

編集の序

インターネットをはじめとする情報通信技術の発達によって、デジタル化されたデータが世界中を飛び回る世界が実現しました。例えば近所のコンビニエンスストアで買い物をしたとき、レジに打ち込まれた商品のデータは即座に集計され、いつ、どの店舗で、どの商品が売れたのか分析することができます。こういったことを一昔前にやろうとすると、全国の各店舗から伝票を集めて集計しなければならず、膨大な手間と時間のかかる作業となります。それが今では商品のバーコードを読むだけで済むのです。このように発達した通信技術によって集められたデータは膨大な容量となります。これも人力で解析していたら大変な作業となるので、高パフォーマンスのコンピュータを使って機械的に解析する必要があります。そこで活躍するのが人工知能（Artificial Intelligence, AI）です。膨大な容量のデータを機械に学習させることで、あたかも人が行うような、あるいは人にはできないようなタスクをコンピュータに行わせ、これまでの解析手法では得られなかった知見を得たり、それをさらに新しい技術開発に利用することが可能になってきています。

通信技術やコンピュータの発達によって起こっている一連の技術革新は、18世紀の産業革命に匹敵するインパクトを我々の社会にもたらすかもしれません。医療の分野もその例外ではありません。近い将来、あらゆる病院でカルテは電子化され、患者のデータにどこからでもアクセスできるようになり、膨大な容量を持つ個人のゲノム情報を臨床の現場で活用できるようになるでしょう。

このように我々の社会は大きな変革のときを迎えています。しかし残念ながら、このような大きな変化に対して人材の育成が追いついていないのがわが国の現状です。デジタルデータを扱える知識や技術を持った人材が、社会のあらゆる分野で不足しています。このような人材を育成することは、一刻の猶予もない喫緊の課題となっています。これに対してさまざまな施策が打ち出されています。今では小学生の一人一人にタブレット端末が配布されるようになりました。自分の子どものころを思い出すと、学校から配布されるのは習字道具やそろばんでしたが、これがタブレット端末に取って代わられたわけです。もちろん、現在でも小学校で習字やそろばんの授業は行われていますが、その意味合いは昔と異なっているでしょう。

その昔、寺子屋で教えることを読み、書き、そろばんと言いました。今の言葉で言えばリテラシーということになります。リテラシーとは、何かを学習するときの前提となる必要最低限度の知識や技術のことです。確かに文字が読み書きできなければ、

何も学習することはできません。現在、デジタルデータ革命とも言える社会の変化を迎え、新しい人材を育成するための新しい読み書きそろばん、すなわち新時代にふさわしい新しいリテラシーが必要となっています。つまり、AI・データサイエンス教育のためのデータリテラシーが必要ということです。

現在ではAIやデータサイエンスそれ自体を研究の対象とする専門家でなくても、あらゆる学問を学ぶときにこのデータリテラシーが必要となります。初等教育の段階からデジタルデータに触れ、それをコンピュータなどを用いて分析するという作業に慣れ親しんでおく必要があります。さらに、高等教育においてもデータリテラシー教育が必要となります。

ここで、医療従事者はAIやデータサイエンスをユーザーとして利用する立場だから、理屈についてそれほど詳細に理解しておく必要はないのではないかと考える人がいるかもしれません。しかし、例えばX線写真について、X線に関する物理学的知識を全く持たず、ただそこに写った白い影が肺であるとして胸部X線写真を見ている医師と、なぜフィルム上に肺が白く写るのか理解したうえで読影している医師とで、そこから得られる情報に差が出ることは明らかでしょう。AIやデータサイエンスについても同じことが言えます。ユーザーとしてこれらを利用する立場であったとしても、必要最低限度の基本的理論について理解したうえで利用すべきです。さらに医療の分野ではインフォームドコンセントの問題があります。あらゆる医療行為は患者に対して十全に説明し、同意を得たうえで実施する必要があります。AI技術を医療分野に応用するにあたって、それを利用する医師はその原理をしっかりと理解しておかなければ、患者に対して説明責任を果たすことができないでしょう。

このように、これからは医療従事者であってもデータリテラシーを身につける必要があるのですが、今では科学技術の発達によって専門知識の高度化、専門化が進んでいます。例えば、さまざまな疾患に関する発病のメカニズムは分子のレベルで解明されています。学ぶべき医学的知識や技術は一昔前と比較して急激に増加しており、6年間の医学教育の期間に収めるのに苦労するほどです。ここにさらに、データリテラシー教育を加えなければなりません。そのために医療従事者を志す者が無理なく効率よく学べる、いわば医療に特化したデータサイエンス教育が求められています。そのための教材として活用してもらうことが本書の目的となっています。

本書の特徴は実際の作業の副読本として利用できることです。何かを学ぶとき、実際に手を動かし、自分で試行錯誤してみることが重要であることは言うまでもありません。AI・データサイエンス教育も例外ではありません。本書では学習用のデモデータやAIプログラムのコードが豊富に提供されています。これらを使えば、自分でAI

プログラムを動かし、データを処理・解析させてみて、AIを体感しながら学ぶことができます。

「AIって難しそう」「捉えどころがなさそう」——そんな印象を抱いている方が、本書によってAIや機械学習、深層学習を少しでも身近に感じていただければ幸いです。もちろん、この本で学ぶのはデータリテラシー、つまりAIやデータサイエンスにおける読み書きの部分です。AIやデータサイエンスの分野は現在も急速に進歩し、応用の範囲を広げています。医療分野でも予想もしなかったような新技術が次々に実用化されていくでしょう。そこにはまだ見ぬ新しい医療の世界が広がっています。本書で身につく知識や技術は、そのような世界への扉を開く鍵となります。本書を通じてAIやデータサイエンスに興味を持った読者は、本書に掲載された参考文献などを手掛かりに、さらに本格的なAI・データサイエンスの世界へ進んでいただきたいと思います。本書がその一助となることを望んでいます。

2024年7月

中林 潤