

グラフ畳み込みネットワークによる推理小説の犯人推定

勝島修平¹ 穴田一¹

概要: 近年、機械学習への社会的な期待が高まっている一方、専門家でも推論過程に対して説明を与えられないことが問題となっている。そんな中、解釈可能性を題材とした推論を行うコンテスト「ナレッジグラフ推論チャレンジ」が開催された。既存研究では、単語の意味を学習するために埋め込みに基づいた手法が提案されているが、小説上の場所、時間、対象物などの同時性を考慮できていない。本研究では、Graph Convolutional Network (GCN)を用いてグラフの構造関係をそのまま学習し、重要となるグラフの関係を Layer-wise Relevance Propagation (LRP)によって明らかにすることによって犯人推定を行う手法を提案する。

キーワード: 知識グラフ推論, 解釈可能性, GCN, LRP

A Criminal Detection of Mystery Novel Using the Graph Convolutional Network

SHUHEI KATSUSHIMA^{†1} HAJIME ANADA^{†1}

Abstract: The interpretability problem, where even experts cannot explain the reasoning process of machine learning, has garnered considerable attention recently. Knowledge Graph Reasoning Challenge 2018, a contest concentrating on interpretability, was conducted in Tokyo. A previous study proposed a method based on word embedding to understand the meaning of the word in the novel. However, the method resulted in ignoring the flow of events. In this study, the graph structure is learned by Graph Convolutional Network, and we explain the important connections on graph by layer-wise relevance propagation.

Keywords: Reasoning on Knowledge Graph, Interpretability, Graph Convolutional Network, Layer-wise Relevance Propagation

1. はじめに

近年、機械学習の発展に伴い、これらの技術への社会的な期待が高まっている一方、専門家でも結果に対して説明を与えられないことが問題となっている。そんな中、説明性を題材とした推論を行うコンテスト「ナレッジグラフ推論チャレンジ」(以下、推論チャレンジ)が開催された[1]。

既存研究では、黒川らが元の小説データをトリプルと呼ばれる構造に変換し、TransE[2]と呼ばれる埋め込み手法を用いることによって単語をベクトルに変換し、ベクトルのトリプル関係から犯人推定を行っている。しかし、小説データをトリプルに変換したことによって、本来同時に学習されるべき場所や時間といった、目的語の同時性に関する情報が失われてしまう。

そこで本研究では、トリプルごとの学習を行う手法ではなく、グラフの構造関係をそのまま学習し、重要となるグラフの関係を明らかにすることによって、犯人推定を行う手法を提案する。

2. ナレッジグラフについて

推論チャレンジでは、場面間の時間経過を考慮するために、内容を場面ごとの最小単位に分割したものに ID を付

与したナレッジグラフと呼ばれるデータ構造で、登場人物やその行動の関係を表現した。図1にナレッジグラフのイメージを示す。

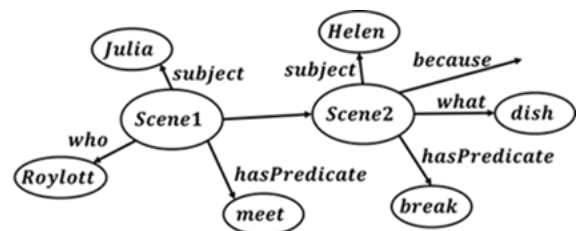


図1 ナレッジグラフのイメージ

図1における場面間の関係(矢印であらわされる遷移)の例を以下に記す。

- Subject: 場面間の記述において主語となる人やモノ
- hasPredicate: 場面間の内容を表す述語
- 場面間の詳細を表す目的語: who(誰), what(何)
- 場面間の関係: then, because など・・・

3. 既存研究

既存研究では、与えられたナレッジグラフ上の単語の意味の学習を、TransE と呼ばれる埋め込み手法を用いて行った。TransE では、トリプルと呼ばれる、<主語・述語・目的語

¹ 東京都市大学大学院 総合理工学研究科
Graduate School of Integrative Science and Engineering, Tokyo City
University Graduate School

>で構成された単語それぞれにベクトルを割り当て、そのトリプルの正誤を判断しながら文章の学習を行う。

既存研究では、単語の学習に TransE を用いるために、ナレッジグラフをトリプルと呼ばれる単位に変換した。しかし、一つの文章からトリプルを生成する際に、2つ以上ある目的語(場所や時間、対象物)のそれぞれを分割し、主語と述語にそれぞれ付け加えることでトリプルを構成している。これではトリプル同士の時間的関係は考慮されておらず、本来単語の持っている同時性の情報が失われている。

4. 提案手法

本研究では、小説上の犯人を、グラフの構造をそのまま学習したベクトルの距離によって推定し、その推定に至った経緯の説明を行う。推定では、GCN を用いて[3]、グラフ構造を考慮したノードの埋め込みベクトルを算出する。そして、小説上の単語 murder と、小説上の人物とのノード間の距離を計算する。予測結果の説明では、LRP [4]を用いることで、出力された埋め込みベクトルに対して、どの構造が影響したかを計算する。

4.1 GCN

GCN とは、深層学習を、ノード(頂点)同士の隣接関係を表すエッジ(枝)によって結びつけられたグラフデータに適用する手法である。GCN では、ノード同士がグラフ上で隣接しているか否かを示す隣接行列 A と、各ノードの特徴ベクトルを表す特徴行列 H を用いる。GCN における、周囲のノードを考慮した特徴量の計算式を以下に示す。

$$H^{(l+1)} = \text{ReLU}\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) \quad (1)$$

ここで、 $H^{(0)}$ は単位行列、 $\tilde{A} = A + I$ は自身のノード情報を考慮した隣接行列の計算式であり、 $W^{(l)}$ はニューラルネットワークの重みパラメータ、 $\text{ReLU}(x) = \max(0, x)$ はランプ関数を表す。(1)式では、ノードの結合関係を表す A と、そのノードの特徴量 H を掛けることで、ノードの構造を考慮した特徴量を生成できる。

4.2 LRP

LRP とは、深層学習における解釈手法の一つである。出力結果に対する層のユニットごとの関係性を、貢献度と呼ばれる値を伝播していくことで、入力データの出力データへの関係性を計算することができる。

$$\begin{aligned} R_i^{(l)} &= \sum_j R_{i \leftarrow j}^{(l+1)} \\ R_{i \leftarrow j}^{(l+1)} &= \frac{z_{ij}}{z_j} R_j^{(l+1)} \end{aligned} \quad (2)$$

ここで、 $R_i^{(l)}$ は層 l のユニット i の出力値に対する貢献度、 $R_{i \leftarrow j}^{(l+1)}$ は層 $l+1$ のユニット j から層 l ユニット i へ逆伝播する貢献度、 z_j は層 $l+1$ ユニットの出力値、 z_{ij} は層 l のユニ

ット i から層 $l+1$ ユニットへ順伝播する値である。

一度、入力データを出力層まで順伝播させた後、(2)式によって層のユニットごとの貢献度を逆伝播して計算することにより、グラフ上のどの隣接ノードを重要視して畳み込んでいるかを算出する。

5. 実験結果

実験に用いたパラメータを表 1 に示す。

表 1 実験に用いたパラメータ等

中間層 1 ユニット数	32
中間層 2 ユニット数	16
学習率	0.01
エポック	200
最適化	Adam
追加知識	ConceptNet

追加知識である ConceptNet [5]とは、Web を介した何千人も協力によって構築された、常識的知識を蓄えたネットワークである。犯人推定は、小説上の人物とこの ConceptNet の追加知識によって得られる単語 murder とのコサイン類似度を計算し、類似度の高い人物を犯人とする。また、犯人推定に至った理由を、LRP による貢献度の逆伝播を行うことで畳み込み時に重要視した隣接ノードを明らかにする。

実験の結果、小説まだらの紐にて犯人推定に必要な ConceptNet の snake を加えた場合の Roylott の犯人推定順位は一位となった。また、小説「僧坊荘園」用の追加知識として、犯人推定のための知識を体系化した犯人推定オントロジーを構築した。

LRP による貢献度及び、独自オントロジーによる推論結果の詳細については、発表にて報告を行う。

参考文献

- [1] 川村隆浩, 江上周作, 田村光太郎, 外園康智, 鶴飼孝典, 小柳佑介, 西野文人, 岡嶋成司, 村上勝彦, 高松邦彦, 杉浦あおい, 白松俊, 張翔宇, 古崎晃司: "第 1 回ナレッジグラフ推論チャレンジ 2018 開催報告—説明性のある人工知能システムを目指して—", 人工知能, 人工知能学会, Vol.34, No.3, pp.396-412 (2019)
- [2] Antoine Bordes, "Translating Embeddings for Modeling Multi-relational Data", In Proceeding of the 26th International Conference on Neural Information Processing Systems(NIPS'13), pp.2787-2795 (2013)
- [3] Thomas N. Kipf, "Semi-supervised classification with Graph Convolutional Networks", In International Conference on Learning Representations (2017)
- [4] G. Montavon, A. Binder, S. Lapuschkin, W. Samek, K.-R. Müller "Layer-wise Relevance Propagation: An Overview" in Explainable AI: Interpreting, Explaining and Visualizing Deep Learning, Springer LNCS, vol. 11700 (2019)
- [5] Robyn Speer, Joshua Chin, and Catherine Havasi. "ConceptNet 5.5: An Open Multilingual Graph of General Knowledge.", In proceedings of AAAI 31 (2017)