

自然言語処理を活用した犯人推測方法

村嶋 義隆

要約

第1回ナレッジグラフ推論チャレンジにおいて、上小田中411チームが「動機を持つか・機会を持つか・方法を持つか」の3つの観点でナレッジグラフを活用するツールを開発した。犯人を推測することができれば、このツールを利用することで裏付けが得られる。そこで犯人を効果的に推測するツールを考えることとした。

推理小説は読者に推理の楽しさを与えるものであり、ストーリー展開にいくつかのパターンがある。例えば、いかにも犯人に違いない、動機のある悪人風の登場人物だがアリバイが崩せない、あるいは犯行方法が不明というパターンはよくある。シャーロック・ホームズシリーズの代表的なパターンを「勧善懲惡」「人情嘶」と名付け、著者は登場人物にそれぞれのパターンにふさわしい振る舞いをさせ、それによって読者の推理を誘導すると仮定した。そしてその振る舞いは、登場人物のセリフの言い回し（例えば無実のアピールやホームズに対する敵対）や、感情を表現するための特定の単語の使い方に現れると考えた。そこで、自然言語処理を活用してセリフから犯人を推測する方法を提案する。

自然言語処理を活用した犯人推測方法

全体像

犯人推測には Naive Bayes Classifier、LSTM など（以下、分類器と表記）を活用し、シャーロック・ホームズの小説を表1のように、そのストーリーのパターン（「勧善懲悪」と「人情嘶」）に分類したうえで、それぞれのパターンごとに登場人物のセリフを抽出する。図1のとおり抽出したセリフで学習させて分類器を作成する。「まだらの紐」を犯人推測のターゲット小説とし、「まだらの紐」のナレッジグラフに図2の例のとおりセリフ・トリプルを追加しておく。SPARQL で登場人物ごとのセリフを抽出して分類器にかけ、もっとも高い確からしさで犯人と分類された登場人物を得る。得られた登場人物を、第1回ナレッジグラフ推論チャレンジ・上小田中411チームのツール（あるいは、今回の推論チャレンジでツール部門へ応募されるであろうツール）で、動機・機会・方法の裏付けを取ることとし、全体で説明可能な推論システムとする。

表1 小説のパターン

タイトル	パターン	分類ポイント
まだらの紐	勧善懲悪	金銭的な動機
踊る人形	人情嘶	キュービット夫人の生い立ち
背中の曲がった男	人情嘶	ヘンリ・ウッドの受難
悪魔の足	勧善懲悪・人情嘶の両方	モーマティは勧善懲悪 スターントン博士は人情嘶
花婿失踪事件	勧善懲悪	金銭的な動機

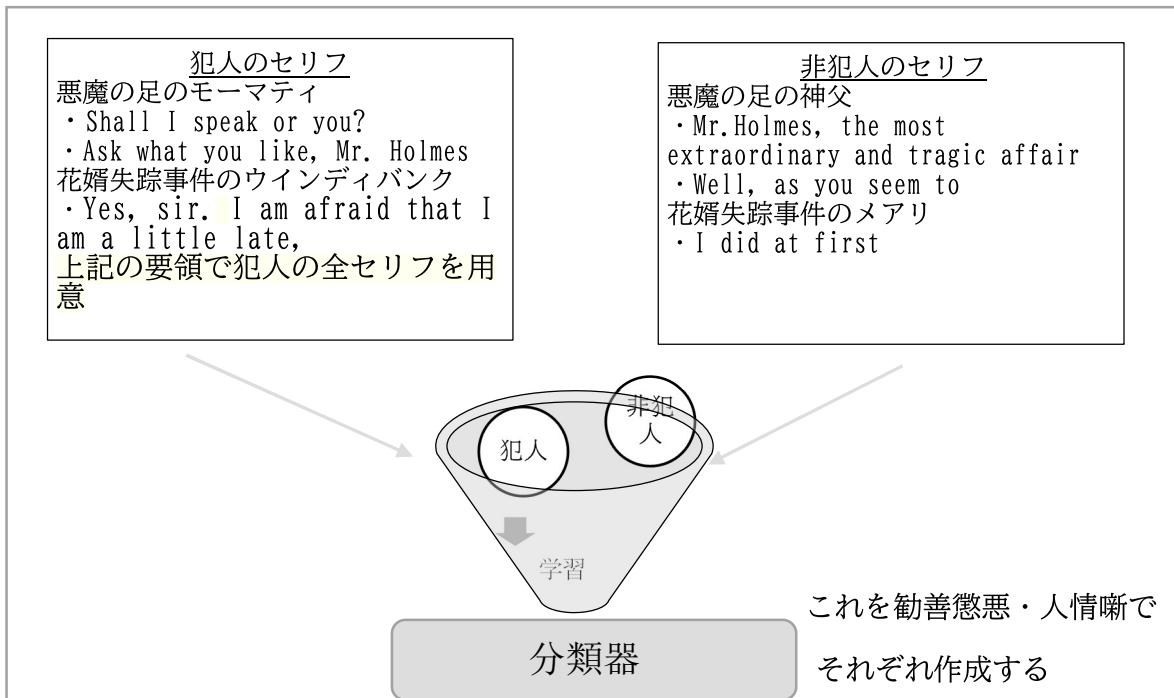


図1 セリフの学習

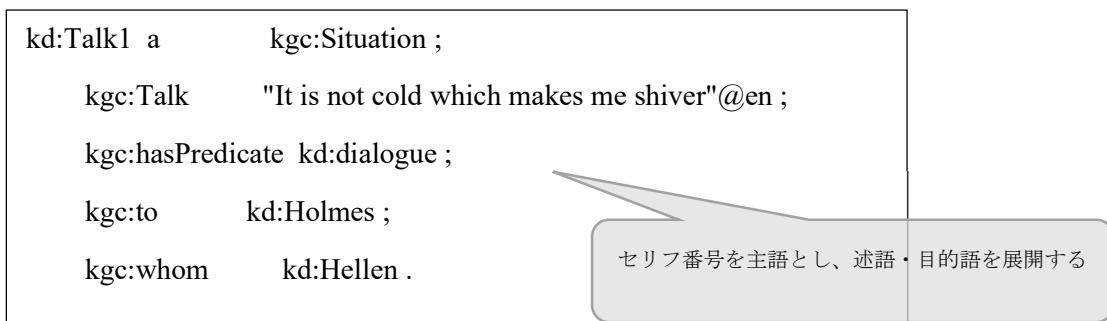


図2 セリフトリプルの例

はじめに

私は IT ベンダーで生命保険会社向けの基幹系業務パッケージシステムを開発してきた。生命保険会社の基幹系システムの大部分は「社内の事務処理」を機械化[克元亮 2004]したものである。急速に進展する機械学習の成果を取り入れて機械化による省力化を実現すべく調査を始めたが、機械学習における注目分野である画像認識・音声認識などの認識系システムや顧客行動・購買傾向分析などのデータ分析系システムを採用できる事務領域がなかなか見いだせない。生命保険会社において機械学習の成果を取り入れている領域は、コールセンター¹やチャットボットでの問合せ対応や保険契約の解約防止・Web マーケティング、給付金の不正請求検知²などである。これらは大規模な「社内の事務処理」のごく一部であり、抜本的な事務の省力化に資する事例は見当たらない。

社内事務は基本的にマニュアル（事務ルール）に基づく。事務ルールのうち、判断が単純なものはプログラムに組み込み、あるいはルールエンジン等を活用してシステム化してきた。しかし、複雑（例えば事務対象の状態が複雑に変化し、その状態によって適用する事務ルールが違う場合など）なケースは人間の判断にゆだねられている。契約が超長期間にわたる生命保険では、契約の状態が複雑になる傾向がある。この複雑な事務ルールをナレッジグラフと自然言語処理と組み合わせることで機械化・自動化できないかと考えて調査を開始し、その過程で第 2 回ナレッジグラフ推論チャレンジに出会った。何事も使ってみて評価するのがビジネスマンの基本である。このような背景で、研究者ではないビジネスマンが当チャレンジに応募することになった。

着想

自然言語処理分野における機械学習の演習で、コーネル大学の感情分析データセット [Bo Pang 2002] を用いたものが多くある。これは映画の評論をポジティブ・ネガティブに分類したもので、Naive Bayes Classifier や LSTM での学習データとして利用することで、会話がポジティブであるかネガティブであるかを分類する。

単純にこの仕組みを、犯人である・犯人でない、の分類に流用する。その前提として、コナン・ドイルは犯人には犯人らしいセリフを喋らせて、謎解きの本題である犯行手段や真相を読者に推理させようとすると考えたからである。

分類器の作成

学習データの準備、それぞれのモデル（今回は Naive Bayes Classifier と LSTM を想定した。現時点ではプログラム例は NBC のみ作成済）による学習の 2 段階の作業により、モデル別の分析器を作成する。

¹ https://www.dai-ichi-life.co.jp/company/news/pdf/2019_012.pdf 第一生命保険ニュースリリース(2019)

² https://www.tis.co.jp/news/2018/tis_news/20180702_1.html TIS ニュースリリース(2018)

学習データの準備

シャーロック・ホームズシリーズを勧善懲悪パターンと人情嘶パターンに分類したうえで、それぞれの小説から犯人のセリフと犯人以外（ただし、ホームズとワトソンを除く）を抽出する。参考として悪魔の足の勧善懲悪部分、すなわちブレンダが死亡し、ふたりの兄弟が発狂した事件の犯人であるモーマティのセリフのファイルとモーマティ以外のセリフのファイルを添付する。

データを準備する際の考慮点として、ひとつのセリフの長さの調整がある。学習する際にすべてのセリフを単語に分解し、 $1 \text{ 行 } n \text{ 列}$ (n は固定。セリフに含まれる最大の単語の数) の行列で定義する。各セリフの単語数が極端に相違すると精度に影響する可能性があるため、いったん 50 単語程度で分割することとした。分割する単語数はチューニング対象とする。

それぞれのモデルによる学習

まずは Naive Bayes Classifier で分類器を作成する。時間的制約により十分なセリフデータが用意できなかつたため、例示したソースコードは python の動作確認程度の意味しかない。比較のために LSTM による分類器、あるいはその他のモデルの分類器を作成することを想定している。

小説のパターン（勧善懲悪、人情嘶）ごとにセリフデータを分割するため、最終的には図 3 のとおり分析器を作成する。

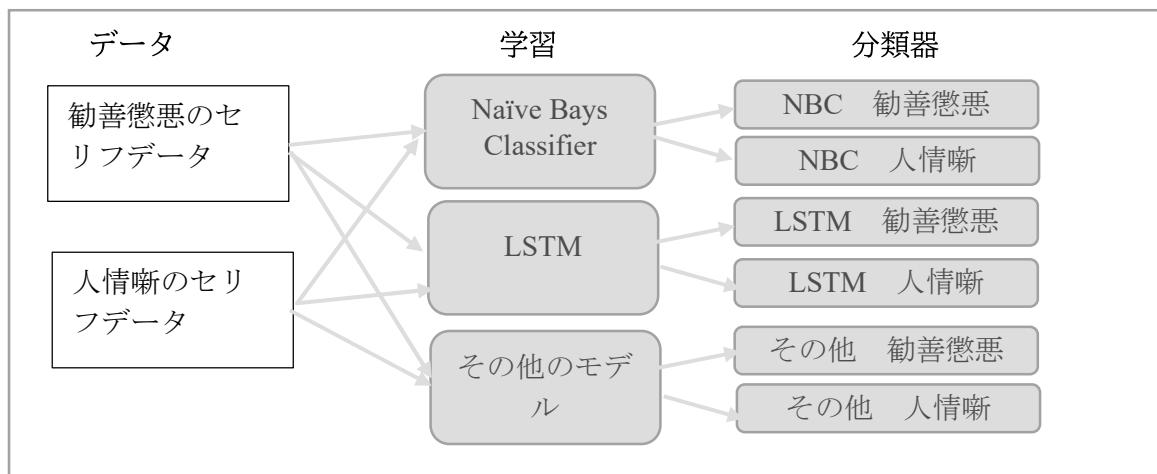


図3 分類器の作成

トリプルの準備

セリフトリプルの構造

主語をセリフの番号とし、表 2 のとおり述語・目的語を定義する。

表 2 セリフトリプルの構造

述語	目的語	説明
Talk	セリフそのもののリテラル	分類器にかける対象
hasPredicate	explanation、flashback、dialogue のいずれか	発したセリフの状況や目的の分類
whom	セリフの発言者	
to	セリフの主たる相手	

述語 hasPredicate では、セリフを発した状況や目的によって 3 つに分類した。 explanation は登場人物が過去の経緯や犯行前後の状況をホームズ等に説明する部分である。この部分は記憶をたどりながら淡々と説明することが多く、感情が入りにくい。そのため犯人推測に有効ではないと考えたため、分析対象から除外できるようにした。また、 flashback は登場人物が回想のなかで別の登場人物のセリフを代弁するものである。虚偽の証言である可能性もあるため、こちらも分析対象から除外できるようにした。 dialogue は主にホームズとの会話であり、このセリフが分析の対象である。

SPARQL

「まだらの紐」におけるセリフデータ[wikisource 2019]抽出例は図 4 のとおり。この例では、ロイロット博士のセリフのうち、dialogue と分類したものを取り出す。実行結果を図 5 に示す。

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX kgc: <http://kgc.knowledge-graph.jp/ontology/kgc.owl#>
PREFIX kd: <http://localhost:3030/2019_01/sparql/>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

select ?s ?p ?o
where
{
  ?s kgc:hasPredicate kd:dialogue.
  ?s kgc:whom kd:Roylott .
  ?s kgc:Talk ?o .
}

```

図 4 SPARQL の例

s	p	o
1 kd:Talk72		"Which of you is Holmes? I am Dr. Grimesby Roylott, of Stoke Moran"@en
2 kd:Talk73		"Roylott to Holmes I will do nothing of the kind. My stepdaughter has been here. I have traced her. What has she been saying to you?"@en
3 kd:Talk74		"What has she been saying to you?"@en
4 kd:Talk75		"Ha! You put me off, do you? I know you, you scoundrel! I have heard of you before. You are Holmes, the meddler"@en
5 kd:Talk76		"Holmes, the Scotland Yard Jack-in-office!"@en
6 kd:Talk77		"I will go when I have said my say. Don't you dare to meddle with my affairs. I know that Miss Stoner has been here. I traced her! I am a dangerous man to fall foul of! See here"@en

図 5 SPARQL 実行結果

犯人推測

SPARQL で抽出したセリフを作成した分類器のエンドポイントに入力して得られた結果を集計する。結果は'cri'(犯人), 'inn'(非犯人)のいずれかであるので、登場人物ごとに集計する。'cri' と分類した確率が最も高い登場人物を犯人と推測する。

評価

セリフデータを作成する時間とナレッジグラフや機械学習に関する能力の制約のため、有効な分類器を作成できなかった。仮にこれらの制約がなかったとしてもシャーロック・ホームズの短編小説では有用な結果は出せないと考える。例えば「まだらの紐」における犯人(ロイロット博士)の dialogue は図 5 のとおり 6 回だけである。データとして過少である。この欠点がなければ、ワトソンとホームズの対話においてホームズが状況を説明する過程で読者に対して犯人像を形成するシャーロック・ホームズシリーズにおいて、犯人のセリフに著者が意図する性格付けの特徴が集中することはあり得るので、推測は可能であると考える。

ビジネスへの応用の観点では、生命保険会社の事務処理において人が判断している対処(例えば何らかの異常状態が発生した際に保険契約者に確認のために連絡をして、その後の手続きを促す)について、機械学習によって推論して対処案を導き出し(当アイデアに置き換えると、犯人推測)、その対処案が当該契約の状態を鑑みたときに事務ルールに適合しているのかについて、事務ルールをナレッジグラフ化した事務ルールトリップル等を利用して確認(当アイデアに置き換えると、動機・機会・方法の裏付け取り)することができるのではないかと考える。この方向で調査を進めることとする。

参考文献

[克元亮 2004]克元 亮、「SE のための金融の基礎知識」、日本能率協会マネジメントセンター (2004)

[Bo Pang 2002]Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. [Thumbs up? Sentiment Classification using Machine Learning Techniques](#). *Proceedings of EMNLP*, 2002. Introduced polarity dataset v0.9

[wikisource 2019] The Adventure of the Speckled Band. (2019, February 6). In *Wikisource*. Retrieved 04:56, November 28, 2019, from https://en.wikisource.org/w/index.php?title=The_Adventure_of_the_Speckled_Band&oldid=9092200
