

Hidden CRFs を用いた極性反転構造のモデル化

貞光 九月 山本 幹雄

筑波大学 システム情報工学研究科, {sadamitsu@mibel.cs,myama@cs}.tsukuba.ac.jp

1 はじめに

近年インターネット上に膨大なテキストデータが蓄積されるようになり、それと同時にテキストの中に含まれる情報を分析し、有効に活用することへの要求が高まっている。ある対象に対する評価を含む文書(評価文書)を、肯定評価・否定評価の2値ラベルに分類する評価文書分類(Pang, Lee, and Vaithyanathan 2002)(乾 奥村 2006)は、その対象に対する評価を定量的に提示できるという点で有益であり、広く一般に用いられている。評価文書分類は単純な機械学習の適用において既に高い性能を示しているが(Pang et al. 2002)、極性が部分的に反転する箇所の扱いは難しく、ヒューリスティクスによって解決されることが多かった。本稿ではHidden Conditional Random Fields(Hidden CRFs)を用いて、従来の単語単位のモデルでは捉えられなかった極性反転の構造を単語または文のシーケンス構造の中に捉え、評価文書分類の精度向上を図る。提案手法は単語情報のみを用いたSVMsと同等の性能を示した。

2 Hidden CRFs を用いた極性反転構造のモデル化

2.1 極性反転構造

評価文書において、部分的な極性の反転が生じているものは少なくない。具体例として、以下にAmazon.com^{*1}のレビューデータを示す。

An example of review

I like this razor very much. It gives a close shave without damaging the skin. However, the only drawback is that you must constantly be holding in the “On” button. My last razor had a button that when it was switched on, it stayed on.

このレビューは対象商品である“razor”についての強い肯定の極性を持っているが、部分的に否定的意見も含まれている。我々はこのような部分的な極性反転の現象をシーケンシャルモデルを用いてモデル化する。モデル化の方法としては、単語による遷移モデルと文による遷移

モデルの2つのモデルを提案する。単語遷移モデルでは大域的な極性の遷移と局所的な極性の遷移といった細かいモデル化が可能である一方、文遷移モデルでは文中の極性の一貫性を仮定するため、単語遷移モデルよりも遷移の揺れが少ないロバストな推定が期待される。

2.2 Hidden CRFs の概要

隠れマルコフモデル(HMMs)はシーケンシャルなデータを生成するモデルの一種であり、HMMsを識別モデルとしての発展させたものがConditional Random Fields(CRFs)(Lafferty, McCallum, and Pereira 2001)である。CRFsは以下のような線形対数モデルとなっており、HMMsに比べ素性設計の自由度が高い。

$$P_{CRF}(\mathbf{q}|d; \boldsymbol{\lambda}) = \frac{\exp\{\boldsymbol{\lambda} \cdot f(\mathbf{q}, d)\}}{\sum_{\mathbf{q}} \exp\{\boldsymbol{\lambda} \cdot f(\mathbf{q}, d)\}} \quad (1)$$

ここで d は対象とする文書、 \mathbf{q} は隠れラベルの系列、 $f(\mathbf{q}, d)$ は素性関数、 $\boldsymbol{\lambda}$ は各素性に対応するパラメータベクトルを表す。CRFsは系列ラベルに対する識別モデルであるが、本稿での目的は各文の系列ラベル \mathbf{q} の推定ではなく、文書全体のラベル推定を行うことである。系列を全て考慮した上での文書ラベルへの条件付き確率を付与する識別モデルは、Hidden CRFsとして提案されている(Quattoni, Collins, and Darrell 2004)(Gunawardana, Mahajan, Acero, and Platt 2005)。Hidden CRFsは以下のように定義される。

$$P_{HCRF}(\phi|d; \boldsymbol{\lambda}) = \frac{\sum_{\mathbf{q}} \exp\{\boldsymbol{\lambda} \cdot f(\phi, \mathbf{q}, d)\}}{\sum_{\phi'} \sum_{\mathbf{q}} \exp\{\boldsymbol{\lambda} \cdot f(\phi', \mathbf{q}, d)\}} \quad (2)$$

ここで ϕ は肯定/否定を表す2値ラベルである。Hidden CRFsは、式1に示した各ラベル間におけるCRFsの分母(全系列の和)の相対比をとった形となっている。Hidden CRFsとHMMsとの違いは、条件付確率 $p(\phi|d; \boldsymbol{\lambda})$ を直接推定する点にあり、識別の観点においてHMMsより優れているといえる。

2.3 単語遷移モデルへのHidden CRFsの適用

はじめに最も単純なHidden CRFsの適用として、1単語毎に状態が遷移する単語遷移モデルを考える。状態数や各素性の重みを変化させることによってモデル化の方法は無数に考えられるが、本稿では主に図1に示した2通りの単語遷移モデルを用いる。図1左側は肯定(Posi.)・否定(Nega.)状態と反転状態(Rev.)を

^{*1} <http://www.amazon.com>

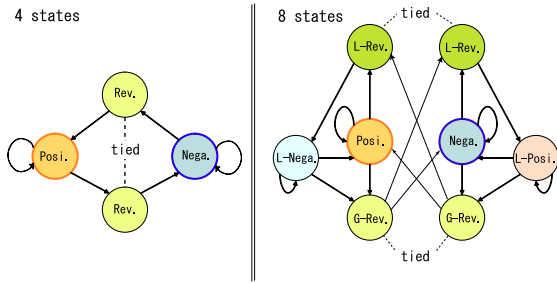


図1 Hidden CRFsによる単語遷移モデル(4状態および8状態)

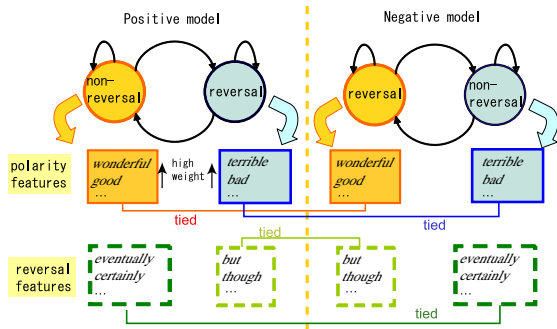


図2 Hidden CRFsによる文遷移モデル

2つ持った4状態モデル、右側は状態をさらに細かく分けて考えた8状態モデルである。8状態モデルでは、notやno等によって生じるの局所的な反転(L-Rev.)と、butやalthough等によって生じるの大域的な反転(G-Rev.)の2種類の反転状態があると仮定し、それに伴う極性状態にも大域的な状態(Posi.,Nega.)と局所的な状態(L-Posi.,L-Nega.)があると仮定している。図中では4状態、8状態それぞれ1つのモデルのみを示しているが、実際にはPositiveモデル、Negativeモデルの2つのモデルを学習する。素性は単語出現重みと、遷移重みの2種のみを考慮する。また反転状態に関しては、前後の状態に依存せず同じ単語が現れやすいと考え、単語出現重みの一部にtiedな制約を設けた。具体的には4状態におけるRev.、8状態におけるG-Rev.,L-Rev.は、それぞれPositive,Negative側のモデルにおいて、共通な単語出現重みを持つと仮定した。図中ではいずれも非エルゴディック型のモデルを示しているが、実験ではエルゴディック型との比較を行った。

2.4 文遷移モデルへのHidden CRFsの適用

文遷移モデルでは文毎に状態が決まり、文中の単語出現はBag-of-Wordsに従うと仮定する。文遷移モデルは単純に肯定状態・否定状態の2状態のみを仮定する(Sadamitsu, Sekine, and Yamamoto 2008)。局所的反

転と大域的な反転を区別しないことから、意味的には単語遷移モデルの4状態モデルに近いモデルになっている。単語遷移型と大きく異なる点は、1つの単語が極性を示す役割と、反転のしやすさに影響を与える役割の二重の役割を持っていると仮定している点である。ただしこれらを制約なしに同等に扱ってしまうと、パラメータが単純に倍増することになるため、Hidden CRFsのモデル化は有効に行えない。そこで、Positive/Negativeモデル間のパラメータを部分的にtiedとする制約を設けた上で各素性を扱う。図2はこれを図示したものであり、以下のように定式化される。

$$P_{HCRF}(\phi|d; \lambda) = \frac{1}{Z} \sum_{\mathbf{q}} \prod_t \exp\{\lambda_{q_t-1, q_t, \phi}^{trans} + \sum_{w \in \mathbf{s}_t} (\lambda_{w, \gamma(\phi, q_t)}^{pn} + \lambda_{w, q_t}^{rev})\} \quad (3)$$

$\lambda_{trans}, \lambda_{pn}, \lambda_{rev}$ はそれぞれ、遷移素性、極性素性としての単語出現、反転素性としての単語出現に対する重みパラメータである。 $\gamma(\phi, q_t)$ は文書ラベルが ϕ の時に、現状態 q_t が非反転の時に ϕ を、反転の時に $\bar{\phi}$ を返す関数である。反転素性のtied構造を意味的に解釈すれば、「反転を起こしやすい単語はPositive/Negativeモデルに関係なく、同様に反転を生じさせやすい」といえる。遷移素性についても同様にtiedにすることは可能であるが、本稿では制約を持たせないままに学習することとした。なお3.2節では、実験において学習された両極性モデルの遷移パラメータが、それぞれ傾向の異なる値を示したことについて述べる。

2.5 損失関数の設定

Hidden CRFsの学習は極性ラベルの条件付確率が大きくなる方向へ進んでいくが、必ずしも条件付確率の最大化が分類精度の向上に繋がるとは限らない。そのため本節では、Hidden CRFsに損失関数を導入する。基本的には(Suzuki, McDermott, and Isozaki 2006)のCRFsに対して損失関数を適用する手法と同様に、分類誤り尺度 D とsigmoid関数による損失関数 F を定義する。

$$F(d; \lambda) = \frac{1}{1 + \exp(-D(d; \lambda))} \quad (4)$$

$$D(d; \lambda) = -\log p(\phi|d; \lambda) + \log p(\bar{\phi}|d; \lambda)$$

$\phi, \bar{\phi}$ はそれぞれ正解ラベルと誤りラベルを、 λ はモデルパラメータ、 $D(d; \lambda)$ は文書 d の分類誤り尺度を示す。パラメータの推定にはGeneral Probabilistic Descent(GPD)(Juang and Katagiri 1992)を用いる。

2.6 関連研究

文書構造のモデル化にHMMsを利用した関連研究として、Barzilay et al. が提案した文構造のHMMsに

よるモデル化が挙げられる (Barzilay and Lee 2004)。Barzilay らの目的はトピックの傾向を捉えることにあり、一方我々の目的は感情極性の傾向を捉えることにあり、かつ識別モデルを用いている点で異なる。Mao et al. は CRFs を用いた局所的な感情の遷移を予測する手法を提案している (Mao and Lebanon 2007)。先に述べたように、CRFs は局所的な感情をラベル付けしたトレーニングデータが必要となる。これに対し我々が用いる Hidden CRFs は教師なしのシーケンシャルデータからの学習が可能で、人手コストを削減することができる。我々の研究にもっとも近いものとしては、McDonald et al. (McDonald, Hannan, Neylon, Wells, and Reynar 2007) と Ikeda et al. (Ikeda, Takamura, Ratinov, and Okumura 2008) のものが挙げられる。McDonald らは我々と同様にシーケンシャルな素性の設計をして文書の極性推定を行っているが、学習データにおける各々の文自体にタグが付与されている必要があり、文毎のタグが必要ない Hidden CRFs への自然な拡張が求められる。一方 Ikeda らは単語レベル、文レベルの両方から極性反転構造に焦点を当てた研究を行っている点において、我々の研究に近い。評価文書ではなく評価文の分類を研究対象としているため単純な比較はできないが、前後の文脈を用いる必要がないためにシーケンシャルな情報は用いておらず、あらかじめ得られた極性表現辞書以外の単語について扱うことを考慮していない。これに対し我々は、全ての単語を反転作用の可能性のある素性としてみなしている点で異なる。

3 実験

3.1 実験条件

評価実験に際し Amazon.com からレビューデータを取得した。Amazon.com のレビューには評点がレビュアーによって既に付与されており、実験では 17 カテゴリ*2のそれぞれにおいて、各評点 200 レビューずつ、計 13,600 レビューを取得した。評点 5,4 のレビューを Positive レビュー、評点 1,2 のレビューを Negative レビューとし、10 交差検定を行った。素性には document frequency が 20 回以上の 1gram 単語、計 4,051 単語を用い、文遷移モデルではその文に単語が含まれているかどうか素性とした。レビューのタイトル、及びレビュアー名はレビューデータに含めていない。学習を行う際には過適応を緩和するために、全データについての更新 1 イテレーションを終える毎に、パラメータの平均化を

*2 現在 Amazon のトップページには 24 個のカテゴリが用意されているが、レビュー数が十分でないことから、うち 7 個のカテゴリについては実験を行っていない

行っている。GPD 学習の更新度合いを決めるパラメータ ρ は 0.0001 に固定している。

Hidden CRFs のパラメータ推定は GPD を用いて行うが、GPD は一般に初期値に強く依存するためパラメータの初期値を以下のように設定する。極性パラメータの初期値はナイーブベイズモデルとし、文遷移モデルにおける反転パラメータ及び単語遷移 8 状態モデルにおける大域反転パラメータの初期値には以下の混合モデルから最尤値を計算する。

$$P(d; \mathbf{p}) = \prod_t \sum_q \frac{p^{pn}(s_t | \gamma(q, \phi))}{\sum_{q'} p^{pn}(s_t | \gamma(q', \phi))} p^{rev}(s_t | q)$$

$$p^{rev}(v | q) = \frac{\sum_{\phi} \sum_d \sum_t \frac{p^{pn}(s_t | \gamma(q, \phi))}{\sum_{q'} p^{pn}(s_t | \gamma(q', \phi))} n_{dvt}}{\sum_{v'} \sum_{\phi} \sum_d \sum_t \frac{p^{pn}(s_t | \gamma(q, \phi))}{\sum_{q'} p^{pn}(s_t | \gamma(q', \phi))} n_{dvt'}}$$

式中 $p^{pn}(s_t | \gamma(q, \phi))$ は、ラベル ϕ の状態 q において、文 s_t が生成される確率を示し、ナイーブベイズに基づくユニグラム確率の積によって計算される。一方、単語遷移 8 状態モデルにおける局所反転パラメータの初期値には “not”, “no”, “neither”, “nor” の 4 単語に等しい値 (0.25) を付与し、他の単語に極めて小さな値 (10^{-10}) を付与した。遷移パラメータの初期値は全て同一としたが、単語遷移型における非エルゴディック型の遷移のない箇所に対しては極めて小さな値 (10^{-10}) を付与し、近似的に非エルゴディック型を表した。エルゴディック型の場合には、全状態に同一の遷移初期パラメータを与えると、複数の Rev. 状態が 1 状態に圧縮されてしまうため、若干のノイズを加えている。なお、全てのパラメータを Hidden CRFs にパラメータの初期値として導入する際には、対数をとることとする (Gunawardana et al. 2005)。

3.2 評価文書分類実験結果

表 1 はベースラインと Hidden CRFs を用いた評価文書分類の実験の結果を示している。ベースラインはナイーブベイズ (表中 NB) と SVMs である。素性には文書中の単語出現 (0/1) を用いた。SVM の学習と評価には $SVM^{light*3}$ を用い、カーネルにはガウスカーネルを用いた。Hidden CRFs (表中 HCRF) 中の各表記は以下を意味する。“wd4” は単語遷移モデルの 4 状態型、“wd8” は単語遷移モデルの 8 状態型、“man” は人手によって遷移が起こらないと構造を仮定した非エルゴディック型モデル、“elg” はエルゴディック型モデル、“sen” は文遷移モデルである。HCRFs を用いたモデルはナイーブベイズの精度を超え、SVMs と同等の性能を示した。単語遷移モデルについて見ると、4 状態と 8 状態を比べてやや 8 状態の方が性能的に上回っている

*3 <http://svmlight.joachims.org/>

表1 Hidden CRFs とベースラインによる評
価文書分類精度の比較 (10 交差検定)

NB	SVMs	HCRFs				
		wd4		wd8		sen
		man	elg	man	elg	
81.54	82.51	82.10	82.10	82.49	82.21	82.74

ことが確認でき、より緻密なモデル化に成功したと考える。またエルゴディック型であっても、人手での遷移構造付与の場合と同程度の性能を示したことから、有効なモデル化が自動的に行われたと言える。ただし、エルゴディック型で学習した場合の Rev. からの遷移重みは、遷移元の状態に戻る方向にも大きく付与されており、人手によって与えた非エルゴディックの構造とは大きく異なる傾向となった。

単語遷移モデルよりも文遷移モデルは若干ながら識別性能が高かった。これはモデルの優劣よりも、素性関数による性能差が Hidden CRFs の結果に反映されてしまっている可能性が高いと考える*4。文遷移モデルの各パラメータについて考察する。表2は遷移パラメータの学習結果である。ここからいくつかの興味深い結果が伺える。まず、遷移先 (transition target) が反転状態 (rev.) である場合に対し、非反転状態 (non-rev.) である場合の方が高く見積もられている。これは人間の直感と同じ結果を示していると言える。また表中 π は初期遷移を表しており、Negative モデルでは反転状態の方が重みが高くなっている。これは表中全ての場合においてこの箇所のみの特徴である。Negative なレビューでは最初に一部 Positive な意見を認めたり、あるいは他人の意見を参考にしたりする傾向があるのではないかと予想される。最後に表3に極性及び反転素性として特徴的な単語を示す。これらの単語はそれぞれの素性に対する重みの差によって順位付けしている。極性素性の上位5単語はその極性に沿った単語が上位に来ていたものの、反転素性の上位の単語は人間の直感にそぐわないものだったため、上位2単語の他に上位50単語から人手で選択した3単語を示している。括弧内の数字は順位である。

4 結論

単語または文を単位とする Hidden CRFs を用いて評価文書中で生じる部分的な極性の反転に対するモデル化

*4 HCRFs の単語遷移モデル、文遷移モデルと比較のため、文における単語出現または単語出現回数を素性関数とした場合のナイーブベイズによる分類精度を調べたところ、それぞれ 80.11% と 79.81% の正解率だった。

表2 文遷移モデルにおける遷移重み (λ^{trans})

		transition source		
		π	non-rev.	rev.
transition target (posi.)	non-rev.	0.02	1.27	-0.688
	rev.	-1.3	-1.14	-2.03
transition target (nega.)	non-rev.	-0.28	0.74	-0.929
	rev.	-0.11	-1.4	-4.04

表3 特徴的な極性単語と反転単語

polarity words (top 5)		reversal words (top 2 and selected from top 50)	
posi.	nega.	non-rev.	rev.
excellent	unusable	unusable(1)	replay(1)
pleased	returned	receipt(2)	lens(2)
pleasantly	disappointing	especially(20)	initially(11)
compliments	disappointed	secondly(34)	vs.(34)
easy	horrible	additionally(50)	e.g.(45)

を行った。単語遷移モデルにおいては、反転を引き起こす単語を大域的な反転と局所的な反転の2種に分ける等、いくつかのモデルを提案し、文遷移モデルにおいては極性と反転という両面における重みを単語毎に二重に付与することで効果的なモデル化を可能とした。今後はより適切な初期値の推定、他の損失関数の適用、さらに Hidden CRFs の状態数を増やした場合の適用法について考えていく。

参考文献

- Barzilay, R. and Lee, L. (2004). “Catching the Drift: Probabilistic Content Models, with Applications to Generation and Summarization.” In *Proc. of the HLT-NAACL 2004 Conf.*, pp. 113–120.
- Gunawardana, A., Mahajan, M., Acero, A., and Platt, J. C. (2005). “Hidden Conditional Random Fields for Phone Classification.” In *Proc. of the INTERSPEECH Conf.*
- Ikeda, D., Takamura, H., Ratinov, L.-A., and Okumura, M. (2008). “Learning to Shift the Polarity of Words for Sentiment Classification.” In *Proc. of the IJCNLP-08 Conf.*
- Juang, B.-H. and Katagiri, S. (1992). “Discriminative learning for minimum error classification.” In *IEEE Trans. Signal Processing*, Vol. 40, pp. 3043–3054.

- Lafferty, J., McCallum, A., and Pereira, F. (2001). “Conditional random fields: Probabilistic models for segmenting and labeling sequence data.” *Proc. 18th International Conf. on Machine Learning*, 282–289.
- Mao, Y. and Lebanon, G. (2007). “Isotonic Conditional Random Fields and Local Sentiment Flow.” In *Neural Information Processing Systems*, Vol. 18.
- McDonald, R., Hannan, K., Neylon, T., Wells, M., and Reynar, J. (2007). “Structured Models for Fine-to-Coarse Sentiment Analysis.” In *Proc. of the 45th ACL Conf.*, pp. 432–439.
- Pang, B., Lee, L., and Vaithyanathan, S. (2002). “Thumbs up? Sentiment Classification using Machine Learning Techniques.” In *Proc. of the ACL-02 Conf. on EMNLP*, Vol. 10, pp. 79–86.
- Quattoni, A., Collins, M., and Darrell, T. (2004). “Conditional Random Fields for Object Recognition.” In *Neural Information Processing Systems*, Vol. 17.
- Sadamitsu, K., Sekine, S., and Yamamoto, M. (2008). “Sentiment Analysis Based on Probabilistic Models Using Inter-Sentence Information.” In *Proc. of the 6th LREC*.
- Suzuki, J., McDermott, E., and Isozaki, H. (2006). “Training Conditional Random Fields with Multivariate Evaluation Measures.” In *Proc. of the 21st COLING and 44th ACL Conf.*, pp. 217–224.
- 乾孝司 奥村学 (2006). “テキストを対象とした評価情報の分析に関する研究動向.” *自然言語処理学会論文誌*, **13** (3), 201–241.