

## まえがき

近年、データサイエンティストブームもあり、データマイニングの理論やツールの情報は書籍やWeb上にあふれている。これらの情報は、データマイニングを実行していくうえで非常に便利で、実際データマイニングを実行するだけであれば困ることはほとんどない。しかし、実際にビジネスではこれらの理論やツールについて知っているだけでは、上手くいかないことが多い。いくら理論が正しくても、ビジネスとして価値がなければ全く意味がない。ビジネス価値の高いデータマイニングを実施していくためには、理論やツールだけでなく、ある程度かかわっているビジネスの戦略についても意識する必要がある。課題や目的、各タイミングに応じた適切な手法の選定・実行、さらに施策活用まで考えなければならない。

ここでいう戦略とは、あるビジネス上の目標を達成するために、何をすべきかを定めることである。今やっていることは、目標を達成するためにどれくらいインパクトがあるのか、インパクトがあったとして、それを何らかのアクションにつなげることができるのかを意識し、インパクトのあるテーマを見つけることが重要である。高度で複雑なモデルを構築しすばらしい精度を出しても、全体の中で占める規模が小さかったり、アクションにつながらなければ、ビジネス価値は乏しくなってしまう。また、このアクションにつながる課題を見つけること自体にデータマイニングの各手法を活用することも可能である。考えられる要因の中から、一番インパクトの大きい要因をデータから見つけ出すことは、非常に価値がある。

あるいは、目標を達成するために、データマイニングをどのような形で活用するかも重要なテーマである。何を実施するのかという意思決定をするための判断材料として使うのか、施策の自動化や最適化に使うためのアルゴリズム開発なのかで意味合いが変わってくる。意思決定の判断材料として使う場合には、人間がわかりやすいアウトプットである必要がある。高度で複雑なモデルを使って、精度の高い結果を出しても、意思決定者に伝わらなければ意味がない。たとえば、ある商品を買っている人と買っていない人の違いを理解して、買う人を増やすための施策を考えたいような場合に、「複雑なモデルを使うことで、99.9%の確率で、買っている人と買っていない人を判別できました」といわれても、意思決定者は「で、結局何をすればよいの?」となってしまうのである。一方、実施するアクションは決まっていて、その対象や量を最適化したり、自動で実行したいような場合には、わかりやすさよりも精度が重視される。わかりやすさのために、精度を犠牲にしてシンプルなモデルを使っても、「作ってもらった予測モデル、全然当たらないんだけど・・・」となり、ビジネス価値が乏しくなってしまう。

ここまでの話を整理すると、ビジネス価値の高いデータマイニング活用とは、以下の3つになり

#### iv まえがき

そうである。

- 解決すべき課題をみつけるためのデータマイニング
- 何をすべきかの意思決定を支援するためのデータマイニング
- アクションを自動化するためのデータマイニング

このように、ビジネス課題の理解とそれに対してデータマイニングをどう活用できるのか、という知見が重要であるが、こういった情報は非常に少ない。本書では、このような実態を考慮し、戦略とデータマイニングの両方を意識した構成になっている。具体的には以下のとおりである。

- 序論
- データ加工
- データ可視化
- 問題発見
- 問題解決：意思決定支援
- 問題解決：自動化・最適化
- 効果測定

序論では、戦略と問題発見・問題解決の概念について触れる。これらは、ビジネスで成果をあげるためには非常に重要なプロセスである。次に、R言語によるデータ加工・データ可視化について解説する。この2つの章は、データマイニングで問題発見・問題解決をしていくうえで必要な知識をまとめたものである。すでにR言語のスキルがある方や、データマイニングの活用だけを知りたい方は次の章から読み始めてもよいだろう。続けて、戦略的データマイニングのメインである、問題発見・問題解決について解説し、最後に、効果測定について述べる。効果測定は、施策を実際に実施した際に、どれくらい効果があったのかを検証するプロセスで、PDCAサイクルをまわしていくために重要なものである。

本書は、著者のこれまでの実務経験をもとに構成されている。エンジニアからデータ分析担当になった時に感じたギャップや、マーケティング部門に配属になったときに得られた知見などを整理したものである。これらが、すでにビジネスでデータ分析を行っている方、あるいはこれからデータ分析者になろうとしている方々にとって参考になれば幸いである。

本書を執筆するにあたって、多大な助言を頂いた市川太祐氏、大城信晃氏（ヤフー株式会社）、倉橋一成氏（iAnalysis 合同会社）、酒巻隆治氏（株式会社ドリコム）、鈴木雅彦（ヤフー株式会社）、高柳慎一氏、森本修氏（株式会社ディー・エヌ・エー）、本書を執筆する機会を与えて頂いた金明哲氏（同志社大学）、横田穂波氏（共立出版）、長期間にわたって執筆を支えてくれた家族に感謝する。

2014年7月

里 洋平