

修士学位論文

Triple-GANによる感情分析に対する
半教師あり学習

令和元年度

茨城大学大学院理工学研究科

情報工学専攻
楊 金成 (18nm741g)

令和 元年度茨城大学大学院理工学研究科情報工学専攻 修士学位
論文

Triple-GAN による感情分析に対する半教師あり学習

著者：楊 金成 (18nm741g)

指導教員：新納 浩幸 教授

論文要旨

本論文では Triple-GAN を用いて感情分析の半教師あり学習を行うことで、GAN の Generator が生成するデータを文書の埋め込み表現とできるかどうかを検討する。

GAN (Generative Adversarial Networks) は深層生成モデルの 1 つであり、生成モデル $p_g(x)$ の形を陽に与えない形であることから、VAE (Variational Autoencoder) や AR (Autoregressive Model) とは全く異なるアプローチを取り、生成するサンプル画像の質が非常に高いという特徴をもつ。現在、GAN は本来の画像生成のタスクを超えて、様々な方面に応用されている。その一つが半教師あり学習への応用である。原理的には GAN を利用して画像とラベルのペアのデータを生成する。これによって大量に教師データを得ることができるため画像認識の精度が向上する。この枠組みの上でラベルなしのデータを利用するものが半教師あり学習である。現在、画像認識の半教師あり学習においては GAN を用いた手法が state-of-the-art となっている。Triple-GAN もそのような半教師あり学習の 1 つである。

GAN は画像の分野では大きな成功を収めているが、自然言語処理の分野での応用は必ずしもうまくいっていない。いくつか原因があるが、1 つは自然言語処理分野で生成するサンプルは基本的に文であり、文が系列的な離散値であることから、Discriminator の設計や学習が困難となる点にある。

本論文ではタスクを文の生成ではなく、文書の識別に GAN を利用する。生成するデータを文書の埋め込み表現とすることで、Discriminator は容易に実現できる。文書の埋め込み表現の Generator の設計が問題であるが、ここでは埋め込み表現の各次元が正規分布から生成されるというモデルを用いる。この枠組みを用いて、本論文では Triple-GAN による感情分析の半教師あり学習を試みる。

実験では Amazon Dataset (ポジティブもしくはネガティブのラベルのついたレビュー文書) を用いて、Triple-GAN の適用を試みた。標準的な半教師あり学習よりも良い結果を得たが、GAN を活用できているとは言い切れず、生成するデータを文書の埋め込み表現とすることにはいくつかの問題があることが判明した。

Semi-Supervised Learning for Sentiment Analysis Using A Triple Generative Adversarial Net

Author : JinCheng Yang (18nm741g)

Adviser: Prof. Hiroyuki Shinnou

Abstract

This study applies a triple generative adversarial net (Triple-GAN) for the semi-supervised learning of sentiment analysis. We try to see whether the data generated by the generator can be used as an embeddings representation of a document.

GANs is one of the deep generative models. Since the shape of the generated model $p_g(x)$ is not given explicitly, we employ an approach different from the variational auto encoder (VAE) and auto regressive (AR) models. The quality of the resulting sample image has an extremely high feature. GAN has currently gone beyond the original image generation task and has been applied in various fields of study. It has been applied in semi-supervised learning. In principle, if such data can be generated using image/label pairs as data, then numerous teacher data can be obtained, improving the accuracy of image recognition. In addition, methods using GAN in semi-supervised learning image recognition has become state-of-the-art. Triple-GAN is also one of such semi-supervised learning.

GAN has successfully been applied in the field of image processing, but not in the field of natural language processing (NLP). This is due to numerous reasons. One of them is that the sample generated in the NLP field contains basically a sentence. The sentence is a series of discrete values that make it difficult to design and train discriminator.

In this study, we use GAN for document identification instead of sentence generation. Discriminator can be easily achieved by utilizing the generated data as an embedded representation of the document. Meanwhile, it is a problem to design the generator of the embedded representation of the document. Here in, we utilize a model, in which each dimension of the embedded representation is generated from a normal distribution. Using this model, we employ semi-supervised learning of sentiment analysis by a Triple-GAN. For our experiments, we attempt to apply Triple-GAN using Amazon dataset.

From the experimental results, we obtain better results when compared with the standard semi-supervised learning. However, the results are insufficient to prove the effective application of GAN. It has been observed that there are some problems in using the generated data as an embedded representation of the document.

目次

1	序論	4
1.1	概要	4
1.2	構成	5
2	GAN	6
2.1	概要	6
2.2	モデル	6
2.3	原理	7
2.4	DCGAN	8
2.5	Conditional Generative Adversarial Nets	9
3	半教師あり学習	11
3.1	Transductive SVM	12
3.2	Label propagation	12
4	文書の埋め込み表現	14
4.1	Doc2vec	14
4.2	Bert	16
4.3	MeCab	16
5	Triple-GAN	18
5.1	モデルのアーキテクチャ	19
5.2	文書埋め込み表現の生成	20
6	実験	21
7	考察	23
7.1	他手法との比較	23
7.2	生成データ数の影響	23
7.3	BERT の利用	23
7.4	今後の課題	24
8	結論	27

1 序論

1.1 概要

本論文では, Triple-GAN を用いて感情分析の半教師あり学習を行うことで, GAN の Generator が生成するデータを文書の埋め込み表現とできるかどうか検証するを目的とし, その実現のためにデータセットを 2 つ手法を利用して文書の埋め込み表現を変換した, またその効果を評価するために, 文書の感情分析の精度を評価手法により評価を行った.

GAN (Generative Adversarial Networks) [3] は深層生成モデルの 1 つであり, 生成モデル $p_g(x)$ の形を陽に与えない形であることから, VAE (Variational Autoencoder) や AR (Autoregressive Model) とは全く異なるアプローチを取り, 生成するサンプル画像の質が非常に高いという特徴をもつ. 現在, GAN は本来の画像生成のタスクを超えて, 様々な方面に応用されている. その一つが半教師あり学習への応用である. 原理的には画像とラベルのペアをデータとして, そのようなデータが生成できれば, 大量に教師データを得ることができるため画像認識の精度が向上する. この枠組みの上でラベルなしのデータを利用するものが半教師あり学習である.

現在, 画像認識の半教師あり学習においては GAN を用いた手法が SOTA となっている. Triple-GAN [1] もそのような半教師あり学習の 1 つである. GAN は画像の分野では大きな成功を収めているが, 自然言語処理の分野での応用は必ずしもうまくいっていない. 1 つ原因は自然言語処理分野で生成するサンプルは基本的に文であり, 文が系列的な離散値であることから, Discriminator の設計や学習が困難となる点にある.

本論文ではタスクを文の生成ではなく, 文書の識別に GAN を利用する. Triple-GAN は識別対象のデータさえ生成できれば, 識別対象は画像以外のデータでも利用できる. 生成するデータを文書の埋め込み表現とすることで, Discriminator は容易に実現できる. 文書の埋め込み表現の Generator の設計が問題であるが, ここでは埋め込み表現の各次元が正規分布から生成されるというモデルを用いる. この場合, VAE [6] で用いられる Reparameterization trick を利用することで学習も可能となる.

実験では Amazon Dataset を用いて, DVD, Books, Music 3 つの領域のデータを Doc2vec や Bert による文書の埋め込み表現に変換する. 訓練データとテストデータそれぞれに 2000 文がある. ラベルなしデータは 10,000 文を利用した. そして, 埋め込み表現に変換したの文書データを利用して, 教師あり学習の SVM, 半教師あり学習の TSVM や LPA と今回の提案手法 Triple-GAN による感情分析を行う. 結果として, 今回の提案手法 Triple-GAN は標準的な半教師あり学習より良い結果を得たが, GAN を活用できているとは言い切れず, 生成するデータを文書の埋め込み表現とすることにはいくつかの問題があることが判明した.

1.2 構成

本論文では感情分析の半教師あり学習に対して、GANの Generator が文書の埋め込み表現生成や利用する方法を検討しました。2章はGAN及び先行研究について説明を行う。3章は半教師あり学習の説明を行う。4章では文書埋め込み表現について説明を行う。5章は今回の提案手法Triple-GANについてを述べる、6章は実験についてを述べる。7章では本研究においての問題点、改善点また今後の課題について述べる。

2 GAN

2.1 概要

敵対的生成ネットワーク (Generative adversarial networks、略称: GANs) は、2014 年にイアン・グッドフェローらによって発表された教師なし学習で使用される人工知能アルゴリズムの一種であり、ゼロサムゲームフレームワークで互いに競合する 2 つのニューラルネットワークのシステムによって実装される。

GAN は深層生成モデルの 1 つであり、基本的には画像生成に用いられる。当初、VAE よりも質の高い画像を生成できるが、学習が困難という問題があった。すぐに逆畳み込みを用いた DCGAN [9] により、安定して学習できるようになった。その後、GAN は単純な画像生成だけでなく、画像変換 [14] や画像編集 [5] にも応用された。また GAN の生成する画像は基本的にランダムであり、どのような画像が生成されるかはわからない。

画像分野で大きな成功を収めた GAN であるが、自然言語処理では画像分野ほど成功しているとは言いがたい。いくつか問題があるが、1 つは生成するサンプルが基本的には文であり、系列情報、つまり離散値であることである。系列的な離散値であるために Discriminator の設計や学習が困難になっている。SeqGAN [12] では強化学習を用いることで Generator を学習し、Discriminator ではモンテカルロ探索を行っている。textGAN [13] では Generator に潜在空間を考慮した LSTM、Discriminator では CNN、学習は MMD (Maximum Mean Discrepancy) の最小化により、SeqGAN よりも質のよい文を生成している。MaskGAN [2] では文生成に使われる encoder-decoder モデルの学習に GAN を利用している。生成器は encoder でマスクされた単語を decoder で復元できるように学習している。また GAN の応用としては、基本的に、文を生成するタスクが対象となるので、また GAN の自然言語処理への応用としては、機械翻訳と対話の研究が活発である。機械翻訳では基本的に人間の訳か機械が作った訳かを敵対的に学習している [11]。対話では強化学習が用いられている [8]。

2.2 モデル

GANs は生成ネットワーク (generator) と識別ネットワーク (discriminator) の 2 つのネットワークから構成される。例として画像生成を目的とするなら生成側がイメージを出力し、識別側がその正否を判定する。生成側は識別側を欺こうと学習し、識別側はより正確に識別しようと学習する。このように 2 つのネットワークが相反した目的のもとに学習する様が敵対的と呼称される所以である。(wiki)

機械学習のモデルは大きく二つの種類に分けられます。生成モデルと判別モデル (Discriminative Model) です。GAN の仕組みでは、生成モデル G (Generator) と判別モデル D (Discriminator) がある

G は、画像を生成するネットワークであり、ランダムなノイズ z を受信し、このノイズによって画像を生成する。

D は、Generator が生成した偽物のデータと本物のデータが与えられ、その真偽を判定する。

この 2 つのネットワークを交互に競合させ、学習を進めることで、Generator は本物のデータ

に近い偽物データを生成できるようになる。

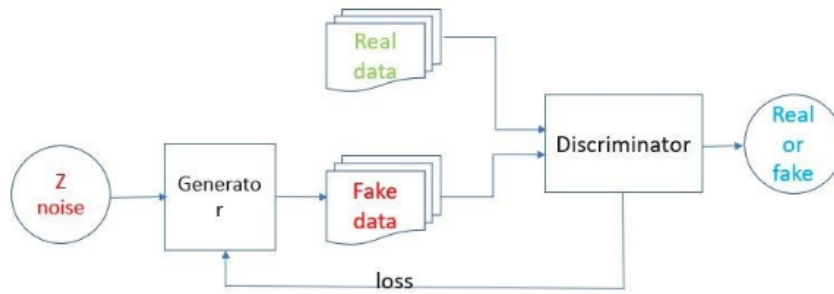


図1 GANの仕組み

2.3 原理

まずは訓練プロセスについて説明します。Generatorは、潜在変数と言われる値(z)を入力値として受け取り、画像データ(img)を出力します。訓練データのデータ分布 $P \sim \text{data}(X)$ から学習します。Generatorの生成分布を $P_G(X; \theta)$ にして、 $P_G(X; \theta)$ はパラメータ θ によって制御される。生成モデルの尤度関数が $L = \prod_{i=1}^m P_G(x^i; \theta)$ となる。

この尤度関数を最大化する、つまりGeneratorはreal画像を生成する確率が最も高くことで、そのため θ^* を見つける必要があります。推定は以下になります：

$$\begin{aligned}
 \theta^* &= \arg \max_{\theta} \prod_{i=1}^m P_G(x^i; \theta) \\
 &= \arg \max_{\theta} \log \prod_{i=1}^m P_G(x^i; \theta) \\
 &= \arg \max_{\theta} \sum_{i=1}^m \log P_G(x^i; \theta) \\
 &\approx \arg \max_{\theta} E_{x \sim P_{\text{data}}} [\log P_G(x; \theta)] \\
 &= \arg \max_{\theta} \int_x P_{\text{data}}(x) \log P_G(x; \theta) dx - \int_x P_{\text{data}}(x) \log P_{\text{data}}(x) dx \\
 &= \arg \max_{\theta} \int_x P_{\text{data}}(x) (\log P_G(x; \theta) - \log P_{\text{data}}(x)) dx \\
 &= \arg \min_{\theta} \int_x P_{\text{data}}(x) \log \frac{P_{\text{data}}(x)}{P_G(x; \theta)} dx \\
 &= \arg \min_{\theta} KL(P_{\text{data}}(x) || P_G(x; \theta))
 \end{aligned}$$

実際の分布におけるすべてのxの期待は、確率積分を求めることと等価であるため、マイナス記号の後の項と $\int_x P_{\text{data}}(x) \log P_{\text{data}}(x) dx$ は関係がないので、付加後はやはり等価である。そして、共有する項目を提出します。maxからminの変換はKL divergenceの形に変換できます。KL divergenceは2つの確率分布の違いを記述しています。

したがって、尤度関数を最大化するように、generatorに最大確率でrealの画像を生成させます。つまり、 θ^* を探して P_G を P_{data} に近づければいいの事です。

どのようにしてより近い分布を見つけるか、これが GAN の貢献です。まず、GAN の公式を与えます。

$$\min_G \max_D V(G, D) = E_{X \sim P_{data}}[\log(D(X))] + E_{X \sim P_G}[\log(1 - D(X))]$$

Discriminator は Generator を固定した上で上式を最大化します。 $E_{X \sim P_{data}}[\log(D(X))]$ は本物データを用いるケースです。この項を最大化するには \log の内部を最大化、すなわち Discriminator の識別結果として 1 を出力させるよう学習させます。

$E_{X \sim P_G}[\log(1 - D(X))]$ では Generator により生成されたデータを示します。 \log の内部を最大化するために、Discriminator の出力、つまり出力 0 を最小化すれば良いです。これは、Discriminator 学習においてのところで触れた入力データと正しいラベルの組み合わせと正解ラベルの組み合わせに一致します。 $P_G(x) = P_{data}(x)$ の場合、G が最適になっている。

2.4 DCGAN

GAN の仕組みを研究するため、先行研究として、画像分野のデータセットを利用した DCGAN を研究した。

DCGAN (Deep Convolutional GAN) は ICLR 2016 で発表された論文 *Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks* で提案された生成モデル。オリジナルの GAN では生成画像がぼやけていたが、DCGAN ではより自然な画像の生成が可能になっている。DCGAN の訓練プロセスは GAN と同じのようになっている。

DCGAN の論文では、畳み込みニューラルネットワーク (CNN) を適用し、かつネットワークを深くした場合であっても安定して学習を進めることができる構成を実証しました。特徴としては、隠れ層に全結合を用いず、畳み込みとする。プーリングの代わりにストライドの畳み込みを用いる。

DCGAN では生成器と判別器両方とも CNN を GAN の全結合層の代わりに利用している、同時にネットワーク全体をマイクロにするために、CNN 中の pooling 層を除去し、さらに全結合層を Global Average Pooling 層に置き換えて計算量を軽減する。そして CNN の重みは Batch Normalization で規格化した、Generator/Discriminator の活性化関数には ReLU/LeakyReLU を利用する。

GAN の学習の不安定性に対処するため、Discriminator の活性化関数には ReLU の代わりに Leaky ReLU が導入されている。通常の ReLU では、入力が 0 未満の場合、出力は 0 になりますが、Leaky ReLU は、入力が 0 未満であっても 0 にならず、負の値を出力します。これにより、学習中に勾配がゼロになることを防ぎ、誤差逆伝搬が滞りにくくなっている。

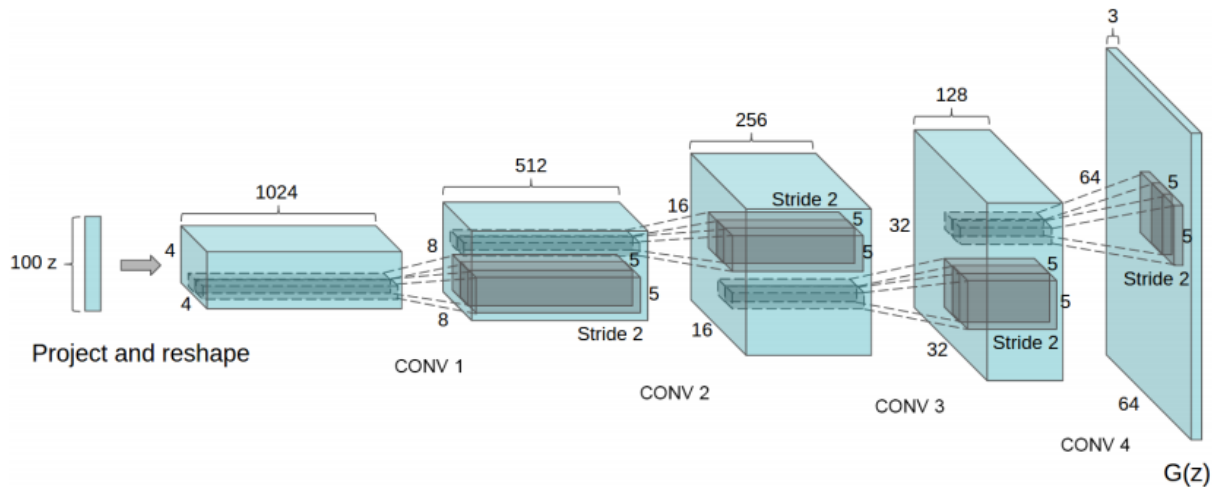


図 2 DCGAN Generator の仕組み

2.5 Conditional Generative Adversarial Nets

CGAN (Conditional GAN) は 2014 年に arXiv で公開された論文 Conditional Generative Adversarial Nets で提案された生成手法です。Generator, Discriminator による GAN の基本構造を踏襲しつつ、条件を与えられるように拡張されている。オリジナルの GAN と CGAN の大きな違いは、Generator の入力にノイズベクトルだけでなく、条件ベクトルも与えている。それに伴い、Discriminator も条件ベクトルに相当する条件データを入力できるよう改良されている。目的関数は以下になっている (two-player minimax game) :

$$\min_G \max_D V(G, D) = E_{X \sim P_{data}(x)} [\log(D(X | Y))] + E_{Z \sim P_Z(Z)} [\log(1 - D(Z|Y))]$$

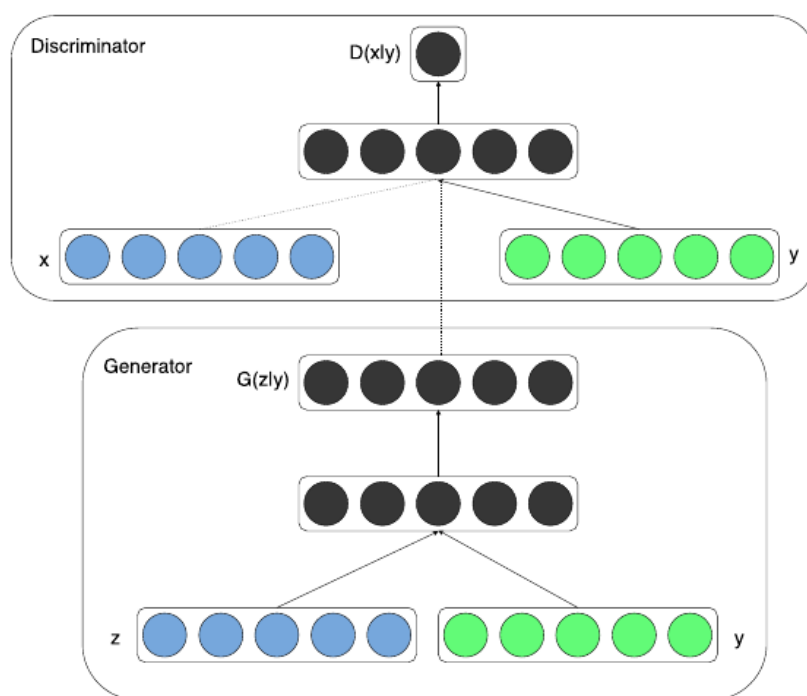


図3 Conditional Adversarial Net

通常の GAN の Generator の入力にはノイズベクトルだけですが、CGAN の入力には、ノイズベクトルに条件ベクトル分の桁を結合したベクトルとなる。つまり、ノイズベクトルが 100 個の場合、条件ベクトルの要素が 10 個だとすると、110 個の要素を持つベクトルが Generator の入力となる。例えば、MNIST データセットの場合は、ノイズベクトルと条件ベクトル $[0, 0, 0, 0, 1, 0, 0, 0, 0, 0]$ 一緒に Generator に入力する。

続いて、Discriminator の入力を説明します。通常、GAN の Discriminator の入力には width × height の実データ (画像) または生成データ (画像) だけですが、CGAN ではデータ生成時の条件も入力する。Discriminator には Generator のように条件ベクトルをそのまま入力することができないため、条件を width × height のデータ (画像) として入力する。これは、Generator に入力する one-hot 表現の条件ベクトルの各要素をそれぞれ 1 枚の条件画像に変換するということ。

3 半教師あり学習

半教師あり学習は教師あり学習と教師なし学習を組み合わせる学習方法のことである。通常、少量のラベル付けされたデータと大量のラベル付けされていないデータを使用して、トレーニングを行う。半教師あり学習は、教師なし学習（ラベル付きトレーニングデータなし）と教師あり学習（完全なラベル付きトレーニングデータ付き）の間にあります。多くの機械学習研究者は、ラベルなしのデータと少量のラベル付きデータと組み合わせて使用すると、学習精度が大幅に向上することを発見しました。学習問題に対するラベル付きデータの取得は、熟練した人間エージェントまたは物理的な実験を必要とする。したがって、ラベル付けプロセスに関連するコストは、完全にラベル付けされたトレーニングセットを実行不可能にする可能性があります。ラベル付けされていないデータの取得は比較的安価です。そのような状況では、半教師あり学習は非常に実用的な価値があります。(wiki)

また感情分析とはレビュー文書（例えば映画のレビュー）が肯定的なものか、否定的なものかを判定するタスクである。これは文書分類の一種であり、通常にはラベル付けされていないデータが手に入れやすいため、半教師あり学習の手法により少量のラベル付け作業と大量のラベルなしデータと一緒に利用して、感情分析の精度が上がると思われている。図4は簡単に半教師あり学習の学習方法を示す。

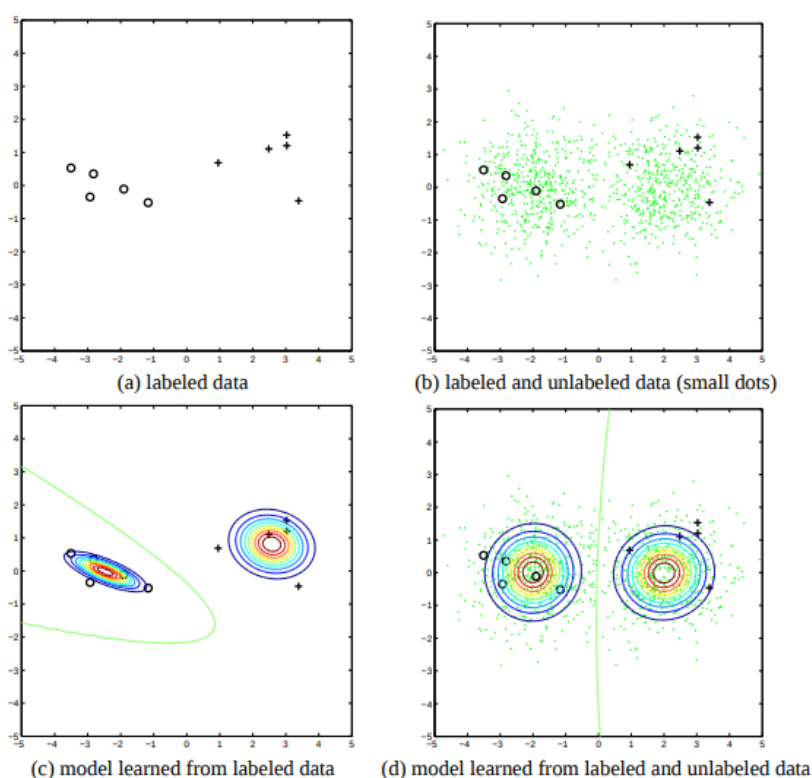


図4 二値分類問題において、各クラスにガウス分布があると仮定すれば、標記されていないデータを使ってパラメータ推定を助けることができます。

3.1 Transductive SVM

SVM (support vector machine) は、教師あり学習を用いるパターン認識モデルの一つである。分類や回帰へ適用できる。SVM は線形二値分類器であり、クラス数が 2 つであるような問題に用いられてきた。二つのクラスは、それぞれ正クラス及び負クラスと呼ばれ、正クラスに属する事例は正例と呼ばれ、負クラスに属する事例は負例と呼ばれる。

サポートベクターマシンは、現在知られている手法の中でも認識性能が優れた学習モデルの一つである。サポートベクターマシンが優れた認識性能を発揮することができる理由は、未学習データに対して高い識別性能を得るための工夫があるためである。(wiki)

Transductive SVM(TSVM) は、半教師あり学習である Transductive Learning を行う SVM である。

Transductive SVM と Inductive SVM の区別では、Inductive SVM ではトレーニングデータとテストデータを利用して重みを学習する。Inductive SVM では label points と unlabeled point を同時に利用して分類器を訓練する手法である Transductive SVM では、トレーニングデータを学習してから、ラベルなしのデータをテストデータにする。つまり、最初はトレーニングデータを使って分類器を訓練する。それに、訓練した分類器にテストデータを入れる。テストデータと訓練データの距離が近く場合、これらのテストデータセットを訓練データセットに見なすことにする。

TSVM の手順は次のようになっている：

1. 少量のラベル付きで分割平面を構築。
2. 分割平面からラベルなしデータを分類する。
3. 平面上に十分近いラベルなしデータのペアを選択。
4. ペアラベルの交換。
5. 終了条件に満たせれば終了。(2 から繰り返し)

Transductive SVM のメリットとしては、少量のテストセットだけでモデルのトレーニングができます。訓練の回数が速く、計算量があまり空間を占めない。

デメリットは：予測モデルにはならず、テストセットを分類するしかできません。パラメータを訓練していないので、新しいテストセットが入るたびにもう一度計算します。

3.2 Label propagation

Label propagation(ラベル伝播) は半教師あり学習の手法であり、半教師あり学習の二つの仮説前提を備えている。1. 隣接するデータのラベルは同じです。この仮定によれば、境界の両側はできるだけ密集したサンプルデータ点を選択することを避け、代わりに、アルゴリズムがラベル付けされていないデータを使用してラベル付けされていない大量のデータを導くことができるように、まばらなデータを選択する。2.2. 同じ流れ構造上の点は同じラベルを持つこと。この仮定によれば、標記されていないサンプルデータは、データ空間をより密にすることができ、局所領域特徴の十分な解析を容易にする。

LPA は、ラベル付けされていないデータの内部構造、分布法則、および隣接データのラベルを使用して、ラベル付けされていないデータのラベルを予測および伝播し、それらをラベル付けさ

れたデータセットにマージする少量のトレーニングラベルガイダンスのみを使用する必要があります。この操作が簡単で、計算量が少ないため、大規模なデータ情報のマイニングと処理に適しています。また、LPA は、隣接ノード間でラベルを渡すことにより分類を学習できるため、データ分布に制限されません。だから、同じクラスのデータの空間分布が類似している限り、データ分布の形状に関係なく、ラベルを介して伝播して同じクラスに分類できます。したがって、音声、ビデオ、画像、テキストの注釈、検索、分類を処理できます。

4 文書の埋め込み表現

埋め込み表現（分散表現）とは、文や文書を高次元の実数ベクトルで表現する技術です。コンピュータが感情分析を行う場合、データは文字列になるので、ここで文書の単語をベクトル表現化することで、文や文書の意味的な表現をコンピュータが扱いやすい数学的表現に変換する自然言語処理でのブレイクスルーだった。

今回の研究では Doc2vec や Bert、二つ埋め込み表現変換手法で同じデータセットに対して感情分析の実験を行った。なお実験モデルにおける入力には各文書が固定長のベクトルが使われている。^{*1}

4.1 Doc2vec

Doc2Vec は 2014 年に Quoc Le と Tomas Mikolov によって発表された文章の埋め込みの手法です。任意の長さの文書をベクトル化する技術で、文やテキストに対して分散表現（Document Embeddings）を獲得することができる。

Doc2Vec には、分散表現を得るための手法が 2 つ存在している。それは、dmpv(図 4.1) と DBoW(図 4.2) である。DBoW は文書中に単語の順序を考慮しないシンプルなモデルで計算効率が高く、dmpv では単語周囲の単語を考慮し、その単語の分散表現を計算する。入力層は " コンテキスト " を表し、出力は予測単語として学習する。dmpv では、文書 ID もコンテキストとして保持する意味合いを持つ。

^{*1} Doc2Vec の場合 200 次元、Bert の場合は 768 次元

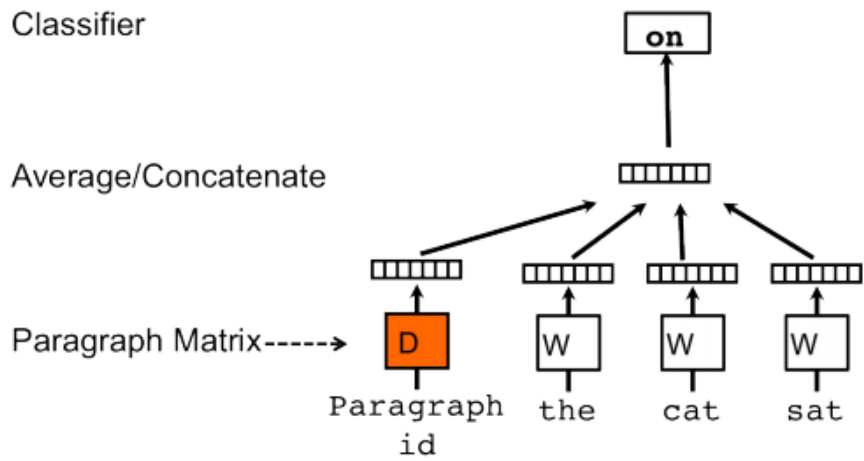


图 5 Doc2vec - dmpv

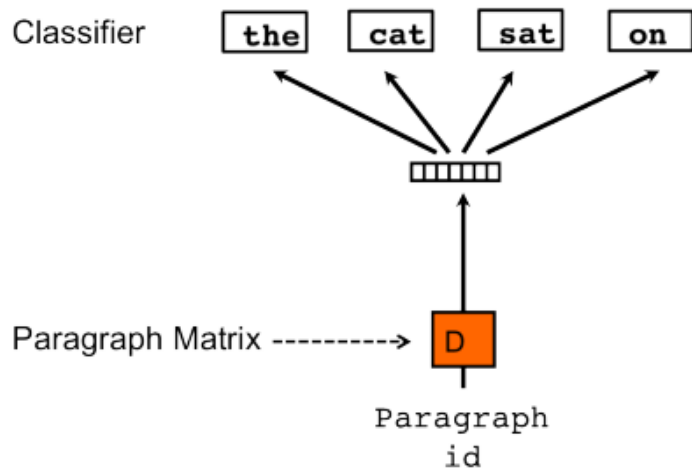


图 6 Doc2vec - DBOW

4.2 Bert

Bidirectional Encoder Representations from Transformers (BERT) は、Google が開発した NLP (Natural Language Processing) の事前学習の手法です。Semi-supervised Sequence Learning、Generative Pre-Training、ELMo、および ULMFitなどをベースとして構築されている、最初の deeply bidirectional(深く双方向) な、教師なしの言語表現です。

事前トレーニングとは、最初にウィキペディアなどの大規模なテキストソースを使用して BERT のトレーニングを行うことを意味します。コアモデルにレイヤーを追加すれば、このトレーニング結果を [question answering] や感情分析などの他の自然言語処理 (NLP) タスクに適用できます。分類タスクや感情分析などは、Transformer の出力に分類層を加えるだけです。

今回は <http://nlp.ist.i.kyoto-u.ac.jp/index.php?BERT%E6%97%A5%E6%9C%AC%E8%AA%9Epretrained%E3%83%A2%E3%83%87%E3%83%AB> 京都大学の黒橋・河原研究室が公開した「BERT 日本語 Pretrained モデル」を用いて文書の埋め込み表現を導出する。

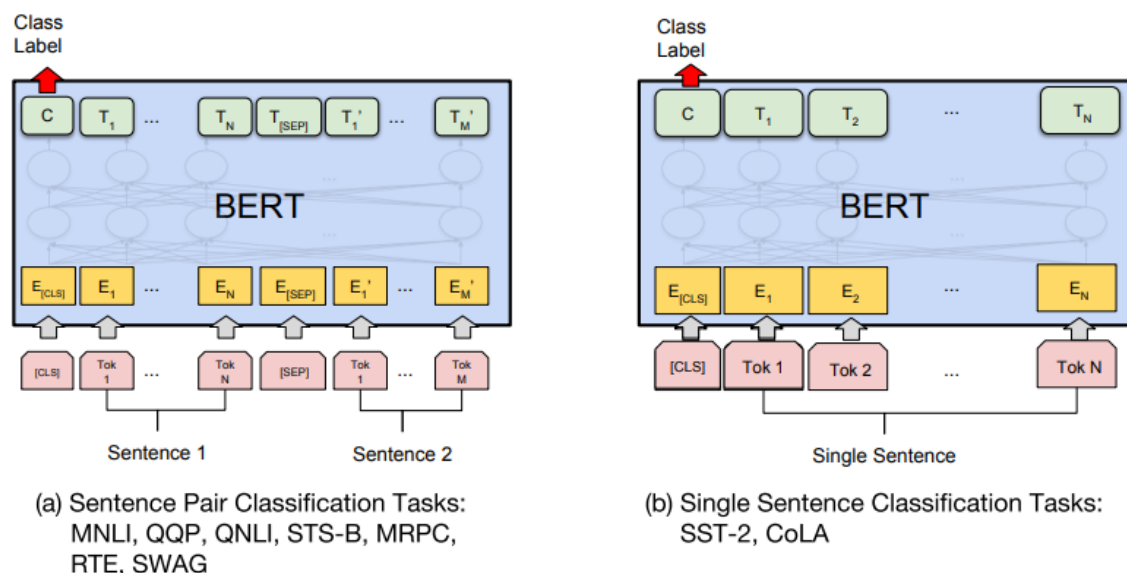


図7 bert model

4.3 MeCab

MeCab は 京都大学情報学研究科 - 日本電信電話株式会社コミュニケーション科学基礎研究所 共同研究ユニットプロジェクトを通じて開発されたオープンソース 形態素解析エンジンです。言語, 辞書, コーパスに依存しない汎用的な設計を 基本方針としています。パラメータの推定に Conditional Random Fields (CRF) を用いており, ChaSen が採用している 隠れマルコフモデルに比べ性能が向上しています。また、平均的に ChaSen, Juman, KAKASI より高速に動作します。

日本語の形態素解析では一般的に次の 2 つのことを行います。

単語分割

品詞付与

次の結果は「私は犬が好きです」を MeCab で形態素解析した結果です。入力文が適切に分割され、適切な品詞が割り当てられています。

```
yjc@hiro2:~$ mecab
%echo 私は犬が好きです
%
私 名詞,サ変接続,*,*,*,*,*
は 助詞,係助詞,*,*,*,*,*は,ハ,ワ
犬 名詞,一般,*,*,*,*,*犬,イヌ,イヌ
が 助詞,格助詞,一般,*,*,*,*,*が,ガ,ガ
好き 名詞,形容動詞語幹,*,*,*,*,*好き,スキ,スキ
です 助動詞,*,*,*,*,*特殊:デス,基本形,です,デス,デス
EOS
記号,空白,*,*,*,*,* , ,
```

図 8 bert model

文書を埋め込み表現手法 (doc2vec、Bert) を行う場合、まずは MeCab で形態素解析を行う、単語分割する。

5 Triple-GAN

一般的な GANs モデルは2つの問題点を抱えている。1つは Generator と Discriminator を同時に学習できない点、もう1つは Generator で生成されるサンプルの意味をコントロールできない点である。この問題点は Discriminator が偽物のサンプルを見分けると同時にそのラベルを予測するという矛盾した役割をし、結果ラベルを考慮することなくデータを推定するところからきている。

この問題にとりくむため Generator と discriminator に加えて Classifier を用いた。Generator と Classifier は画像とそのラベル間の条件付き確率分布を特徴化し、Discriminator は偽物の(画像-ラベルの)ペアか否かを判断する。この仕組みにより state-of-the-art な分類性能を達成した。

Triple-GAN は画像識別をタスクとした半教師あり学習の手法であり、Conditional-GAN の改良手法と見なせる。Conditional-GAN はラベルと画像のペアがデータである。つまり Generator はラベルと画像のペアのデータを生成し、Discriminator はラベルと画像のペアが本物がどうかを識別する。Conditional-GAN によりラベル付きの画像を大量に得ることができるので、画像認識の精度を上げることができる。ただし Conditional-GAN の問題点として、Discriminator がそのデータを偽物と判断したとき、画像の質が悪くて偽物と判断したのか、ラベルと画像の対応が悪くて偽物と判断したのかがわからない点がある。また学習時にはラベル付きデータのみしか使えないことも問題である。Triple-GAN ではこれらの問題点に対処するために、通常の Generator と Discriminator の他にラベルの識別を行う Classifier をモデルに組み込む。

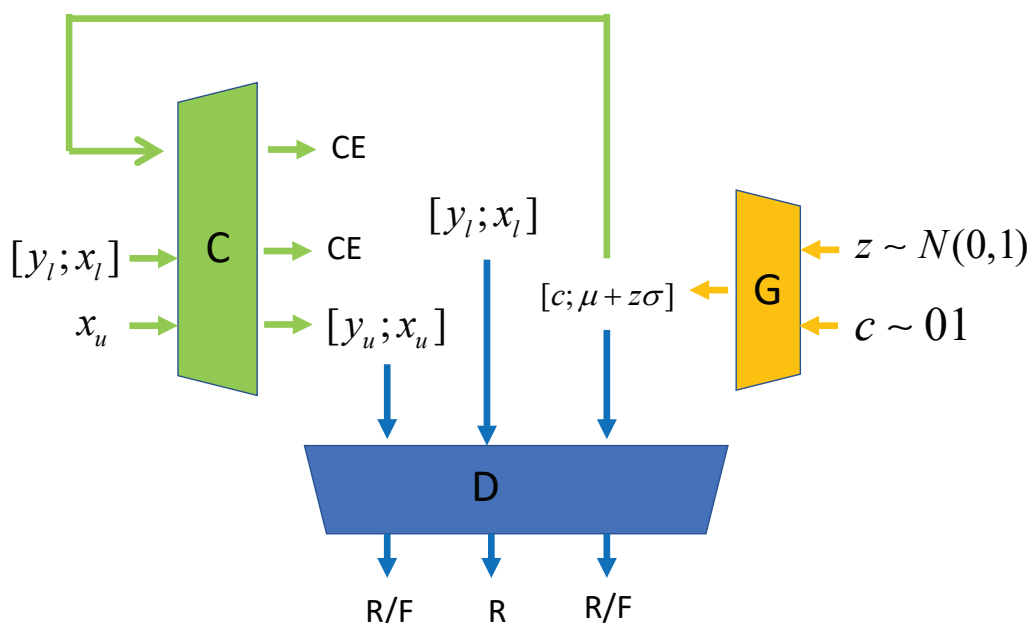


図9 Triple-GAN Net

5.1 モデルのアーキテクチャ

Generator は一様乱数などに従うノイズ $Z_g \sim p_z(Z)$ とそのラベル $Y_g \sim p(Y)$ を入力データとして、 (X, Y) のペア $(X_g, Y_g) \sim p_g(X, Y)$ が生成される。通常の GAN による敵対的学習を行う。

Classifier は 3 種のデータから学習する。第 1 はラベル付きデータである。これは通常の教師あり学習である。学習には CE (Cross Entropy) を用いる。第 2 はラベルなしデータである。この場合、現在の Classifier からラベルを推定して、そのラベルとデータをペアにして、Discriminator に渡す。Discriminator ではこのラベルとデータをペアが本物 (Real) であることから、誤差逆伝播で Classifier の学習も行える。ただし本論文の実験では、この部分で推定したラベルの信頼度が 0.8 以上であった場合を Real、それ以外を Fake として Discriminator と Classifier の学習に利用する。第 3 は Generator から生成されたラベルとデータのペアデータである。このデータに対しても教師あり学習を行う。ここでも CE で学習する。

Discriminator は 3 種のデータからの敵対的学習を行う。第 1 はラベル付きデータであり、ラベルとデータのペアの形にする。第 2 は Generator から生成されたラベルとデータのペアのデータである。第 3 は Classifier から生成されたラベルとデータのペアのデータである。これら 3 種のデータに対して敵対的学習を行う。

Triple-GAN の目標関数：

$$\min_{C,G} \max_D U(C, G, D) = E_{(x,y) \sim p(x,y)} [\log D(x, y)] + \alpha E_{(x,y) \sim p_c(x,y)} [\log(1 - D(x, y))] \\ + (1 - \alpha) E_{(x,y) \sim (p_g(x,y))} [\log(1 - D(G(y, z), y))] + \mathcal{R}_C$$

第 1 項目が正しい画像とラベルのペア。第 2 項目が classifier で推測されたラベルと画像のペア。第 3 項目が generator で生成された画像とラベルのペア。第 4 項目が正しいペアによる classifier の学習部分。

5.2 文書埋め込み表現の生成

Triple-GAN は識別対象のデータさえ生成できれば、識別対象は画像以外のデータでも利用できる。ここでは識別対象のデータの対象を文書の埋め込み表現に設定する。この場合、文書の埋め込み表現の Generator が問題となる。特に学習可能な形で Generator を構築する必要がある。そのためここでは文書の埋め込み表現の各次元が正規分布から生成されるというモデルを設定する。

離散的なデータに対して微分不可能であり、勾配を伝えることができないことに対して、VAE [6] で用いられる Reparameterization Trick を利用することで学習も可能となる。

具体的に示す。まず各実験で用いた埋め込み表現の次元数について、doc2vec は 200、Bert は 768 としている。各次元 i に対して平均 μ_i と分散 σ_i^2 がパラメータとなるが、本論文ではクラス 0 (ネガティブ) とクラス 1 (ポジティブ) のそれぞれに対して平均 $\mu_i^{(c)}$ と分散 $\sigma_i^{(c)2}$ のパラメータを設けることにする。さらにラベル付きデータから $\sigma_i^{(c)2}$ の値を求め、Generator のパラメータを平均 $\mu_i^{(c)}$ のみとする。doc2vec の場合、次元数は 200、クラス数は 2 なので、400 個のパラメータとなる。Generator ではまずランダムにクラス c を発生させ、さらに 200 次元の標準正規分布から 200 次元の乱数を発生させる。 i 次元目の乱数を z_i とすると、Reparameterization trick を利用して生成される埋め込み表現の i 次元目の値は $\mu_i^{(c)} + z_i \cdot \sigma_i^{(c)}$ となる。この 200 次元のベクトルにクラス c を連結させた 201 次元のベクトル $[c; \mu^{(c)} + z \cdot \sigma^{(c)}]$ が生成されたデータとなる。

Discriminator では上記の Generator から生成されたデータを敵対的に学習させる。またクラス c を連結するときに、そのクラスを反転させたクラスを連結したデータも作成し、このデータに対しては Discriminator は正しく Fake を出力するように学習させる。

6 実験

まずは実験に使用したデータについての解説を行います．実験で使用したデータセットは，以下のサイトで公開されている Amazon のレビュー文書である．評価の 4,5 を positive ，評価の 1,2 を negative とした感情分析データとして利用できる

<https://webis.de/data/webis-cls-10.html>

このデータは books (B), dvd (D), および music(M), の 3 つの領域をもち、それぞれの領域に含まれる文書数は表の通りである．どの領域でも positive データと negative データはそれぞれ 1,000 個あり、test データは 2000 個あり、またラベルなしデータとしては 10000 文書を利用する．

予め、これら文書全体を学習データとして、各文書を doc2vec [7] により 200 次元の埋め込み表現に変換した．次に教師データの埋め込み表現を利用して、クラス毎の各次元の偏差値 $\sigma_i^{(c)}$ を算出した．1 エポックで Generator により生成されるデータ数は、ラベルなしデータと合わせて 10,000 個とした．100 エポックまで学習し、各エポックが終了する毎に、テストデータを Classifier で識別しその正解率を測る．各領域における実験結果を図 10、11、12 に示す．各図中の baseline は SVM で正解率を示す．学習が進む毎に徐々に正解率は上昇しているため、本手法が有効に作用していることが確認できる．

表 1 各領域データ数

	trian data	test data	unlabel data
books	2000	2000	10000
DVD	2000	2000	10000
music	2000	2000	10000

precision

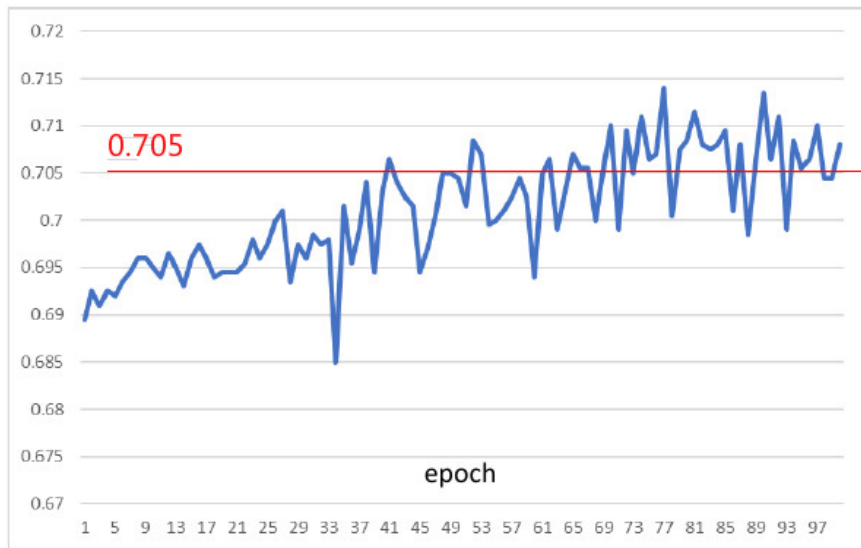


図 10 doc2vec による music 領域の実験精度

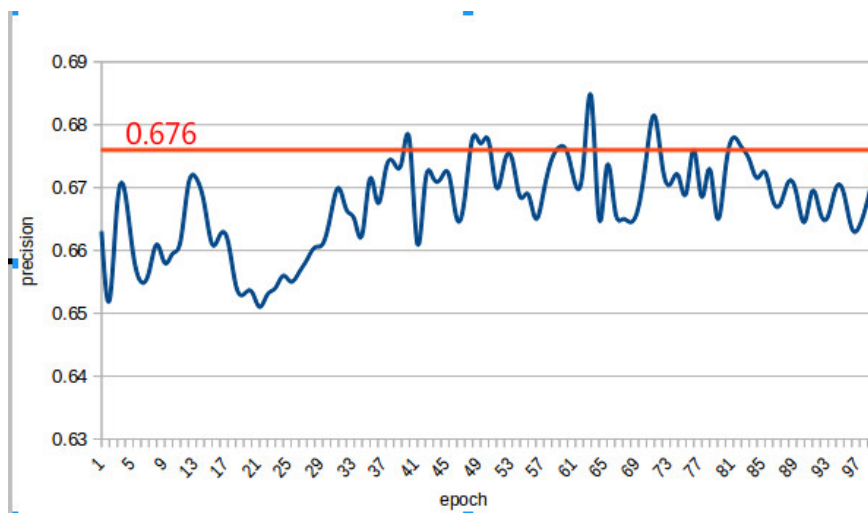


図 11 doc2vec による books 領域の実験精度

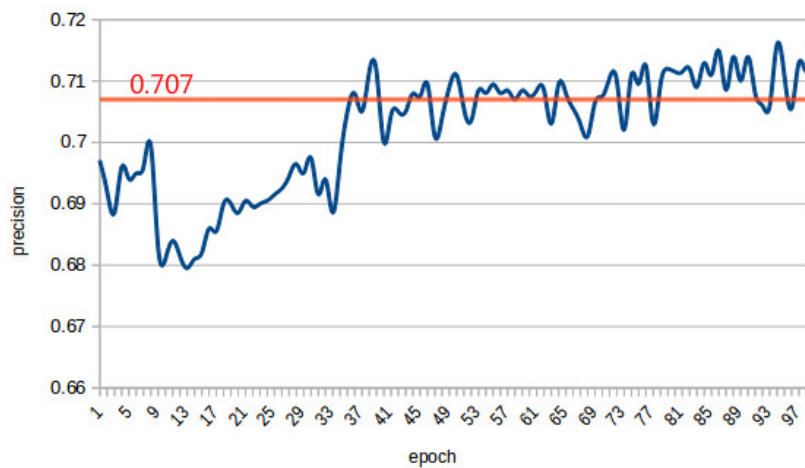


図 12 doc2vec による DVD 領域の実験精度

7 考察

7.1 他手法との比較

実験結果の評価のためにラベルつきデータを用いない、つまり訓練データだけを用いた SVM による結果と他の半教師あり学習である T-SVM (Transductive SVM) [4] と LP (Label Propagation) ^{*2} の結果を表 2 に示す。

表 2 比較手法の結果 (正解率)

	books	DVD	music	Average
SVM	0.676	0.707	0.717	0.700
T-SVM	0.692	0.694	0.704	0.697
LPA	0.580	0.580	0.582	0.581
Triple-GAN	0.680	0.710	0.720	0.703

T-SVM や LP と比べると本手法の方が正解率では勝っている。本手法の正解率は SVM の正解率よりもわずかではあるが、ラベルなしデータを利用した効果もあったと言える。また T-SVM や LP はラベルなしデータを利用した効果がほとんどない。ただし本手法であってもラベルなしデータが効果的に利用出来ているとは言えず、この点については今後の課題で述べる。

7.2 生成データ数の影響

本実験では 1 エポックで Generator により生成されるデータ数は、ラベルなしデータと合わせて 10,000 個としている。この数を 1,000 あるいは 100,000 とした場合の music 領域における各エポック後に学習できたモデルの正解率を調べた。結果を図 16 に示す。

傾向として、生成数が少ない場合よりも多い方が正解率が高い。ただし生成数が多すぎる場合、高い正解率を出す部分もあるが、正解率が安定していない。

7.3 BERT の利用

本実験では文書の埋め込み表現を得るために doc2vec を利用した。文書の埋め込み表現の質を高くすることで精度向上が期待できる。ここでは BERT を利用して文書の埋め込み表現を構築し、それを用いてそれを確認した。具体的には BERT により文書に対する単語埋め込み表現列を得た後に、それらの平均ベクトルを求めることで文書の埋め込み表現を求めた。日本語の BERT モデルとしては、以下で公開されているモデルを利用した。

<http://nlp.ist.i.kyoto-u.ac.jp/index.php?BERT%E6%97%A5%E6%9C%AC%E8%AA%9Epretrained%E3%83%A2%E3%83%87%E3%83%AB>

BERT による文書埋め込み表現を用いた方が doc2vec を用いたものよりも正解率が約 14% も高い。明らかに、本手法においては、より質の高い文書埋め込み表現を用いることが効果的であ

^{*2} https://scikit-learn.org/stable/modules/label_propagation.html

る. またラベルなしデータの利用の効果は、doc2vec の場合と同様、わずかであった.

表 3 データ数及びベクトル長さ

	doc2vec	Bert
labeled data	2000*200	2000*768
unlabeled data	10000*200	10000*768
test data	2000*200	2000*768

表 4 doc2vec と BERT との比較 (正解率)

	Triple-GAN (doc2vec)	SVM (BERT)	Triple-GAN (BERT)
books	0.680	0.838	0.836
DVD	0.710	0.840	0.841
music	0.720	0.844	0.854
Average	0.703	0.841	0.844

7.4 今後の課題

Triple-GAN を本タスクに対して更に効果的に利用するためには、まず、モデルの改良が必要である. ここでは埋め込み表現の各次元が独立な正規分布であることを仮定したが、その仮定が正しくないことは十分にあり得る.

また Generator の評価手法を考える必要がある. 生成するデータが画像や文であれば、主観的ではあっても Generator が正しく学習できているかどうかを確認できる. 本手法のように埋め込み表現を生成する場合、正しく生成が行われていることは、目的のタスクでの評価でしか確認できない^{*3}. Generator がどの程度正しく文の埋め込み表現となっているかが分からないと、モデルの改良が困難である.

今後は上記の点を考慮しながら、生成するデータとして埋め込み表現を利用する方法を更に考えてゆきたい.

^{*3} Discriminator と Classifier は単調に損失値が減少してゆくことは確認できている.

図 11 中の baseline は SVM で正解率を示す

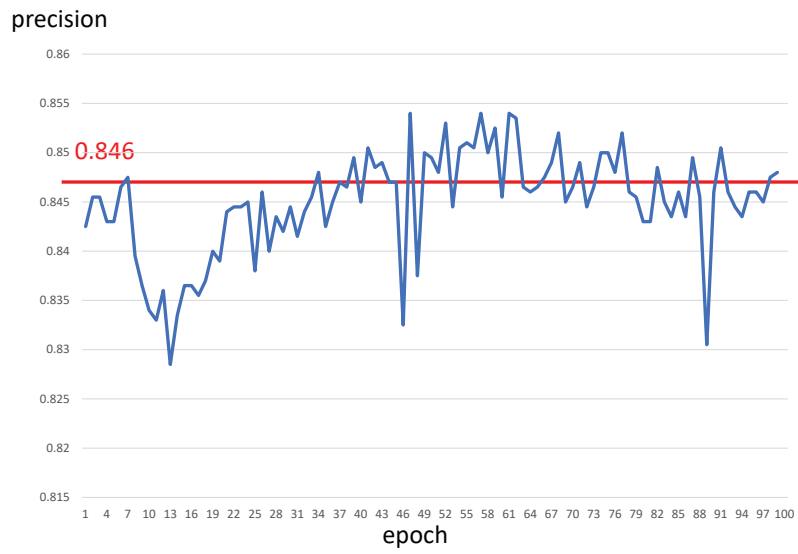


図 13 bert による music 領域の実験精度

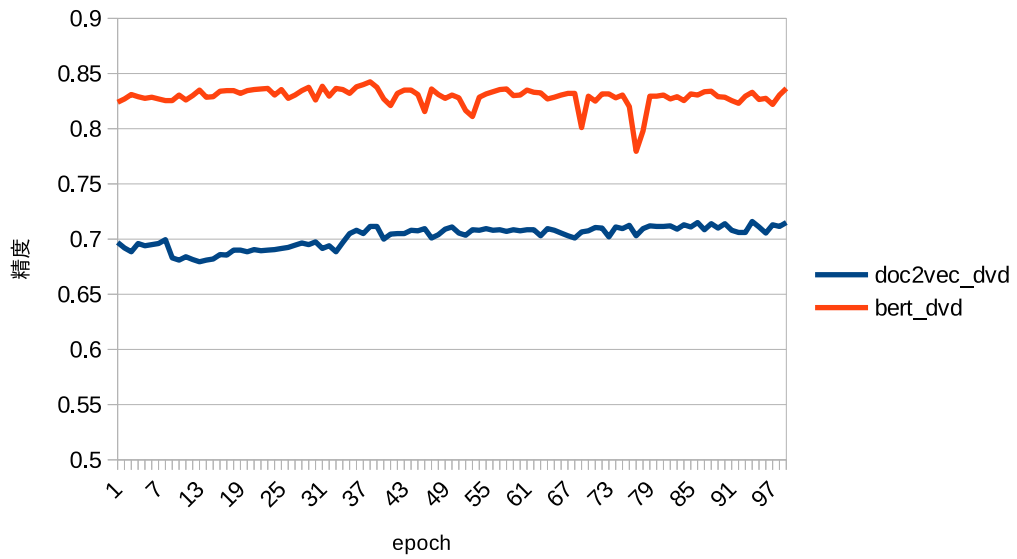


図 14 dvd 領域実験結果

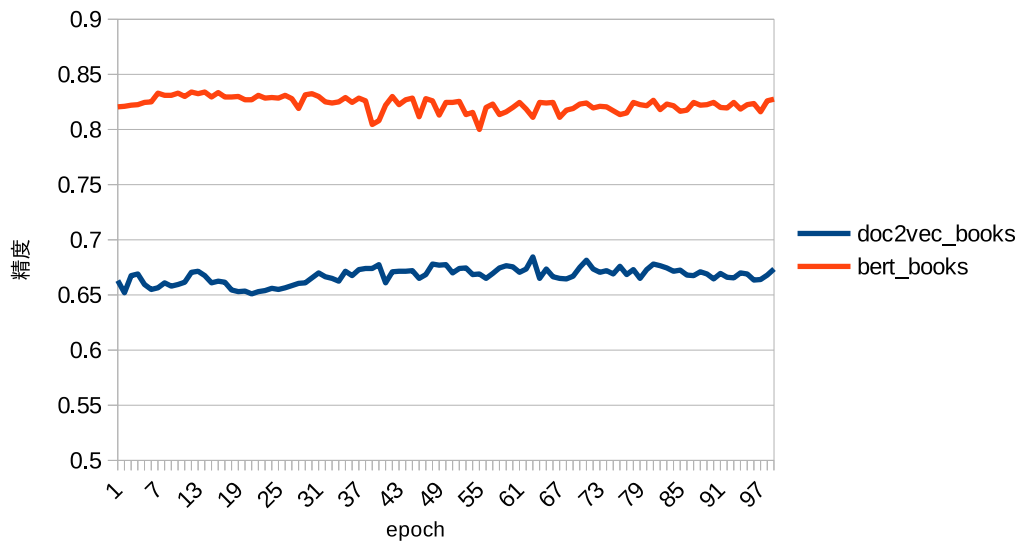


図 15 books 領域の実験結果

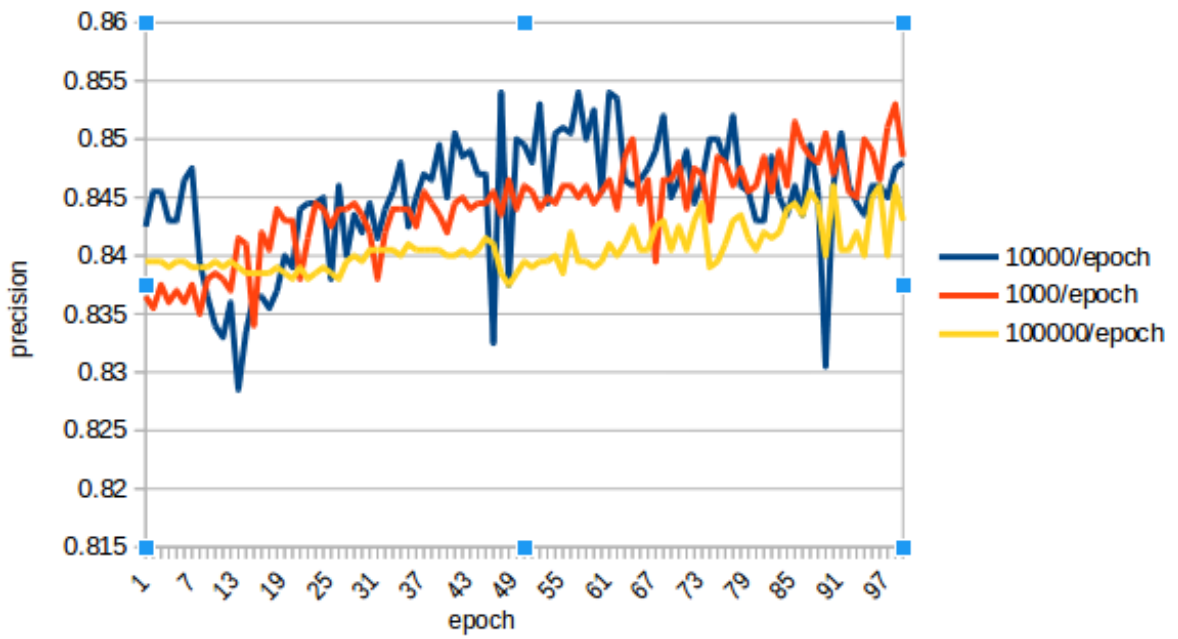


図 16 music 領域各 epoch 生成数の実験精度 (Bert)

8 結論

本論文では Triple-GAN を用いて感情分析の半教師あり学習を行うことで, GAN の Generator が生成するデータを文書の埋め込み表現とできるかどうかを検討した.

文書の埋め込み表現をデータとした場合, Generator の設計が問題となるが, ここでは埋め込み表現の各次元が正規分布に従うというモデルと Reparameterization trick を用いることで対処した.

実験では Amazon Dataset を用いて, doc2vec と Bert2 つ文書の埋め込み表現変換手法を用いて Triple-GAN の適用を試みた. 標準的な半教師あり学習よりも良い結果を得た, 実験結果から見ると埋め込み表現自体の質が高ければ精度もたかくなる. GAN を活用できているとは言い切れず, 生成するデータを文書の埋め込み表現とすることにはいくつかの問題があることがわかった. 今後はこれら問題を詳細に調べ, 生成するデータとして埋め込み表現を利用する方法を更に考えてゆきたい.

謝辞

この研究を修士論文として形にすることが出来たのは、指導教員の新納浩幸教授の厳しくも優しい指導を賜りました。新納先生に深謝致します。また、日常の議論の際、多くの知識やヒント新納研究室の皆様へ感謝致します。

参考文献

- [1] LI Chongxuan, Taufik Xu, Jun Zhu, and Bo Zhang. Triple generative adversarial nets. In *Advances in neural information processing systems*, pp. 4088–4098, 2017.
- [2] William Fedus, Ian Goodfellow, and Andrew M Dai. Maskgan: better text generation via filling in the.. *arXiv preprint arXiv:1801.07736*, 2018.
- [3] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pp. 2672–2680, 2014.
- [4] Thorsten Joachims. Transductive learning via spectral graph partitioning. In *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*, pp. 290–297, 2003.
- [5] Takuhiro Kaneko, Kaoru Hiramatsu, and Kunio Kashino. Generative attribute controller with conditional filtered generative adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6089–6098, 2017.
- [6] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.
- [7] Jey Han Lau and Timothy Baldwin. An empirical evaluation of doc2vec with practical insights into document embedding generation. *arXiv preprint arXiv:1607.05368*, 2016.
- [8] Jiwei Li, Will Monroe, Tianlin Shi, Sébastien Jean, Alan Ritter, and Dan Jurafsky. Adversarial learning for neural dialogue generation. *arXiv preprint arXiv:1701.06547*, 2017.
- [9] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015.
- [10] Jincheng Yang, Rui Cao, Jing Bai, Wen Ma, and Hiroyuki Shinnou. Semi-supervised learning for sentiment analysis by using triple-gan. In *IEICE Tech. Rep.*, Vol. 119 of *NLC2019-26*, pp. 99–102, Tokyo, Sept. 2019. Fri, Sep 27, 2019 - Sat, Sep 28 : Future Corporation (NLC, IPSJ-DC).
- [11] Zhen Yang, Wei Chen, Feng Wang, and Bo Xu. Improving neural machine translation with conditional sequence generative adversarial nets. *arXiv preprint arXiv:1703.04887*, 2017.
- [12] Lantao Yu, Weinan Zhang, Jun Wang, and Yong Yu. Seqgan: Sequence generative adversarial nets with policy gradient. In *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.
- [13] Yizhe Zhang, Zhe Gan, Kai Fan, Zhi Chen, Ricardo Henao, Dinghan Shen, and Lawrence Carin. Adversarial feature matching for text generation. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pp. 4006–4015. JMLR. org, 2017.

- [14] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pp. 2223–2232, 2017.
- [15] 新納浩幸. Chainer v2 による実践深層学習 (ディープラーニング). オーム社, 2017.
- [10] [15]