

AI クラスルーム（FM 文理融合科目）レポート

未来型医療創造卓越大学院プログラム生 F グループ

1. 授業前の常識・知識

AI (Artificial Intelligence) の話題が毎日のようにメディアに登るようになったと感じる人は、科学者やコンピューターに精通する一部の人では無くなったように思う。有名な例では、自動運転技術の導入や PONANZA などの将棋 AI がプロ棋士を倒したなど、世間でも非常に注目を集めるテーマである。私自身、AI についての授業前の知識は、一般人に少し毛が生えた程度のもので、ドラえもんや鉄腕アトムのように人間のように動くロボットが「汎用人工知能」と呼ばれるのに対して、上で挙げた例に加えて、商品の購入履歴から、自動的にオススメの商品を紹介してくれるシステムのような我々が普段意識せずとも使用している「特化型人工知能」の2種類が存在するという事だけだった。本講義で学習した「機械学習・深層学習」については、耳にしたことがある程度で、予備知識は皆無であったと言っても良いと思う。

2. 授業の目的

本講義の目的は、データサイエンティストでない医療従事者が、未来型医療の key となる深層学習の基本的概念を学び、データ科学者と医療者を繋ぐ優れた人材を育成することである(植田琢也教授(東北大学大学院医学系研究科 画像診断学分野)による第1回の講義資料から引用)。

3. 達成目標

学習の到達目標は、以下に挙げる3つである。

- 1) 機械学習・深層学習の基本概念を理解する
- 2) 数理的・統計学的背景を理解する
- 3) 必要なプログラミングの基礎を学ぶ
(同じく第1回の講義資料から引用)

4. 授業内容

オムニバス授業を始める前に、オンライン講座 Udemy (キカガク流人工知能・機械学習 脱ブラックボックス講座 https://www.udemy.com/course/kikagaku_blackbox_1/) での事前学習を前提とし、オンライン講座では、初級編として、4つの Lecture を通して、機械学習に必要な数学的概念(導関数、偏微分)と、最も単純なモデルである単回帰分析の原理を理解し、最後に Python の基本的な構文を学習した。オンライン講座中級編 (https://www.udemy.com/course/kikagaku_blackbox_2/) では、Numpy を用いた線形代数演算をプログラミングで実装し、次に Scikit-learn を用いた機械学習の実装を行なった。平均・標準偏差・正規分布といった統計の数式と、それらの適切な使用方法を理解することができた。

計4回(1回90分×2コマ)のオムニバス授業では、株式会社キカガクの今西講師を中心に、(入力 x と出力 t を用意する)教師あり学習の回帰・分類の実装から学習した。回帰のなかで最も基本的な手法である重回帰分析の実装から始め、全てのデータから訓練データと検証データでの決定係数を求めた。その後、PLS (Partial Least Squares: 部分的最小二乗回帰) や Support Vector Regression (SVR) などの手法を用いて、精度が向上するか検証した。分類も同様に、グリッドサーチとクロスバリデーションを使用してハイパーパラメーター(学習の実行前に人間が設定する値)を調整しながら、SVC や決定木、Neural Network などの手法を練習した。

授業の中盤では、Chainer というディープラーニングに特化したフレームワークの基礎を学習し、実際にボストンの家賃をモデルに回帰を実装した。その際に、Chainer で使用できるように

Numpy を用いてデータを準備したり、モデルを定義したりするなど学習を行う上での基本技術を習得することができた。

最後の講義では、画像処理を行うために、CNN (Convolutional Neural Network : 畳み込みを行う Neural Network) を学習した。画像を加工し、クラス分類を行うために、OpenCV や pillow といったライブラリを使用して、手書き数字の格納や犬猫の画像分類などの実装を通して、GPU の使用方法やモデルの保存方法などの周辺技術も学ぶことができた。講義の前半から後半に向かうに従って、より実用的な構文 (や技術) を体験することができ、自らで学習を進める非常に良いきっかけとなった。

5. 研究や仕事などに活かせる点、影響を受けたことなど

私の専門とする構造生物学分野で、近年最も注目される革新的な技術の 1 つとして、2017 年にノーベル化学賞を受賞したクライオ電子顕微鏡 (単粒子解析) がある。この技術の優れているところは、これまで X 線結晶構造解析や NMR の技術で、立体構造の決定が困難とされていた溶液中の生体分子 (タンパク質複合体などの大きな分子や膜タンパク質などの結晶化が困難な試料) を高分解能で構造決定できるようになった点である。クライオ電子顕微鏡を用いた単粒子解析は、撮影された電子顕微鏡像から、機械学習・深層学習のアルゴリズムを用いて、(様々な方向を向いて薄氷中に固定された) 分子をピックし、二次元画像での分類および三次元像での分類を行うことで、分子の密度マップを得るというものである。そのため、観察分子の構造を崩すことなく、高いコントラストの movie を得ることが非常に重要である。しかしながら、それらの技術の多くがブラックボックスになっており、基本原理の理解がしばしばユーザーの手の出しにくいものとなっている。この講義全体を通して、機械学習・深層学習の最も基本的な概念を学習したことで、ブラックボックスの中身をより深く理解しようという学習のきっかけになっただけでなく、実際に解析に役立てることのできるヒントを多く学ぶことができた。

6. 来年度以降の改善点、授業の限界

全ての講義は、非常に系統立てられており、今西講師の説明も非常に分かりやすく、1 回 90 分 × 2 コマの講義も短く感じるほどだった。いつでも閲覧が可能な、配布資料には説明される全ての構文が細かく書かれており、授業中に追いきれなかった箇所を補ったり、講義後に復習したりするのに非常に役立った。

8 時間のオンライン講座と 4 回の講義だけで python を使った機械学習・深層学習の全てを網羅することは、不可能なことだが、この講義のボリュームで (予習と講義後の宿題の時間を含めても) 吸収できる内容は非常に多く、授業全てに大変満足している。来年も本講義が開講されることを切に願っている。

7. まとめ

本講義の最終目的である、未来型医療の key となる深層学習の基本的概念を学び、データ科学者と医療者を繋ぐ優れた人材を育成するために、文理の枠組みを超えた初学者に対する学習の第一歩となるという点で非常に良い講義だったと思う。