

## データサイエンスの教育が大学では必須になりつつあります

### 1. はじめに

データサイエンスというのは比較的最近使われるようになった言葉です。

コンピュータが登場した際には、データ処理（データプロセッシング）が情報処理の基本的なテーマでした。情報処理に適したきちんと整理されたデータを扱っていました。いわゆるデータベースはそのようなデータを意味します。その頃は情報処理システムのことを普通に EDPS（電子式データプロセッシングシステム）とっていました。第1次 AI（人工知能）の時代です。

ほぼ第2次 AI ブームに呼応しているある時期からデータ工学（データエンジニアリング）という言葉がはやりはじめました。多様で膨大なデータ（従来では取扱いを諦めていたものや、捨てていたようなものに範囲を広げた）、あるいは曖昧なデータ、分類しにくいデータや感性データを取り扱うようになりました。第2次 AI の特徴である知識や推論を情報処理に取り込んだのもその時代の特徴です。＜当時私は川崎重工業(株)に在籍しており、高信頼度知能化船高度自動運航システム：言葉を変えれば船長のエキスパートシステムの開発に取り組んでいました＞

第3次 AI ブームを迎えている今では深層学習を装備したニューラルネットワークやデータマイニング（データ発掘、データ採掘）という言葉が流行語になってきました。

そしてそれらを俯瞰したデータサイエンスという言葉が流行しています。大学の教育においても、データサイエンスを看板に掲げる学部が出てきています。岐阜県に近いところでは、名古屋市立大学に2025年に学部の新設予定が報じられており、またすでに滋賀大学では2017年にデータサイエンス学部ができて、卒業生が出ています。学科の新設（改称を含む）や授業内容に反映させるところは枚挙にいとまがないほどです。

文部科学省でもデータサイエンス教育が大学生全般に極めて重要であるという認識が出てきたようで、「大学の数理・データサイエンス教育強化方策について」（2016年12月 文部科学省）が出され、「第5期科学技術基本計画に依り、幅広い分野で活躍する新しい人材の供給により、Society 5.0の実現とその中での競争力の向上を目指す」と謳われています。そのおかげで、文系理系を問わず大学においては、必須事項として数理・データサイエンス教育に取り組むようになってきています。2019年に2月に着任した名古屋工業大学でもその流れを受けて、真正面から取り組んでいました。

## 2. 数理情報、数理情報基礎、数理情報概論

ひょんなことから着任早々(2019年4月)に、数理・データサイエンスに取り組む科目を担当することになり、教材づくりから始めてぶっつけ本番で授業に取り掛かりました。

ただしその科目は創造工学教育課程<sup>注)</sup>の非情報系の学生向けのもので、文部科学省が推進する全学共通の科目ではありませんでした。しかしながら、トライアル的な科目としてとらえてよいと思います。

2019年にスタートしたのは、科目名は「数理情報」という科目です。講義を4回程、別の情報系の先生が担当し、その後の12回の演習を私が担当しました。私の担当分はすべて演習で、コンピュータ演習室を利用した授業でした。

2020年には、「数理情報」という科目はなくなり、名前を「数理情報基礎」という科目名に変わり、現在に至ります。しかしながら、基本的な考え方は踏襲され、全16回を一人の先生が担当し、また全面的に演習科目としてすべての授業をコンピュータ演習室で実施しています。

一方、名古屋工業大学の全学共通でかつ必修科目として「数理情報概論」という科目が2021年度から正式にスタートしました。全学に対応する必修科目なので、内容の審議とコンセンサスをとるのが大変で、カリキュラム改定にはかなり時間がかかりました。

正式科目の「数理情報概論」は全学共通の科目であるため、専門学科の教員ではなく数学教室の教員がメインとなって担当しています。また講義中心の科目であり、演習は極めて少なくなっています。そのため、私が担当した演習中心の「数理情報基礎」とは若干内容は異なっています。その違いを紹介したいと思います。

注) 名古屋工業大学の創造工学教育課程は、新しい枠組みの教育課程であり、工学分野における学部・大学院<博士前期課程>6年一貫の課程である。ここでは、『主軸となる分野の専門的な知識・経験を持ちながらも、様々な分野を横断的に俯瞰し新しい「価値」を創造する人材を育成している』ところが特徴である。2021年度に第1期の修了生が誕生する予定である。なお、従来からの課程<学部、4年生課程>は高度工学教育課程と称している。

## 3. 数理情報基礎の主な内容

ここでは、大学教育への数理・データサイエンス教育強化の一例として、「数理情報基礎」のシラバスを紹介します。

---

### [ 日本技術士会岐阜県支部 会報の情報連絡先 ]

〒509-0108 各務原市須衛町1-179-1 テクノプラザ5F  
TEL : 0583-79-0580 FAX : 0583-85-4316 Email : gcea9901@ybb.ne.jp

表1に「数理情報基礎」のシラバス(2020年度実施したもの)を示します。

全16回を田中が担当し、また全面的に「演習科目」としてすべての授業をコンピュータ演習室で実施しています。

基本的な確率統計の話からはじめ、データの解析、データの認識、データの活用という展開になっています。

演習には当時コンピュータ演習室に標準装備されていたRというデータサイエンス向けの言語があったため、それを利用しました。Pythonは当時標準には装備されていなかったため、演習には利用していませんでした。

ニューラルネットワークやデータマイニングもさわりだけではありますが、一応俯瞰しています。ただ急ごしらえの教材であり、手持ちのネタがC言語やJavaで書かれたものであったため、教材としては非情報系の学生には適していなかったと反省しています。

コロナ禍の関係もあって、対面の期末試験を実施しづらかったため、課題だけで成績評価をしています。

表1 数理情報基礎のシラバス、抜粋(2020年度)	
授業科目名	数理情報基礎
担当教員名	田中 秀和
学科・年次	工学研究科博士前期課程(新) 工学専攻 1年次
科目区分	専門教育科目、単位数1
授業の目的・達成目標	
	1. 情報化社会においては、データの収集、解析、認識、活用 の技術が必須であり、そのベースとなる技術を、演習を通じて 学ぶ。
	2. データの収集及び解析については、Rというデータサイエン スソフトウェアを利用する。
	3. データの認識、活用あるいはデータマイニングについては、 その基礎を学ぶ。
授業計画	
	1. イントロダクション、数理情報とは、データサイエンスとは
	2. Rとデータサイエンス、データの入出力
	3. マトリックスの取り扱い
	4. データグラフィックス(データの視覚化)
	5. 確率・統計と確率分布(正規分布)
	6. 確率分布(二項分布ほか)
	7. 推定
	8. 検定、二項検定、t検定、カイニ乗検定
	9. z検定、母比率の検定、分散分析
	10. 予測・判別・分類、回帰分析
	11. 判別分析
	12. 主成分分析、因子分析
	13. クラスタ分析
	14. サポートベクターマシン
	15. ニューラルネットワーク
	16. データマイニング

#### 4. 数理情報概論の主な内容

名古屋工業大学の全学共通でかつ必修科目としてスタートした「数理情報概論」という

[ 日本技術士会岐阜県支部 会報の情報連絡先 ]

〒509-0108 各務原市須衛町1-179-1 テクノプラザ5F  
TEL: 0583-79-0580 FAX: 0583-85-4316 Email: gcea9901@ybb.ne.jp

科目の概要を表2に示します。

「講義科目」であり、演習は4回程度に収めており、「数理情報基礎」と比べて、極めて少なくなっています。また学部の低学年向けであるため、推定問題に重点が置かれています。

大学の科目の特徴ではありますが、「演習科目」は授業時間に比べて単位数は少なく(1単位)、「講義科目」の単位数は多くなっています(2単位)。

＜講義科目には、予習時間+授業時間+復習時間を基本として数えるため、単位数が多くなる。演習科目はその予習や復習を含まないため単位数が少なくなる＞

## 5. おわりに

ここでは、大学教育への数理・データサイエンスの潮流を紹介しました。内容が情報工学部門のキーワードに限定されてしまったようではありますが、これからの技術の主要分野としてのこの流れを汲み取っていただければ幸いです。

現在、日本技術士会では技術士倫理綱領の改定に向けての実務作業がスタートしています。まず技術士倫理綱領の改定をすべきかどうかを検証する審議からはじまりました。

その議論に関するワーキンググループ(WG)「倫理綱領検証WG」が2019年9月にスタートした際に、私は、縁あって、「倫理綱領検証WG」の委員に選出され、倫理綱領の検証作業に携わってきました。

2021年2月の理事会では、WGの検証報告を受けて、技術士倫理綱領は改定すべきであるという結論を出しました。そしてその後WGは「倫理綱領改定WG」という名前に変わりました。WG委員は横滑りで全員がそのまま委員として残り、統括本部の倫理委員会のメンバーの出入りはありましたが現在、改定作業に入っています。

ただ改定の結論はまだ先のことになるため、次の機会にでも議論の内容を紹介したいと考えています。

表2 「数理情報概論」シラバス (2021.3.29、名古屋工業大学における数理情報履修プログラムの紹介セミナーから)

第1回	はじめに
第2回	線形モデルと最小二乗法
第3回	確率モデル
第4回	統計的推測・Pythonによるデータ分析演習1
第5回	行列・ベクトルを用いた最小二乗法の定式化
第6回	最小二乗法の幾何学的解釈
第7回	最小二乗法と最尤推定法
第8回	さまざまな確率分布と統計的推測・Pythonによるデータ分析演習2
第9回	ロジスティック回帰分析
第10回	クラス分類・パターン認識
第11回	線形モデルの解釈・Pythonによるデータ分析演習3
第12回	特徴選択とモデルの正則化
第13回	モデルの選択と評価
第14回	非線形モデリングとニューラルネットワークの基礎・Pythonによるデータ分析演習4
第15回	期末試験
第16回	データ利活用上の心得