

ディープラーニングの流域圏研究への応用にあたって

中根英昭（高知工科大学）

→

1. ディープラーニングとは何かどのように利用すべきか

ディープラーニング（深層学習）はより大きな概念である人工知能の一部であるが、「AI 白書 2017」の帯に「ディープラーニングがすべてを変えた、」とあるように、2010年代の人工知能の爆発的な発展の鍵となる技術である。従来から使われていたニューラルネットワークの層を深くしたものには違いないが、「過学習（over fitting）」の問題を解決するため、あるいは計算を高速化するため等の様々な技術の組み合わせ（例えば、活性化関数 Relu, dropout、各種の optimizer、GPU による並列計算）とビッグデータの活用等が相まって、高い費用対効果（時間費用も含む）が得られるようになったのである。これにより、コンピュータシミュレーションによって飛躍的に進歩した演繹法と、ディープラーニングによってコンピュータの大幅な利用が可能になった帰納法がそろって、人間の認識能力を大幅にかつバランスのあるものとして大きく発展させたものと考えている。ディープラーニングを一言で表現すると、「分類または回帰を行う帰納的モデル自動生成のための巨大なフィッティング装置」である。

2. ディープラーニングをどのように利用すべきか

—上流雨量を入力とした河川水位の推定を例に

ディープラーニングが入力データと出力データを結びつける「分類または回帰を行う演繹的モデル自動生成のための巨大なフィッティング装置」であるとすれば、入力と出力の間どのような関係がある場合に最も有効か、について考慮して利用すべきである。入力と出力の間の因果関係が明確に式で表される場合には、ディープラーニングを使う必要はない。因果関係がある、複雑で、詳細なプロセスを数式化するのが困難、近似の誤差が大きくなってしまふようなプロセスによって入力データと出力が結びついている場合にこそディープラーニングが有効であろう。典型的な例が、雨量を入力とした河川水位、水量の推定である。

図2に示されるように、流域圏の降水は地表面を流れて河川に流入するだけでなく、様々な深さの地層を流れる地下水となり、長い時間をかけて湧水として河川に流れ込む。この過程は、タンクモデル等によって近似されるが、タンクの数や分岐比等は実測データに合わせて決定せざるを得ず、演繹的な手法のみによってモデル化することは困難である。従って、図3のようなニューラルネットワーク（多層パーセプトロン）を用い、雨量等の入力と河川水位あるいは流量等を出力とした回帰モデルの自動生成が期待される。隠れ層が一つの「浅層ニューラルネットワーク」による研究は2000年代前半から行われ、最近では隠れ層が2以上の深層ニューラルネットワーク（深層学習）が使われるようになってきている³⁾。また、深層ニューラルネットワークと分布型モデルのハイブリッドモデルも提案されている⁴⁾。これらの研究では、洪水時のデータのみを用いている。それに対して、高知工科大学中根研究室で開発されたモデルは、1時間分解能の4-5日間の雨量時系列データと1か月の時間分解能の1年間の雨量時系列データを入力として、豪雨による短時間の水位上昇と「緑のダム」の効果の双方をモデル化した点に特長がある^{1,2)}。すなわち、深層学習のみによって水文過程を含む上流雨量—河川水位の過程のモデル化を行い、洪水時だけでなく、渇水時も含め、一年を通した切れ目のないモデリングを可能にしたのである。



図1 人工知能、機械学習、中根研究室で用いている深層学習の関係¹⁾

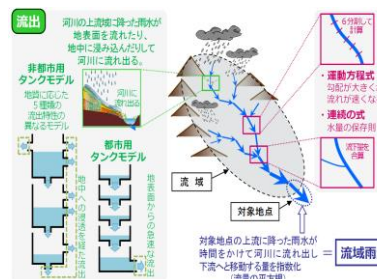


図2 気象庁の流域雨量指数計算において考量されているプロセスと計算式²⁾

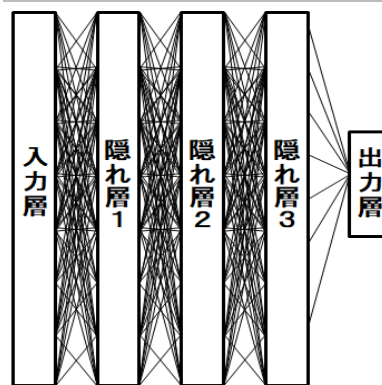


図3 多層パーセプトロン概念図

3. 2018 年度前期の進展—ディープラーニングによる四万十川と物部川のモデリング

2018 年度に入って、高知工科大学では、ダムの影響の少ない四万十川とダムの影響が重要な物部川についての、深層学習によるモデリングを進めている。これらの研究では、図 3 の隠れ層 3 層の深層ニューラルネットワークを用いている。さらに、1 時間分解能と 1 か月分解能のデータを入力データとするのではなく、連続的に分解能が変化する 1 年間あるいは 2 年間の上流雨量時系列データを用いている。それぞれの河川の具体的な研究内容は、口頭発表 1, 2 で紹介される。

4. ディープラーニングに用いた計算機資源及びソフトウェア

ディープラーニングを用いたモデリングが、地方大学の研究室が外部資金無しにある程度実行可能であることについて述べたい。中根研究室では、以前からデータ解析に用いていた「手作り PC」のスロットにゲーム用の GPU (Graphics Processing Unit) を挿入して並列計算機として使用している。表 1 に諸元を示す。

深層学習に用いるフレームワークとして、Google 社の Keras (バックエンドは Tensorflow) を用いている。プログラムの大半は入力データの処理と結果の表示に費やされており、ディープラーニングのモデルは数行のみである。プログラミングは容易である。また、フレームワークは無料で公開されている。

表 1 深層学習用計算機諸元

CPU	Core i7-7700
GPU	GeForce GTX1070 or GTX1080Ti
OS	Windows 10
フレームワーク	Keras/Tensorflow
プログラム言語	Python 3.6 (Anaconda 4.4.0)

5. 河川水位推定以外の応用

高知工科大学中根研究室では、この約 1 年の間に、以下のテーマに取り組んだ²⁾。

- ・早明浦ダム上流平均雨量を入力としたダム流入量の推定モデル (多層パーセプトロン; MLP)
- ・ひまわり 7 号雲画像を用い降水の有無の判別 (畳み込みニューラルネットワーク; CNN)
- ・ひまわり 7 号雲画像を用いた全天日射量推定 (CNN による回帰)
- ・3 種のユリ科植物 (ウバユリ, ヤマユリ, オニユリ) の判別 (Xception を用いた転移学習)
- ・ドローンによる森林写真を用いた樹種判別 (Xception を用いた転移学習)

これらはある意味において流域圏関連研究であるが、現在は水位推定モデル開発に集中している。

6. おわりに

ディープラーニングを用い、上流の雨量時系列を用い入力データとして、河川水位の推定を容易に、低いコストで実行できることを示した。これは一例に過ぎず、ディープラーニング自身は「分類または回帰を行う帰納的モデル自動生成のための巨大なフィッティング装置」であるため汎用性が高く、様々なテーマに応用できる。利用にあたって重要なことは、入力データと出力データ、及びそれらの関連、そして利用目的についての検討であると考えている。

参考文献

- 1) 中根英昭 若槻祐貴.(2018):環境分野への深層学習応用研究の立ち上げについて, 高知工科大学紀要 15 巻(1), 111-120.
- 2) 中根英昭.(2018):環境分野の深層学習を通じた人材育成, 高知工科大学紀要 15 巻(1), 69-78.
- 3) 岡田 治, 柿澤 一弘, 田所 正, 中安 正晃, “AI による洪水予測と水位データの異常検知”, 平成 29 年度河川情報シンポジウム講演集, pp.3_1-3_16, 2017.
- 4) 一言正之, 櫻庭雅明 “深層ニューラルネットワークと分布型モデルを組み合わせたハイブリッド河川水位予測手法”, 土木学会論文集 B1 (水工学), Vol.73, No.1, pp.22-33, 2017.