

企画セッション（「シグマ」特別専門委員会主催、核データ部会共催）

「『シグマ』特別専門委員会 2017、2018 年度活動報告」

(2) 放射線計測分野への機械学習の応用

— 核データ測定とガンマ線分析の例 —

九州大学大学院

総合理工学府 先端エネルギー工学専攻

金 政浩

kin@aees.kyushu-u.ac.jp

tadahirokin@kyudai.jp

1. はじめに

本企画セッションにおいて、近年急速に様々な分野で応用が進んでいる機械学習を核データ測定や放射線計測分野に取り入れる活動について紹介した。講演当日は発表時間の都合上割愛した内容も含めて紹介する。あらかじめ断っておくが、私は情報系の専門家ではないため、厳密に言えば正しくない表現なども含まれていることはご容赦頂けると幸いである。

2. 機械学習の歴史

現在は第3次機械学習ブームと呼ばれている。逆に言うと過去に2回の衰退の時期があったことになる。機械学習（あるいは人工知能）の歴史を概観することは、これらの盛衰を知ることと同義であり、数学的な背景や技術を身に着ける一助となることはもちろんのこと、今回のブームに乗ることが果たして我々のような基礎科学に近い分野にも有益なことかどうかを把握することができる。貴重な時間を割いて拙文をご覧いただいた皆さまが有益だと判断して頂き、勉強・研究仲間になっていただければ嬉しく思う。それでは本題に入ろう。

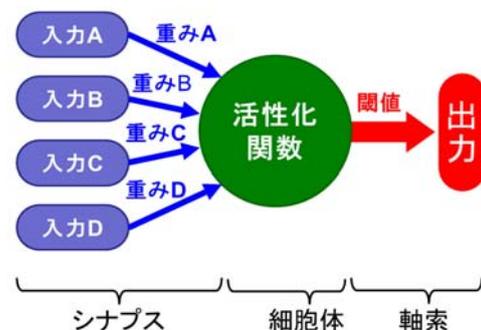


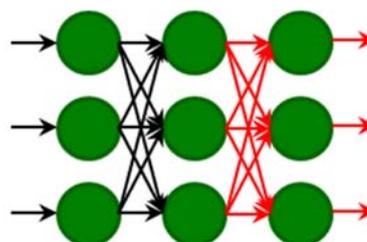
図1 McCullochが1943年に提案した人工ニューロンモデル [1]。現在も色褪せることなく活躍中である。

そもそもの発端は「神経の動作を数式で表そう」という McCulloch [1] による人工ニューロンモデル(図 1)の提案であった。1943 年に発表されたこのモデルは、現在も色褪せることなく、ほとんどの機械学習モデルの最も基礎的な数理モデルとなっている。機械学習を学ぶ者はまずここからスタートせねばならない。神経細胞は、シナプスを介して他の多くの神経細胞から電気信号を受け取る。これらは図中の「入力」に該当する。これらの電気信号電位を $x_1, x_2, \dots, x_i, \dots, x_m$ としよう。また、シナプスはこれらの信号を後段の神経細胞に伝達する構造であり、誤解を恐れずに言えば回路の電気抵抗のようなものである。電気抵抗の逆数に比例して入力信号の減衰が決まるため、この重みを $w_1, w_2, \dots, w_i, \dots, w_m$ としよう。前段の神経細胞と自分自身の相関が強ければ強いほど抵抗値が小さく(重みが大きく)なり、前段の電気信号を減衰させずに受け取る。逆に、前段の神経細胞からの情報がさして重要でない場合は高い抵抗値(小さい重み)を与えた形となり、減衰された電気信号が伝達される。前段の神経細胞からの電気信号すべての合計がある閾値(b)を超えた場合のみ、その神経細胞は軸索を通じて後段の神経細胞に電気信号を送る。初期の形式ニューロンでは、閾値を超えたかどうかだけを判断するため、この判定に用いられる関数はステップ関数だった(近年では用いられない)。この関数のことを「活性化関数」と呼ぶ。まとめるとこの数理モデルは以下のように表すことができる。

$$y = \varphi \left(\sum_{i=1}^m x_i w_i + b \right)$$

この式と図 1 の関係だけわかれば人工ニューラルネットワークの基礎理論はほとんど理解したようなものである。

さて、ここまでは重要なので要点だけかいつまんでもこれだけ長くなってしまったが、ここからは細かい数式は出さないのが気楽に読んでいただきたい。このようにして発案された人工ニューロンモデルだが、実は「知能」として欠かすことのできない「学習」について未着手だったため、これから 15 年はなんら「ブーム」とはならなかった。この状況を打開したのが Rosenblatt による



黒:乱数等の固定重み
赤:学習で更新する重み

図 2 Rosenblatt が提案したパーセプトロンの提案[2]である。図 2 のように、人工ニューロンモデルを複数並べたものである。彼は「誤り訂正学習」という学習方法まで提案したため、ついに「人工知能」の歴史が拓けたのである。ただし、注意してもらいたいのは図中の赤で示されたシナプスしか学習は行わないモデルとなっていることである。これは提案された学習方法

に欠陥があったからである。もし黒で示されたシナプスも学習しようとする、実は最適解を導出することができない。つまり、近年はやりのワードである「深層学習」という表現にならえば、「単層学習」と言えるであろう。ただし、学習手法が提案されたことによって、様々な問題に適用する研究者が現れ始めて、徐々に第一次…結果として「一時」的なものだったが…機械学習ブームが訪れたのである。

当時はコンピュータに真空管が用いられていた時代である。情報科学に関与する研究者の絶対数もそう多くはなかった。そのため現在ならばすぐに気づかれたであろう「誤り訂正学習」の問題点が明らかになるには約10年の歳月を費やした。1970年頃、複数の数学者が「パーセプトロンは線形分離可能な問題しか解くことができない」ということを数学的に証明した。もし興味があれば手計算して頂きたい。数学を生業としない我々でもすぐにこのことは証明ができる(学生への宿題に丁度よい)。いずれにせよ詳細は述べないことにして、結果だけ言おう。「単層」であることが、この数学モデルの表現力を「線形」に固定していたのである。では、今こそ「深層」にすべきでは？多くの情報科学の研究者もそう考えた。しかし、「誤り訂正学習」では出力の直前のシナプスしか学習させられなかったのである！こうして、第一次機械学習ブームは急速に終息し、このあと20年近く日の目を見ることは無くなった。

応用分野の研究者は、誰も人工知能だとか機械学習だとか、そういう分野から興味がなくなってきた頃、その時は突然訪れた。RumelhartらがNature誌に投稿した「誤差逆伝播法」[3]による学習手法の提案である。これを用いることで、出力に含まれる誤差(正解値との差)を前段だけでなく、それよりも前の段の層に伝播させ、学習ができるようになったのである。キーは活性化関数を連続関数にし、微分を可能にしたことである。この時に提案されたのはシグモイド関数である。原子核物理分野の皆さんならウッズサクソン分布の形だと言った方がピンとくるかもしれない。いずれにせよ $\pm\infty$ で ± 1 に収束する関数である。さて、このようにして「深層化」が可能になり、線形分離不可能な問題も解けるようになった。このように多層化させたパーセプトロンは「人工ニューラルネットワーク」と呼ばれている。人工ニューラルネットワークは「万能性」が証明されており、あらゆる形の関数を近似できることが分かっている。なんでも解けるではないか！多くの研究者は狂喜乱舞し、この手法に飛びついた。私でもおそらく飛びついただろう。第二次機械学習ブームのはじまりである。

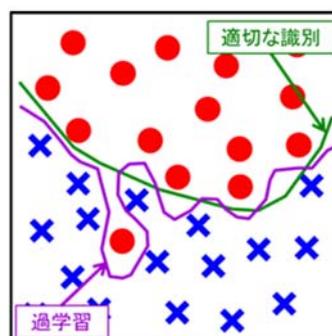


図3 「過学習」のイメージ図

実に面白いことがおこる。あまりに多くの研究者がこの手法に着手したことにより、パーセプトロンより高性能だった人工ニューラルネットワークの問題点が明らかになる

にはほんの数年しかかからなかった。1990年から2005年頃まで、様々な問題点が指摘され、機械学習に取り組む研究者は一人、また一人と去っていった。大きな問題点は2つである。ひとつは、先述した万能性により、「過学習」を起こすことである。図3に概要図を示す。赤や青(あるいは○と×)で示されるデータを識別する際、我々は直感的に図中の緑線のような区切り方をするはずだ。そして、それが識別としては正しい。外れ値に引っ張られてはならないわけである。しかし、人工ニューラルネットワークはその万能性から図中の紫線のような識別も可能である。これは全く避けられない問題だった。なお、過学習については今日の機械学習でも必ず注意しなければならない、そこに言及しなかったことによって reject された多くの論文が存在していることは想像に難くない。さて、もうひとつの問題は「勾配消失問題」である。誤差逆伝播法では、出力側の誤差を活性化関数の微分値でどんどん前段に伝播させていく。シグモイド関数のように±∞で微分値がかぎりなくゼロに近づくような場合、それより前段では伝播させる誤差もほとんどゼロになり学習が全く進まなくなるのである。これらの問題が明らかになると、あまり複雑な問題は解けないうえ、過学習による結果の信頼性への疑念を拭い去ることはもはや不可能になってきた。人工ニューラルネットワークの研究自体が「イロモノ」的にみられることすらあった。

もう機械学習の時代はこないのか。第二次機械学習ブームがあつという間に終焉を迎えて多くの研究者はそう思っただろう。しかし、夜はまた明けたのである。2006年から徐々に、この2大問題が解決されていった。Hinton らによる「教師無し事前学習」[4]、Nair らによる新たな活性化関数「ReLU」の提案[5]、Srivastava による「ドロップアウト法」[6]…などによって2014年には機械学習に再び日が差したのである。

閑話休題。[3-6]の著者をよく見て欲しい。全てに Hinton が主著あるいは共著に含まれている。彼は、自分たちが提案した[3]の手法を、自分たちで[4-6]のように解決していったのである。非常にかっこいいではないか。また、最初こそ Nature 誌や Science 誌などに論文が掲載されていたが、[5]の文献などは国際会議の Proceedings である。原子力学会企画セッション当日は「被引用数が6684回で、おそらく Proceedings としては世界記録だろう」と紹介したが、本稿の執筆時点で8328回。まだまだとどまることを知らない。良い論文は雑誌を選ばず引用されるのである。その他、文献検索をすれば気が付くことがもうひとつある。機械学習自体(応用ではなく)を研究している論文の多くは Proceedings にしか記事が載っていない。[5]で味をしめたのだろう…というのは冗談で、機械学習の研究があまりに急激に進んでいるため、従来の「投稿から掲載まで数カ月から半年」というスタイルになじまなくなっているのだろう。皆さんも機械学習の論文検索をする場合はプレプリントサーバーものぞいてみることをお勧めする。そこにしかない文献がかなりあるうえに、それらは往々にして大変インパクトのある論文だったりするのだ。

さて、これで重要な機械学習の歴史は全部おさらいした。皆さんはどう感じただろう

か。あくまで私見ではあるが、第三次機械学習ブームが終焉することは無いように思える。今際の際に「ああ、第三次も終焉しちゃったな」と思いながら本稿を読み返すことがあったなら、それはそれで笑い話として語り継いで頂ければよい。ただ、将来がどうなろうと一度は触れてみる価値があることは疑いようがないだろう。

3. 核データ測定・放射線計測への適用例

前置きだけで 4 ページも消費してしまい、大変申し訳ないが今しばらく付き合ってください。この章では、私が取り組んできた核データ測定や放射線計測への機械学習の適用例を示したいと思う。はじめに、なぜ機械学習を研究に取り入れようと思ったのか、について述べる。

ひとつは、前章で述べたように機械学習が技術として成熟していると感じたというのが理由である。4 ページも駄文に筆を走らせたのでこの点はもはや言うこともない。ただし、もうひとつの点はぜひともここで述べておきたい。最近の機械学習の応用例を見ていて、多くの方々は「機械がものを考えている」という誤解をしているように思う。実際は機械学習が行っているのは、高度な関数フィッティングであり、「自動販売機にお金をいれてボタンを押せばジュースが出る」という例とほとんど変わらない。自動販売機は何かを考えてジュースを提供したわけではないのである。顔をうかがって、コーラのボタンをおした人に、「健康のために緑茶を飲めよ、ガチャリ！」というような動作は決してしないわけである。実際に機械学習が行っている「顔をうかがう」という動作は、肌の色だとか、無精ひげの伸び具合だとか、そういうものを特徴量として抽出して機械的に判断しているにすぎない。ただ、「何を特徴量とするのか」というのは機械学習の応用を考える上で極めて重要である。「トラ」と「ネコ」をどうやって区別するのか。特徴量として何をいれれば良いのか。「トラっぽさ」は簡単には数値化できないわけである。一方で放射線計測分野はどうだろうか？ 入力として与えるものは実験データである。いわずもがなほとんどの場合は数値データだろう。そして導きたいもの、すなわち出力は物理量であり、「コーラ」でも「トラ」でもなく、これもまた数値で表すことが可能な

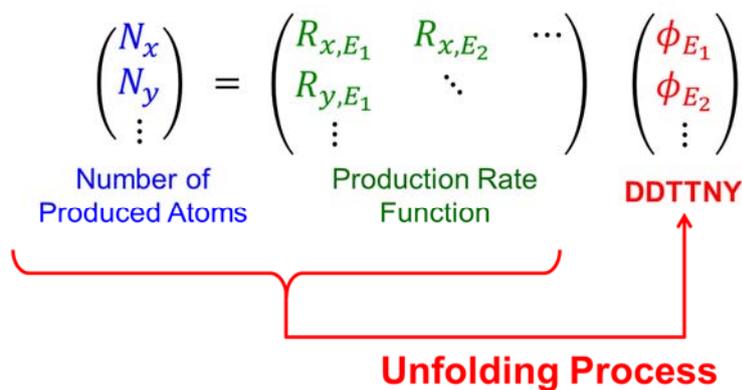


図 4 多重箔放射可能による二重微分中性子収量 DDTTNY の導出

のである。機械学習の応用上、大きなハードルである入力と出力の数値化ははじめからできているのである。これは使わない手はない、と考えたのがもうひとつの理由である。皆さんも一番最近行ったデータ解析を思い出して欲しい。数値以外のものはどれくらいあったらうか。デジタル時代である。我々の扱いたいデータは全て数値化されていると言っても過言ではないだろう。

さて、私の実施した応用例を示そう。言いたいことはほとんど述べたあとなので、ここからは結果を示すだけにしたい。

3.1 ニューラルネットワークを用いた中性子スペクトルアンフォールディング [7]

我々は加速器中性子源を用いた放射性同位元素製造技術に資するため、 $C(d,n)$ および $Be(d,n)$ 反応による厚いターゲットからの二重微分中性子収量DDTTYを系統的に測定している。限られたマシンタイムの中で一気にデータをとるために、我々は多重箔放射化法を用いた測定を適用している。図4に解析法の概要図を示す。実験で得られる値は、放射化生成物の原子数 N_x, N_y, \dots である(x, y, \dots は核種)。また、生成率関数($R_{x,E1}$ など)は実験の照射条件と核データから得られる断面積で決まる値である。多くの場合、着目している放射化反応の数は、求めたい

DDTTYのエネルギー群数より極めて少なく、解析的には求めることができない。よって、何らかの方法で逆問題を解かねばならないことになるが、その時用いられるのがアンフォールディングである。これまでに様々なアンフォールディングコードが開発されてきたが、基本的にやっていることは関数フィッティングである。それならば人工ニューラルネットワークにもできるはずだ。

学習用のデータは、フォールディングによって作成した。適当な中性子収量分布を与えてあげれば、生成率関数が既知なので核種生成量も計算できる。これで、入力(実験データ)と出力(中性子収量分布)のデータセットをいくらでも作ることができる。

この研究は、私が機械学習を初めて

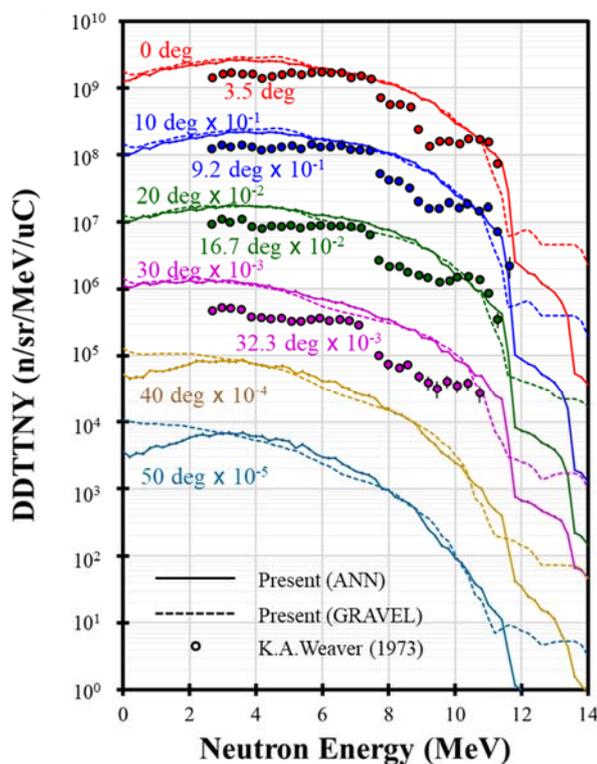


図5 人工ニューラルネットワーク(ANN)およびGRAVELによるアンフォールディング結果の比較

適用した例だったので、全て勉強することにした。要は便利なライブラリは用いなかったということだ。実験データは重陽子エネルギー12 MeV の $C(d,n)$ 反応による DDTTNY である。機械学習によるアンフォールディング結果を良く用いられる GRAVEL コードによる解析結果と比較したものを図 5 に示す。参考に Weaver らによる過去の実験データも載せているが、注目してもらいたいのは両アンフォールディング結果である。いずれもほぼ同等の解析結果が得られていることがわかるだろう。ただし GRAVEL と異なって、アンフォールディングの収束条件や初期推定スペクトルを人手で決める必要はないため、その点では機械学習の解析結果が有利と言える。

予想通りアンフォールディングには機械学習が適用可能だという結果が得られたという点が重要である。ライブラリを全く使わなかったため、ハイパーパラメータチューニング(どこまで深層化するか、ユニットの数をどう選択するか、など)はあまりできていないため、さらに高性能にできるだろう。

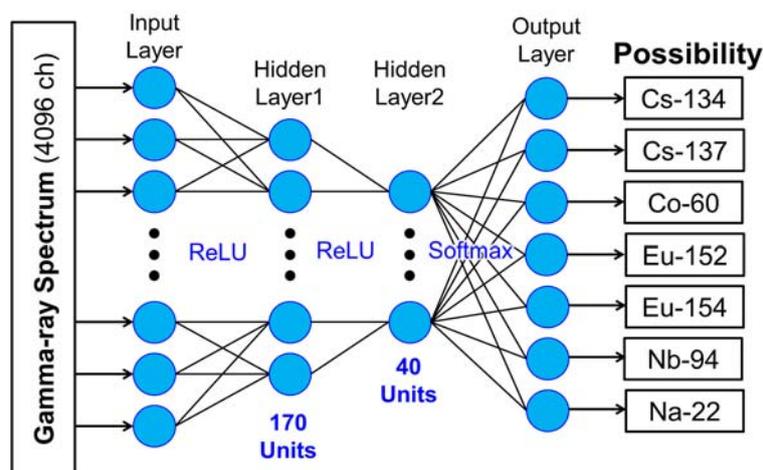


図 6 ガンマ線スペクトルから核種を推定する機械学習モデル

3.2 ガンマ線スペクトルからの核種推定 [8]

アンフォールディングに機械学習が適用可能なら、ガンマ線スペクトロメトリによる放射能分析にも使用可能なはずである。従来の光電ピークだけを解析する手法に比べ、コンプトン成分も使えるため、計測時間を短縮できるというのがモチベーションである。本稿では、その最初のステップとして「統計の悪いガンマ線スペクトル」を入力することで、「どの核種から放出されたガンマ線か」を識別する機械学習モデルを作った。今回は、便利なライブラリを用いた。最近は色々なライブラリが用意されているが、もっとも人気のある Tensorflow+Keras を用いた。Preliminary に決めた(しかし性能の高い)機械学習モデルを図 6 に示す。

学習データはGeant4を用いて新潟大の後藤さんに作って頂いた。図中の出力に示す7核種について、様々な統計の単核種スペクトルを作成した。ハイパーパラメータは粗くチューニングしただけだが、学習に用いていないスペクトル(検証用スペクトル)による核種同定精度は95%を超えた。なお、過学習が起きた場合、学習に用いた

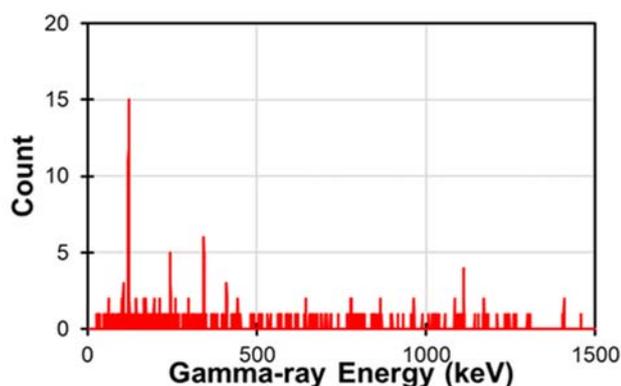


図7 統計の悪いガンマ線スペクトルの例(^{154}Eu)

スペクトルに比べ、検証用スペクトルに対する核種同定精度が高くなる現象がおこる。もちろん過学習が起きていないことはチェック済みである。ヒト知能として、私が核種同定を行った結果を優に超えていた。例えば、図7に示すような統計の悪いスペクトルに対して、これを瞬時に ^{154}Eu のスペクトルだと判断できる人が果たしてどれだけいるだろうか。今回の機械学習モデルではこのスペクトルは正しく ^{154}Eu と判定された。

4. まとめ

機械学習の歴史を概観し、第三次機械学習ブームの終焉は来ない…あるいは当分来ないことを示した。また、放射線計測分野と機械学習の親和性の高さについて紹介し、私が実施した例を2つ示した。中性子スペクトルアンフォールディングの例では、初めて機械学習を触ったとしても、何十年も使われている伝統的なコードと同等の性能を出せることを示せた。また、人力で解析されることが多いガンマ線スペクトルに関しても機械学習で(少なくとも)核種同定をできることを示した。核種同定だけでも応用先はあると考えているが、今後は放射能推定まで拡張し、複数核種が混在する例にもチャレンジしていく予定である。

謝辞

我々の研究結果のうち前者については、共に勉強・議論を行い、立派な修士論文をまとめて下さった卒業生 三仙 幸将さんの多大なる努力の結晶です。この場をお借りして感謝いたします。また、後者の成果については、日本分析センター 大島 真澄 さま、新潟大学 後藤 淳 先生のご協力なしには達成し得なかった結果です。今後もお世話になることと思いますが、この場を借りて感謝申し上げます。加えて九州大学工学部4年生の高橋龍平くんとは今年度毎日のように機械学習について議論・研究を行ってきました。私自身の知識も深めることができ、深く感謝しています。修士課程でも共に頑張りましょう。

参考文献

- [1] W. McCulloch and W. Pitts: “A logical calculus of the ideas immanent in nervous activity” *Bulletin of Mathematical Biophysics* **7** (1943) 115.
- [2] Frank Rosenblatt: “The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain” *Psychological Review* **65** (1958) 386.
- [3] David E. Rumelhart, Geoffrey E. Hinton & Ronald J. Williams: “Learning representations by back-propagating errors” *Nature* **323** (1986) 533.
- [4] G.E. Hinton and R.R. Salakhutdinov: “Reducing the Dimensionality of Data with Neural Networks” *SCIENCE* **313** (2006) 504.
- [5] V. Nair, G.E. Hinton: “Rectified Linear Units Improve Restricted Boltzmann Machines” *Proc. of 27th ICML'10* (2010) 807.
- [6] N. Srivastava, G.E. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov: “Dropout: A simple Way to Prevent Neural Networks from Overfitting” *J. of ML Res.* **15** (2014) 1929.
- [7] T. Kin, Y. Sanzen, et al.: “Artificial Neural Network for Unfolding Accelerator-based Neutron Spectrum by Means of Multiple-foil Activation Method” Conference Record of 2017 IEEE NSS/MIC (2017) 8532892.
- [8] T. Kin, J. Goto, and M. Oshima: “Machine Learning Approach for Gamma-ray Spectra Identification for Radioactivity Analysis” Conference Record of 2019 IEEE NSS/MIC (to be published).