

# 慶應義塾大学理工学部におけるデータサイエンス教育について

田中敏幸  
慶應義塾大学 理工学部

キーワード：データサイエンティスト検定、ビジネス力、教育理念

## 1 はじめに

近年、多くの企業や組織でデータサイエンティストが必要とされている。それを踏まえて、文部科学省が、数理・データサイエンス・AI教育プログラム認定制度を作り、全国の大学に参加を呼び掛けている。この教育プログラムに対して、各大学での対応は異なっており、それが大学の特色とも関連している。本稿では、慶應義塾大学理工学部として、どのように対応しているかについてまとめてみることにした。

## 2 データサイエンスとはなにか

大学におけるデータサイエンス教育についてまとめる際に、そもそもデータサイエンスとはどのようなものなのかを考えておく必要がある。データサイエンスと聞いた時、AIを活用するということについては、全ての人からコンセンサスを得られると思われるが、それ以外にはどのように考えればよいのだろうか。

データサイエンスという言葉は、「統計学と情報工学などの手法を組み合わせ、ビッグデータから問題解決に必要な知見を引き出す研究分野」と定義されているようである。しかし、マスコミなど取り上げられているデータサイエンス教育とデータサイエンティスト教育というものが必ずしも同じではないように思われる。まずはその点から考えていく必要がある。次の節で、「データサイエンティスト教育」と「データサイエンス教育」についてどのような違いがあるかについて少しまとめてみることにする。

## 3 データサイエンティスト教育

まず、データサイエンティストとデータサイエ

ンスの違いについて考えてみることにする。データサイエンティストとしての知識と能力を評価する試験として、データサイエンティスト検定がある。この検定試験では、データサイエンス力、データエンジニアリング力、ビジネス力の三つの能力を評価している。次に、それぞれについてどのようなものか簡単に説明することにする。

### 3.1 データサイエンティスト力

データサイエンス力というのは、主にデータサイエンスに必要な情報処理、人工知能、統計学など情報科学計算を扱う能力である。データ分析を行う上で必要な機械学習の理解もデータサイエンス力に含まれる。線形代数や確率統計など数学の基本的な能力がデータサイエンティスト力の中心となっている。

必要な数学および能力として、統計数理基礎、線形代数基礎、微分・積分基礎、集合論基礎、データ理解、回帰・分類、評価、推定・検定、性質・関係性の把握、因果推論、サンプリング、データクレンジング、データ加工、特徴量エンジニアリング、表現・実装技法、意味抽出、時系列分析、機械学習、深層学習、自然言語処理、画像認識、映像認識、音声認識、パターン認識などがある。

### 3.2 データエンジニアリング力

データエンジニアリング力というのは、プログラムの設計や実装など直接的にシステムを作る力である。データサイエンス力で記述した機械学習の具体的な設計の工程などもここに含まれる。データサイエンティスト力とデータエンジニアリング力は車の両輪のようなもので、両者があって初めて力を発揮できる。

重要なスキルとして、Python などのプログラミングスキル、システム企画、システム設計、アーキテクチャ設計、クライアント技術、通信技術、データ抽出、データ収集、データ構造の基礎知識、テーブル定義、分散技術、クラウド、フィルタリング処理、ソート処理、結合処理、前処理、マッピング処理、サンプリング処理、集計処理、変換・演算処理、データ出力、データ展開、データ連携、プログラミング、アルゴリズム、SQL、IT セキュリティ、攻撃と防御手法、暗号化技術、認証などが挙げられる。

### 3.3 ビジネス力

ビジネス力というのは、データ分析を通じてビジネスに貢献する力である。分析結果を出力して終わりではなく、仮説や検証を繰り返して、実際にビジネスの改善に関与する技量が重要である。データサイエンスは、ビジネスを改善することによって、会社の業績を上げていくことが目的なので、分析を行うだけでは十分ではない。ゼロから会社や事業を創り出す人をアントプレナー（起業家）というが、アントプレナー力がなければせっかくの分析結果が無駄になってしまう。

他の必要な能力として、ビジネスマインド、データ・AI 倫理、コンプライアンス、契約、構造化能力、言語化能力、ドキュメンテーション、説明能力、AI 活用検討、スコーピング、データ入手、分析アプローチ設計、データ理解、意味抽出・洞察、評価・改善方法、リスクマネジメントなどが挙げられる。

### 3.4 データサイエンス教育

前節まで、データサイエンティスト教育に必要な 3 つの能力について説明を行った。それでは、データサイエンス教育は、データサイエンティスト教育とどのように違うのだろうか。簡単に言うと、データサイエンティスト教育のビジネス力に関する教育を除いた内容がデータサイエンス教育となる。

ビジネス力まで含めた教育を行うためには、かなり特殊な授業が必要となる。本塾大学理工学部では、管理工学科がそのような授業科目を設置し

ている。アントプレナー力を養うためには、座学の授業だけでは無理で、フィールドワーク的な内容を含む必要がある。授業の一環として理工学部の文化祭（矢上祭）に模擬店などを出店し、どのような店にするかについての検討、ビジネスモデルとしたときにどのような成果があったのか、改善するにはどのようにしたらよいかなどを現実の問題として扱っているようである。少人数教育で行う必要もあり、学部全体で行うことは難しい。理工学部で行うことができるのは、データサイエンス教育までと考えたほうがよい。ビジネス力については、各学生が独自に身に付けていく必要がある。社会で必要としているのはデータサイエンティストのようであるが、ビジネスセンスを養うのはかなり難しく、時間がかかると思われる。

## 4 慶應義塾におけるデータサイエンス教育

### 4.1 SFC におけるデータサイエンス教育

慶應義塾大学湘南藤沢キャンパスの総合政策学部・環境情報学部では、文理融合・学際的研究のためのキャンパスとして、数理・統計科目の習得を重要視している。このキャンパスでは、社会の問題を的確に認識し、解決していく手法データビジネスにおける利用データサイエンス人材の育成に積極的に取り組んでいる。

データサイエンス人材を育てるためには、問題発見・問題解決能力の育成が不可欠であり、そのためには数学・統計学が極めて必要となる。SFC はいろいろな大学入試形態をとっており、入試科目で数学を選択せずに大学に入学している可能性もある。これらの学生も含めてデータサイエンス教育を行うには、データサイエンス科目の必修化が必要だと言われている。もともとが分離融合型の学部なので、近年の話題となっているデータサイエンス教育には対応しやすいものと考えられる。

### 4.2 理工学部におけるデータサイエンス教育

次に慶應義塾大学理工学部におけるデータサイエンス教育についてみてみることにする。理工学部は表 1 に示すように、大学入学時に 5 つの学門のいずれかを選んで、入試を受けることになる。近年、学門制の変更があり、学門 A, B, C, D, E と

いう名前になっている。それぞれの学門は表に示す分野で分けられている。学門 C が、情報・数学・データサイエンス分野となっており、データサイエンスに興味のある高校生は、この学門を受験するものと思われる。学門 C に割り当てられている学科は、情報工学科、数理科学科、管理工学科、生命情報学科である。

表 1：理工学部の学門制

学門	対応学科
学門 A (物理・電気・機械分野)	物理学科、物理情報工学科、電気情報工学科、機械工学科
学門 B (電気・情報分野)	電気情報工学科、情報工学科、物理情報工学科、システムデザイン工学科
学門 C (情報・数学・データサイエンス分野)	情報工学科、数理科学科、管理工学科、生命情報学科
学門 D (機械・システム分野)	機械工学科、システムデザイン工学科、管理工学科
学門 E (化学・生命分野)	化学科、応用化学科、生命情報学科

大学 1 年時は、全学科に対してほぼ同様（若干の違いはあるが）の数学・物理・化学の教育が行われている。大学 2 年以降の授業については、各学科に任されており、同じ名前の授業があったとしても、教える内容については異なっている可能性がある。

#### 4.3 理工学部の教育理念

慶應義塾大学理工学部は、王子製紙の会長だった藤原銀次郎氏が私財を投じて「藤原工業大学」を設立したことから始まる。藤原工業大学は、後に慶應義塾に寄付され、現在の慶應義塾大学理工学部となっている。藤原工業大学が寄付されたとき、藤原銀次郎氏としては社会（会社）で役に立つ即戦力の教育をしてほしかったようである。しかし、寄付を受けた際の慶應義塾の小泉信三塾長の「すぐに役に立つものは、すぐ役に立たなくなる」「すぐ役に立たないようなことを教えれば、生涯ずっと役に立つ」という言葉は印象的なもので

ある。この考え方が、今のリベラルアーツにもつながっているように思われる。

即戦力になる人物をスペシャリスト、即戦力ではないが基礎学力を基にして自分自身で開発していく人物をジェネラリストとすると、慶應義塾大学理工学部ではジェネラリスト教育を目指していることになる。そのようなことから、学生が各自でデータサイエンスに対する意識を持ち、自分自身が目指すデータサイエンティストに向けて自分自身を鍛えていくことになる。つまり、慶應義塾大学理工学部では、学部としてデータサイエンティストに特化したカリキュラムを作るということは理念としてふさわしくないように思われる。

#### 4.4 学生主体のデータサイエンス教育

先にも述べたように、理工学部の学門制の学門 C に「データサイエンス分野」というものがあり、確かにデータサイエンスの基礎になる数学、プログラミング能力、ビジネス力を強化する科目などが多い学門となっているが、学生が受け身になっていてはデータサイエンティストになることはできない。個々の学生の意識というものが重要な要素となっている。また、学門 C 以外でもデータサイエンスに必要な数学科目やプログラミング科目は多数用意されている。さらに、各学門の中に情報系の研究室は多数存在するので、研究室の研究テーマによっては、データサイエンティストの素養を身に着けることは十分に可能となっている。

プログラミングは、データサイエンスのために重要な能力である。慶應義塾では、AI・高度プログラミングコンソーシアムを設立し、義塾の学生に対して講習会やプログラミングコンテストなどのサービスを提供している。また、国際大学対抗プログラミングコンテストへの出場なども積極的に勧めている。ビジネス力については、現時点で対応できているのは、管理工学科だけのように思われる。いずれにしても、自分をどのように鍛えて能力を身に着けていくかは、学生次第となる。

#### 5 おわりに

本稿では、慶應義塾大学理工学部としてどのようなデータサイエンス教育を行っているかについて

で説明した。まず、慶應義塾大学理工学部では、データサイエンスのスペシャリストではなくジェネラリストとしての教育を行っている。これは、慶應義塾大学理工学部設立当時の理念にも関係している。各学科で、データサイエンティスト教育に必要なデータサイエンス力、データエンジニア力に関連する科目は用意されている。ビジネス力までを考慮した科目設置としては、管理工学科が最も対応できている。

今回はデータサイエンティスト検定に必要な教育を中心に記述したが、今後のデータサイエンスがどのようなものかについては検討が必要である。現在、各企業でデータサイエンティストが必要となっていることは、マスコミが報道しているとおりである。今後も同じタイプのデータサイエンティストが必要になるのか、また違ったタイプのサイエンティストが必要になるのかについては、常に観察が必要である。また、即戦力になるかどうかを企業が判断するとき、検定試験や認定試験などの結果が必要になるが、どのような認定試験が出てくるのかについても注意しておかなければならない。