

職務経歴書

2022年4月15日

菊田俊平

■職務要約

東京大学工学系大学院卒業後、株式会社リクルートに新卒入社し、ML施策やML基盤の開発・運用に携わる。①ダッシュボード自動生成ツールの開発や②GKE上でのGPUノード導入を実行し、コストや工数の削減、開発体験の向上に貢献。①では問題提案から技術選定、設計、開発まで行い16%のコスト削減、②ではGPU利用の必要性の提案からコスト試算、設計、開発まで行い1.3億円の売上増に貢献。

株式会社リクルート退職後、株式会社タイミーにフリーランスとしてジョインし、データ基盤の開発運用に携わる。データ転送のETLからELT化やデータ基盤の質を向上させるdbtやLookerの導入、Github Actions上でのTerraform CI/CDの開発・運用など、データ品質の向上や安全性、開発体験の改善に貢献。

副業としてジョインしていたAnytech株式会社では、AIエンジニアとして顧客の課題の明確化や問題設計から解決案の提案、実装まで幅広くリード。

■職務経歴

株式会社タイミー (2021年9月～現在)

契約形態: 業務委託(フリーランス)

職種: データエンジニア

業務内容: データ基盤の開発・運用、BIツールの運用、CI/CDの開発・運用

使用技術: Embulk, GCP, Terraform, DataDog, Github Actions

株式会社 リクルート (2020年4月～2021年8月)

契約形態: 正社員

職種: 機械学習(ML)エンジニア/ML基盤エンジニア

業務内容: ML基盤の開発・運用、MLを利用した施策の実装

使用技術: GCP, Kubernetes, Terraform, DataDog, Github Actions

Anytech株式会社 (2019年9月～2022年3月)

契約形態: 正社員

職種: AIエンジニア

業務内容: クライアント課題を解決するAIの提案および実装、運用

使用技術: PyTorch, TensorFlow, GCP

■経験・資格

要件に対する経験

- ・ GCPのインフラ構築およびサービス運用経験: 2年
- ・ Docker、Kubernetesなどのコンテナ環境の構築・運用経験: 1.5年
- ・ CloudWatch、Datadogなどのモニタリングツール利用経験: 2年
- ・ Goの開発経験: 1.5年
- ・ TerraformなどのIaCツールの実務経験: 2年

その他の経験

- ・ Python の開発経験: 4年
- ・ ML(特に Computer Vision)の実務経験: 4年

職務経歴書

2022年4月15日

菊田俊平

- ・ データ基盤の開発・運用経験: 2年
- ・ ML 基盤の開発運用経験: 1.5年
- ・ AWS のインフラ構築およびサービス運用経験: 0.5年
- ・ CI/CD の開発・運用経験: 2年

資格

Certified Kubernetes Administrator(CKA) , Certified Kubernetes Application Developer(CKAD)

■主な成果とプロセス

各データ施策用のDataDogダッシュボードを自動生成するCLIツールの開発

背景/ 課題	データ施策用のオンデマンド環境やJob、APIを動作させるML基盤が所属チームで開発・運用されていた。施策利用者にてセルフサービスのJobやAPIの開発が可能だったが、施策利用者によるインフラ監視ができる仕組みがなく、過剰なリソースが消費されていた。その結果、施策利用者の想定より高コストが各施策に請求されていた。
打ち手	各施策に特化したダッシュボードを自動生成するCLIツールを開発し、施策登録時に自動でCIで実行される仕組みを構築した。
プロセス	<u>技術選定</u> <ul style="list-style-type: none">・ モニタリングツール: REST APIが提供されCLIで操作可能なDataDogを選定・ 言語: DataDogのAPIクライアントを持ち、CLIツールとも相性の良いGolangを利用 <u>設計</u> <ul style="list-style-type: none">・ ①CLIツールを自動ビルド/デプロイするリポジトリと②施策登録時に自動でCLIツールが実行されるリポジトリに分けて実装・ ①ではマスタのダッシュボードに対して各施策の名前空間を埋め込んだ上で施策用のダッシュボードを作成するCLIツールをCobraで実装し、PR時にビルド、マージ時にタグをつけてGCRにpushする仕様・ ②では①でビルドされたCLIツールを実行するCIをmerge時にrunする仕様 <u>開発時の工夫</u> <ul style="list-style-type: none">・ マスタのダッシュボードが改良された場合に自動で全施策のダッシュボードが更新される仕組みにし、継続的にモニタリング項目を増やせる仕様にした・ CLIツールでdry-runオプションを開発し、事前に作成するダッシュボードをJsonで吐き出せるようにした
結果	各施策のインフラコストを施策利用者が把握することが可能に。 約100個の施策担当者が自らコスト管理可能になり、平均約16%コスト削減。

GKE上のGPU Job 基盤の開発・運用

背景/ 課題	ML基盤はGKE上に構築されており、コスト懸念からGPUノードの導入は避けられてきた。しかし画像や言語の深層学習には必須のインフラなため、施策利用者は自力でGCEにて環境構築、ジョブ・APIの開発・運用を行なってきた。自力での運用はML基盤と開発体験が異なるため、施策利用の妨げになっていた。
-----------	--

職務経歴書

2022年4月15日

菊田俊平

打ち手	KubernetesのTaintやGKEのAutoScaling機能、プリエンプティブルVMを駆使し、コストを最低限に保った上でGPUノードをML基盤に導入
プロセス	<u>設計</u> <ul style="list-style-type: none">・ 施策登録時にGPUを必要とすることを登録させ、登録した場合のみGPUノードを使用可能に・ GPUを必要とするコンテナ数に応じて、自動でスケールイン・アウトする仕様・ GPUを必要とするコンテナにのみ、GPUノードにスケジュールするTaintを利用 <u>開発時の工夫</u> <ul style="list-style-type: none">・ ジョブやオンデマンド環境下では高可用性は不要なため、プリエンプティブルVMでコスト削減・ DataDogにてGPUノードの利用時間が長い場合にSlackへアラートを飛ばす仕組み
結果	施策利用者がインフラの知識なしにGPUを利用したML施策実行を可能に。 2022年8月までに7件の施策で利用され、1.3億円の売上増加に貢献。

Github Actions 上での Terraform CI/CD の開発・運用

背景/ 課題	Terraform管理されたGCPインフラが、ローカルでの手動Plan/Applyにて運用されていた。手動運用では、サービスアカウントキー流出や誤操作でインフラの削除、開発の属人化のリスクが懸念されていた。
打ち手	CI上でPR時にPlan、マージ時にApplyが自動的に実行される仕組みを構築。
プロセス	<u>設計</u> <ul style="list-style-type: none">・ ローカルでの手動Plan/Apply権限をなくするため、IAMを絞り込み・ PR 作成時に validateやplan, lint、マージ時に applyをCI上で実行 <u>開発時の工夫</u> <ul style="list-style-type: none">・ Github PR にてPlanやLintの結果を確認するため、結果をpushする仕組み
結果	安全性や開発体験が改善し、工数削減に貢献。