

卒業論文 2020 年度 (令和 02 年)

カーネル関数を用いた汎用的な活性化関数の開発

慶應義塾大学大学院 環境情報学部
神保和行

カーネル関数を用いた汎用的な活性化関数の開発

近年機械学習では、ニューラルネットワークにおける活性化関数として、シグモイド関数や ReLU 関数などの一般的に用いられてきた。ニューラルネットワークにおける活性化関数は、その種類や問題に応じて最適な活性化関数を経験則に基づいて調整していた。一方、統計学の分野では、リンク関数が未知の場合には、カーネル関数を用いてノンパラメトリックに推定するという手法が推定されている。そこで本論文は事前に関数の形を指定しないカーネル関数を用いた活性化関数を提案する。さらに、実際のデータセットを用いて、ニューラルネットワークの出力層を本論文で提案する方法に置き換えることにより従来の活性化関数と同等かそれ以上の精度で予測できること示した。

キーワード:

1. ディープラーニング, 2. 活性化関数, 3. ノンパラメトリック, 4. カーネル関数

慶應義塾大学大学院 環境情報学部
神保和行

In recent years, activation functions such as sigmoidal and ReLU functions have been commonly used in deep learning. Those activation functions can achieve high accuracy under certain conditions, but whether or not they are optimal in all situations is less discussed. On the other hand, in the field of statistics, when the link function is unknown, a non-parametric estimation method using kernel functions has been estimated. Therefore, this paper proposes a deep learning activation function that does not specify the form of the function in advance. Furthermore, we show that it is possible to predict a deep learning activation function on an actual dataset with the same or better accuracy than conventional activation functions.

Keywords :

1. Deep lerning, 2. Activation Function, 3. Non parametric, 4. Kernel Function

Keio University Faculty of Environment and Information Studies

Kazuyuki Jimbo

目次

第1章	序論	1
1.1	はじめに	1
1.2	本論文の構成	1
第2章	背景	2
2.1	活性化関数	2
2.2	ニューラルネットワークの	2
2.3	統計学における位置付け	2
2.4	ノンパラメトリックモデルとカーネル法	2
2.5	実社会における学習の問題点	2
第3章	提案手法	3
3.1	概要	3
3.2	ノンパラメトリック	3
3.3	活性化関数	4
3.4	kernel 活性化関数	4
3.5	アルゴリズム	4
第4章	実装	5
4.1	実装環境	5
4.2	比較データ	5
4.3	実装手法	6
4.4	活性化関数	6
4.5	6
4.6	6
第5章	評価	7
5.1	評価内容	7
第6章	結論	8
6.1	本研究のまとめ	8
6.2	本研究の課題	8
6.3	将来的な展望	8

图 目 次

表 目 次

第1章 序論

本章ではまず、本研究を取り巻く社会の背景について述べる。そして本研究の解決する課題及び課題を解決する意義、解決するための手法を提示する。最後に本論文の構成を外観し、序論を締める。

1.1 はじめに

慶應義塾大学 SFC では、卒業要件として卒業論文の執筆が必要とされている。近年、多くの学生が提出間近になってから卒論を執筆することが多くなっている。そうした学生の多くは、残留を繰り返し、魔剤を飲みながらデスレースを実施することとなる。

その中でも、 \LaTeX の理解は執筆において不可欠であり避けられない。しかしながら、多くの学生は WIP/TERM で予稿の執筆を怠り、いざ執筆を始めようとしても \LaTeX を用いて論文を執筆することが難しい。

そこで、本研究では RG の学生に向けて心優しい博士課程として、RG の卒業論文のスタイルに合った形であると言われているテンプレートを整理し、提供する。本テンプレートでは、基本的な章立ての中で、 \LaTeX の使い方を概説し、このクソみたいな文章を削除し、卒業論文を執筆するにあたって基本的な記法を理解できることを期待する。

なお、Bitcoin [1] は関係ない。

1.2 本論文の構成

本論文における以降の構成は次の通りである。

2 章では、ノンパラメトリックモデル、カーネル法などといった本研究へとつながる背景の解説し、これらの手法における課題を洗い出す。?? 章では、本研究における問題の定義と、解決するための要件の整理を行う。3 章では、本研究におけるカーネル法を用いた活性化関数についての解説を行い、提案手法の解説の詳細を述べる。4 章では、3 章で述べた手法の実装及び、実装における留意点と。5 章では、?? 章で求められた課題に対しての評価を行い、考察する。6 章では、実験の結果に対する考察を行い、本研究を行う上で浮上した提案手法の限界を示し、今後の研究方針についてまとめる。

第2章 背景

本章では本研究の背景について述べる。まず機械学習における活性化関数の役割について明確にする。活性化関数について概説し、現在の機械学習における活性化関数の抱える問題点を明らかにする。

次に、活性化関数の他に、ニューラルネットワークにおける精度を向上させるいくつかの構成要素について述べる。

また、統計学において活性化関数というのがどのような

次に、ノンパラメトリックモデルを

最後に、実社会において機械学習を行う上での問題点や課題を述べ、本研究が取り組むべき課題を明確にする。

2.1 活性化関数

2.2 ニューラルネットワークの

・ラーニングレート・初期値、・レギュライザー (l1 ノルムなど) ・optimizer この辺について

2.3 統計学における位置付け

2.4 ノンパラメトリックモデルとカーネル法

2.5 実社会における学習の問題点

第3章 提案手法

本章では提案手法について述べる。

3.1 概要

現在、深層学習に利用できる活性化関数が研究されている。最近では、中間層に ReLU を用い、最終的な出力層にシグモイドを用いた組み合わせがよく用いられている。しかし、これらの組み合わせは経験的なものだけでなく、データに対する人間の知識が事前に必要とされる。本研究では、統計学の世界で使われている SIM のノンパラメトリック法を用いて、活性化関数に我々のリンク関数推定法を適用しました。そうすることで、より精度の高い結果を導き出すことができます。関数の形式はカーネル関数であり、入力 of 出力を一つの式で表現することができます。これにより、深層学習に利用できる程度に計算コストを削減することができます。また、今回の実験では、活性化関数を既存の関数から選択するのではなく、状況に応じた関数、つまり活性化関数の形を作ることができる。そうすることで、関数全体から逆算して最適な関数を見つけることができる。これにより、これまでディープラーニングで課題とされてきた活性化関数の選択の問題を解決することが可能となり、新たなアプローチが可能になるのではないのでしょうか。

3.2 ノンパラメトリック

現状ではディープラーニングに活かせるようなノンパラメトリックに推定する活性化関数は研究されておらず、経験的に中間層では Relu、最終的なアウトプット層ではデータセットに合わせて Sigmoid が使われることが多い。しかしながらこれらの組み合わせは経験的であるだけでなく、データに対する人知見が事前に必要である。本研究では、統計の世界で使われていた SIM でのノンパラメトリックな手法を用いて行われていたリンク関数の推定方法を活性化関数に応用する。そうすることにより、経験的な知見による活性化関数の選択という行為を行わずともより高い精度の結果を導けるのではないかということである。関数の形式はカーネル関数を用いることで、入力に対しての出力を一つの式で表せるようにする。そうすることでディープラーニングでも使えるぐらいの少ない計算コストが実現できる。また、この実験により状況に応じた適切か活性化関数の形を既存のものから選択するのではなく、関数全体の中から逆算できると考えている。それにより、ディープラーニングの課題であった、活性化関数の選択問題という課題も新しいアプローチで解決できると考えている。

Algorithm 1 K-AF-tron

Input: data $\langle (x_i, y_i) \rangle_{i=1}^m \in \mathbb{R}^d \times [0, 1]$, $u : \mathbb{R} \rightarrow [0, 1]$.
 $w^1 := 0$;
for $t = 1, 2, \dots$ **do**
 $h^t(x) := u(w^t \cdot x)$;
 $w^{t+1} := w^t + \frac{1}{m} \sum_{i=1}^m (y_i - u(w^t \cdot x_i)) x_i$;
end for

3.3 活性化関数

本節では既存の活性化関数の問題点を具体的な事例を交えて考える。

3.4 kernel 活性化関数

本論文で私が提案する活性化関数を以下の数式で表現する。

$$G(X_i w) = \frac{\sum_{i \neq j} K \left(\frac{X_j^{calc} w - X_i w}{h_{calc}} \right) Y_j^{calc}}{\sum_{i \neq j} K \left(\frac{X_j^{calc} w - X_i w}{h_{calc}} \right)} \quad (3.1)$$

これは ichimura(1993) の手法を改良したものになっている。ichimura ではデータセットの数だけで表現していたが、一部を省略することにより少ない変数で表現することに成功した。

3.5 アルゴリズム

第4章 実装

本章では本研究における実装環境, 提案手法の実装, 提案手法の評価に用いるデータセットについて述べる. その後, 実験に用いたハイパーパラメータ, 実装における留意点について解説を行う

4.1 実装環境

本研究において利用した実装環境を Table 4.1 に示す. 提案手法の実装は Pytorch 及を用いた. PyTorch, Chainer は計算グラフの自動微分ライブラリであり, 深層ニューラルネットワークの研究や開発にも用いられる. Pytorch を用いた理由は実装コストが低く研究領域に従事できるところにある。

ソフトウェア	バージョン
Python	3.6.2 or above
CPU	intel core i7 Tensorflow 2.1.0-rc0
PyTorch	6

4.2 比較データ

他の活性化関数と適当に比較するために、以下の条件を比較して実験を行う。・ラーニングレート・初期値、・レギュライザー (l1 ノルムなど) ・optimizer ・テストデータ既存のものと比較している”

ラーニングレート

初期値

初期値は

レギュライザー

レギュライザーは L1 ノルム, L2 ノルム等の比較ができる。

optimizer

テストデータ

各テストデータの特徴を以下の表にまとめる。

データセット名	出力層
Python	3.6.2 or above
CPU	intel core i7 Tensorflow 2.1.0-rc0
PyTorch	6

4.3 実装手法

各データセットにおける中間層は以下のパラメータで固定した。

4.4 活性化関数

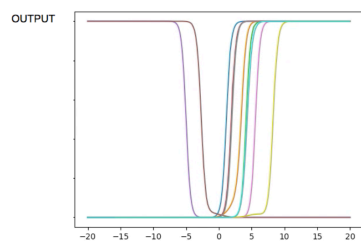
4.5

4.6

第5章 評価

本章では，提案システムの評価について述べる．

5.1 評価内容



第6章 結論

カーネルを使った汎用的な関数でニューラルネットの最終層を置き換えることで、実際に制度の向上を図ることができた。

6.1 本研究のまとめ

6.2 本研究の課題

スカラー値が大きなデータセットにおいては、その推論の精度が低下するだけでなく、Nan 値に陥ってしまうことがわかる。これらを解消するために、適切なニューラルネットの構成をより一層研究するだけでなく、それらが起こる原因を探求する必要がある。

6.3 将来的な展望

最終層だけではなく、中間層も置き換える。

謝辞

本論文の執筆にあたり、ご指導頂いた慶應義塾大学環境情報学部村井純博士、同学部教授中村修博士、同学部教授楠本博之博士、同学部准教授高汐一紀博士、同学部教授三次仁博士、同学部准教授植原啓介博士、同学部准教授中澤仁博士、同学部準教授 Rodney D. Van Meter III 博士、同学部教授武田圭史博士、同大学政策・メディア研究科特任准教授鈴木茂哉博士、同大学政策・メディア研究科特任准教授佐藤 雅明博士、同大学 SFC 研究所上席所員齊藤賢爾博士に感謝致します。

特に齊藤氏には重ねて感謝致します。研究活動を通して技術的視点、社会的視点等の様々な視点から私の研究に対して助言を頂き、深い思考と学びを経験させて頂くことができました。これらの経験は私の人生において人・学ぶ者として、素敵な財産として残りました。博士の指導なしには、卒業論文を執筆することは出来ませんでした。

徳田・村井・楠本・中村・高汐・バンミーター・植原・三次・中澤・武田合同研究プロジェクトに所属している学部生、大学院生、卒業生の皆様に感謝致します。研究会に所属する多くの方々が各々の分野・研究で奮闘している姿を見て学んだことが私の研究生活をより充実したものとさせました。

異なる分野同士が触れ合い、学び合う環境に出会えたことを嬉しく感じます。また、NECO 研究グループとして多くの意見・発想・知見を与えてくださった、慶應義塾大学政策メディア・研究科 阿部涼介氏、卒業生 菅藤佑太氏、在校生 島津翔太氏、宮本眺氏、松本三月氏、梶原留衣氏、渡辺聡紀氏、木内啓介氏、後藤悠太氏、倉重健氏、九鬼嘉隆氏、内田溪太氏、山本哲平氏、吉開拓人氏、金城奈菜海氏、長田琉羽里氏、前田大輔氏に感謝致します。

皆様には、私の研究に対する多くの助言や発想を頂いただけでなく、研究活動における学びを経験させて頂きました。多くの出会いと学びの環境である SFC に感謝致します。多様な学問領域に触れ、学生同士で議論し思考することが出来ました。幸せで素敵な時間でした。

最後に、これまで私を育て、見守り、学びの機会を与えて頂いた、父 良昭氏、母 ちや子氏、兄 良行氏 に感謝致します。

参考文献

- [1] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. <http://www.cryptovest.co.uk/resources/Bitcoin%20paper%20original.pdf>, 2008.