# Spatial Pyramid Pooling を用いた Gated CNNによる評判分析 

Sentiment Classification with Gated CNN using Spatial Pyramid Pooling

岡田 真＊1<br>Makoto Okada<br>柳本 豪一＊1<br>Hidekazu Yanagimoto<br>橋本 喜代太＊2<br>Kiyota Hashimoto

${ }^{* 1}$ 大阪府立大学 ${ }^{* 2}$ Prince of Songkla University Phuket Campus<br>Osaka Prefecture University


#### Abstract

The paper proposes a neural sentiment classification system with Gated CNN applied Spatial Pyramid Pooling． We apply Gated CNN to feature generation from text but not language model construction．Moreover，to generate the final single feature vector the proposed method uses max pooling and Spatial Pyramid Pooling．In experiments we evaluate the proposed method with real datasets：amazon product reviews．We confirmed the proposed method achieves the same performance as SVM with linear kernel without any customization．


## 1．はじめに

評判分析はインターネット上の大量の文書情報に含まれる筆者の意見や評価（好評，不評）を理解するために重要かつ代表的な手法の一つとなっている。自然言語処理を用いた評判分析 は多くの文書の内容を正確に理解する手法としてその重要性は高待っている。

自然言語処理分野では文書を単語の集合としてあらわす Bag of Words（BOW）モデルがよく用いられている．BOW モデ ルではそれぞれの単語は独立に扱われ，単語の意味的なつな がりは失われる．そのため，潜在的意味解析（latent semantic analysis，LSA）等の手法を用いて意味の近い単語を集約する手法が従来用いられてきた。

近年，文書コーパスから単語の意味を表すベクトルを生成 する手法が注目されている。それにより構築されたベクトル を単語の分散表現ベクトルと呼ぶ。分散表現ベクトルは単語 の意味を計算により比較，評価することが可能なため，自然言語処理関連分野では自動翻訳やテキスト生成などさまざまな分野で用いられている。また，再帰的ニューラルネットワーク （reccurent neural network，RNN）を用いて，文の分散表現を作成し，文間の内容を比較するなどの応用も提案されている．

近年の自然言語処理分野ではニューラルネットワークを用い た言語モデルが提案されており，さまざまな分野で用いられてい る．Long Short－term Memory（LSTM）や Gated Recurrent Unit（GRU）といったRNN を用いた言語モデルが代表的で ある．RNN が用いられているモデルは距離の近い単語の関係 を考慮して処理することが可能なため，言語処理と相性が良い と言われている。

畳み込みニューラルネットワーク（convolutional neural net－ work，CNN）は，人間の視覚処理を基にしたニューラルネット ワークモデルであり，その性質上，画像処理分野での成果が多 い。CNN は畳み込み層とプーリング層から構成されている。畳み込み層では入力データを小さな領域に分けて，その中の要素を重み付きで集約する。重み付き小領域用のデータをフィル タ（カーネル）と呼ぶ，プーリング層では，重要な特徴を最大 プーリングや平均プーリングなどのさまざまな戦略に基づいて選出する．CNNをGPGPUを利用した並列計算処理と相性 がよく，データ処理や学習を高速におこないやすい。その特性

連絡先：岡田 真，大阪府立大学，599－8531 大阪府堺市中区学園町 1－1，072－252－1161（代表），okada＠cs．osakafu－u．ac．jp

から近年では自然言語処理分野でも CNNを用いた言語モデル が注目を集めている。単語間の距離を考慮する RNN と違い， CNN では入力データは距離に関係なくフラットに扱う特徴が ある。
文書中のある種の単語は，その文の評価情報に強い影響を与えると仮定される。文書中のどの単語が評価情報に強い影響 を与えているかを予測するのは難しい。我々は CNN を用いた言語モデルがそれらの予測を可能にするのではないかと考えて いる．重要な単語を適切に抽出するためには，不必要な単語を うまく無視する必要がある。RNN を用いた処理では各文中の どの位置の単語が意味的に重要かという情報をゲート構造を用 いて扱えるようにしている。CNN を用いた言語モデルにおい て，ゲート構造を用いた手法が注目されている。
本論文では，我々は Gated CNN を用いた評判分析を提案 する．Gated CNN はゲート構造を通常の畳み込みニューラル ネットワーク構造に組み达んだモデルである。Gated CNN は言語モデルを構築する際に用いられることがあるが，提案手法 では Gated CNNを入力データから特徴ベクトルを生成する際に用いる。 さらに，プーリング手法の一つである Spatial Pyramid Pooling［8］を組み込み，その有効性を実験により検証した。
2 章では Gated CNN による評判分析について説明する． 3章ではSpatial Pyramid Poolingについて述べる．4章で評価実験について説明し，最後にまとめと今後の課題について述 べる。

## 2．Gated CNN による評判分析

提案手法である Gated 畳み込みニューラルネットワーク （Gated convolutional neural network，Gated CNN）［7］を用 いた評判分析について説明する。 Gated CNN は文書中の連続 した単語列をまとめて意味のベクトルを生成する。さらに，最大プーリングを文中の単語から生成された意味ベクトルに適用 することで提案手法は文全体の意味ベクトル表現を得ることが できる．また，提案手法は文中の単語列の長さを処理に適切な長さで自由に扱うことが可能である。

## 2．1 Gated 畳み込みニューラルネットワーク

本提案手法の重要部である Gated CNN について説明する ［7］．図 1 では，Gated CNN は文中の単語 2 個ずつ選び，文中の単語すべてに処理が適用され，意味ベクトルが生成されて


図 1：Gated CNN による評判分析の模式図


図 2：Gated CNN の模式図

いる。図2 では，gated CNN 自体の模式図を示す。自然言語処理分野ではLSTM を用いた言語モデルが用いられることが多い。LSTM は3個のゲートで情報の流れを制御して，適切 な言語モデルを構築している。Gated CNN には，LSTM で用いられているようなゲート構造が従来の CNN に導入されて いる。

通常の CNN と同様に，Gated CNN においても入力として複数のベクトルからなる行列を受け取る。入力された行列は 2個の異なるカーネル（テンソル）により， 2 個の別々のベクト ルに変換される。ベクトルの内， 1 つは入力データの意味情報 を表し，もう 1 つは Gated CNN 内部のコントロールゲート の情報を表す。この処理は次の数式で示される。

$$
\begin{equation*}
\mathbf{X}^{\prime}=\sigma\left(\mathbf{X} * \mathbf{F}_{g}\right) \otimes\left(\mathbf{X} * \mathbf{F}_{c}\right) \tag{1}
\end{equation*}
$$

ただし， $\mathbf{X} \in \mathbb{R}^{N \times m}$ は Gated CNN への入力， $\mathbf{F}_{g} \in \mathbb{R}^{k \times m \times n}$ と $\mathbf{F}_{c} \in \mathbb{R}^{k \times m \times n}$ は畳み込み処理のカーネル，$\sigma(\cdot)$ はシグ モイド関数，$*$ は畳み込み処理，$\otimes$ はベクトル間の要素ごとの積を表す。したがって出力は $\mathbf{X}^{\prime} \in \mathbb{R}^{k \times(N-n+1)}$ となる。出力長はデータでパディングがされない場合は入力 X よりも短 くなる。ここでは， $\mathbf{F}_{g}$ はコントロールゲートの情報であり， $\mathbf{F}_{c}$ は入力データの畳み込み後の意味情報を表す。シグモイド関数の出力は 0 から 1 までの範囲なので，入力データの値に かかわらず出力もその範囲内に収まる。パラメータと $\mathbf{F}_{g}$ と $\mathbf{F}_{c}$ は訓練データにより学習される。

## 2．2 Gated CNN による評判分析

本来の Gated CNN では，文書コーパスから言語モデルを生成することを提案している。であるので，入力は単語列であ り，出力もデータの系列となる。Gated CNN を用いた評判分

析のためには，入力は単語列であり，それらの中の意味情報を基にして素性ベクトルが出力として生成されなければならな い。出力された素性ベクトルは好評不評などの分類器への入力 として扱われる。

まず，入力データを生成する．文中の単語を埋め込み行列 $\mathbf{D} \in \mathbb{R}^{m \times|V|}$ と参照して単語の埋め込みベクトルを得る。こ こで $|V|$ はコーパス全体の出現単語数（語彙数），$m$ は分散表現のベクトルの次元数を表す。したがって， $\mathbf{w}_{i}$ は 1－of－N コーディングベクトルとなり，埋め込み処理は以下の式で示さ れる。

$$
\begin{equation*}
\mathbf{e}_{i}=\mathbf{D} \mathbf{w}_{i} \tag{2}
\end{equation*}
$$

文は単語の系列なので，埋め込みベクトルの系列に変換され る。本論文ではベクトルの系列を X で表す。したがって X は Gated CNN への入力となる。

$$
\begin{equation*}
\left(\mathbf{w}_{1}, \mathbf{w}_{2}, \cdots, \mathbf{w}_{N}\right) \Rightarrow\left(\mathbf{e}_{1}, \mathbf{e}_{2}, \cdots, \mathbf{e}_{N}\right)=\mathbf{X} \tag{3}
\end{equation*}
$$

Gated CNN が文に適用されると，単語の情報は集約される。例えば2個の単語で考えると， $\mathbf{e}_{1}$ と $\mathbf{e}_{2}$ が Gated CNN で集約される。

$$
\begin{equation*}
\mathbf{m} \otimes\left(\mathbf{k}_{1}^{\mathrm{T}} \mathbf{e}_{1}+\mathbf{k}_{2}^{\mathrm{T}} \mathbf{e}_{2}\right) \tag{4}
\end{equation*}
$$

$\mathbf{m}$ は式1中の $\sigma\left(\mathbf{X} * \mathbf{F}_{g}\right)$ で計算される。 $\mathbf{k}_{1}$ と $\mathbf{k}_{2}$ は重みで あり，式1中の $\mathbf{F}_{c}$ から求められる。

入力 X を Gated CNN で処理したのち，ベクトル列は短 くなっている。出力されたベクトル列から素性ベクトルが構築 される．その素性ベクトルを用いて好評不評といった文書のポ ラリティを予測する。系列から素性ベクトルを生成する手法は数多く提案されている。本論文では単語出現順の最大要素を用 いて素性ベクトルを生成する。

$$
\begin{equation*}
\mathbf{f}=\max _{1 \leq j \leq N-n+1} \mathbf{X}^{\prime} \tag{5}
\end{equation*}
$$

式 5 は要素の順序に従う。

$$
\begin{equation*}
f_{i}=\max \left(X_{i 1}^{\prime}, X_{i 2}^{\prime}, \cdots, X_{i(N-n+1)}^{\prime}\right) \tag{6}
\end{equation*}
$$

特徴ベクトルは 3 層のニューラルネットワークに入力され， その出力がポラリティ（好評不評など）となる。

$$
\begin{gather*}
\mathbf{h}=\sigma\left(\mathbf{W}_{1} \mathbf{f}\right)  \tag{7}\\
\mathbf{o}=\operatorname{Softmax}\left(\sigma\left(\mathbf{W}_{2} \mathbf{h}\right)\right) \tag{8}
\end{gather*}
$$

提案手法の学習の際に，ロス関数を定義する。本論文ではク ロスエントロピーを用いて，予想と正解のポラリティとの差異 を求めた．t．

$$
\begin{gather*}
\mathbf{t}= \begin{cases}(1,0) & \text { (negative) } \\
(0,1) & \text { (positive) }\end{cases}  \tag{9}\\
E(\mathbf{o}, \mathbf{t})=t_{1} \log o_{1}+t_{2} \log o_{2} \tag{10}
\end{gather*}
$$

提案手法で用いるすべてのパラメータは以下の通りである。 $\theta=\left(\mathbf{D} ; \mathbf{F}_{g} ; \mathbf{F}_{c} ; \mathbf{W}_{1}, \mathbf{W}_{2}\right)$ ．システムは誤差逆伝搬法を使って学習する。

$$
\begin{equation*}
\theta \leftarrow \theta-\epsilon \frac{\partial E(\mathbf{o . t})}{\partial \theta} \tag{11}
\end{equation*}
$$

| 表 1：レビューデータセツトの内容 <br> Vビュー情報 |  |
| :--- | :--- |
| 好意的評価レビュー数 | 10,000 |
| 否定的評価レビュー数 | 10,000 |
| 訓練データ数 | 18,000 |
| テストデータ数 | 2,000 |
| 平均単語数 | 59.1 words |

## 3．Spatial Pyramid Pooling

Spatial Pyramid Pooling（SPP）は［8］プーリングの一種 であり，CNN による画像処理分野で用いられている手法であ る ${ }^{* 1}$ 。通常のものとの違いはプーリングの際に対象領域を分割 し，階層的に処理することにある．SPP のパラメータにレベ ルがあり，レベルによって対象領域を分割する個数が変わる。 レベル 1 では通常通り，レベル 2 では 2 個に分割，レベル $n$ では $2^{\prime n-1)}$ 個に分割される。それぞれの領域でプーリング処理をする．そしてどのレベルの結果までを用いるかを事前に決 めて，レベル 1 からそのレベルまでの出力結果を結合する。そ の結果を最終的な出力とする。画像処理では拡大縮小に対して頑健な処理が可能になるということで用いられることが多い。本論文では文に適用する。この場合，プーリングの対象範囲が文を先頭から $2^{(n-1)}$ 分割して，それぞれの領域から情報を取得する．これにより，文中のポラリティに関係する語がどのよ うな位置にあってもその情報を適切に取得して，素性ベクトル を適切に作成できることを期待する。今回のシステムでは，図 1 の最上部 Max Pooling の部分を SPP に置き換えたシス テムを提案する。

## 4．実験

提案手法の有効性をデータセットを用いた実験により検証 し，その結果に対して考察する。機械学習手法として一般的な サポートベクターマシン（SVM）による実験結果をベースライ ンとして用いる。

## 4.1 データセット

今回，データセットしては UCSD の Julian McAuley ［9，10］が配布している Amazon 商品レビューを用いた。こ のデータセットには Amazon の商品レビューとメタデータが含まれている。取得期間は1996年5月から2014年7月まで， レビュー総数は 1 億 4280 万レビューとなっている。各レビュー には（評価値，レビュー文書，役に立ったかどうか（投票））の情報と商品のメタデータ（商品説明文，商品カテゴリー，値段， メーカー名，画像）とリンクの情報が付与されている。

実験では，データセットから抽出された小データセットを実験用データとして用いた。これらは 24 個のカテゴリーに分け られており，今回は5個のカテゴリー（Electronics，Home and Kitchen，Sports and Outdoors，Health and Personal Care， Video Games）を実験用データを選択した。各小データセット から 2 万件のレビュー（好評 1 万件，不評 1 万件）をサンプリ ングした。 5 段階評価の評価 5 がつけられているものから好評のデータを抽出し，評価 1 がつけられているものを不評の データとして抽出した。それらの $10 \%$ に当たるレビュー千件 ずつ 2 千件をテストデータ，残りの 1 万 8 千件のレビューを

[^0]表 2：提案手法のパラメータ

| パラメータ名 | 値 |
| :---: | :---: |
| 単語の最低出現頻度 | 3 |
| 埋め込みのサイズ $(m)$ | 200 |
| カーネルのサイズ $\left(\mathbf{F}_{g}, \mathbf{F}_{c}\right)$ | $200 \times 200 \times 2$ |
| ミニバッチのサイズ | 500 |
| 分類器（3層 MLP） | 200－100－2 |
| 最適化手法 | AD |



図 3：Video Game レビューテストデータを用いた際の提案手法のロスと精度

訓練データとして用いた。各レビューの本文は千文字以下とし た。表？？にレビューデータの情報を示す。

## 4.2 実験用モデル

提案手法である Gated CNN による評判分析システムを Chainer＊2 で構築した。表2に提案手法で用いたパラメー夕を示す。
単語の最低出現頻度は 3 とした。 3 回以下の単語は未知語と考えられる．埋め込みのサイズは推定制度に強く影響する。今回は 200 次元とした。最適な次元数は今後の課題とする。カー ネルのサイズは $200 \times 200 \times 2$ とした。今回は 2 単語間の関係のみを考える． 2 単語を集約したのちの素性ベクトルは同じ次元数となる。提案手法では，最後の分類器は 3 層のニューラ ルネットワークとした。分類器も重要ではあるが，その比較調査は今後の課題とする。

## 4.3 実験結果

Amazon 商品レビューを用いた評判分析の結果を以下に示 す．今回の実験では，文書データに前処理を適用せず，そのま ま用いた。
図 3 に Video Game のデータセットを用いた際の訓練デー タのエラーとテストデータの精度を示す。提案手法の学習が進 むにつれて，エラーが減り，精度が向上している，30エポッ クまでに過学習の兆候は見られなかった。

サポートベクターマシン（SVM）の実験結果をベースライン として，提案手法と比較した。言語は Python（version 3．5）， ライブラリは scikit－Learn SVM（LIBSVM）を用いた。パラ メータは以下の通りである：カーネル＝線形， $\mathrm{C}=1.0$ ．提案手法の Gated CNN（GCNN）によるシステムの実験結果と比較した。SPP のレベルは 2 とした。
実験結果を表？？に示す。

[^1]表 3：実験結果（ F 値）（Amazon 商品レビューデータセット）

| カテゴリー | SVM <br> （Linear） | GCNN <br> （no SPP） | GCNN <br> （SPP） |
| :--- | :--- | :--- | :--- |
| Electronics | 0.91 | 0.91 | 0.89 |
| Home \＆Kitchen | 0.91 | 0.92 | 0.87 |
| Sports \＆Outdoors | 0.90 | 0.91 | 0.91 |
| Health \＆Personal Care | 0.88 | 0.88 | 0.90 |
| Video Game | 0.87 | 0.88 | 0.87 |

提案手法の精度（ F 値）は，ベースラインと同等か $1 \%$ 高い ものとなった。これにより前処理や調整なしでも提案手法が有効に働くことが示された。一般に，SVM で文書データを扱 う際に，線形カーネルが有効であると言われる。事前の実験 でも RBF カーネルよりも線形カーネルの方が有効であったた め，今回我々はSVM に線形カーネルを用いた。SVM と提案手法はどちらも高い精度を示しており，その有効性は明らかと いえる。提案手法の調整は今後の課題である。

また，今回は単語の集約単位を 2 個としている。これらも文書の長さに合わせて変更することで，精度が向上する可能性 がある．それらの適切な値などについても今後の課題とする。

SPP を用いた場合，精度が向上したものと低下したものが見受けられた。SPPを用いることにより，文はいくつかの領域に分けられてプーリングされる。その際に好評不評を評価す る単語が含まれる個所が適切に評価される場合とされない場合 で制 d の差が出たのではないかと考えられる。適切なレベル の値についても今後の課題である。

## 5．まとめと今後の課題

本論文では，ゲート機構付き畳み込みニューラルネットワー ク（Gated convolutional neural network，Gated CNN）に よる評判分析システムを提案した．Gated CNN を通常言語処理で系列データの生成によく用いられる。提案手法では，そ の系列データを用いて素性ベクトルを生成し，それらを訓練 データとして分類器を生成した。Amazon review datasetを用いた実験により，提案手法の有効性を確認した。サポートベ クターマシン（SVM）をベースラインとして比較し，前処理な しに同等の性能を得ることができた。

今後の課題として，提案手法の構造の改良がある。例えば埋 め込みベクトルやカーネルの大きさ，単語の出現頻度の最低値，複数の分類器による性能の比較などである．カーネルの大 きさは一度に集約する単語数に関連するため，重要な課題の一 つである．非線形なデータに対応可能な機構の検討も重要な課題の一つである。また，SPP の適切なレベルの値についても今後の課題の一つである。

## 参考文献

［1］V．Hatzivassiloglou and K．R．McKeown，Predicting the Semantic Orientation of Adjectives，Proc．of the 35 Annual Meeting of the Association for Computa－ tional Linguistics and Eight Conference of the Euro－ pean Chapter of the Association for Computational Linguistics，pp．174－181， 1997.
［2］R．Socher，J．Pennington，E．H．Huang，A．Y．Ng， and C．D．Manning，Semi－supervised recursive autoen－ coders for predicting sentiment distribution，Proc．of
the Conference on Empirical Methods in Natural Lan－ guage Processing，pp．151－161， 2011.
［3］R．Socher，B．Huval，C．D．Manning，and A．Y．Ng，Se－ mantic compositionally through recursive matrix－vector spaces，Proc．of the Conference on Empirical Methods in Natural Language Processing，pp．1201－1211， 2012.
［4］R．Socher，A．Perelygin，J．Wu，J．Chuang，C．D．Man－ ning，A．Y．Ng，and C．Potts，Recursive deep models for semantic compositionality over a sentiment tree－ bank，Proc．of the Conference on Empirical Methods in Natural Language Processing，pp．1631－1642， 2013.
［5］S．Hochreiter and J．Schmidhuber，Long short－term memory，Neural Computation，Vol．9，No．8，pp．1735－ 1780， 1997.
［6］Kim Y．：Convolutional neural networks for sentence classification，Proc．of the Conference on Emprical Methods in Natural Language Processing，pp．1746－ 1751 （2014）．
［7］Dauphin Y．N．，Fan A．，Auli M．，and Grangier D．： Language Modeling with Gated Convolutional Net－ works，arXiv：1612：08083（2016）．
［8］He K．，Zhang X．，Ren S．and Sun J．：Spatial Pyra－ mid Pooling in Deep Convolutional Networks for Vi－ sual Recognition，journal arXiv：1406：4729（2014）．
［9］R．He and J．McAuley，Ups and downs：Modeling the visual evolution of fashion trends with oneclass collab－ orative filtering，Proc．of International Conference on World Wide Web（WWW＇16），pp．507－517， 2016.
［10］J．McAuley，C．Targett，J．Shi，and A．van den Hen－ gel，Image－based recommendations on styles and sub－ stitutes，Proc．of the 38th International ACM SIGIR Conference on Research and Development in Informa－ tion Retrieval，pp．43－52， 2015.


[^0]:    ＊1 http：／／image－net．org／challenges／LSVRC／2014／
    slides／sppnet＿ilsvrc2014．pdf

[^1]:    ＊2 https：／／chainer．org

