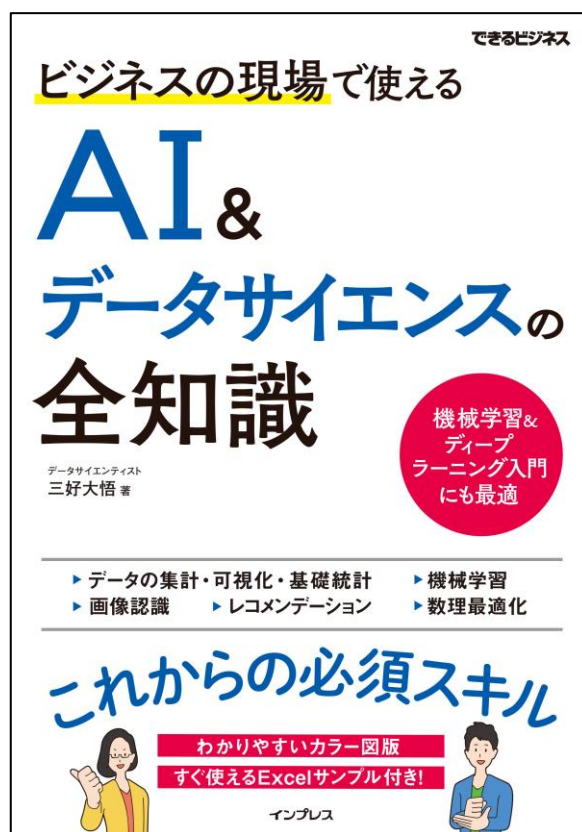


各 位

2022年3月15日  
株式会社インプレス

ビジネス事例からデータサイエンスがしっかりわかる  
『ビジネスの現場で使える AI&データサイエンスの全知識（できるビジネス）』を2022  
年3月15日（火）に発売  
早期購入特典ダウンロードキャンペーンも実施

インプレスグループでIT関連メディア事業を展開する株式会社インプレス（本社：東京都千代田区、代表取締役社長：小川 亨）は、ビジネス事例を通じてAI&データサイエンス技術を丁寧に解説した『ビジネスの現場で使えるAI&データサイエンスの全知識（できるビジネス）』を2022年3月15日（火）に発売します。



### ■ビジネススキルという観点で解説したAI&データサイエンス入門書

データサイエンスは、ビジネスのさまざまな現場で役立つスキルです。たとえばECサイトの運営にはおすすめ商品を提案してくれる「レコメンデーション」の仕組みが欠かせず、顧客ごとに最適な施策を打つには「ターゲティング」が欠かせません。ほかにも、需要予測にもとづく販売計画を立てたり、最適な商品単価を設定したり、効率的な配送ルートを導き出したりなど、データサイエンスと無縁のビジネス領域はないといっても過言ではありません。また、社会全体を見渡しても、2021年にはDXやIT化を国を挙げて推進するためデジタル庁が設置され、2023年度からは高校でAIやデータサイエンスを学ぶ「情報II」の授業がスタートします。データサイエンスはこれからの必須スキルなのです。

そんなデータサイエンスですが、本格的に扱うには数学や統計学の知識、またプログラミングスキルが必要であり、一般のビジネスパーソンにとって取り組みづらいイメージがあるのも事実です。本書はそのハードルをできるだけ低くして、知識ゼロでも「データサイエンスでできること・仕組み」を学べる解説書です。

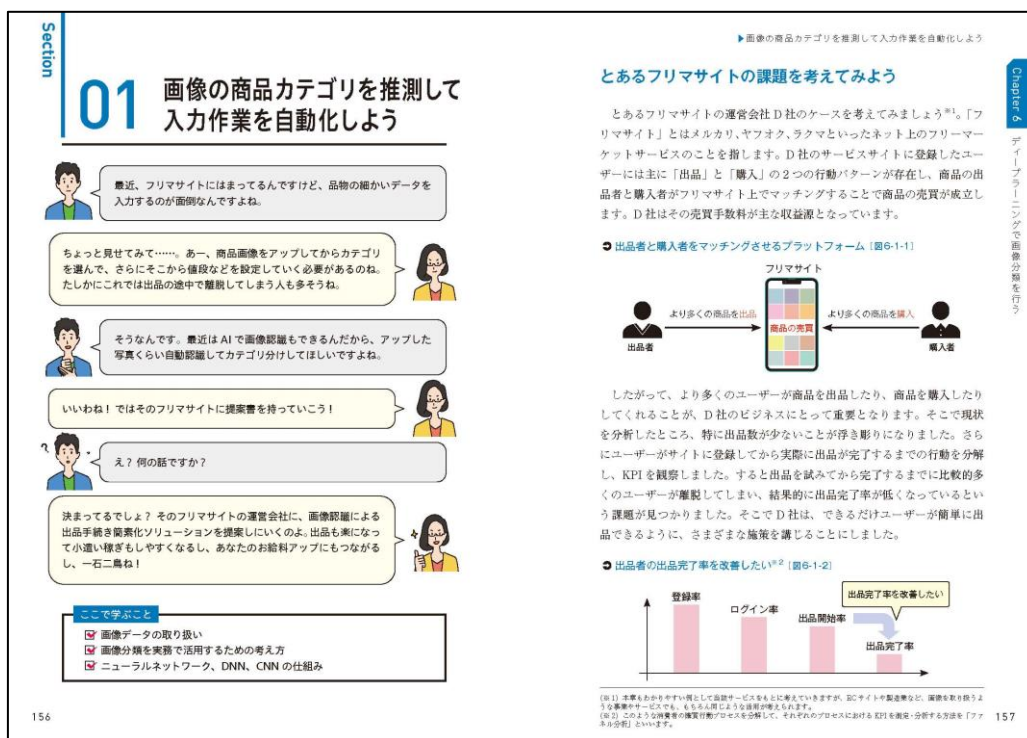
### ■身近なビジネス事例&Excelで自分ゴトとして理解できる

本書では、まずAIやデータサイエンスの定義、ビジネス活用事例を紹介しつつ、基本的な可視化や統計手法といったデータ分析の基礎知識を身につけます。そのうえで、需要予測、ユーザーターゲティング、画像分類といったビジネス上の課題をAIやデータサイエンスの技術を通じて解決していくという構成になっています。各トピックにサンプルデータを可視化・分析したExcelファイルをつけてあるので、実際にデータに触れて学べるのも本書の特徴の1つです。ビジネス課題を題材とすることで、データサイエンス技術を自分ゴトとしてしっかり消化できます。

<本書の特徴>

- ・学んだ内容を実際に試せるExcelサンプルファイル付き
- ・オールカラー&豊富なイラスト図解で難しい概念が理解しやすい
- ・AIやデータサイエンスに関する用語集つきでステップアップにも最適

### ■紙面イメージ



オールカラー&イラスト図解で読みやすい

▶ 実装：宿泊予約サイトのユーザーデータを活用しよう

練習用ファイル: chap05\_mail\_targeting / test\_result.xlsx

### 実践 教師あり学習 (分類問題) の予測結果を確認する

今回は、学習 (train) データ 11,258 行のデータ、すなわち 11,258 ユーザーのデータに対してロジスティック回帰モデルを適用・学習させた後、テスト (test) データ 4,825 行 (ユーザー) に対して予測し、その精度を確認してみましょう。学習モデルをテストデータに対して適用した予測結果を、「test\_result.xlsx」の「予測結果\_閾値別予測フラグ」シートの A~C 列に格納しています (実際は、Python によるデータ処理や学習をして、その結果を Excel に出力している格好となります)。

test\_result.xlsx には、テストデータ (test.csv) の全 user\_id それぞれの「予測確率 = CV してくれそうな確率」が付与されています。これは、学習データ = train.csv でロジスティック回帰モデルを学習させ、そのモデルをテストデータに適用することで得られた値になります。その予測確率と、正解の値である実測フラグとを比較することで、精度を評価します。

● モデルを学習させ、テストデータにモデルを適用し、予測結果を得る [図 5-4-7]

user_id	CV
1	0
2	1
3	0
4	1
5	0
6	1
7	0
8	1
9	0
10	1
11	0
12	1
13	0
14	1
15	0
16	1
17	0

user_id	予測確率	実測フラグ
1	0.397212	1
2	0.242536	0
3	0.684747	1
4	0.146538	0
5	0.146538	0
6	0.432399	1
7	0.163575	0
8	0.598747	1
9	0.453866	0
10	0.340807	0
11	0.513434	1
12	0.346716	0
13	0.549093	0
14	0.179732	0
15	0.355303	0
16	0.396281	1
17	0.343703	0

学習データにより「学習」したモデル

学習器を使って「予測」

テストデータにモデルを適用し、予測確率を付与する

---

練習用ファイル: chap05\_mail\_targeting / test\_result.xlsx

### 実践 閾値ごとに予測フラグを計算する

今回は、より実践的に考えてみましょう。先ほど、予測確率から閾値に応じて予測フラグを算出すると述べました。そこでこの演習では、出力された**予測確率から、閾値を 0 から 0.9 まで 0.1 ずつずらした際の予測フラグを算出**しましょう。「予測結果\_閾値別予測フラグ」シートを開いてください。B 列にユーザー ID ごとの予測確率があります。またセル F1 からセル O1 にかけて、閾値があります。この各閾値と各予測確率とを照らし合わせて、[図 5-4-8] のロジックを IF 関数で表現します。

● 予測値から予測フラグへ変換する [図 5-4-8]

- 「予測確率 > 閾値」であれば、予測フラグ = 1
- 「予測確率 < 閾値」であれば、予測フラグ = 0

閾値が 0.2 の場合の式: =IF(SB2>=H\$1,1,0)

● 閾値に応じて、予測確率を予測フラグに変換する [図 5-4-9]

	A	B	C	D	E	F	G	H	I	J	K
H2		予測確率	閾値								
		=IF(SB2>=H\$1,1,0)									
	user_id	予測確率	CV	閾値							
	1	0.397212	1	0.2	0.3	0.4					
	2	0.242536	0	0.2	0.3	0.4					
	3	0.684747	1	0.2	0.3	0.4					
	4	0.146538	0	0.2	0.3	0.4					
	5	0.146538	0	0.2	0.3	0.4					
	6	0.432399	1	0.2	0.3	0.4					
	7	0.163575	0	0.2	0.3	0.4					
	8	0.598747	1	0.2	0.3	0.4					
	9	0.453866	0	0.2	0.3	0.4					
	10	0.340807	0	0.2	0.3	0.4					

「SB2」が予測確率、「H\$1」が閾値に相当します。これを全予測確率、全閾値で計算できます。試みに、閾値が 0.4、0.5 の部分を空欄しておいたので計算してみてください<sup>15)</sup>。当然ですが、閾値が上がるほど、予測フラグも 1 が少なく、0 が多くなっていくはずですよ。

(注 15) Excel の詳細は「Test\_result\_answer.xlsx」ファイルとして随時提供します。

ダウンロード特典のExcelファイルで実際の分析結果を確認できる

## ■早期購入キャンペーンを実施

本書の発刊を記念して、2022年3月15日(火)~2022年4月14日(木)の期間中、全国の書店で本書をご購入いただいた方に、本書の電子版(PDF)をダウンロードにてプレゼントいたします。詳細は以下のURLをご参照ください。

[https://book.impress.co.jp/readers\\_entry/9784295013631cp.html](https://book.impress.co.jp/readers_entry/9784295013631cp.html)

## ■本書の構成

- Chapter1 データサイエンスをビジネスで活用する
- Chapter2 データサイエンスの手法を理解する
- Chapter3 基本的な可視化・統計手法を理解する
- Chapter4 線形回帰モデルで需要予測を立てる
- Chapter5 ロジスティック回帰モデルでユーザーターゲティングを行う
- Chapter6 ディープラーニングで画像分類を行う
- Chapter7 教師なし学習でユーザーセグメントを精緻化する
- Chapter8 レコメンデーションエンジンを実装する
- Chapter9 数理最適化で利益の最大化を図る

## ■書誌情報



ビジネスの現場で使えるAI&データサイエンスの全知識 (できるビジネス)

著者：三好大悟

発売日：2022年3月15日(火)

ページ数：272ページ

判型：A5判

定価：1,980円(本体1,800円+税10%)

電子版価格：1,980円(本体1,800円+税10%) ※インプレス直販価格

ISBN：978-4-295-01363-1

◇インプレスの書籍情報ページ：<https://book.impress.co.jp/books/1121101015>

◇Amazonの書籍情報ページ：<https://www.amazon.co.jp/dp/4295013633>

◇書影（高解像度）ダウンロード：<https://dekiru.net/press/501363.jpg>

## ■著者プロフィール

三好 大悟（みよし・だいご）

慶應義塾大学理工学部で金融工学を専攻。大学卒業後、株式会社データミックスにてデータサイエンティストとして、統計学や機械学習を用いたデータ分析・アルゴリズム開発を中心としたコンサルティングに従事。2020年7月からは株式会社セブン&アイ・ホールディングスにて、小売や物流・配送などの事業におけるデータ・AI活用を推進。一方で兼業としても活動し、データ分析やAI開発など、データサイエンスに関するアドバイザーや受託開発を行うかたわら、講師活動なども行う。

以上

---

### 【株式会社インプレス】 <https://www.impress.co.jp/>

シリーズ累計7,500万部突破のパソコン解説書「できる」シリーズ、「デジタルカメラマガジン」等の定期雑誌、IT関連の専門メディアとして国内最大級のアクセスを誇るデジタル総合ニュースサービス「Impress Watch シリーズ」等のコンシューマ向けメディア、「IT Leaders」、「SmartGrid ニュースレター」、「Web 担当者 Forum」等の企業向け IT 関連メディアブランドを総合的に展開、運営する事業会社です。IT 関連出版メディア事業、およびデジタルメディア&サービス事業を幅広く展開しています。

### 【インプレスグループ】 <https://www.impressholdings.com/>

株式会社インプレスホールディングス（本社：東京都千代田区、代表取締役：松本大輔、証券コード：東証1部9479）を持株会社とするメディアグループ。「IT」「音楽」「デザイン」「山岳・自然」「航空・鉄道」「モバイルサービス」「学術・理工学」を主要テーマに専門性の高いメディア&サービスおよびソリューション事業を展開しています。さらに、コンテンツビジネスのプラットフォーム開発・運営も手がけています。

### 【本件に関するお問合せ先】

株式会社インプレス 広報担当：丸山

E-mail: [pr-info@impress.co.jp](mailto:pr-info@impress.co.jp) URL: <https://www.impress.co.jp/>

※弊社はテレワーク推奨中のため電話でのお問い合わせを停止しております。メールまたは Web サイトからお問い合わせください。