

自動発注問題を題材としたビッグデータ・AI 技術に対する実践的人材育成コースの設計

神田 哲也 福安 直樹 佐伯 幸郎 市川 昊平 中村 匡秀 楠本 真二

ビッグデータ・AI・クラウドの技術を用いて社会の課題を解決できる実践的な人材教育が求められている。「成長分野を支える情報技術人材の育成拠点の形成 (enPiT)」のビッグデータ・AI 分野は、enPiT AiBiC として人材教育活動を行っており、学部生向けのコースを提供している。本稿では、AiBiC 関西地区における、3 技術に対する包括的かつ実践的な知識の習得を目指したコース設計について報告する。

1 はじめに

文部科学省が推進する「成長分野を支える情報技術人材の育成拠点の形成」(通称: enPiT [4]) は、産学協働の実践教育によって高度 IT 人材を育成する取り組みである。enPiT では、図 1 のフレームワークに従って教育プログラムを実施している。このフレームワークの下、ビッグデータ・AI 分野、セキュリティ分野、組込システム分野、ビジネスシステムデザイン分野の 4 分野において実践的な人材教育を目指している。

ビッグデータ・AI 分野では enPiT AiBiC [3] として、ビッグデータ・AI 技術を活用したシステム開発のできる人材育成プログラムの実施に取り組んでい

る。クラウド技術の発展に伴い、様々な情報がビッグデータとして蓄積され、これらの情報を元に様々な付加価値を生み出す AI 技術の利活用が急速に広がっている。このような社会的背景を鑑み、本分野では、様々な社会的課題をビッグデータ処理技術、AI 技術、クラウド技術等の基盤技術を用いて解決すると共に、新たなビジネスや価値の創出を行える人材の育成を目指している。また、産学の協働ネットワークを構築し、多くの優秀な学部学生を育成すると共に、実践的情報教育の知見を蓄積し、学部教育に広く普及させることを目的としている。

本稿では、2017 年度に、関西圏の大学に在籍する学部学生を対象として実施している enPiT AiBiC 関西地区の教育活動について報告する。enPiT AiBiC 関西地区では、本人材育成プログラムにおける一つの取り組みとして、機械学習による小売店の大量データの分析に基づく自動発注問題を題材とした教育コースの開発を行っている。本稿では、enPiT が目指す教育プログラムのフレームワークの下で、どのように基礎知識学習や PBL 演習等のコース設計を行っているのかについて紹介し、現時点での実施状況について報告する。

2 enPiT AiBiC

enPiT AiBiC は現在 AiBiC 東日本・AiBiC 関西・AiBiC 九州の 3 地域で教育プログラムを展開してい

Course Design for Training AI and BigData Experts Using Automatic Ordering Problem

This is an unrefereed paper. Copyrights belong to the Authors.

Tetsuya Kanda, Shinji Kusumoto, 大阪大学大学院情報科学研究科, Graduate School of Information Science and Technology, Osaka University.

Naoki Fukuyasu, 和歌山大学システム工学部, Faculty of Systems Engineering, Wakayama University.

Sachio Saiki, Masahide Nakamura, 神戸大学大学院システム情報学研究科, Graduate School of System Informatics, Kobe University.

Kohei Ichikawa, 奈良先端科学技術大学院大学情報科学研究科, Graduate School of Information Science, Nara Institute of Science and Technology.

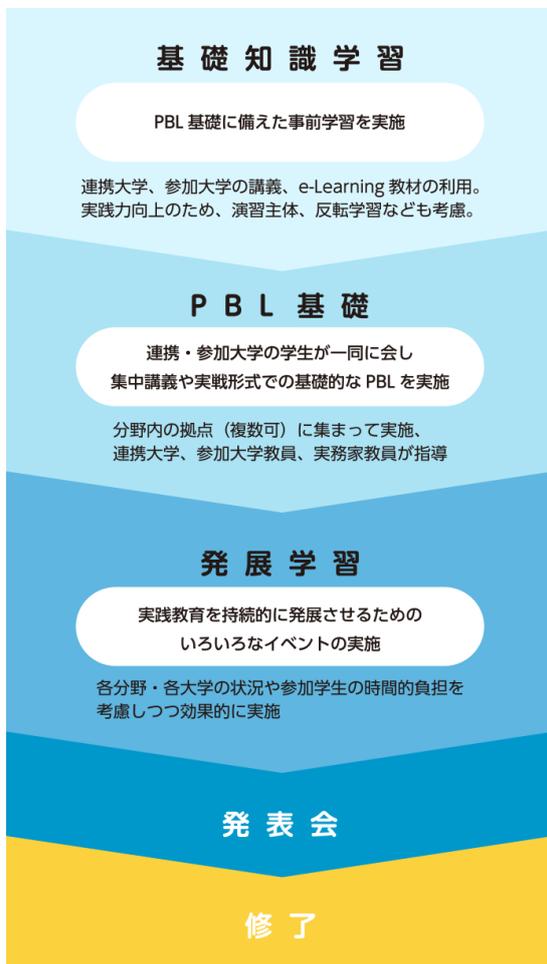


図 1 enPiT のフレームワーク

る。AiBiC 関西[1]は、関西圏の大学を中心として、ビッグデータ・AI・クラウド技術を活用する企業と連携して実施している取り組みである。また、対象を学部学生だけでなく、高等専門学校を学生の試行的に受け入れている。

AiBiC 関西の教育プログラムは、図 1 のフレームワークに基づき、各大学で実施されるビッグデータ・AI・クラウド技術に関連する基礎知識学習、夏季に 5 日間程度の集中講義形式で実施される PBL 基礎、PBL 基礎で行った内容をもとに、発展的な内容を受講生同士が分散した状態で行われる PBL 発展から設計されている。基礎知識学習は主に PBL 基礎に必要な知識習得を目的とするが、各大学の実施内容は

厳密には同一とならないため、PBL 開始時の前提知識の共通化を目的とし、前期期間にひと月に 1 度の割合で基礎知識学習補助講義を開講する。また、基礎知識学習補助では、それぞれの分野に関する技術を先進的に取り入れている連携企業より、実際の企業での利活用例やその技術説明などを企業セミナーとして開催する。AiBiC 関西では、これらの講義、PBL 演習、企業セミナーを通して、ビッグデータ・AI・クラウド技術に関して、基礎知識から、実践的な技術力、および応用知識までを広く包括的に習得した高度な人材を育成することを目指している。

3 自動発注問題を題材とした PBL 課題

PBL 基礎・発展ではビッグデータ・AI 技術を活用したシステム開発のできる人材育成を目的として、小売店における商品発注業務を自動化するプログラムの開発を PBL 課題として選定した。

商品発注においては、商品の在庫と需要予測から、以下の 2 つのロスができるだけ小さくなるように適切な発注数を決定する必要がある。

- 商品が売り切れ、顧客の購買機会を損失する (チャンスロス)
- 商品が売れ残り、廃棄や割引による損失が生じる (実質ロス)

多くの小売店では、これらを考慮した発注業務を担当者の経験や勘で行っているが、PBL では、POS データから得られる過去の販売実績や、天気などから機械学習により需要予測モデルを構築し、そのモデルに基づいて小売店の利益ができるだけ大きくなるように自動的に発注業務を行うシステムを開発する (図 2)。

3.1 データ

本 PBL において用意した実データは以下の通りである。

- POS データ
- カレンダーデータ
- 気象データ

POS データには、スーパーマーケット 116 店舗分の日別販売実績データを用意した。ある商品が 1 日に売れた個数とその売上げを店舗ごとにすべての商

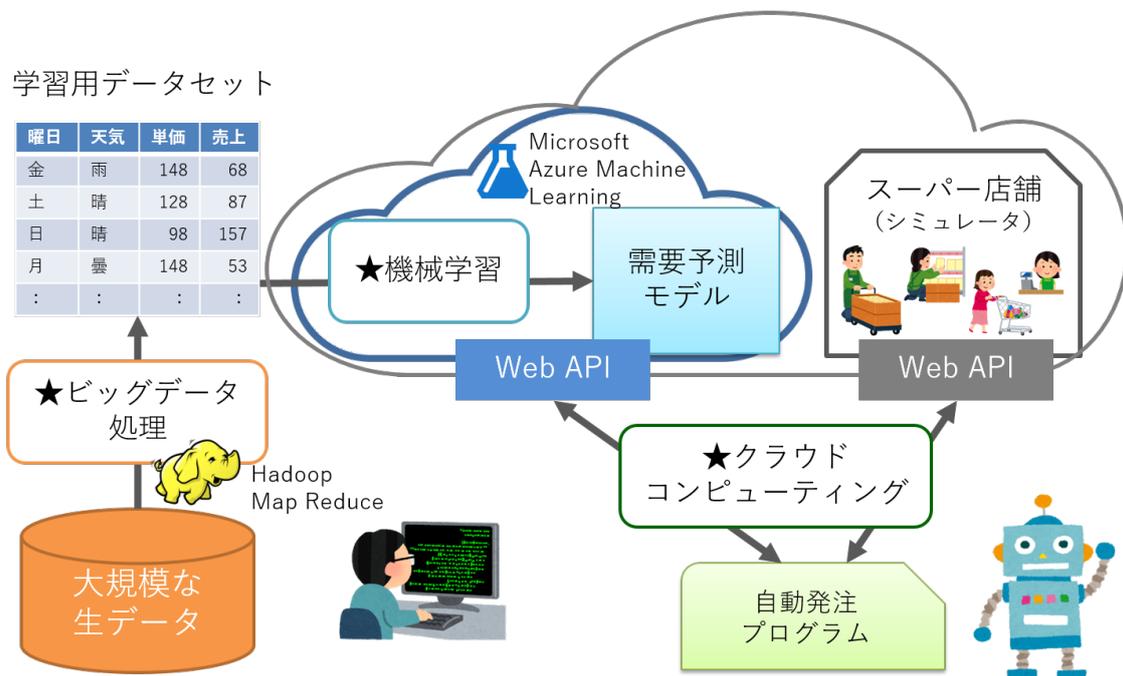


図 2 PBL 課題設定の全体像

品について 2009～2013 年の 5 年間分記録したデータであり、合計 3.1 億レコードである。また上記 POS データの期間に対応して、各日の曜日、祝日などのカレンダーデータ、天気、気温、湿度、降水量などの気象データも合わせて用意した。

3.2 自動発注システム

PBL においてチームごとに実装するシステムであり、需要予測モデルの作成と自動発注プログラムの実装に大きく分けられる。

3.2.1 需要予測モデルの作成

与えられたデータに基づいて機械学習を行うことで、需要予測モデルを作成する。目的変数はある商品の販売個数であり、説明変数には与えられたデータであれば何を利用してよい。需要予測モデルの作成環境には、Azure Machine Learning Studio[5][7] (Azure ML) を利用する。作成した需要予測モデルは Web サービスとしてデプロイすることが可能であり、自動発注プログラムから呼び出すことで利用される。

3.2.2 自動発注プログラムの実装

需要予測モデルに基づいて、明日の需要予測を行い、後述する店舗シミュレータに対して発注業務を自動で行うプログラムである。需要予測モデルで予測した商品の需要予測数と現在の商品の在庫数から発注数を決定し発注業務を行う。このとき、必ずしも需要予測数と在庫数の差により発注数を求める必要はなく、商品の需要の特性や廃棄までの期限を考慮したヒューリスティックがチームごとに工夫されることを期待している。

自動発注プログラムは Python により実装する。店舗シミュレータを実行する Python コードを雛形として与え、ここに

- 需要予測モデルの呼び出し (Azure ML による自動生成)、
 - モデルに与える入力の作成、
 - ヒューリスティックによる発注個数の決定
- のそれぞれを行うコードを追加することで実装する。

3.3 自動発注コンテスト

PBL では、チーム対抗による自動発注システムの能力向上を一つの目標とする。ここで自動発注システムの能力とは、指定された期間で発注業務を行った際の店舗の利益であり、期間中の総売り上げ額から総仕入れ額を引いたものとする。

ただし、コンテストとして利益の数値だけを比較することは PBL の趣旨にそぐわないものとする。特に、偶然良い結果が得られた場合、その経験には再現性がないため、後の活動に活かせる可能性は低いからである。

本 PBL における各チームの工夫のポイントとしては、

- 機械学習アルゴリズムの選定
- 機械学習アルゴリズムに与えるパラメータ
- 機械学習における説明変数の選択
- 需要予測に対するヒューリスティック

などが挙げられる。PBL では、これらについてさまざまな組み合わせを試行することで、より良いシステムを実装することが期待される。そこで、各試行においてこれらの記録を残し、比較することをもう一つの目標とし、成果発表会では比較に基づくプレゼンテーションも評価の対象にすることを検討している。

3.3.1 店舗シミュレータ

本来であれば PBL で開発したシステムを実際の発注業務で活用した上で評価をすることが望ましいが、結果の検証に実時間を要するため、ここでは小売店の振る舞いを 3.1 節のデータに基づいて再現するシミュレータ上で発注業務を行ってシステムを評価することにした。シミュレータの機能は以下の通りであり、PBL で開発する自動発注システムは、あらかじめ用意した REST API を通じてこれらの機能にアクセスする。

- 特定の期間と店舗を指定してシミュレータ（店舗）のインスタンスを生成する
- 数量を指定して商品を発注する
- 過去の日付を指定して店舗の状況（在庫、販売個数など）を取得する
- 過去の日付を指定してその日の気象データを取得する

- 翌日の天気予報を取得する
- 日付を指定してその日のカレンダーデータを取得する

問題をあまり複雑にしすぎないため、今回用意したシミュレータでは、以下の制約を設けている。

- 店舗は毎日閉店する（24 時間営業ではない）
- 店舗閉店後に翌日分の発注業務を行う
- 発注した商品は、翌日の開店前までに納品される
- 仕入れ金額はリアルデータがないので、商品ごとに固定である
- 商品ごとに消費期限が設定されており、消費期限を過ぎた商品は廃棄される
- 商品の消費期限に応じた割引などは行わない（販売価格は実データに基づく）

店舗インスタンスを生成すると、シミュレータ上の日付が指定された期間の前日に設定され、翌日の発注待ちの状態になる。発注を行うと、日付が進められ、商品が納品される。また、POS データに基づいて店舗インスタンスでの販売個数、在庫が更新され、会計処理が行われる。

自動発注システムは、シミュレータの REST API を通じて在庫や販売個数、気象データを取得することができる。このとき取得できるのは、シミュレータ上での過去および現在のデータと、翌日の天気予報である。ただし、天気予報は実データを返しているため、厳密には予報データではない。

4 自動発注問題に基づくコースウェアの実装と 2017 年度の実施状況

4.1 概要

AiBiC 関西では、2017 年度現在、8 大学・1 高専から 52 名の受講生を受け入れて実施している。2017 年度の年間スケジュールを表 1 に示す。受講生は、大阪大学中之島センターに集まり、個別技術を習得するとともに、チームによる活動のためのファシリテーションスキルを獲得する。また、5~6 名の 9 チームに分かれて 3 節で述べた PBL 課題を実施する。

表 1 2017 年度スケジュール

日程	講義タイトル
5月27日	クラウド技術
6月10日	ビッグデータ処理技術
7月1日	AI 技術
8月5日	総合演習
9月4日～8日	PBL 基礎
10月14日	PBL 発展
11月11日	PBL 発展
12月9日	最終成果発表会

4.2 基礎知識補助講義

クラウド技術

クラウド技術の講義では、クラウド技術の発展に関してその歴史的な背景を含めて解説する。特に、クラウドサービスを実現する上で要素技術となる仮想計算機技術に関して詳しく解説し、クラウドサービスが有する柔軟さや、迅速な拡張性などの特徴が仮想計算機技術によってどのように実現されるのかを学習する。また、この講義の後に実施されるビッグデータや機械学習の講義で使用する Amazon Web Services (AWS) [2] の Elastic MapReduce や、Azure ML などのクラウド上のビッグデータおよび機械学習プラットフォームサービスに関して概説する。演習では、クラウドを構成する仮想計算機が如何に容易かつ迅速に生成できるのか、計算機リソースを制御するソフトウェア技術に関して理解を深めるため、AWS 上の EC2 サービスを利用した仮想計算機環境構築の演習を実施する。

クラウド技術に関する企業セミナーは、2017 年度は株式会社 NTT データ・楽天株式会社により開催され、自社内で開発されたクラウドを活用した設計ツールの紹介や、自社サービスを開発する際にクラウドを活用することによる利点などについて講演が行われた。

ビッグデータ処理技術

ビッグデータ処理技術の講義では、ビッグデータの定義や活動事例を紹介した後、MapReduce の考え方について解説する。データ処理の流れを理解し実践するために、基礎部分の説明の後は演習中心の

設計となっている。演習では、コーディングの前に MapReduce の考え方を身に着けるため、まずアナログなグループ演習として、グループメンバーをワークノードに見立てた記号計数ゲームを実施する。その後、使用するツールをローカル環境の Apache Hadoop, Amazon Web Service 上の Elastic MapReduce へと順次移行し、本格的なビッグデータ処理の演習を行う。

ビッグデータ技術に関する企業セミナーは、2017 年度は富士フイルム ICT ソリューションズ株式会社により開催され、自社およびグループ会社でのビッグデータ分析基盤構築やビジネスへの展開例などについて講演が行われた。

AI 技術

AI の講義は、人工知能に関する包括的な知識習得を目的とした 4 コマの講義である。講義では、近年特に注目されている機械学習に基づく人工知能だけに限らず、旧来のエキスパートシステムによる人工知能や人工知能分野発展の歴史的な解説、さらにディープラーニングなどの先進的な内容について講義設計を行っている。特に考慮した点として、なるべく平易に理解を進め、かつ実践的に活用できるよう、概念としての理解を主目的として、アルゴリズムやその計算式、導出法などについては意図的に内容から排除している。AI 技術に関する演習として、不動産オープンデータを用いた予測・分類問題を行う。演習環境としてはクラウド上に展開されたプログラミング環境である jupyter [6] を使い、python による演習を行う。

AI 技術に関する企業セミナーは、2017 年度は日本アイ・ビー・エム株式会社・株式会社日本総合研究所により開催され、IBM Watson やそのシステムの実際の活用例について講演が行われた。

総合演習

総合演習では、これまで学習したクラウド・ビッグデータ・AI 技術の PBL での活用を目的とし、Amazon Web Service 上の Elastic MapReduce を利用したビッグデータ処理結果を用い、Azure ML 上で機械学習によるモデル構築を行う。講義では予測モデルの評価指標についての説明を行った後、Azure ML 上でのモデル構築と評価手順を説明する。その後、各自でモデル構築を工夫する演習を行う。

4.3 PBL

4.3.1 PBL 基礎

PBL 基礎では、最初に個人の演習として店舗シミュレータを操作するための API を一通り説明した後、チームごとに自動発注システムを実装する。練習として、ヨーグルトの販売実績データを対象に 3.3 節に挙げた工夫のポイントをチームで相談しながら実装を進める。機械学習の教師データには、5 年間のデータのうち 1~3 年目のものを利用する。

ここでの主な目標は、チームでの進め方を体験することである。一通りチームで実装したところで、中間報告会として、各チームの工夫点を発表してもらう。また 4 年目のデータを利用して実際に店舗シミュレータを実行し、各チームの自動発注システムの性能を共有する。

4.3.2 PBL 発展

PBL 発展では、販売実績の傾向の異なる 6 つの商品を対象を広げて総合的な結果の向上を目指す。また、結果が偶然に左右されにくいように、試行の過程を記録として残すことを各チームの課題として設定する。

最終成果発表会では、5 年目のデータを対象として店舗シミュレータを実行した結果とともに、各チームが実施した試行を比較させる予定である。

4.4 実施結果

AiBiC 関西の取り組みは今年度から始まり、実施途中であるため教育効果や受講生からの評価といった教育コースとして評価は難しいが、これまでに行っている基礎知識補助講義で得られた自由記述による感想では、包括的な知識学習や実際のクラウドサービスを活用した演習には概ね高い評価が挙がっており、企業セミナーに対しても、「サービスを提供する側、される側の意見を聞くことができ、他にはない体験だった。また、技術的な面などだけではなく、どのように普段働いているかや周りの人物はどのような人かな

ど、具台的に解説してくれたため、会社の中身がイメージしやすかった。」など非常に好意的な回答が多く得られた。この結果から、AiBiC 関西の目標としている「ビッグデータ・AI・クラウド技術に関して、基礎知識から、実践的な技術力、および応用知識までを広く包括的に習得した高度な人材を育成」実現に向け、成果を出しつつあると言える。

5 おわりに

本稿では、ビッグデータ・AI・クラウドの技術を包括的かつ実践的に利活用できる人材育成を目指す enPiT AiBiC における AiBiC 関西で実施している、機械学習による小売店の大量データの分析に基づく自動発注問題を題材とした教育コースの開発について報告を行った。今年度は PBL 開始前となるため、教育コース全体の有効性について議論は行えていないが、今後の施行を通じ教育効果やその改善について検討・検証を行っていく予定である。また、今年度は特に、演習環境のネットワークに起因するトラブルも発生したため、演習自体に対する学生の印象も低下したことが考えられるが、この問題については今年度の施行を通じ適宜改善活動を行いながら、次年度以降への実施につなげて行く。

謝辞 本教育コースの実施にあたりご尽力頂いた AiBiC 関西参画大学、連携企業の関係者各位ならびに本教育コースの受講生各位に感謝します。

参考文献

- [1] AiBiC 関西地区: <https://aibic-spiral.enpit.jp/>.
- [2] Amazon Web Services: <https://aws.amazon.com/jp/>.
- [3] enPiT AiBiC: <https://aibic.enpit.jp/>.
- [4] enPiT2 成長分野を支える情報技術人材の育成拠点の形成: <http://www.enpit.jp/>.
- [5] Microsoft Azure Machine Learning: <https://azure.microsoft.com/ja-jp/services/machine-learning/>.
- [6] Project Jupyter: <https://jupyter.org/>.
- [7] 脇森浩志, 杉山雅和, 羽生貴史: クラウドではじめる機械学習: *Azure ML* でらくらく体験.