

経済テキストデータを用いた極性概念辞書構築とその応用

伊藤 友貴^{*1} 坪内 孝太^{*2} 山下 達雄^{*2} 和泉 潔^{*1}
Tomoki Ito Tatsuo Yamashita Tatsuo Yamashita Kiyoshi Izumi

^{*1}東京大学大学院工学系研究科 ^{*2}ヤフー株式会社
School of Engineering, The University of Tokyo Yahoo! JAPAN Research

現状、単語単位での極性辞書自動構築に関する既存研究はあるものの、似たような意味の単語がひとまとまりになった形での極性辞書、極性概念辞書の構築に関する研究はほとんどされていない。本研究では、極性概念辞書の構築手法を提案すると共にその有用性を検証した。まず、ロイターニュース記事・ヤフーファイナンス掲示板のテキスト情報、専門家によって与えられた約 100 単語の極性値、及び構造化されたニューラルネットワークモデルを用いて意味的にクラス分けされた極性概念辞書を作成した。その後、極性概念辞書を用いて市場動向の分析を試みた。

1. はじめに

情報通信技術の発達に伴い、金融テキストマイニングの技術に対する個人投資家及び機関投資家からの関心が高まってきている [1]。金融テキストマイニングとは、投資に有用な情報を SNS や記事のような大規模のテキストデータから抽出する技術である。この分野においては、極性辞書を使うことが有用であることがわかっている [2]。極性辞書とは単語の極性情報に関する辞書である。特定の文脈においてポジティブな意味を持つ単語にはプラスの極性値を、ネガティブな意味を持つ単語にはマイナスの極性値を与える形で各単語にその極性値を割り振っている辞書である。

通常、極性辞書は人手によって作成されるが、経済用語・ネットスラングについては十分な極性についての情報を極性辞書からは手に入れない。経済用語のポジネガ極性付与についての研究もいくつかされているが [3, 4]、今のところ確立された手法はない。また、これらの方法により作られるポジネガ辞書や現在世間に出回っている極性辞書からは、単語単位でのポジネガ情報は抽出できるものの、単語間の類似度は取ることができないため、異なる語群からなるテキストの情報を同じ指標による分析が難しい。このような分析には、似た意味をもつ極性単語がひとまとまりの概念になっている形での極性辞書が必要だが、そのような辞書の自動獲得手法は確立されていない。

自動獲得に関わる機械学習の手法として、近年深層学習の技術発展が著しいが、ニューラルネットワークモデルはその変数の多さ故、中身の分からないブラック関数に近い扱いをされており、モデルの解釈が難しい。

本研究の目的は市場動向分析に有用な単語の極性概念辞書を半自動的に構築するモデルの開発、モデルの解釈が可能でかつ予測精度が高いニューラルネットワークモデルの開発、語彙の異なる複数種類のテキストを同一の指標で分析することが可能な極性概念辞書の作成である。

まず、word2vec [5] を用いてニュース記事に出てくる単語にベクトルを与え、それをもとに各記事に特徴量ベクトルを与えた。これらの単語のうちの一部には経済専門家の手によってつけられた極性スコアが付与されている。その後、極性が不明な単語の極性を求めるために、各ニュース記事の特徴量と株価動向、及び Yahoo!Finance 掲示板上の投稿とその極性タグの対応の関係を分析した。機械学習の学習の過程で極性辞書に含まれる単語の極性スコアが極性辞書外の単語にも伝播することが期待できる。本手法によって得られる特徴量を用いた株価動

向・極性スコアの予測結果と既存手法によって得られる特徴量を用いた予測結果を比較することで、提案手法が与える極性値の妥当性を検証した。さらに、ロイターニュースとヤフーファイナンス掲示板から得られた二つの極性辞書をマージし、両種類のテキストを同時に分析できる極性概念辞書を作成した。

2. 手法

本節では単語への極性値付与に関して本研究で新たに提案する手法 II (importance infiltration) algorithm を紹介する。

2.1 Word classification and document representation 法 (CDR 法)

まず、word2vec によって単語に与えた分散表現を利用した文書の特徴量の生成手法、CDR 法 [6] の紹介をする。CDR 法は提案手法 II algorithm のベースとなった手法である。CDR 法では、意味の近い単語が同じクラスになるようにクラスタリングし、文書内に出現する各クラスの出現回数によって文書の特徴量を生成する。まず、word2vec [5] を使い、ニュース記事に出現する各単語にベクトル表現を与える。その後、クラス数 K を決めた後、K-means 法 [7] により単語のクラスタリングを行う。 K 個のクラスターを得た後、各文書の特徴量 $V_{\text{document}} (\in \mathbb{R}^K)$ を文書中に出現する各クラスの単語の回数をを用いて求める手法である。

2.2 Importance Infiltration propagation algorithm

word2vec で単語をクラスタリングすると、対義語同士が同じクラスに入ってしまうことがある。これは word2vec では各単語がどの言葉と組み合わせで使われるかによって単語の分散表現を獲得するためである。しかし、これは市場動向の分析をする上では望ましくない。そこで、これを改善するために、本論文ではニューラルネットワークモデルを利用したモデル (II (importance infiltration) algorithm) [8] を用いて極性辞書外単語に極性を与え、文書のベクトルを生成した。文書のベクトルを生成するにあたり、金融機関に所属する機関投資家が人手で作成した経済用語極性辞書の情報を用いた。II algorithm では図 1、図 2 のように表現されるニューラルネットワークモデルを用いた。入力層の次元は単語の数で、各単語に対応するノード (1 層目の各ノード) は、それぞれ属するクラスを表すノード (2 層目のノード) のみに結合するという構造のニューラルネットワークモデルである。ここで、 $W_{\text{polarity}} \in \mathbb{R}^{K \times m}$ (式 (5) にて定義、 K はクラス数)、 $W_3 \in \mathbb{R}^{2 \times K}$ 、 $W_3^* \in \mathbb{R}^{K^2 \times K}$ 、 $W_4^* \in \mathbb{R}^{2 \times K^2}$ は重み行列、 K^2

は層の数が4の場合における第3層目のノード数, $b_0 \in \mathbb{R}^2$ はバイアスペクトル, $y_{cls} \in \{0 \text{ (ネガティブ)}, 1 \text{ (ポジティブ)}\}$ は出力層の値である. y_{cls} は文書につけられるタグに該当する値である. 文書番号 j の文書内に出現する単語の頻度からなるベクトル V_{BOW_j} を式 (1) のように定義する.

$$V_{BOW_j} := Z_j^{(1)} = [Z_{j,1}^{(1)T}, Z_{j,2}^{(1)T}, \dots, Z_{j,K}^{(1)T}]^T \quad (1)$$

$Z^{(l)}$ は l 層目の出力を意味する. $Z_{j,k}^{(1)} \in \mathbb{R}^{n(k)}$, $Z_{j,k}^{(1)}[i]$ ($k = 1, 2, \dots, K$) は単語 $w_{k,i}$ ($w_{k,i}$ は単語のクラスが k でクラス内単語 ID が i の単語とする.) の文書番号 j の文書における頻度, $n(k)$ はクラス k である単語の数である. このとき, II algorithm で用いるニューラルネットワークモデル (図 1, 図 2) はそれぞれ3層時には式 (2), 4層時には式 (3) で y_j を表し, (4) のように y_{cls} を表すことで表現できる.

$$y_j = f_3(W_3(\tanh(W_{\text{polarity}} V_{BOW_j}) + b_0)) \quad (2)$$

$$y_j = f_4(W_4^*(\tanh(W_3^*(\tanh(W_{\text{polarity}} V_{BOW_j}) + b_0)))) \quad (3)$$

$$y_{cls_j} = \operatorname{argmax} y_j \quad (4)$$

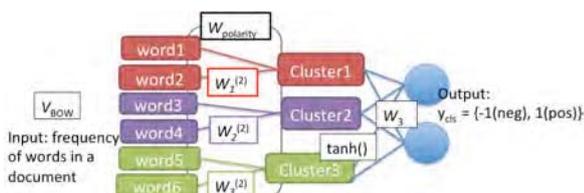


図 1: 本研究で用いたニューラルネットワークモデル (3 層)

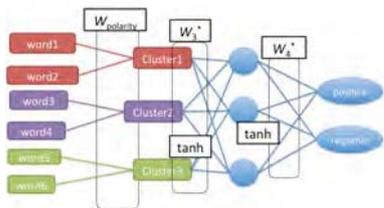


図 2: 本研究で用いたニューラルネットワークモデル (4 層)

活性化関数 f_3, f_4 には Softmax 関数を用いた. また, 学習時の損失関数には Softmax cross entropy 関数を用いた. 過学習を防ぐために学習時には, Dropout 法 [9] を用いた. $W_{\text{polarity}} (\in \mathbb{R}^{K \times m})$ を以下のように定義する.

$$W_{\text{polarity}} := \operatorname{diag}(W_1^{(2)T}, W_2^{(2)T}, \dots, W_K^{(2)T}) \quad (5)$$

$W_k^{(2)} \in \mathbb{R}^{n(k)}$ であり, $W_k^{(2)}[i]$ は $w_{k,i}$ の極性値に対応する.

次に $W^{(2)}$ の初期値の与え方について説明する. 単語 $w_{k,i}$ が既存の極性辞書内の単語であり, 極性値が事前に専門家の手によって与えられている場合, 単語 $w_{k,i}$ の極性辞書値 $PS(w_{k,i})$ を用いて $W_k^{(2)}[i]$ の初期値を以下のように与える.

$$W_k^{(2)}[i] = \begin{cases} PS(w_{k,i}) & \text{(単語 } w_{k,i} \text{ が極性辞書単語のとき)} \\ 0 & \text{それ以外} \end{cases}$$

とする. 教師あり学習の過程で, W_{polarity} の値が更新されていく. 学習後に W_{polarity} の値を取り出すことで既存極性辞

書外単語の極性値を得ることができる. $W_k^{(2)}[i]$ の値が伝搬後の単語 $w_{k,i}$ の極性値に対応する.

以下に理論解析から判明した II algorithm (3 層) の性質を述べる.

2.3 定義

学習時のミニバッチサイズ N を用いて入力値 $X (\in \mathbb{R}^{m \times N})$, j^+, j^-, p^+, p^-, U, Z を以下に定義する.

$$X := [V_{BOW_1}, V_{BOW_2}, \dots, V_{BOW_N}]$$

$$p^-(w_{k,i}) := p(w_{k,i} \text{ 出現文書がネガティブ} | w_{k,i} \text{ が文書に出現})$$

$$p^+(w_{k,i}) := p(w_{k,i} \text{ 出現文書がポジティブ} | w_{k,i} \text{ が文書に出現})$$

$$U^{(2)} := W_{\text{polarity}} X, Z^{(2)} := \tanh(U^{(2)})$$

$$U^{(3)} := W_3(Z^{(2)} + b_0)$$

便宜的に j を文書番号 j の文書がポジティブである場合に j^+ , 文書番号 j の文書がネガティブである j 場合に j^- と表記する. また, $w_{k,i^+}, w_{k,i^-}, w_{k,i^{neu}}$ を

$$w_{k,i} = \begin{cases} w_{k,i^+} & p_{w_{k,i}^+}^+ > p^{t^+} \\ w_{k,i^-} & p_{w_{k,i}^-}^- > p^{t^-} \\ w_{k,i^{neu}} & \text{otherwise} \end{cases}$$

と定め,

- $n^+(k)$: クラス k 内の単語で $p_{w_{k,i}^+}^+ > p^{t^+}$ である単語の数,
- $n^-(k)$: クラス k 内の単語で $p_{w_{k,i}^-}^- > p^{t^-}$ である単語の数,
- $n^{neu}(k)$: 上記二つの条件を満たさない単語の数

と定める.

2.4 極性伝搬条件 (II algorithm (3 層))

II algorithm (3 層) について以下に述べる単語への極性付与に関する性質が成り立つ.

- (i) ポジティブ単語 w_{k,i^+} に十分大きい正の極性値, ネガティブ単語 w_{k,i^-} に十分小さい負の極性値が付与
- (ii) $p^-(w_{k,i^-}), p^+(w_{k,i^+}), n^-(k), n^+(k)$ が十分に大きい
- (iii) $W_3[0][k] < 0$ 十分に小さく, $W_3[1][k]$ 十分に大きい
- (iv) ミニバッチサイズ N が十分に大きい

このとき, 任意の i^-, i^{neu}, i^+ について条件 (i)–(iv) を満たすように初期値が与えられるという理想的な場合には

$$E[\partial W_k^{(2)}[i^+] < 0, E[\partial W_k^{(2)}[i^-] > 0$$

となり, ポジティブ単語 w_{k,i^+} にはプラスの極性値が与えられ, ネガティブ単語 w_{k,i^-} にはマイナスの極性値が与えられる.

この主張は条件 (i)–(iv) が成り立つという条件下での誤差逆伝搬の様子を計算して求めることで示すことができる. 具体的な説明については, 付録 A にて紹介する.

2.5 極性伝搬条件 (II algorithm (4 層))

II algorithm (4 層) について以下に述べる単語への極性付与に関する性質が成り立つ. $U^{(3)}$ についての

$$U^{(3)} := W_3^*(Z^{(2)} + b_0)$$

と定義し直すと, 以下が成立する.

$$\begin{aligned} ((\operatorname{diag}(1 - (\tanh(U_j^{(3)}))^2) W_3^*)^T W_4^{*T})[0] < 0 \\ ((\operatorname{diag}(1 - (\tanh(U_j^{(3)}))^2) W_3^*)^T W_4^{*T})[1] > 0 \end{aligned} \quad (6)$$

が満たされ、かつ 2.4 節の条件 (i)–(iii) が任意の i^- , i^{neu} , i^+ について成り立つとき、単語の極性は

$$E[\partial W_k^{(2)}[i^+]] < 0, E[\partial W_k^{(2)}[i^-]] > 0$$

を満たすように伝搬する。また、条件 (i)–(iii), 及び「 $W_3^* > 0$ かつ $W_3^* > 0$ が十分に大きく、 $W_4^*[0] < 0$ が十分に小さく、 $W_4^*[1] > 0$ が十分に大きい (このとき式 (6) は成立)」を満たすように初期値が与えられるという理想的な場合にはポジティブ単語 w_{k,i^+} にはプラスの極性値が与えられ、ネガティブ単語 w_{k,i^-} にはマイナスの極性値が与えられる。

この主張も上記の条件が成り立つという条件下での誤差逆伝搬の様子を計算して求めることで示すことができる。具体的な証明については、長くなるので本稿では省略する。

3. 人工データを用いた極性伝搬テスト

提案手法の妥当性を示すために、人工データを用いた単語へのポジネガ付与に関する実験をおこなった。本実験は 2.4, 2.5 の確認を目的として行う実験である。

3.1 実験設定

まず、人工データを以下のように生成した。

- クラスタ数 $K = 200$, クラスタ内単語数 $n(k) = 30$ ($k = 1, 2, \dots, 200$)
- $k \leq 100$ の場合: $p^+(w_{k,i}) = \eta$ ($1 \leq i \leq 10$), $p^-(w_{k,i}) = \eta$ ($11 \leq i \leq 20$)
- $k > 100$ または ($k \leq 100 \wedge 20 < i \leq 30$) の場合: $p^+(w_{k,i}) \sim \text{Norm}(0.5, 0.05)$
- 単語 $w_{k,i}$ の文書への出現頻度 $\text{freq}(w_{k,i}) = 0.2$
- データ数: ポジティブ文書 800, ネガティブ文書 800
- ミニバッチサイズ $N = 10$
- $W_k^{(2)}$ の初期値は

$$\begin{cases} W_k^{(2)}[1] \sim U(0, 1) & (k \leq 60) \\ W_k^{(2)}[20] \sim U(-1, 0) & (21 \leq k \leq 80) \\ W_k^{(2)}[i] = 0 & ((k \geq 81) \vee (k \leq 80 \wedge i \neq 1 \wedge i \neq 20)) \end{cases}$$

で生成。

$\text{Norm}(u, \sigma)$ は平均値 u , 標準偏差 σ の正規分布, $U(a, b)$ は区間 $[a, b]$ 上の一様分布を指す。

人工文書データセットを η の値を

$\eta \in \{1.0, 0.9, 0.8, 0.7, 0.6, 0.5\}$ について変えながら生成した。その後 それぞれのデータセットについて II algorithm (epoch 数 10) を用いて各単語にポジネガスコアを与えた後、各 $k \leq 100$ について

- $(W_k^{(2)}[i] > 0) \wedge (1 \leq i \leq 10)$ を満たす $w_{k,i}$ の数 (= $n_{\text{pred}}^+(k)$ とする。),
- $(W_k^{(2)}[i] < 0) \wedge (11 \leq i \leq 20)$ を満たす $w_{k,i}$ の数 (= $n_{\text{pred}}^-(k)$)

を求めた。 $n_{\text{pred}}^+(k)$, $n_{\text{pred}}^-(k)$ によって、以下の妥当性の指標

$$AC = \frac{\sum_{k=1}^{100} (n_{\text{pred}}^+(k) + n_{\text{pred}}^-(k))}{2000},$$

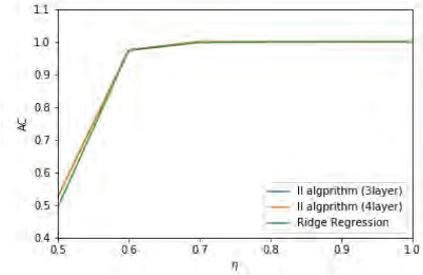


図 3: 極性伝搬の検証 (AC の平均値)

の試行回数 5 回での平均値を求め、妥当性を検証した。2.4 節, 2.5 節における理論解析の結果から、AC の値は η が大きくなれば 1.0 に近づくことが期待できる。

ここで、ニューラルネットワークモデルのパラメータ値 W_3 に初期値を II algorithm (3 層) を使う場合には

$$W_3[0][k] \sim U(-0.01, 0), W_3[1][k] \sim U(0, 0.01)$$

に従って与えた。また、II algorithm (4 層) を使う場合にはパラメータ値 W_3^* , W_4^* に初期値を

$$W_3^* \sim U(0, 0.01), W_4^*[0][k] \sim U(-0.01, 0),$$

$$W_4^*[1][k] \sim U(0, 0.01)$$

に従って与えた。

これは 2.4 節, 2.5 節を踏まえてのものである。II algorithm を用いた極性値付与手法の妥当性を確かめるため、リッジ回帰モデル [3] を用いた極性値付与手法による結果と比較する。ここで、リッジ回帰の L2 正則化のパラメータ値はグリッドサーチによって $\{0.1, 1.0, 10.0\}$ の範囲でチューニングして決めた。

3.2 実験結果

図 3 が実験結果である。II algorithm (3layer) が II algorithm (3 層), II algorithm (4layer) が II algorithm (4 層), Ridge Regression がリッジ回帰モデルをそれぞれ表す。 $\eta = 0.5, 0.6, 0.7, 0.8, 0.9$, の範囲で動かしたとき、期待通り、 η の値が大きいくほど AC の値は大きくなることを確認できた。この結果より、II algorithm による単語への極性値の付与が「 $p^+(w_{k,i})$ が大きい単語はポジティブ単語であり、 $p^-(w_{k,i})$ が大きい単語はネガティブ単語である」という観点のもとでは妥当であることを検証できた。

4. 実データを用いた極性伝搬の検証

実データを用いた検証によって、「II algorithm が付与する単語の極性値の有用性」と「II algorithm の予測モデルとしての有用性」を検証した。

4.1 データセット

有用性を調査するにあたり以下の実験をした。

4.1.1 株価動向分析

本実験においては 2013 年 1 月から 2015 年 12 月までの間に配信されたトムソンロイターの経済ニュース記事のうち、銘柄コードの入りの記事を用いた。これらの記事が個別銘柄の株価動向に与える影響を予測した。まず、ニュース記事の配信日を d , 記事に最初に銘柄コードが出てくる個別銘柄の時刻 t_p

の株価を p_{t_p} 、及び時刻 $t_a (> t_p)$ の株価 p_{t_a} (表 1 のパターン参照) から算出される株価リターンを用いて以下の y_{cls} で定めるタグを各記事につけた。

$$\Delta p(d) = \frac{p_{t_a} - p_{t_p}}{p_{t_p}}, y_{cls} = \begin{cases} 1 & (\Delta p(d) > 0.01) \\ -1 & (\Delta p(d) < -0.01) \end{cases}$$

その後、記事のタグの予測を行い、F 値の 10 交差検定平均スコアを算出し評価した。

表 1: 株価リターンの設定 (ロイターニュース)

パターン名	p_{t_p}	p_{t_a}
open_nowday_to_close_nextday	$d-1$ の始値	$d+1$ の始値
close_nowday_to_close_nextday	d の終値	$d+1$ の終値
close_nowday_to_open_nextday	d の終値	$d+1$ の始値
open_nextday_to_close_nextday	$d+1$ の始値	$d+1$ の終値
close_previousday_to_open_nextday	$d-1$ の終値	$d+1$ の始値
close_previousday_to_close_nextday	$d-1$ の終値	$d+1$ の終値

4.1.2 ヤフーファイナンス掲示板ボジネガタグ予測

Yahoo!Finance 掲示板*1 の投稿に付与されるタグの予測をおこなった。タグの種類は $\{1, 2, 3, 4, 5\}$ である。1 は「強く買いたい」、5 は「強く売りたい」を意味する。

以下の二つのケースについて実験を行った。

- 短期間における多種多様な銘柄の投稿に関する予測: 2014/9/1 ~ 2014/9/30 までの全銘柄の投稿について投稿日の古いタグ 1 とタグ 5 の投稿 10000 件ずつを抽出し、タグを予測し、その F 値の 5 交差検定平均スコアをもとに評価した。
- 長期間における少数銘柄の投稿に関する予測: 2014/11/18 ~ 2016/6/15 まで銘柄番号 998407, 9501, 4777, 8462, 4564, 6871 のタグが 1 または 5 の投稿を抽出し、2015/6/1 ~ 2016/5/31 について各月の投稿のタグを検証データ、各月の前月以前のタグが 1 と 5 の投稿を各 20000 件ずつを訓練用データとして予測精度を検証した。月毎に F 値を算出し、12 回の平均スコアをもとに評価した。

4.2 前処理

本実験において、形態素解析には MeCab [11] を用いた。

特徴量を生成するにあたり、実験 4.1.1 項の「株価動向分析」では動詞、形容詞、名詞を抽出 (全部で 21286 語) し、また、実験 4.1.2 項の [短期間における多種多様な銘柄の投稿に関する予測] においては動詞、形容詞、名詞を中心に単語 (全部で 25218 語) を抽出し、各単語の頻度から文書の特徴量を生成した。実験 4.1.2 項の [長期間における少数銘柄の投稿に関する予測] においては形態素を抽出した上で、極大部分文字列 [12] を抽出する手法により、文書全体にて 100 回以上出現する極大部分形態素列 16627 表現を抽出し、それらの頻度をもとに文書の特徴量を生成した。

4.3 ベースライン

4.3.1 II algorithm が付与する単語極性値の有用性

Label [13], SSL [4], Ridge Regression [3] の各既存手法及び、II algorithm (3 層) (II algorithm (layer3)), II algorithm (4 層) (II algorithm (layer4)) によって得られた各単語の極性

値 ($PS^*(word_l)$ を単語 $word_l$ とする) と CDR 法 (2.1 節) を用いて以下のように $W_{polaritybase}$ を生成し、 $V_{BOW} W_{polaritybase}$ の値を文書の特徴量とした。

$$W_{polaritybase} = (\delta_{ij}^*)^T$$

$$\delta_{ij}^* = \begin{cases} PS^*(word_l) & (word_l \in class_j) \\ 0 & \text{else} \end{cases}$$

これらの特徴量による予測結果に加え、CDR 法 (任意の l について $PS^*(word_l) = 1$ である場合に相当)、AntSyn [14] を用いた CDR 法 (AntSyn)、 V_{BOW} (BOW) により生成された特徴量による予測結果を比較した。予測モデルには線形 SVM を用いた。II algorithm の初期値付与に利用する単語極性値には金融機関に所属する機関投資家が人手で作成した経済用語極性辞書の値を利用した (約 200 単語)。II algorithm の精度を求めるにあたっては、5 回の試行平均値をもとに算出した。II algorithm (3 層・4 層) の 2 層目の次元を $K = 500$, II algorithm (4 層) の 3 層目の次元数、4.1.2 項における epoch 数 (学習回数) は訓練データ内におけるチューニングにより決定し、4.1.1 項では epoch 数 (学習回数) を 50 とした。

4.3.2 II algorithm の予測モデルとしての有用性

II algorithm を用いた場合の予測精度と多層パーセプトロンモデル [15] の 3 層モデル (MLP3)、及び 4 層モデル (MLP4) を用いた場合の予測精度の比較をおこなった。

4.4 実験結果

図 4 の上図が 4.3.1 項における「株価動向分析 (4.1.1)」の結果、下図が 4.3.2 項における「株価動向分析 (4.1.1 項)」の結果である。この結果からロイターニュースはパターン close_previousday_to_open_nextdayII.algorithm の株価リターンに最も影響を与えたとわかる。また、パターン close_previousday_to_open_nextdayII.algorithm の株価リターンの場合のより詳細な結果は表 2 の通りである。表 3 が実験 4.3.1 項における「ヤフーファイナンス掲示板ボジネガタグ予測 (4.1.2 項)」の結果、表 4 が 4.3.2 項における「ヤフーファイナンス掲示板ボジネガタグ予測 (4.1.2 項)」の結果である。既存手法との結果の比較により、II algorithm が与える単語の極性値から生成される特徴量から作られる予測モデルの方が他の手法から生成される特徴量により作られる予測モデルに比べて F 値を指標とした場合に高い予測力を持つことがわかった。また、II algorithm (3 層) が MLP3 に比べ、また II algorithm (3 層) が MLP 4 に比べそれぞれ F 値を指標とした場合に高い予測力を持つことがわかった。

本節では II algorithm を用いたマルチタイプ極性概念辞書の作成手法の説明、及びマルチタイプ極性概念辞書の出力結果の紹介を行う。

4.5 極性辞書マージ手法紹介

4.5.1 手法概要

ヤフー辞書内の単語クラスとロイター辞書内の単語クラスについて、クラス単位でのマージを行い、マージ後のクラスが似た意味の単語群で構成されることを目指す。

4.5.2 マージ手順

具体的なマージ手法の手順を以下に記す。

- ロイター辞書、ヤフー辞書のそれぞれの各クラスターのセントロイドを求め、 \cos 類似度をもとに類似度行列 $W_{ruiter} \in \mathbb{R}^{K_{ruiter} \times K_{ruiter}}$ と $W_{yahoo} \in \mathbb{R}^{K_{yahoo} \times K_{yahoo}}$ をそれぞれ作成 ($K_{ruiter} = 500, K_{yahoo} = 1000$) する。ここで、

*1 <http://textream.yahoo.co.jp/category/1834773>

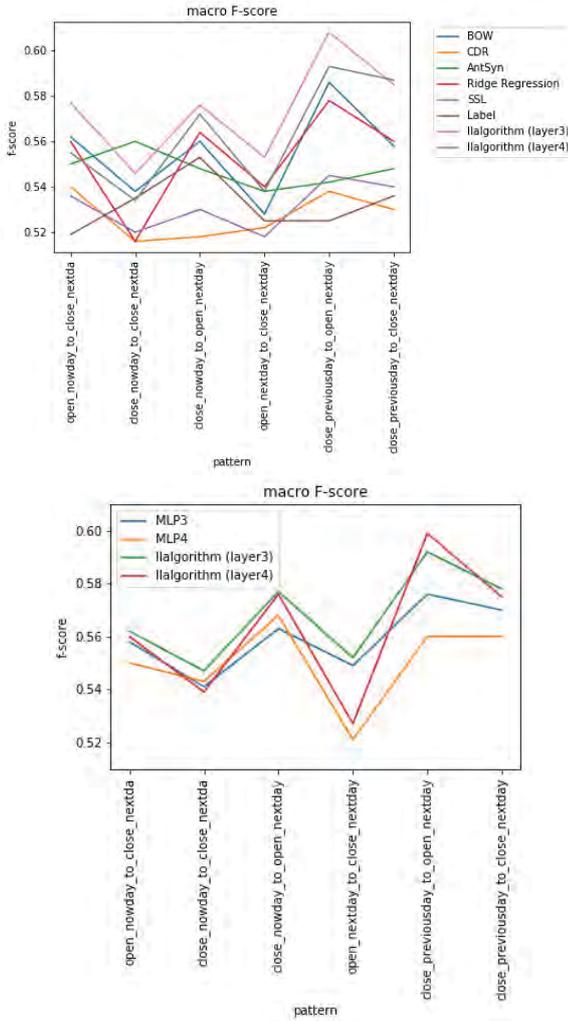


図 4: 株価動向分析結果 (F 値)

表 2: パターン close_previousday_to_open_nextday の結果

Methods	F 値 ($\pm\sigma$)
BOW	0.586
CDR	0.538
AntSyn	0.542
Ridge Regression	0.578
SSL	0.545
Label	0.525
II algorithm (layer3)	0.608 (± 0.005)
II algorithm (layer4)	0.593 (± 0.004)

K_{ruiter} , K_{yahoo} はそれぞれロイター辞書, ヤフー辞書におけるクラスタ数である.

- 類似度行列 2 つを対角に並べて W_{merge} を作成

$$W_{merge} = \begin{pmatrix} W_{yahoo} & 0 \\ 0 & W_{ruiter} \end{pmatrix}$$

表 3: ヤフーデータ掲示板ボジネガタグ予測

Methods	F 値 ($\pm\sigma$) (短期)	F 値 ($\pm\sigma$) (長期)
BOW	0.735	0.790
CDR	0.684	0.724
Ridge Regression	0.754	0.786
AntSyn	0.684	0.700
SSL	0.570	0.660
Label	0.633	0.651
II algorithm (layer3)	0.763 (± 0.001)	0.795 (± 0.001)
II algorithm (layer4)	0.760 (± 0.001)	0.791 (± 0.001)

表 4: ヤフーデータ掲示板ボジネガタグ予測

Methods	F 値 ($\pm\sigma$) (短期)	F 値 ($\pm\sigma$) (長期)
MLP3	0.742 (± 0.004)	0.774 (± 0.002)
MLP4	0.743 (± 0.004)	0.772 (± 0.002)
II algorithm (layer3)	0.775 (± 0.001)	0.800 (± 0.001)
II algorithm (layer4)	0.771 (± 0.001)	0.792 (± 0.001)

- ロイター辞書内の単語とヤフー辞書内の単語と人手極性辞書の共通単語群 $\{w_{com}[0], w_{com}[1], \dots, w_{com}[n-1]\}$ を利用して Algorithm 1 のようにロイター単語クラスとヤフー単語クラスをつなげる.

Algorithm 1 エッジ挿入

Input: W_{merge}

- for $i \leftarrow 0$ to $n-1$ do
- $W_{merge}[yclass(w_{com}), rclass(w_{com}) + K_{yahoo}] \leftarrow 1$;
- $W_{merge}[rclass(w_{com}) + K_{yahoo}, yclass(w_{com})] \leftarrow 1$;
- end for
- W_{merge} を返す

$yclass(w_{com})$ は単語 w_{com} のヤフー辞書におけるクラス番号であり, $rclass(w_{com})$ は単語 w_{com} のロイター辞書におけるクラス番号である.

- SVD [16] により欠損値の補完を行う.
- W_{merge} の成分のうち, 閾値 $T = 0.9$ より小さい値は 0 にし, それ以外は 1 にする.
- W_{merge} について SpectralClustering [17] を用いてマージを行う.

4.6 数値実験

4.6.1 実験設定

まず, II algorithm (3 層) によりヤフー辞書, ロイター辞書の作成をおこなった. その後, 4.5 節の方法でヤフーファイナンス掲示板から作成した極性辞書とロイターニュースから作った極性辞書のマージを行った. マージ後のクラスタ数は 100 とした. ロイター辞書にはパターン close_previousday_to_open_nextday の株価リターンとニュース記事の対応から IIalgorithm (layer3) を用いて作成した極性辞書を用いた. ヤフー性辞書には「ヤ

フーデータ掲示板ポジネガタグ予測 (短期) から II algorithm (layer3) を用いて作成した極性辞書を用いた。

4.6.2 極性概念辞書出力結果

次に、実験 4.6.1 によって作成された極性概念辞書の一部を紹介する。

表 5-6 は辞書のマージ後に「上昇」と同じクラスの単語とその極性値を、マージ前のロイター辞書の単語クラス、ヤフー辞書の単語クラス別に表示した結果の一部である。例えば、表 5 から「上昇」のプラスの極性値が II algorithm によって同じクラスの「急騰」や「急上昇」に伝搬し、逆に「下落」や「下降」のマイナスの極性値が II algorithm によって同じクラスの「急降下」や「暴落」に伝搬している様子がわかる。

5. まとめ

本論文において、提案手法の性質を理論的に解析すると共に、提案手法が市場動向分析において他の既存手法に比べ有用であることを実験的に示した。また、ロイターニュースとヤフーファイナンス掲示板の両方を同じ指標で分析可能な極性概念辞書を作成した。今後の課題として、II algorithm (4 層) を利用したフレーズ単位での極性情報抽出手法の考案、極性辞書の情報が少ない場合や少数のラベルつきデータしか手に入れない状況でも安定的で高い予測精度を出すための改良法の開発、マルチタイプ極性概念辞書を用いた複数種類のテキストデータの分析方法の開発などが考えられる。

参考文献

- [1] A. K. Nassirtoussi, S. Aghabozorgi, T. Y. Wah, D. C. L. Ngo, "Text mining for market prediction: A systematic review", Journal of Expert Systems with Applications, Vol. 41, Issue 16, pp. 7653-7670, 2014
- [2] W. Ye and F. Ren, "Learning sentimental influence in twitter", ICFCSEA, 2011
- [3] K. Tsubouchi and T. Yamashita, "Positive / Negative Detection for Finance Contents via Stock Bulletin Boards Data", JSAI 2014, 2014
- [4] H. Yanagimoto, "Improvement of Sentiment Dictionary Using Neural Network Language Model", JSAI 2014, 2014
- [5] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, J. Dean, "Distributed Representations of Words and Phrases and their Compositionality", NIPS 2013, pp. 3111-3119, 2013
- [6] Y. Yuan, L. He, L. Peng, Z. Huang, "A New Study Based on Word2vec and Cluster for Document Categorization", Journal of Computational Information Systems, Vol. 10, Issue 21, pp. 9301-9308, 2014
- [7] K. Hornik, I. Feinerer, M. Kober, C. Buchta, "Spherical k-Means Clustering", Journal of Statistical Software, Vol. 50, Issue 10, pp. 1-22, 2012
- [8] T. Ito, K. Izumi, K. Tsubouchi, T. Yamashita, "Polarity propagation of financial terms for market trend analyses using news articles", CEC 2016, 2016

表 5: クラス 551 (ヤフー辞書)

単語	極性値 (伝搬後)	(極性辞書値)
上昇	0.328	(0.5)
反発	0.232	(0.5)
動意	0.297	(0.0)
動き	0.276	(0.0)
急落	-0.525	(0.0)
反転	0.498	(0.0)
下落	-0.655	(-0.333)
下降	-0.818	(-0.333)
急騰	0.184	
下げ	-0.307	
短期間	0.007	
急上昇	0.278	
反騰	-0.015	
上下	0.137	
値動き	0.095	
急降下	-0.411	
上げ	0.017	
調整	0.366	
盛り上がり	0.096	
乱高下	0.081	
上下動	0.333	
上がり	0.208	
リバウンド	-0.240	
上げ過ぎる	-0.160	
暴落	-0.791	
暴騰	0.245	
連騰	0.240	
相場	0.083	
上げ下げ	0.245	

表 6: クラス 304 (ロイター辞書)

単語	極性値 (伝搬後)	(極性辞書値)
続伸	1.357	(1.118)
急伸	0.678	(0.5)
反発	0.594	(0.5)
上昇	0.443	(0.5)
急落	-0.070	(0.0)
推移	0.167	(0.0)
下落	-0.453	(-0.333)
低下	-0.384	(-0.375)
続落	-1.229	(-0.929)
反落	-1.250	(-0.938)
急騰	-0.088	
下げ	-0.120	
下押し	-0.096	
下押す	0.118	
鬼門	0.072	
下がる	-0.052	
伸び悩む	-0.034	
軟化	-0.171	
値上がり	0.025	
値下がり	0.060	
安	0.069	
買われる	0.060	
乱高下	0.261	
安い	0.155	
下げる	-0.178	
急上昇	0.089	
上げ	-0.160	
下げ止まる	0.200	
ドル安	-0.033	
伸し	0.082	

- [9] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting", Journal of Machine Learning Research, Vol. 15, pp. 1929–1958, 2014
- [10] D. P. Kingma, J. L. Ba, "Adam: A Method for Stochastic Optimization", arXiv:1412.6980, 2014
- [11] T. Kudo, K. Yamamoto, Y. Matsumoto, "Applying Conditional Random Fields to Japanese Morphological Analysis", EMNLP 2004, pp. 230–237, 2004
- [12] D. Okanohara, J. Tsujii, "The Categorization with all Substring Features", SDM 2009, pp. 838–846, 2010
- [13] D. Zhou, O. Bousquet, T. N. Lal, J. Weston, B. Scholkopf, "Learning with Local and Global Consistency", NIPS 2003, pp. 321–328, 2003
- [14] K. A. Nguyen, S. S. im Walde, N. T. Vu, Integrating Distributional Lexical Contrast into Word Embeddings for Antonym-Synonym Distinction, ACL 2016, 2016
- [15] 岡谷貴之, "機械学習プロフェッショナルシリーズ 深層学習", pp12–14, 講談社, 2015
- [16] K. Yu, S. Zhu, J. Lafferty, Y. Gong, "Fast Nonparametric Matrix Factorization for Large-scale Collaborative Filtering", SIGIR 2009, 2009
- [17] U. V. Luxburg, "A tutorial on spectral clustering", Journal of Statistics and Computing, Vol. 17, Issue 4, pp. 395–416, 2007

付録 A II algorithm (3層) 理論解析

本節では、提案手法である II algorithm (3層) によってどのようにして単語に極性値が付与されるのかについて理論的な側面から説明し、II algorithm によって妥当に各単語に極性値が付与されることを説明する。

A1 定義

まず、出力値 y , ミニバッチ内文書番号 j の文書の目的値 d_j について以下のように定義する。

$\text{freq}(w_{k,i})$: 単語 $w_{k,i}$ が文書に出てくる頻度確率

$$\mathbf{y} := (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N)$$

$$= \text{softmax}(\mathbf{W}_3(\tanh(\mathbf{W}_{\text{polarity}} \mathbf{X}) + \mathbf{b}_0)) \in \mathbb{R}^{2 \times N}$$

$$\begin{cases} d_j = (0, 1)^T & (\text{文書番号}j\text{の文書がポジティブの場合}) \\ d_j = (1, 0)^T & (\text{文書番号}j\text{の文書がネガティブの場合}) \end{cases}$$

また、簡単のため、 $\mathbf{Z}_{j,k}^{(1)}[i] \in \{0, 1\}$ であると仮定する。

A2 パラメータの更新

$\mathbf{W}^{(3)}, \mathbf{W}_k^{(2)}$ ($k = 1, 2, \dots, K$) は次のように更新される。

$$\Delta_j^{(3)} := \mathbf{y}_j - d_j \quad (7)$$

$$= \begin{cases} (|\Delta_j^{(3)}[0]|, -|\Delta_j^{(3)}[1]|)^T & (d_j = (0, 1)^T) \\ (-|\Delta_j^{(3)}[0]|, |\Delta_j^{(3)}[1]|)^T & (d_j = (1, 0)^T) \end{cases} \quad (8)$$

$$\Delta^{(2)} := (1 - (\tanh(\mathbf{U}^{(2)}))^2) \odot (\mathbf{W}_3[1]^T \Delta^{(3)}[1] + \mathbf{W}_3[0]^T \Delta^{(3)}[0]) \quad (9)$$

$$\partial \mathbf{W}^{(3)} = \frac{1}{N} \Delta^{(3)} \mathbf{Z}^{(2)T}$$

$$\partial \mathbf{W}_k^{(2)} = \frac{1}{N} \sum_{j=1}^N \Delta_j^{(2)} [k] \mathbf{Z}_{j,k}^{(1)}$$

$$\mathbf{W}^{(3)} \leftarrow \mathbf{W}^{(3)} - \epsilon \partial \mathbf{W}^{(3)}, \mathbf{W}_k^{(2)} \leftarrow \mathbf{W}_k^{(2)} - \epsilon \partial \mathbf{W}_k^{(2)} \quad (10)$$

\odot はアダマール積(要素積)を指し、 $\epsilon (> 0)$ は更新時に Adam [10] によって計算される。

また、 $E_{j+} [|\Delta_{j+}^{(2)}[k]| \mathbf{Z}_{j+,k}^{(1)}[i] = 1]$ を $E_{j+} [|\delta_{j+}^{(2)}[k]|]$, $E_{j-} [|\Delta_{k,j-}^{(2)} [|\mathbf{Z}_{k,j-}^{(1)}[i] = 1]|]$ を $E_{j-} [|\delta_{j-}^{(2)}[k]|]$ と表記することにする。

A3 極性伝搬条件

このとき、2.4節で述べた主張が成り立つ。これは、2.3節における p^{t+}, p^{t-} がそれぞれ十分に大きい場合に、ポジティブ単語 $w_{k,i+}$ にはプラスの極性値が与えられ、ネガティブ単語 $w_{k,i-}$ にはマイナスの極性値が与えられることを保証するものである。

この主張が成り立つ理由について以下の A4, A5 にて説明する。まず、 $\mathbf{W}_3[0][k] < 0, \mathbf{W}_3[1][k] > 0$ の場合に十分大きい p^{t+}, p^{t-} が与えられたときに

$$E[\partial \mathbf{W}_k^{(2)}[i^+]] < 0, E[\partial \mathbf{W}_k^{(2)}[i^-]] > 0$$

が成り立つことを示す。次に条件 (i)–(iv) を満たすように初期値が与えられた場合に更新後も $\mathbf{W}_3[0][k] < 0, \mathbf{W}_3[1][k] > 0$ であることが維持されることを示す。これら二つを示すことによって、条件 (i)–(iv) を満たすように初期値が与えられるという理想的な場合にはポジティブ単語 $w_{k,i+}$ にはプラスの極性値が与えられ、ネガティブ単語 $w_{k,i-}$ にはマイナスの極性値が与えられることを説明できる。

A4 $\mathbf{W}_3[0][k] < 0, \mathbf{W}_3[1][k] > 0$ の場合の解析

$$\Delta^{(2)} = (1 - (\tanh(\mathbf{U}^{(2)}))^2) \odot (\mathbf{W}_3[1]^T \Delta^{(3)}[1] + \mathbf{W}_3[0]^T \Delta^{(3)}[0])$$

について

$$\Delta_{j+}^{(2)}[k] < 0, \Delta_{j-}^{(2)}[k] > 0 \quad (11)$$

であることを示す。(8)より

$$\begin{aligned} & (\mathbf{W}_3[0][k]^T \Delta^{(3)}[0])_j \\ &= \begin{cases} \mathbf{W}_3[0][k]^T |\Delta_j^{(3)}[0]| < 0 & (d_j = (0, 1)^T) \\ -\mathbf{W}_3[0][k]^T |\Delta_j^{(3)}[0]| > 0 & (d_j = (1, 0)^T) \end{cases} \quad (12) \end{aligned}$$

である。同様に、

$$\begin{cases} (\mathbf{W}_3[1][k]^T \Delta^{(3)}[1])_j < 0 & (d_j = (0, 1)^T) \\ (\mathbf{W}_3[1][k]^T \Delta^{(3)}[1])_j > 0 & (d_j = (1, 0)^T) \end{cases} \quad (13)$$

である。ここで、 $(1 - (\tanh(\mathbf{U}^{(2)}))^2) > 0$ なので、(9)(12)(13)より、(11)が言える。このとき、

$$E[\partial \mathbf{W}_k^{(2)}[i]] = E \left[\frac{1}{N} \sum_{j=1}^N \Delta_j^{(2)} [k] \mathbf{Z}_{j,k}^{(1)}[i] \right] = E_j [\Delta_j^{(2)} [k] \mathbf{Z}_{j,k}^{(1)}[i]]$$

$$= \text{freq}(w_{k,i})(-p^+(w_{k,i})E_{j+}[\delta_{j+}^{(2)}[k]] + p^-(w_{k,i})E_{j-}[\delta_{j-}^{(2)}[k]])$$

$$\therefore \begin{cases} E[\partial \mathbf{W}_k^{(2)}[i]] < 0 & \left(\frac{p^+(w_{k,i})}{p^-(w_{k,i})} > \frac{E_{j-}[\delta_{j-}^{(2)}[k]]}{E_{j+}[\delta_{j+}^{(2)}[k]]} \right) \\ E[\partial \mathbf{W}_k^{(2)}[i]] > 0 & \left(\frac{p^+(w_{k,i})}{p^-(w_{k,i})} < \frac{E_{j-}[\delta_{j-}^{(2)}[k]]}{E_{j+}[\delta_{j+}^{(2)}[k]]} \right) \end{cases} \quad (14)$$

となる．すなわち， p^{++}, p^{+-} がそれぞれ十分に大きい場合，ポジティブ単語 $w_{k,i+}$ には正の極性値，ネガティブ単語 $w_{k,i-}$ には負の極性値が与えられるようパラメータは更新される．

次に，任意の単語の頻度 $\text{freq}(w_{k,i})$ が一定値であり， p^{++}, p^{+-} が十分に大きいという場合にポジティブ単語 $w_{k,i+}$ ，ネガティブ単語 $w_{k,i-}$ に与えられる極性値の絶対値の大きさがニュートラル単語 $w_{k,i^{neu}}$ に比べて大きくなることを示す．

$$E[\partial \mathbf{W}_k^{(2)}[i]] = E \left[\left| \frac{1}{N} \sum_{j=1}^N \Delta_j[k] \mathbf{Z}_{j,k}^{(1)}[i] \right| \right]$$

$$= E[|\Delta_j^{(2)}[k] \mathbf{Z}_{j,k}^{(1)}[i]|]$$

$$= \text{freq}(w_{k,i}) - p^+(w_{k,i})E_{j+}[\delta_{j+}^{(2)}[k]] + p^-(w_{k,i})E_{j-}[\delta_{j-}^{(2)}[k]]$$

$$= \text{freq}(w_{k,i})|p^-(w_{k,i})2E_{j-}[\delta_{j-}^{(2)}[k]] - E_{j+}[\delta_{j+}^{(2)}[k]]|$$

故に，任意の単語の頻度 $\text{freq}(w_{k,i})$ が一定値の場合，

$$T_{pt} := \frac{E_{j+}[\Delta_{j+}^{(2)}[k]|\mathbf{Z}_{k,j+}^{(1)}[i] = 1]}{2E_{j-}[\Delta_{j-}^{(2)}[k]|\mathbf{Z}_{k,j-}^{(1)}[i] = 1]}$$

とすると

$$\begin{cases} E[|\partial \mathbf{W}_k^{(2)}[i^+]|] > E[|\partial \mathbf{W}_k^{(2)}[i^{neu}]|] & (pt - < T_{pt}) \\ E[|\partial \mathbf{W}_k^{(2)}[i^-]|] > E[|\partial \mathbf{W}_k^{(2)}[i^{neu}]|] & (pt + > T_{pt}) \end{cases} \quad (15)$$

が成り立つ．

A5 中間層と出力層の間のエッジの解析

$$E_{j+}[\Delta_{j+}^{(3)}[0]|\mathbf{Z}_{j+}^{(2)}[k]]$$

$$= E_{j+}[\Delta_{j+}^{(3)}[0]|\tanh(\sum_{i=1}^{n(k)} \mathbf{Z}_{j+,k}^{(1)}[i] \mathbf{W}_k^{(2)}[i])]$$

$$\sum_{i=1}^{n(k)} \mathbf{Z}_{j+,k}^{(1)}[i] \mathbf{W}_k^{(2)}[i] = \sum_{i^+=1}^{n^+(k)} \mathbf{Z}_{j+,k}^{(1)}[i^+] \mathbf{W}_k^{(2)}[i^+]$$

$$+ \sum_{i^{neu}=1}^{n^{neu}(k)} \mathbf{Z}_{j+,k}^{(1)}[i^{neu}] \mathbf{W}_k^{(2)}[i^{neu}]$$

$$+ \sum_{i^-=1}^{n^-(k)} \mathbf{Z}_{j+,k}^{(1)}[i^-] \mathbf{W}_k^{(2)}[i^-]$$

である．ここで，

$$E[\mathbf{Z}_{j+,k}^{(1)}[i]] = \text{freq}(w_{k,i})p^+(w_{k,i}) = \text{freq}(w_{k,i})(1 - p^-(w_{k,i}))$$

なので， $\mathbf{W}_k^{(2)}[i^+] > 0$ ， $\mathbf{W}_k^{(2)}[i^+] \gg \mathbf{W}_k^{(2)}[i^{neu}]$ であり，かつ $p^-(w_{k,i-})$ ， $p^+(w_{k,i+})$ ， $n^+(k)$ の値が十分に大きいとき，

$$E_{j+}[\Delta_{j+}^{(3)}[0]|\mathbf{Z}_{j+}^{(2)}[k]] > 0$$

となる．同様に， $\mathbf{W}_k^{(2)}[i^-] < 0$ ， $\mathbf{W}_k^{(2)}[i^-] \ll \mathbf{W}_k^{(2)}[i^{neu}]$ であり， $p^-(w_{k,i-})$ ， $p^+(w_{k,i+})$ ， $n^-(k)$ の値が十分に大きいとき，

$$E_{j-}[\Delta_{j-}^{(3)}[0]|\mathbf{Z}_{j-}^{(2)}[k]] > 0$$

よって，(i)(ii) が任意の i^+ ， i^{neu} ， i^- について成り立つならば，式 (12) より，

$$E[\partial \mathbf{W}_3[0][k]] = E \left[\frac{1}{N} \cdot \sum_{j=1}^N \Delta_j^{(3)}[0] \mathbf{Z}_j^{(2)}[k] \right] > 0$$

となる．同様に，(i)(ii) が任意の i^+ ， i^{neu} ， i^- について成り立つならば，

$$E[\partial \mathbf{W}_3[1][k]] < 0$$

となる．よって， $\mathbf{W}_3[1][k] > 0$ が十分に大きく， $\mathbf{W}_3[0][k] < 0$ が十分に小さいならば，

$$E[\mathbf{W}_3[1][k] - \epsilon \partial \mathbf{W}_3[1][k]] > 0, E[\mathbf{W}_3[0][k] - \epsilon \partial \mathbf{W}_3[0][k]] < 0.$$

が成り立つ．ここで， u ， σ^2 を

$$u := E[\Delta_j^{(3)}[0] \mathbf{Z}_j^{(2)}[k]] (> 0), \sigma^2 := \text{Var}[\Delta_j^{(3)}[0] \mathbf{Z}_j^{(2)}[k]].$$

と定義すると，大数の法則より，

$$P(\epsilon \partial \mathbf{W}_3[0][k] \leq -\mathbf{W}_3[0][k]) \leq \frac{\epsilon^2 \sigma^2}{2N \cdot (eu + \mathbf{W}_3[0][k])^2} \left(\xrightarrow{N \rightarrow \infty} 0 \right)$$

となる．よって，学習時のミニバッチサイズ N が十分に大きく， $\mathbf{W}_3[0][k] (< 0)$ が十分に小さいならば，

$$\mathbf{W}_3[0][k] < 0 \quad (16)$$

は更新後も維持される．

同様に，学習時のミニバッチサイズ N が十分に大きく， $\mathbf{W}_3[1][k] (> 0)$ が十分に大きいならば，

$$\mathbf{W}_3[1][k] > 0 \quad (17)$$

は更新後も維持される．故に，(i)-(iv) が任意の i^+ ， i^{neu} ， i^- について満たされるならば，

$$\mathbf{W}_3[0][k] < 0, \mathbf{W}_3[1][k] > 0$$

は更新後も維持される．

A4 における式 (14) 及び A5 における解析結果から，2.4 節で主張したことが説明でき，II algorithm がボジネガ付与の手法として妥当であることを保証できた．