

## Twitterからの違法薬物売買ユーザの抽出

松波 琴未† 安藤 一秋‡

香川大学工学部† 香川大学創造工学部‡

## 1. はじめに

警察庁が公表した「令和元年における組織犯罪の情勢 [1]」によると、平成 26 年以降、大麻事犯の検挙人員が増加し続けている。年齢層別では、特に若年層の検挙数が急増しており、危機的状況になっている。

現状を把握すべく、四国厚生支局麻薬取締部の捜査担当者にヒアリングを実施した。その結果、違法薬物は、SNS 上で多数取引されており、特に Twitter 上では、大麻をはじめ、覚せい剤、コカインなどを売買するユーザが存在していること、また、若年層ユーザが多い Twitter を利用することで、時と場所を問わず、好奇心から気軽に購入できる環境が構築されていることを確認した。また、SNS 上での捜査においては、薬物売買に関連したユーザを検索し、コミュニティを把握することで、捜査対象を絞り込むが、この流れは全て人手で行われるためコストを要する。さらに、Twitter 上での取引は詐欺事例も多く、それを見分けることが難しいことも確認した。

本研究では、SNS 上での違法薬物売買の捜査に寄与するシステムの実現を目指す。本稿では、Twitter から違法薬物売買ユーザを抽出することを目的に、違法薬物売買を示唆する投稿を分類する手法について提案する。さらに、違法薬物売買ユーザの特徴から違法薬物売買ユーザと詐欺の可能性が高いユーザを分類する手法も提案する。

## 2. 関連研究

安彦らの研究[2]では、違法薬物を売買するユーザの抽出を目的に、キーワードの出現頻度の傾向を学習した Random Forest を用いて違法薬物売買を示唆する投稿を分類する手法を提案している。未知データに対する実験により、薬物に関する投稿 2,094 件とその他の無害な投稿 5,870 件の計 7,964 件で学習した Random Forest により、0.876 の F 値で違法薬物売買を示唆する投稿を判定できたと述べている。しかし、安彦らの手法は、単なる BoW (Bag of Words) しか利用していないため、判定性能を改善できる可能性がある。

## 3. 違法薬物売買を示唆するツイートの分類

## 3.1 違法薬物売買を示唆する投稿の分析

Twitter API を用いて、「手押し」や「野菜」などの違法薬物売買に関する隠語 (4 語) を含む 2,000 件の投稿を収集し、違法薬物売買投稿の特徴を分析する。

分析の結果、2,000 件中、地域名を含んでいる投稿が 75.5%、数字を含んでいる投稿が 58.1%存在した。よって、違法薬物売買に関する投稿には、数字や地域名が含まれやすい傾向があることを確認した。また、野菜や果物などの絵文字を 1 個以上含んでいる投稿が 62.1%存在した。よって、絵文字も含まれやすいといえる。

違法薬物売買ユーザは、検索やフィルタリングなどが

ら逃れるために、単語ではなく、絵文字を利用していると考えられる。そこで本研究では、このような潜在的な違法薬物売買ユーザの抽出を目指し、単語に加えて絵文字にも注目した手法を提案する。

## 3.2 提案手法

違法薬物売買を示唆する投稿の判定には、Support Vector Machine (SVM) と、従来研究[2] で用いられた Random Forest (RF) に加え、勾配ブースティングに基づく手法の 1 つである Light Gradient Boosting Machine (LightGBM) を用いて、各モデルの性能を比較する。

前処理として、URL の削除、代表的な薬物隠語 (4 語) の表記を統一する。本研究では、絵文字を素性に利用するため、原則、絵文字は削除しないが、分析を通じて、顔絵文字は特徴的ではないことを確認したため削除する。さらに、違法薬物売買を示唆する投稿には、地域名や数字が多く使われるため、前処理でカテゴリ化する。

形態素解析には、mecab-ipadic-NEologd で辞書拡張した MeCab を利用する。既存の研究[2]では、名詞のみを素性、その頻度を素性値として、各投稿をベクトル化しているが、本研究では、名詞、動詞、形容詞、副詞、記号からなる単語群をベース素性に、素性値には TF-IDF を用いる。

さらに、絵文字に関する素性をベース素性に追加する。事前分析により、違法薬物売買を示唆する投稿に含まれる絵文字の特徴として、薬物隠語を意味する絵文字や、果物の絵文字が多く使われていることを確認した。そこで、絵文字の有効性について検証するため、違法薬物を売買していると思われるユーザの投稿を分析し、人手で判断した 28 個の薬物隠語を意味する絵文字を利用する。

絵文字に関する素性としては、以下の 4 素性の組み合わせをベース素性に追加して有効性を確認する。

- ① 絵文字の出現頻度
- ② 薬物隠語絵文字の出現頻度+果物絵文字の有無
- ③ 絵文字ごとの出現頻度
- ④ 薬物隠語絵文字と果物絵文字ごとの出現頻度

## 3.3 評価実験

## 3.3.1 データセットによる性能評価

違法薬物売買を示唆する投稿で頻繁に用いられている 4 種の薬物隠語をキーワードに、TwitterAPI を利用して、薬物売買に関連する投稿を収集する。そして、収集した 7,539 件の投稿を人手によりラベル付けし、データセットを構築する。本研究で構築したデータセットには、違法薬物売買を示唆する投稿が 2,941 件、それ以外の投稿が 4,598 件含まれている。

10 分割交差検証における適合率、再現率と F 値を評価指標に、各素性の有効性と分類に最適な素性の組み合わせを確認する。なお、パラメータには、グリッドサーチで求めた最良値を利用する。

安彦らの手法 (Baseline) , SVM, RF, LightGBM による分類器において、最も分類性能が高い素性の組み合わせの結果を表 1 に示す。表 1 からわかるように、LightGBM の結果が最良であった。すべての分類器におい

Extraction of Illegal Drug Dealers from Twitter

† Kotomi Matsunami, Faculty of Engineering, Kagawa University

‡ Kazuaki Ando, Faculty of Engineering and Design, Kagawa University

て、絵文字に関連した素性は効果があり、特に、すべてのモデルに共通して、薬物隠語や果物の絵文字に絞った④素性に効果があることがわかった。

表1. 最良結果を得た素性の組み合わせによる結果

	追加素性	適合率	再現率	F 値
Baseline		0.948	<b>0.963</b>	0.956
SVM	④	0.979	0.900	0.938
RF	③+④	<b>0.991</b>	0.942	0.966
LightGBM	③+④	0.984	0.954	<b>0.969</b>

### 3.3.2 未知データによる性能評価

Baseline 手法と 3 分類器の最良モデルを用いて、未知データに対する分類性能を評価する。データセットの構築時に利用していない薬物隠語を利用して、新たに収集した 1,066 件の投稿を未知データとして利用する。

実験結果を表 2 に示す。表 2 からわかるように、LightGBM モデルの F 値が最も高く、未知データに対しても高い性能で分類できることが確認できた。

エラー分析の結果、出現頻度が少ない薬物隠語や新しい薬物隠語を含む投稿に対する推定ミスが確認できた。今後は、新しい隠語を取り扱う手法を検討する。また、テキストではなく、画像を利用した投稿も確認できた。よって、画像解析結果を特徴に利用することで、分類性能の向上が期待できる。

表2. 未知データに対する結果

	適合率	再現率	F 値
Baseline	0.989	0.859	0.920
SVM (④素性)	0.997	0.633	0.775
RF (③+④素性)	<b>0.998</b>	0.814	0.896
LightGBM (③+④素性)	0.991	<b>0.876</b>	<b>0.930</b>

## 4. 違法薬物売買ユーザと詐欺ユーザの分類

未知データに対する Precision が高いことから、本稿では、違法薬物売買を示唆する投稿をしたユーザを違法薬物売買ユーザ候補とする。そして、この候補の中から違法薬物を売買しているユーザか、詐欺の可能性のあるユーザかを分類する手法を提案する。以降、本稿では、前者を薬物売買ユーザ、後者を詐欺ユーザと呼ぶ。

### 4.1 ユーザの収集と素性検討

素性検討のため、ユーザを収集する。薬物売買ユーザは、複数のユーザから売買が成立したというレビューの存在を基準に収集する。詐欺ユーザは、他ユーザから詐欺アカウントとして晒されていることを基準に収集する。

収集した各 100 ユーザの投稿内容を分析し、薬物売買ユーザと詐欺ユーザを分類するための素性を検討する。

#### ① 他ユーザとの関係

薬物売買ユーザは、客との取引を重ねることで信頼を得ているため、買い手からの支持が厚い。また、詐欺ユーザよりもネットワークが広く、投稿に対する反応も大きいと仮定する。そこで、ユーザのフォロー・フォロー数、投稿数、他ユーザからのリプライ数、ユーザがリツイート・いいねされた総数を素性に利用する。

#### ② レビュー

薬物売買が成立すると、購入ユーザから商品レビューの品質や対応の良さなどをリプライすると仮定する。そこで、対象ユーザに対するリプライ内容に対し、日本語評価極性辞書 [3][4]を用いて、リプライ本文に含まれる名

詞と用言の極性値の総和を求め、ユーザごとの評価極性の平均値を素性に利用する。

#### ③ アカウントの削除・凍結

詐欺ユーザは、他ユーザにアカウントを晒されてしまうため、登録日から日が浅いまま、すぐに削除、凍結する傾向があると仮定する。そこで、登録経過日数を素性に利用する。

#### ④ 薬物売買に関連する投稿

薬物売買ユーザは、一定の顧客がついている場合が多いため、薬物売買に関する内容を頻繁には投稿しない傾向がある。一方、詐欺ユーザは、多くの人からお金を搾取することが目的であるため、頻繁に宣伝する傾向があると仮定する。そこで、薬物売買に関する投稿の割合を素性に利用する。

## 4.2 評価実験

Twitter 上から収集した 307 ユーザ（薬物売買ユーザ 156、詐欺ユーザ 151）の情報をデータセットに利用する。

詐欺ユーザの判定には、Support Vector Machine (SVM) を使用し、4.1 で述べた 4 つの観点から作成した各素性の有効性と分類に最適な素性の組み合わせを、10 分割交差検証における適合率と再現率、F 値で確認する。

最も分類性能が高い素性の組み合わせ上位 3 件の結果を表 3 に示す。表 3 からわかるように、①と②素性を軸に③と④素性の組み合わせに、効果があることがわかった。次に、エラー分析の結果、詐欺ユーザではあるが、最初は取引が成立していた期間があったため、他ユーザとの関係が総合的に薬物売買ユーザの特徴に似たことでエラーとなったユーザを確認した。今後はこのような他ユーザからの評判の変化への対応や、評判は良さそうにみえてもサクラが潜んでいる可能性があるユーザへの対応なども検討する。

表3. 最良結果を得た素性の組み合わせによる結果(上位3件)

素性	適合率	再現率	F 値
①+②+③	<b>1.00</b>	0.950	<b>0.974</b>
①+②+③+④	0.951	<b>0.975</b>	0.963
①+②+④	0.977	0.870	0.920

## 5. おわりに

本研究では、まず、違法薬物売買を示唆する投稿を分類する手法について検討した。実験により、絵文字に関連した素性の有効性を確認できた。また、未知データに対しても薬物売買に関するツイートを高い性能で分類できることを確認した。さらに、薬物売買ユーザと詐欺ユーザを分類する手法についても検討し、高い性能で分類することができた。今後は、これらを利用して違法薬物売買の捜査に寄与するシステムを開発する。

## 参考文献

- [1] 警察庁, “令和元年における組織犯罪の情勢”
- [2] 安彦他, “機械学習を用いた薬物売買におけるサイバーパトロールシステムの開発”, 情報処理学会論文誌, Vol.61, No.3, pp.535-543, 2020.
- [3] 小林他, “意見抽出のための評価表現の収集”, 自然言語処理, Vol.12, No.3, pp.203-222, 2005.
- [4] 東山他, “述語の選択選好性に着目した名詞評価極性の獲得”, 言語処理学会第 14 回年次大会論文集, pp.584-587, 2008.