

## 野口 怜 研究室

経営学部 ビジネスエコノミクス学科 准教授

のぐち れい  
野口 怜 先生



取材に参加した3年生の大原さんと朝倉さんにコンペの手応えを聞く

# テキストや数値などのデータから 機械学習技術を駆使し、新たな知見を発掘する

### 課題や価値を設計するビジネス力が重要

企業が最善かつ戦略的な意思決定を行うには、特定のデータを統計的・定量的手法を用いて分析する「ビジネス分析」がより重要となっており、ビジネスエコノミクス学科では経済学、経営学、金融工学、統計学、データ科学を幅広くかつ専門的に学べる、ビジネス分析に長けた先端的教育を行っている。同学科に属する野口怜研究室では、このビジネス分析の一種である、テキストマイニングやデータマイニング、医療経営データの分析を研究している。

テキストマイニングとは、SNSの投稿や商品の口コミ、ニュース記事など、インターネット上に溢れるテキストデータから、統計学やAIの分析手法を駆使し、新しい知見を発掘する技術である。

「例えば、企業の業績は財務諸表や決算書の数値などから測れますが、こうした資料が公表される前にその企業に関するニュース記事などから業績を読み取れる可能性があります。あるいは災害が起こったとき、報道機関が被害状況を伝える前にX(旧Twitter)などのSNSではすでに様々な情報が飛び交っているケースがあります。テキストデータは、リアルタイムに、かつ現場の状況を細かく描写できるという点で大きな可能性を秘めたデータといえます」と野口先生は話す。

決まりきったアプローチ方法があるわけではないため、“マイニング(採掘)”の言葉の通り、分析方法などは手探りで試行錯誤することが求められる。研究室ではデータサイエンティストに求められる3つのス

キル、①データサイエンス力、②データエンジニアリング力、③ビジネス力をバランスよく育む実践的な教育に力を入れている。

「データサイエンス力とは、統計や分析手法の知識など数理的な力。データエンジニアリング力とは、作ったモデルをアプリに実装する、データベースにアクセスして必要なデータを持ってくるなどのIT力。そしてビジネス力は、『何が課題なのか』を把握して、実際に分析を行うことで『どういう価値を提供できるのか』を示せる、課題と価値を設計する力といえます。生成AIへ指示を出して最適な結果を引き出すプロンプトエンジニアリングでデータ分析ができる時代になった今、人間だけが持つ過去の経験や知識の価値が高まっています。そのため教育においても、ビジネス力をいかに鍛えるかを重視しています」

分析手法を考え、その精度を高めていくことだけでなく、自由度の高いテーマを与えて「この企業の課題は一体何か」「課題を解決するためにはどんなソリューションが必要か」を考えて課題を設計する。こうしたビジネス力を育む機会を意識的に作り出しているという。

### データサイエンスを活用し 医師の負担や見落としを減らす

野口研究室ではもう一つの軸となる研究として、データサイエンスを活用した効果的な医療経営・施策の実現に取り組んでいる。

例として挙げられるのが、検査によって診断が下る前に、医師が書いた電子カルテのテキストデータから患者の候補病名を推測するAIの開発である。これにより、別の診療科の医師が診療する夜間救急外来でも、医療技術の均てん化が図れるのではないかと考えている。

また、大学病院の医師と連携し、電子カルテ確認作業を効率化するAIの開発も進めている。病院では、検査をして予期せぬ別の病変が見つかったときに、別の診療科に紹介してから、その診療科の医師が患者の電子カルテを確認したか、新たに検査をしたかなど、適切な対応が行われたかをトレースして確認する作業が行われている。医師はこの確認作業に多くの時間と労力を費やしているため、この確認作業を自動化することで、医師の働き方改革への支援につなげようとする研究だ。

「分析手法は、分野によって特色があります。経済なら計量経済学で生まれた分析手法が、医療なら医療統計学で使われている分析手法があり、それらの手法は互いの分野では活用されてきませんでした。しかし、手法は分野をまたいで使えるだろうというのが私の一つの信条で、医療分野で適応してきたAIは経営・経済の分野でも適応できるのではないかと考えています」

## 引き出しを増やすべく、リアルな体験を積み重ねる

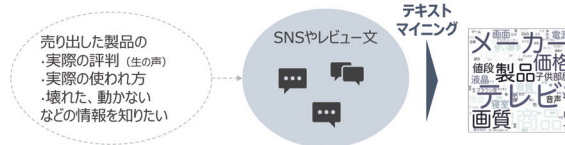
野口研究室は2024年に発足し、ゼミではデータサイエンス関連の輪読会やロジカルシンキング・仮説思考の演習、課題解決ベースのデータ分析演習などを行っている。大学の演習で用いられるのはサンプルデータが多いため、「リアルなデータを使って、リアルな課題にあたる」経験を積むべく、学生を対象に開かれる企業のデータ分析コンペなどへの自主参加も歓迎している。

学部3年生の朝倉拓生さんと大原蒼生さんは、2024年夏に企業が主催したデータ分析コンペにそれぞれ友人とチームを組んで参加した。コンペのテーマは年齢や性別、家族構成、年収などの基本情報のデータから旅行パッケージの成約率を予測するというものだ。

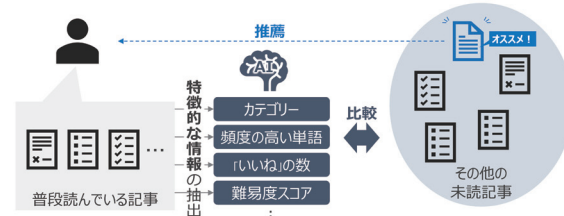
「分析の手法を前期で学び、それを活用しながらコンペに挑みました。データの前処理を行い、自分たちが持っている分析手法の引き出しに当てはめてみて、

## ✓ テキストマイニングによる口コミや新聞・ニュース記事等からの知見抽出

例：SNSやレビューによる製品の市場評価・市場品質の分析



例：ユーザーの関心や難易度に合った記事の推薦AI構築



テキストマイニングを活用した研究例

どの分析手法にするかを選んでいくのですが、どのデータをモデルに組み込めば精度が上がるのかを見極めるのは闇雲な作業で、正答率を0.1%上げるのもかなり大変でした」（朝倉さん）

「前処理と予測の手法は色々なやり方がありますが、慣れてきたら直感で分かってくると思いますが、僕たちの段階だとまだ総当たりで取り組む状態。こうすれば正答率が上がるというのは決まっていないので、手当たり次第試していくことになります。こうした作業は1人だとつらいので、友人と集まって話しながら進める時間と個人で集中してプログラミングする時間のバランスをとることを日頃から意識しています」（大原さん）

王道の手法がない分、分析手法の引き出しをたくさん持つことが大切と野口先生は常に学生に話しているという。

「トップのデータサイエンティストでも行き詰まるときは必ずあります。そのときにどれだけ自分が引き出しを持っているか、そしてどの引き出しをいつ引いたらいいかという感覚を持っていることが大切です。学生のときは引き出しを増やせる絶好のチャンスだと思うので、コンペに参加したりして、リアルな体験や苦勞を積み重ねて、引き出しをどんどん増やしてほしいと思います。研究室でも産学連携に積極的に取り組むつもりですので、リアルなデータ、リアルな課題に触れる機会を増やしていきたいです」と野口先生は語ってくれた。

城市 奈那（ジェイクリエイト）