 リーダー・開発&データエンジニア (3人体制)
- 【開発言語】 Golang
- 【スタック】 docker、GCP (Bigquery、CloudRun、Logging)、Slack
- 【役割】 要件定義、開発、テスト

6. 日常業務

- 【依頼対応】 データの抽出、マート&Looker Explorer作成、GCP周りの権限付与
- 【ドキュメント】 データカタログ整理、文書化 (notion)、各種マニュアル作成
- 【新技術調査】 GCPのトレンド技術に対する調査 (DataPlex、Bigquery Omni またはAI関連のものなど)、ゼミの参加及びレポートの作成
- 【その他】 データ基盤ポリシー策定、バグ対応、全体予算削減の施策、全社のSQL・Looker操作の研修実施

2020/04/01 ~ 2021/11/30

コムチュア株式会社

1. 複合機オプション品の予測案件

- 【内容】 複合機オプション品の売上予測機械学習モデル構築、予測結果評価、可視化
- 【職業】 データサイエンティスト
- 【開発言語】 Python
- 【アルゴリズム】 決定木、RF、重回帰分析
- 【BI】 Tableau
- 【役割】 データクリーニングと前処理、モデル評価と最適化、可視化

2. ビール予測モデル開発案件

- 【内容】 スーパーマーケットのビールの売上予測モデルの開発、更新及び運用・保守
- 【職業】 データサイエンティスト
- 【開発言語】 Python, SQL
- 【アルゴリズム】 : LSTM
- 【役割】 データクリーニングと前処理、データ分析と探索的データ解析、モデル評価と最適化、可視化