

3倍のパフォーマンス改善と9割のコスト削減で、 新規事業やプロジェクトの体制が整う。



Supership株式会社は、「ミライリアルな幸せを、デジタルの力で創る。」をGroup Purposeとして掲げている、Data Technology Companyである。「マーケティングテクノロジー」と「データテクノロジー」から構成されるデジタルトランスフォーメーション事業を展開しており、データ分析、マーケティングコンサルティング、アドテクノロジーなどによって企業のデジタルトランスフォーメーションをワンストップでサポートできるのが特徴である。ビジネスの特性上、データは最も重要な資産の一つであり、データを活用したソリューションやサービス展開を大きな強みとしており、今後は5GやIoTの普及によって爆発的な増加が予想されるデータを軸にさまざまな企業へ新たな価値を提供することを目指している。グループ全体で、約100名のデータ・サイエンティストをかかえており、まさに国内トップクラスのDX支援企業と言える。

採用したプロダクト



ハイライト

パフォーマンスの
大幅な向上により
コストを
9割以上削減!

40名の
プロジェクトが
2名で運用可能に!

数日かかっていた
インフラの準備は
1日以内に!

課題

データサイエンティスト、データエキスパートのリソースが
付加価値の低いタスクに割かれている。

目下のビジネス拡大を期待している注力分野は、アナリティクス/AIとDMPである。2PB(ペタバイト)を超えるデータを元に機械学習を用いた広告配信の効率化や顧客理解を促進させるサービスを提供している。しかしながら、当該分野に関連するプロジェクトを進めていく上で、2つの大きな課題に

直面していた。

1つめの課題は、データサイエンスをするためのプラットフォーム構築に多大な工数が発生する点である。当初はオンプレミス環境のHadoopベースの分析環境で分析をしていたが、データサイエンティストが増えるに連れてリソースが逼迫する



(写真・右) Supership株式会社
プロダクト開発本部
データアナリスト部
プリンシパルデータサイエンティスト
吉原 トマス氏

(写真・左) Supership株式会社
データソリューション事業領域
ソリューション開発部
部長
稲田 暁人氏

ようになり生産性が低下し始めたため、よりスケーラビリティが高い環境を求めて、クラウドベースの分析基盤をスクラッチで構築することにした。ネットワークの設定からアプリケーションの実装まで、非常に多大な工数がかかった。

2つめの課題は、機械学習によるモデリングに付随する付加価値の低いタスクに、データエキスパートのリソースの多くが割かれてしまっていた点である。本来であれば機械学習

を用いたモデリングなどの付加価値の高い業務にフォーカスすべきだが、日々の運用やトラブルシューティング、パフォーマンスの最適化に非常に多くのデータエキスパートのリソースを費やしていた。分析基盤が安定せず数時間データサイエンティストが稼働できないということも頻繁に発生していた。

比較検討

パフォーマンス、コスト、安定性、そして、多機能の観点で、Databricksの採用を決定した。

比較検討にあたっては、これまでのオンプレ環境で大きな課題となっていた迅速性、柔軟性、セキュリティの観点でクラウド環境が大前提であった。複数のテクノロジーの比較の結果、パフォーマンス、コスト、安定性、そして、多機能の観点で、Databricksの採用を決定した。

Databricksのユニファイド・プラットフォームは、他社ソリューションに対して、パフォーマンスが3倍以上であった。また、SQL、Scala、R、python、Java など、複数の言語をサポートしているノートブックを提供しており、BIへ直接アウトプット可能なAPIやODBCコネクタも提供していることも魅力的だった。

さらに、mlflowによって、データサイエンティストが構築した機械学習を用いた予測モデルのライフサイクルを効率的に管理することができる。Apache Sparkの将来性の大きさも、今回のDatabricks採用に寄与した。Apache Sparkは、Databricksの創設者が作ったオープンソースの分散処理フレームワークであり、現在5,000以上の世界のリーディング企業で採用されている。約1ヶ月ほど行ったPOCにおいて、パフォーマンスは他社と比べて3倍良く、導入前のテスト期間に行ったレビューでは10名のプロジェクトメンバー全員が、Databricksに最高得点である5点をつけた。

ビジネスインパクト

Databricksの導入によって、3倍のパフォーマンス改善と、その他付随する効果を含め、最終的に9割のコスト削減を実現

Databricksの導入によって、先の述べた3倍のパフォーマンス改善と、その他付随する効果を含め、最終的に9割のコスト削減を実現した。また、プロジェクトに必要なプラットフォームの準備期間は数日から1日以内に短縮され、これまで約40人

が関わっていたプロジェクトは現在2名で進めることが可能になっている。これにより、積極的に新規事業やプロジェクトを開始できる体制が整ったと言える。