2A-03

位相的データ解析に基づく言語モデルが生成する埋め込みベクトルの特徴抽出

佐藤 哲 †

パーソルキャリア株式会社
テクノロジー本部 デジタルテクノロジー統括部 †

1. はじめに

自然言語処理の研究分野において，単語や文章を数値で表すことは重要な課題である．近年では，ニューラルネットワークを用いた言語モデルによる文章のベクトル化が，大きな注目を集めている．しかし，ニューラルネットワークを用いた言語モデルが出力するベクトルは次元が高く解釈が難しいため，そのベクトルからの特徴抽出手法が研究されている．本研究では，言語モデルが出力する高次元ベクトルに対し，適切な前処理と位相的データ解析を適用することで，効果的に特徴量を抽出する手法を提案する．

2. 自然言語の埋め込みベクトル生成

人間が使う自然言語を計算機で処理できるようにするためには，自然言語を数値で表さなければならない．自然言語が持つ意味を考慮して生成した数値の列を，埋め込みベクトルと呼ぶ．

埋め込みベクトルは，TF-IDF など単語を数えることにより求められる特徴量を計算して生成する手法に加え，近年では言語モデルを利用した手法が利用されるようになってきた．言語モデルを利用した手法には，(1) 埋め込みベクトルを出力するように訓練した言語モデルを使う方法 [2] や，(2) ニューラルネットワークの次元圧縮に基づく手法（Universal Sentence Encoder[1]，(3) 汎用的な言語モデルの内部状態を埋め込みベクトルとして利用する方法など多くの研究がある．本発表では (3) のアプローチを使い，RWKV モデル [3] を用いて作成した埋め込みベクトルを扱う．

一般に，言語モデルにより作成した埋め込みベクトルは，自然言語の多くの情報が含まれているため次元が高く，人間が理解することは難しい．そこで本研究では，高次元ベクトルを多数の低次元ベクトルで表現するアプローチを試みる．

3. 低次元ベクトルの集合による埋め込みベクトルの表現

高次元ベクトルを理解するための方法の一つは，高次元ベクトルを分割し，複数の低次元ベクトルの集合として表すことである．例えば高次元ベクトルを 3 次元空間で表現するためには，(1) 決められた規則に基づき 3 成分ずつサンプリングしていく手法や，(2) 先頭から順に 3 成分ずつ抽出していく手法が考えられる．本発表では，(1) のアプローチである Sliding Window[4] を採用した実験結果を紹介する．

Sliding Window は，データ系列

$$x_k, x_{k+1}, x_{k+2}, x_{k+3}, x_{k+4}, \cdots$$

に対し，L を定数とし，例えば 3 次元であれば順に k，$k+L$，$k+2L$ とデータを取り出すことで，次のような 3 次元データ系列を生成する：

$$(x_k, x_{k+L}, x_{k+2L}), (x_{k+1}, x_{k+1+L}, x_{k+1+2L}), \cdots.$$

この手法により高次元ベクトルを 3 次元空間内に表すことができ，視覚的にも理解がしやすくなる．しかし依然として高次元の大量のベクトル成分データが存在することには変わりなく，分析が困難である問題は解決されていない．そこで次節にて，ここで得られた 3 次元ベクトルの集合から特徴量を抽出する手法を導入する．

4. 位相的データ解析による特徴抽出

3 次元ベクトルデータの集合からの特徴量抽出手法として，位相的データ解析 [5] を導入する．3 次元ベクトルデータ集合に対し位相的データ解析におけるパーシステンスホモロジーを計算することで，特徴量を 2 次元ベクトルの集合として得ることができる．パーシステンスホモロジーは，ベクトルの集合に対し，ベクトルの部分集合が作る空間の中の「穴」という幾何的な特徴を抽出する手法である．ベクトルが一様に分布しており特徴が無いような場合は情報は抽出できないが，埋め込みベクトルのように「文章同士の意味が似ていれば埋め込みベクトル同士も似ている」というような構造が期待できる場合，ベクトルの集合に幾何的な特徴があることが予想されるため，パーシステンスホモロジーにより意味のある特徴量を抽出できる可能性がある．

5. 実験例

機械学習関連のモデルの評価に使われる GLUE データセット [6] の，STS-B タスクデータを用いて評価実験をする．STS-B タスクデータは，文章の類似度を評価するためのデータセットで，文章の文字列の類似度ではなく意味的な類似度を考慮したデータセットであるので，文章の意味を考慮した文章の埋め込みベクトルの評価に適していると考えられる．

Feature Extraction of Embedding Vectors Generated by Language Models Using Topological Data Analysis
†Tetsu R. Satoh, PERSOL CAREER CO., LTD.

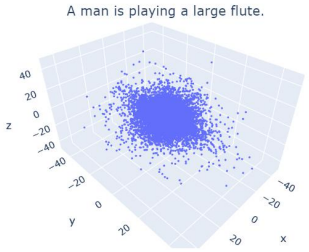

図 1: 文章データの埋め込みベクトルの 3 次元ベクトルの集合による表現例

まず，データセットを RWKV 言語モデルにより埋め込みベクトル化し，Sliding Window により 3 次元空間に射影した結果を図 1 に示す．埋め込みベクトルの次元はモデルにより異なるが，本稿での実験では 409600 次元である．Sliding Window の定数は $L = 1$ とし，その場合に得られる 3 次元ベクトルの数は 409598 個である．埋め込みベクトルの一つの成分が最大 3 個の 3 次元ベクトルに現れ，情報の冗長性があるため，3 次元ベクトル集合から 50 サンプリングしデータ削減を図っている．3 次元空間に射影した埋め込みベクトルから，位相的データ解析の導入によりパーシステンスホモロジーを計算した結果をパーシステンス図に表すと図 2 のようになる．

GLUE データセットの STS-B タスクデータでは，文章が意味する内容が似ているか似ていないか判断できるような自然言語の文章がペアで用意されている．ペアの類似度が 5 段階でラベル付けされており，ラベルと，文章の埋め込みベクトルから計算した類似度あるいは距離との関連性を調べることにより，計算結果の妥当性を評価することができる．本研究では，STS-B タスクデータの文章のペアに対し，それぞれ埋め込みベクトルを求め，Sliding Window とパーシステンスホモロジーを適用して得られた 2 次元ベクトル集合に対し，2 次元ベクトル集合同士の距離を計算することで，文章のペアの類似性を推定する．2 次元ベクトル集合同士の距離としては，Wasserstein 距離を使った結果を紹介する．比較対象としては，文章のペアに対しそれぞれ埋め込みベクトルを求め，埋め込みベクトル同士のコサイン類似度を計算した結果を用いる．

表 1 に，STS-B タスクデータのラベルと，埋め込みベクトル同士のコサイン類似及び 2 次元ベクトル集合間の Wasserstein 距離との相関係数を示す．ラベルは類似度であるため，コサイン類似度とは正の相関があれば計算結果は妥当であり，距離とは負の相関があれば妥当であると考えられる．表に示すように，単純に埋め込みベクトルのコサイン類似度を計算しただけでは，正の相関は得られていない．しかし，埋め込みベクトルから Wasserstein 距離を求めた結果はラベルとの負の相関があり，文章の意味を捕らえられている可能性がある．

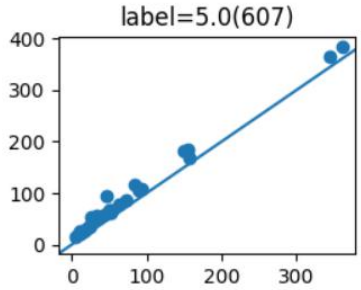

図 2: 文章データの埋め込みベクトルのパーシステンス図による表現例

表 1: GLUE のラベルとの相関係数

	相関係数
コサイン類似度	-0.154630
Wasserstein 距離	-0.202714

以上の実験は，Python3.9 を用いて実装し，Google Cloud の a2-highgpu-1g インスタンス (メモリ 85G，GPU メモリ 40G，12vCPU) 上で実行した．パーシステンスホモロジーの計算には，Homcloud[7] を用いている．

6. おわりに

本研究では，自然言語の言語モデルによる埋め込みベクトルが高次元ベクトルで理解が難しい問題を解決するために，高次元の埋め込みベクトルを多数の低次元ベクトルで表現し，低次元ベクトル空間で位相的データ解析におけるパーシステンスホモロジーを計算することで，埋め込みベクトルを 2 次元ベクトルの集合で表現する手法を提案した．

参考文献

[1] , Universal Sentence Encoder, D. Cer el. al, arXiv 1803.11175, 2018.

[2] , SimCSE: Simple Contrastive Learning of Sentence Embeddings, T. Gao, X. Yao and D. Chen, year=2022, arXiv 2104.08821, 2022.

[3] B. Peng et. al., RWKV: Reinventing RNNs for the Transformer Era, arXiv 2305.13048, 2023.

[4] J. Perea and J. Harer, Sliding Windows and Persistence: An Application of Topological Methods to Signal Analysis, arXiv 1307.6188, 2013.

[5] 平岡裕章, 位相的データ解析とパーシステントホモロジー, 数学, Vol. 68, pp. 361–380, 2016.

[6] , GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding/, A. Wang et. al., Proc. ICLR., 2019.

[7] I, Obayashi ,T. Nakamura and Y. Hiraoka, Persistent Homology Analysis for Materials Research and Persistent Homology Software: HomCloud, J. Physical Society of Japan, Vol. 91, No. 9, pp. 091013, 2022, doi: 10.7566/JPSJ.91.091013.