

# ディープラーニングを用いた LA ワークショップ開催報告

児玉靖司\*、加藤利康\*\*、久保田純\*\*\*、ジェイソン・ゲディーズ\*\*\*\*

\* 法政大学  
\*\* 日本工業大学  
\*\*\* フォトロン  
\*\*\*\* デジタル・ナレッジ

## A Report on the LA Workshop using Deep Learning

Yasushi Kodama\*, Toshiyasu Kato\*\*, Jun Kubota\*\*\*, Gettys Jason\*\*\*\*

\* Hosei University  
\*\* Nippon Institute of Technology  
\*\*\* Photron Limited  
\*\*\*\* Digital Knowledge Co., Ltd

\* yass@hosei.ac.jp

概要: 学習分析学会主催により Learning Analytics ワークショップを2017年2月25日、26日に開催したので報告する。今回は、最近話題となっているディープラーニングを学習分析に応用するための新しい手法を開発しようと企画したものである。全国から22名の参加者があり、2日間かけてディープラーニングツール Tensorflow の使い方から新しい分析方法の開発までを行った。最初に4名がチュートリアルを行い、参加者をグループ分け後、最後に新しい分析方法について成果発表会を行った。

Abstract: This paper reports the Learning Analytics workshop sponsored by Japan Society for Learning Analytics (JASLA) on February 25<sup>th</sup> - 26<sup>th</sup>, 2017. Our objective is to develop new analysis methods using deep learning that is recently getting popular. We have a hackathon within 2 days where they teach how to use the Tensorflow tools and develop new methods each other. At first, 4 persons present the tutorials. After we divide 22 persons into group of six, we hold the final meetings to present new approaches.

キーワード: ハッカソン、ディープラーニング、学習分析  
Keywords: Hackathon, Deep Learning, Learning Analytics

### 1. はじめに

学習分析学会主催により「ディープラーニングを用いた Learning Analytics (LA) ワークショップ」を開催した。2017年2月25日、26日の2日間、ワークショップ(以下、LAハッカソンという)の設備が整い実施可能な伊東温泉(山喜旅館)にて開催した。

最近のAIの発達により機械学習、さらにディープラーニング(Deep Learning)を用いた大規模データ分析が注目さ

れている。ディープラーニングに関してもTensorflow (Google)や、Chainer (Preferred Networks)などプログラミング言語Pythonのフレームワークとして提供され簡単に記述することができるツールが揃っている。我々は、これらツールを用いて学習データを詳細に分析できるのではないかと考えた。(Kato, 2016)等の従来研究(統計分析手法による学習データ分析)を応用し、新しい分析方法を開発しようと企画したものである。Webサイト (<https://sites.google.com/view/hack-la-ws/>)で広く告知し研究者、技術者らを集め、LAハッカソンを実施した。

会場設備の都合により参加者(定員 20 名)を募集したが数日で定員に達してしました。昨今のディープラーニングに関する関心の高さが窺われる。

## 2. LA ハッカソン概要

LAハッカソンの概要は、以下の通りである。

テーマ: プログラミング行動の履歴に対するディープラーニング分析

Webサイト: <https://sites.google.com/view/hack-la-ws/> (告知)

<http://hackathon.site44.com/> (資料)

さらに学習データはUSBで各グループに配布した。参加者の概略は以下の通りである。

参加者: 22名(スタッフを含む)

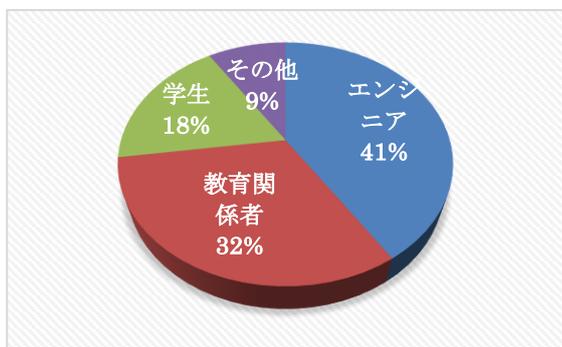


図1 参加者の属性

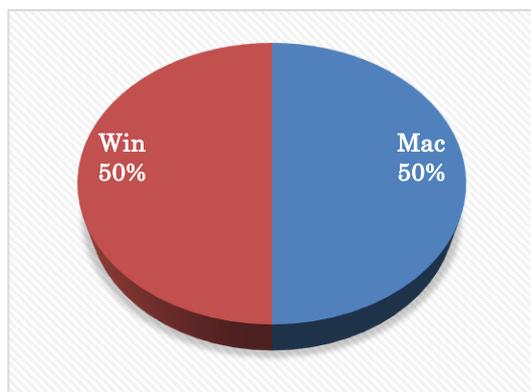


図2 所有するOS

全国から広くエンジニア(企業の研究員等)、教育関係者(大学教員等)らが集まった(図1)。しかし、多くの参加者がディープラーニングの初学者であったため、ディープラーニング学習も目的の一つとした。所持しているOSはWindowsとMacOSが半々であった(図2)ため、配布したデータやプログラムも各OSで動作するものを用意した。

LAハッカソンで用いるデータは、日本工業大学のJavaプログラミング科目において実施された演習時のプログラミング行動の履歴である。

開催スケジュールは以下の通りである。

[第一日]

13:00-14:00 受付

14:00-15:30 会長挨拶

注意事項説明

LA 実務の流れについて

ディープラーニング LA チュートリアル

ディープラーニング LA 研究事例報告

学習データサンプル説明

自己紹介

15:30-19:00 グループ分け

各グループ作業

19:00-深夜 夕食

各グループ作業続行

[第二日]

7:30-8:30 朝食

8:30-11:00 各グループ作業

11:00-12:00 成果報告会

12:00-13:00 昼食

解散

参加者のレベル、参加の目的等に差異があり、グループ分けなどの方針をあらかじめ詳細に決めることができず、開催スケジュールについても臨機応変に行った。



図3 成果発表会の様子(1)



図4 成果発表会の様子(2)

LA ハッカソンでの学習データ分析は、プログラミング行動のデータを特徴別にクラスタリングすることである。従来の統計手法によるクラスタリングでの結果と比較を行い、向上することを目指す。

最初に1.チュートリアルを行い、その後、2.グループ分け(5グループ)、3.各グループ作業、4.成果発表の流れで実施した。

グループ作業が始まってから、久保田より、雛形となるPython プログラムを配布し、プログラムのパラメータを調整するなど、カスタマイズする形で作業を行った。

成果発表会ではグループ毎に10分程度の発表が行われた。

### 3. チュートリアル

最初にジェイソンより、LA 実務の流れについて説明があった。続いて児玉よりTensorFlowを説明するテキストを中心に事例を説明した。最後に加藤よりディープラーニング LA 研究事例報告を行った。

本チュートリアルでの資料は、<http://hackathon.site44.com/>にて配布した。

#### 3.1 ラーニングアナリティクス

ジェイソンより、「ラーニングアナリティクス」として以下の説明をした。

1. データサイエンス・パイプライン
2. Epicycles of Analysis について
3. 機械学習・パイプライン
4. データサイエンティストからの一言

データサイエンティストとして、学習データを分析する一

般的な方法論、流れについて説明した。

#### 3.2 TensorFlow チュートリアル

児玉より、テキスト「中井悦司、TensorFlow で学ぶディープラーニング入門、マイコム」を使って実際にプログラミングの仕方について説明した。

1. 簡単な使い方(プログラミング方法)
2. 平均気温の近似
3. ロジスティック回帰による二項分類器
4. 単層ニューラルネットワークによる二項分類器
5. 多層ニューラルネットワークによる二項分類器

TensorFlow の簡単な使い方から、プログラミングと多層ニューラルネットワークによる分類器までを、違いを明確にしながらか説明した。TensorFlow は、多層ニューラルネットワークのような、ディープラーニングに必要な処理だけでなく、基本的な近似処理(回帰計算等)行うことができるので、処理の違いを明確にしながらか、段階を追って説明した。

#### 3.3 研究事例

加藤より、LA ハッカソンで扱うために参考となる LA 研究事例報告を行った。さらに、テーマにあるディープラーニング分析に必要な学習データを配布し、データについて説明した。

1. プログラミング演習の特徴と問題
2. 指導者の役割(TA など)
3. 従来研究
4. 配布する学習データの説明

### 4. グループ分けおよびグループ作業

部屋割りに基いてグループ分けを行った。TensorFlow に対する知識の違いがあったが概ね部屋ごとグループ分けすることで互いに教えあうなどの良い効果が期待できた。最初は、個々の能力に合わせてTensorFlow の使い方、プログラミングを学び、久保田より「ディープラーニングチュートリアル」資料を基に説明をし、雛形となる Python プログラムを配布し、グループ作業では多くのグループが本 Python プログラムをカスタマイズする形で作業を行った。

結果として、ディープラーニングを用いて大規模学習データ分析するためには、概ね、以下の手順が必要であることが確認できた。

1. 学習データをクレンジングする。
2. Python プログラム(4.1)のデータとしてプログラミングする(DNN, Deep Neural Networks を構成)。
3. パラメータを調整(プログラムをカスタマイズ)し、学習者をK種類にクラスタリングするための、ニューラルネットワーク層の最適な数を求める。
4. 学習者を K 種類にクラスタリングし、従来の統計手法によるクラスタリング結果と比較、検証する。

以上の作業を Python プログラムをカスタマイズしながら最適なパラメータを調節することがよいという事が分かった。

ただし、上記の手順ではなく、独自のプログラミングをし、新しい分析方法を開発したグループもあった。

#### 4.1 ディープラーニングチュートリアル

ひと通り、TensorFlow の簡単な使い方、プログラミングが理解できたあたりで、久保田より、各グループへディープラーニングを実行することができる雛形となる Python プログラムを配布し、説明した。LA ハッカソンで対象としたディープラーニングでは、大規模学習データを分析し、学習者を K 種類にクラスタリングするために、ニューラルネットワークを n 層設ける必要があると考えられる。本 Python プログラムは、K と n の値を簡単なパラメータとして与え、プログラムを修正することで実行できるプログラムであった。

### 5. 成果発表

一日目午後より、二日目午前にかけてグループ毎に作業を行った。一部のグループは徹夜作業であった。

結果として、以下のような成果があった。

1. Python プログラムをカスタマイズし、学習者のクラスタリングを行った。従来の統計手法によるクラスタリング結果と比較、検証した。
2. 処理の様子を図示しデータの可視化を行った。
3. 大規模学習データのクレンジング(前処理)に

注目し処理を行った。さらに、データの可視化を行った。問題点を洗い出し、考察を行った。

4. ディープラーニングの学習に終始した。

2日目午前 11:00 より、成果発表会を開催した(図3、図4)。

### 6. おわりに

今回の LA ハッカソンは、最近話題となっているディープラーニングを用いた学習分析を扱い、新しい分析方法を開発するという試みであった。学習データは、日本工業大学の Java のプログラミング科目において実施された演習時のプログラミング行動履歴であるが、加藤が用意した。学習データが実際の授業で実施されたプログラミング行動の履歴データであり、新しい分析方法を開発、考察するイベントであったため、全国から応募がありすぐに定員に達してしまった。

ディープラーニングを扱うツールも現在では整備されており、今回扱った TensorFlow は、Jupyter Notebook と組み合わせることで簡単にデータの可視化もできる。LA ハッカソンは、最新の開発環境を鑑みたアプローチであった。

参加者のレベル、目的等に差異があったものの、成果発表会では、十分に考察された内容の発表も多く、(加藤, 2017) や (Jason, 2017) 等の論文として発表されたアプローチもあった。

今後の課題としては、今回の LA ハッカソンは短期間で企画し会場も小規模であったため、今後はさらに海外から参加者を募集し、大規模な LA ハッカソンを考えたい。2017 年 8 月の LASI-Asia にて同企画を実践したいと考えている。

**謝辞** 本研究の一部は独立行政法人日本学術振興会 科学研究費補助金(課題番号:26240008, 15K01094)の助成を受けたものである。

#### 参考文献

加藤利康, 久保田純, 卯木輝彦, 梅澤健二 & 児玉靖司 (2017). プログラミング行動の履歴に対する Deep

Learning 分析, 第1回学習分析学会研究会.

Jason Gattys, 堀越泉, 田村恭久 & 小林建太郎 (2017),  
学習行動の未来予想を目指したディープラーニング, 第  
1回学習分析学会.

Google, TensorFlow, <http://www.tensorflow.org>

Preferred Networks, Chainer, <http://chainer.org>

Toshivasu Kato, Yasushi Kambayashi and Yasushi  
Kodama (2016), Data Mining of Students'  
Behaviors in Programming Exercises, Proc. of 3<sup>rd</sup>  
International KES Conference on Smart Education  
and E-Learning (KES-SEL-16), Springer—Verlag.