

サイバーパトロール活動効率化へ向けた 機械学習による投稿の有害判定モデルの開発

横山未有¹ 角田裕²

概要：近年，SNS に起因する児童被害が社会問題化している．対策として，SNS 上の有害投稿を SNS 事業者や警察に通報するサイバーパトロール活動が行われているが，参加者は限られた活動時間で多くの有害投稿を発見する必要があり，活動の効率化が必要である．そこで我々は，煩雑な作業の自動化による効率化と共に，参加者のスキマ時間での活動を促すプッシュ型配信を活用したサイバーパトロール活動促進システムの開発に取り組んでいる．このシステムは，SNS 上で自動でキーワード検索を行い，有害と思われる投稿をプッシュ配信して情報の違法性・有害性の最終的な判別を参加者に依頼する．有害投稿は，一般的な単語を隠語として用いることも多く（"円"を援助交際の"援"の隠語とするなど），隠語をキーワードとした検索では，一般的な投稿も多く収集されてしまうという課題がある．本研究では，キーワード検索により得られた投稿から一般の投稿を取り除くため，Twitter のテキストデータを利用した機械学習による有害判定モデルの開発を行った．本稿では，有害判定モデル開発にあたって考慮した点と実装方法，有害判定の結果について報告する．

キーワード：サイバーボランティア，サイバーパトロール，Twitter，機械学習

1. はじめに

スマートフォン等の容易にインターネット接続が可能な機器の普及に伴い，Social Networking Service（以下，SNS）に起因する性被害を受ける児童が増加している．警察庁が公開している統計情報[1]によれば，2021 年の SNS に起因する事犯の被害児童数は 1,812 人であった．被害児童数は 2012 年以降増加傾向にあり，現在も高止まりしている．このような現状を受け，各都道府県警察に登録されたサイバー防犯ボランティアによる児童性被害防止のためのサイバーパトロールが行われている．

サイバー防犯ボランティアとは，2010 年に警察庁の総合セキュリティ対策会議[2]で提唱された，サイバー空間における規範意識の改善を目指して活動を行うボランティアである．警察庁のマニュアル[3]によれば，活動内容として「教育活動」，「広報啓発活動」，「サイバー空間の浄化活動（サイバーパトロール）」の 3 つが示されている．2021 年末時点での全国のサイバー防犯ボランティアの団体数は 264 団体，そのうち学生団体は 154 団体であり全体の 58.3%を占める[4]．宮城県警察においても，2014 年に初めて大学生のサイバー防犯ボランティアを組織し，筆者らの所属する大学の学生からも毎年参加者が出ている．

サイバーパトロールの一環として，SNS 等に起因する児童性被害防止活動が行われている．この活動では，児童性被害に繋がる援助交際誘引等の有害投稿を探索し，通報を行う．宮城県警察のサイバー防犯ボランティア活動では，特に児童の夏季休暇周辺期間中の SNS に起因する性被害防止のために，Twitter での有害投稿通報活動を行っている．ボランティアのメンバーは各自キーワード検索等により有

害投稿を探索し，Twitter 社の通報フォーム[5]へ通報・削除依頼を行っている．ただし，有害投稿の件数が莫大である上，参加者が活動できる時間は限られているため，活動効率化の取り組みが必要とされている．

そこで筆者らは，煩雑な作業の自動化による効率化と共に，参加者のスキマ時間での活動を促すプッシュ型配信の活用に着目したサイバーパトロール活動促進システム（以下，プッシュ配信システム）の提案と開発を行っている[6]．提案システムは，SNS 上で自動でキーワード検索を行い，有害と思われる投稿をプッシュ配信して情報の違法性・有害性の最終的な判別を参加者に依頼する．ただしこのシステムには，キーワード検索により有害と思われる投稿を収集する際に実際には有害でない投稿も多く含まれてしまうという問題点があった．そこで，その問題を解決するために，機械学習による投稿の有害判定モデルの開発を行い，システムに組み込むこととした．本論文では，有害判定モデル開発にあたって考慮した点と実装方法，有害判定の結果について報告する．

2. 関連研究

2.1 サイバーパトロールに関する先行研究

先に述べたように，サイバー防犯ボランティアで現在行われている児童性被害防止のための主な活動は，Twitter での有害投稿通報活動である．Twitter は公式 API により比較的容易に情報を収集可能であり，テキスト情報が主で処理が容易であることから，サイバーパトロールに関する研究も積極的に行われている．

安彦らの研究[7]では，機械学習を用いて薬物売買を行うユーザの自動抽出・コミュニティの可視化を行っている．この研究では，Twitter 上の薬物売買を対象に，有害性の高いユーザを自動抽出し，コミュニティを包括的に把握可能なシステムを開発することで，有害ユーザのネットワーク

¹ 東北工業大学 大学院 工学研究科
Graduate School of Engineering, Tohoku Institute of Technology

² 東北工業大学 工学部 情報通信工学科
Department of Information and Communication Engineering, Tohoku Institute of Technology

の可視化を実現している。

羽田らの研究[8]では、時と共に変遷する隠語を自動的に検出する手法を提案している。この研究では、有害投稿に含まれる可能性が高い隠語、および隠語と共に出現する傾向が高い単語を推定することを目的とし、二つのコーパス間の同じ単語の用途の差異に着目することで隠語等を検出する方法について提案している。

池辺らの研究[9][10]では、サイバー防犯ボランティアのためのサイバーパトロール支援システムの開発を行っている。これらの研究では、情報収集の効率的な枠組みを構築し、Webシステムとしてサイバー防犯ボランティアに提供することで活動を支援している。

千葉らの研究[11]では、サイバーパトロールにおけるグループでの協調活動を支援する情報共有プラットフォームの開発を行った。この研究では、有害投稿を自動検索してチームで共有することにより、サイバーパトロールの効率化に貢献した。開発には、RSS フィード生成サービス Queryfeed（現在はサービス停止）とビジネスチャットサービス Slack[12]を利用した。

筆者らの従来研究[13]では、ブラウザ自動操作により有害投稿通報活動を効率化するシンプルな GUI アプリケーションの開発を行った。この研究では、Twitterの通報フォームへの入力作業を簡略化し、通報一件当たりの手間を削減した。通報フォームの自動操作には、ブラウザ自動操作フレームワーク Selenium[14]を利用した。

2.2 本研究の位置付け

前述した関連研究は、有害情報の検索作業や通報作業のそれぞれを自動化・システム化することでサイバーパトロール活動の効率化を目指している。また、参加者へ積極的な利用を促し、サイバーパトロール活動の活性化を図るようなシステムの開発は筆者の知る限り行われていない。そこで我々は、活動効率化のために有害情報の検索や通報といった一連の作業を包括的にシステム化し、参加者へ利用を促すためにプッシュ型配信を利用したシステムを開発した[6]。本研究では、そのシステムへの組み込みを目的とした、機械学習による投稿の有害判定モデルの開発を行った。なお、対象 SNS は Twitter とした。

3. 開発したプッシュ配信システムの概要

3.1 システム開発の狙い

図 1 に従来の活動の流れを示す。従来の活動では、参加者それぞれが SNS にアクセスして検索を行い、有害な投稿を探して通報を行っていた。この方法では、投稿の検索や通報の操作が煩雑で、まとまった活動時間が取れないと活動が難しいという問題点がある。また、同じ投稿を重複して通報してしまう、個人で活動するためモチベーションを保ちにくいなどの課題もある。

図 2 に提案するプッシュ配信システムの概要を示す。シ

ステムはまず、指定したキーワードに基づいて自動で投稿検索・収集を行い、収集した投稿を一定数ずつ各参加者へプッシュ配信する。参加者が、配信された投稿を閲覧し有害か無害かの判断結果をシステムへ登録すると、有害と判定された投稿が自動で通報される仕組みになっている。このシステムにより、煩雑な検索や通報の作業が自動化され、参加者はプッシュ配信された情報を見て判断を行うだけで良くなるため、まとまった活動時間が取れない場合にもスキマ時間で活動することが可能になる。また、投稿を各参加者に分配することで通報の重複を避け、プッシュ配信で依頼する形を取ることで参加者に活動の動機づけを与えることを狙う。

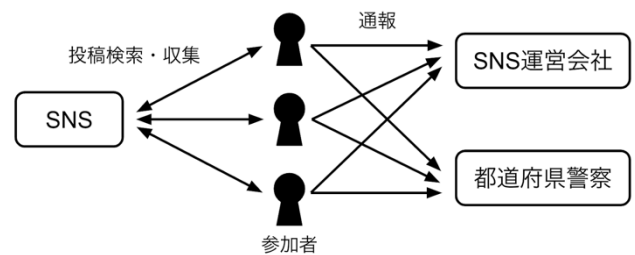


図 1 従来の活動の流れ

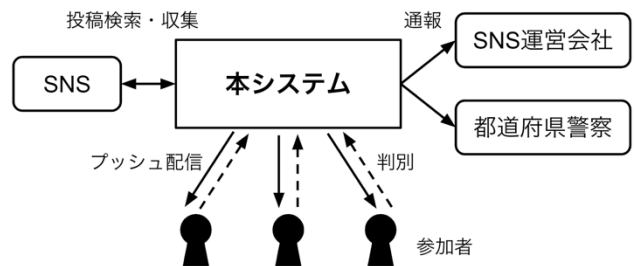


図 2 提案システムの概要

3.2 各機能の設計

図 3 にプッシュ配信システムのブロック図を示す。システムのメインとなる機能は、次の 3 機能である。

- (1) 投稿収集機能
- (2) プッシュ配信機能
- (3) 自動通報機能

ここで、メインとなる機能とそれに付随する各機能について詳しく説明する。

(1) 投稿収集機能

投稿収集機能では、キーワード検索により SNS 上から有害と思われる投稿を収集する。ただし、有害な投稿では隠語として一般的な単語を用いることが多いため、収集される投稿には実際には有害ではない投稿（ノイズ）も多く含まれている。例えば、援助交際を意味する隠語として「円」

があるが、これをキーワードとして検索すると為替相場についての投稿なども多く収集されてしまう。そのため、ノイズ除去機能により明らかに有害ではない投稿を取り除く必要がある。

(2) プッシュ配信機能

プッシュ配信機能では、投稿データを見やすい形に成形し、各参加者へプッシュ配信する。参加者が判別結果を返すと、判別結果集約機能により結果が集約される。

(3) 自動通報機能

自動通報機能では、集約された判別結果から通報すべき投稿を抽出し、SNS 運営会社と警察へ自動で通報する。SNS 運営会社への通報は専用の通報フォーム[5]を利用し、警察への通報はメールにて行う。

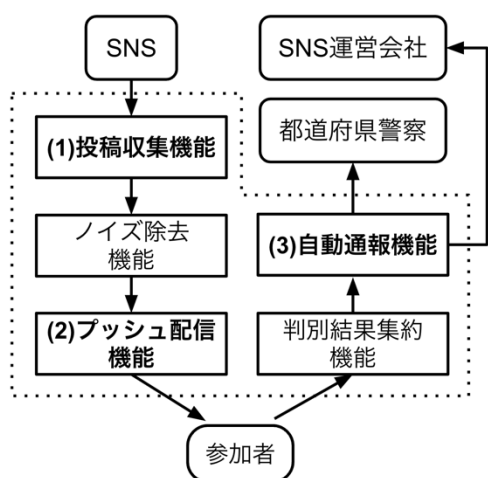


図 3 プッシュ配信システムのブロック図

3.3 有害判定モデルの必要性

キーワード検索により収集される投稿に、実際に有害ではない投稿がどの程度含まれているか確認することを目的に、調査を行った。まず、システムでキーワードとして利用する隠語・ハッシュタグを用いて、Twitter 上で投稿の検索・収集を行った。検索に用いた 9 つの隠語・ハッシュタグを表 1 に示す。

表 1 キーワードとする隠語・ハッシュタグ

えん	サポ	円
援	#jk	#えん
#サポ	#円	#援

なお、これらの隠語・ハッシュタグは、サイバー防犯ボランティア参加者に配布される活動マニュアルに記載の隠語一覧を参考に決定した。投稿の検索・収集には Twitter の公式 API を利用し、投稿の ID・本文・日時や投稿者のユーザー名・プロフィールなどのデータを収集した。調査対象は 2022 年 2 月 1 日に収集した投稿とし、投稿は研究室内の

ボランティアの協力のもと、有害投稿か有害でない投稿かに手動で分類した。2022 年 2 月 1 日に収集した投稿は全部で 20,666 件あったが、全てを手動で分類するとボランティアの負担が大きくなるため、後にモデルの学習に使うことも鑑みて有害投稿が 200 件程度現れるのを目安に分類作業を行った。分類の結果を図 4 に示す。

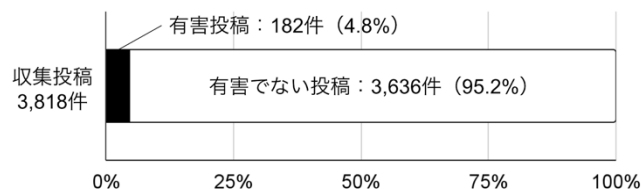


図 4 キーワード検索により収集した投稿の内訳

図 4 から、分類を行った 3,818 件の投稿のうち有害な投稿は 4.8%しか含まれておらず、有害ではない投稿が 95.2%を占めていることが確認できる。有害でない投稿がプッシュ配信に多く含まれてしまうと、その分有害投稿が含まれにくくなってしまいます。また、プッシュ配信に対応しても有害投稿の発見・通報に繋がらないと参加者が感じてしまい、モチベーションの低下にもつながる恐れがある。そのため、有害ではない投稿（ノイズ）を除去する必要がある。ただし、ノイズ除去を行うことで本来プッシュ配信されるはずだった有害投稿が除かれてしまうことは避けたい。したがって、ノイズ除去機能は、以下の要件を満たす必要がある。

- 有害投稿が取り除かれることがない
- ノイズをなるべく多く取り除く

これらの要件を満たしたノイズ除去機能を実現するために、本研究では機械学習による投稿の有害判定モデルを開発することとした。

4. モデルの開発手法

有害判定モデルの開発には、利用できるライブラリ等の豊富さから Python を用いることとした。

4.1 前処理

モデルの学習に利用するデータの前処理については、以下の手順で行った。

(1) キーワードフィルタ

3.3 の調査において投稿の分類を行った際、有害ではない投稿にのみ多く含まれる単語の存在が確認された。そこで、有害ではない投稿 3,636 件に多く確認された 8 つの単語について、その単語が投稿本文またはプロフィールに含まれる投稿数・有害投稿数の調査を、3.3 で分類した 3,818 件の投稿を対象に行った。調査の結果を表 2 に示す。

表 2 から、これらの単語は有害でない投稿にのみ多く含まれていることが確認できる。よって、これらの単語が投稿本文またはプロフィールに含まれる投稿は有害投稿では

ないとみなし、キーワードフィルタにより取り除く。

表 2 有害ではない投稿に多く見られた単語

単語	単語が含まれる投稿の数	単語が含まれる有害投稿の数
Amazon	87	0
楽天	471	0
ふるさと納税	263	0
Bot	220	0
為替	262	0
ドル円	570	0
円安	88	0
円高	82	0

(2) テキストのクリーニング

テキスト中に含まれる HTML タグ・改行文字・タブ文字など、学習に関係のない文字を除去する。

(3) 形態素解析

形態素解析ライブラリ janome[15]を利用して、投稿の文章を形態素に分割する。デフォルトの辞書による解析では新語に対応しておらず、今回対象とする投稿に多く含まれる「パパ活」などの単語がうまく分割されなかったため、そのような単語を辞書に登録した上で分割を行った。辞書に登録する単語を決める際には、事前に収集した有害投稿の本文に対しデフォルトの辞書による形態素解析を行い、適切に分割されない単語を確認した。辞書に登録した 12 個の単語を表 3 に示す。

表 3 形態素解析の辞書に登録した単語

p 活	p 活	P 活
ぴかつ	びーかつ	パパ活
ばば活	ばばかつ	ホ別
円募	えんぼ	身バレ

また、モデルの学習には、分割した形態素のうち名詞のみを用いることとした。

(4) 単語のベクトル表現

文字列を機械学習アルゴリズムに読み込ませやすい形にするために、単語をベクトルに変換する処理を行う。単語のベクトル変換には、Bag of words (BoW) という形式を用いた。BoW は、単語の出現回数のみを考え、語順は考慮せずに文書をベクトル変換する形式である。BoW において単語の出現回数を考える際には、単語の出現有無のみを考える one-hot 表現と単語の出現回数を数え上げる Count 表現とがある。有害投稿の中で同じ単語をハッシュタグ等で繰り返し記述する場合は確認されたため、本研究では Count 表現を利用しベクトル化を行うこととした。

4.2 機械学習アルゴリズム

機械学習のアルゴリズムの実装には、機械学習ライブラリ scikit-learn[16]を利用する。有害判定モデルでは、有害投

稿とノイズとに分類を行う必要がある。そのため、機械学習のアルゴリズムとしては、二値分類を行うアルゴリズムの 1 つであるロジスティック回帰を利用することとした。ロジスティック回帰では、結果はクラスラベルそのものではなく各クラスに該当する確率が出力され、確率が閾値以上であるか閾値未満であるかによって二値分類を行う。

分類の閾値は、3.3 で示した 2 つの要件を満たすように決定する必要がある。そこで、閾値の選定基準としては適合率 (Precision) と再現率 (Recall) を利用することとした。ここで、適合率はモデルが有害と予測した投稿のうち実際に有害な投稿の割合を、再現率は実際に有害な投稿のうちモデルが有害と予測した投稿の割合を表す。「有害投稿が取り除かれることがない」という要件を満たすためには、再現率は 1 である必要がある。また、「ノイズをなるべく多く取り除く」という要件を満たすためには、適合率はなるべく高くする必要がある。なお、適合率と再現率はトレードオフの関係にあるため、再現率 1 の条件下で適合率がなるべく高くなるように閾値を決定する必要がある。

5. 学習と評価

5.1 学習データの偏り

3.3 で分類した投稿のデータを利用し、モデルの学習を行う。ただし、この投稿データは有害でない投稿の方が多く不均衡なものであるため、これを全て学習に用いると有害でない投稿に分類されやすいモデルになってしまう危険性がある。そこで、データの不均衡をなくすため有害でない投稿を有害投稿の数に合わせアンダーサンプリングし、有害投稿と有害でない投稿それぞれ 182 件ずつ、計 364 件の投稿を学習に利用した。

モデルの学習後、過学習や未学習が起きていないか、交差検証を用いて学習曲線を描くことにより確認を行った。描いた学習曲線を図 5 に示す。

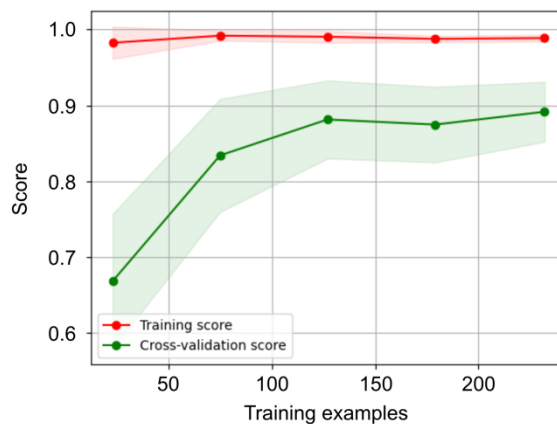


図 5 学習曲線 (学習データ : 364 件)

図 5 を見ると、サンプル数 (Training examples) が増えても訓練データの予測精度 (Training score) が 1 に近い値

のままとなっていることから、過学習が起きている可能性があることが読み取れる。原因としては、学習に用いる投稿データの総数・バリエーションが少ないことが考えられる。投稿データのバリエーションについては、有害投稿を行う投稿者が同じ日に同じような投稿を繰り返す場合があるため、ある1日（2022年2月1日）の投稿のみの利用では不十分であったと考えられる。

また、正則化のパラメータを最適化しなかったことも過学習の原因の一つと考えられる。scikit-learnによるロジスティック回帰では過学習防止のために正則化が行われているが、正則化のためのパラメータがデフォルト値のままでは最適な正則化を行うことができない。正則化のパラメータとしては、正則化の方法（L1正則化、L2正則化）を決定するpenalty、正則化の強さを決定するCがある。なお、Cの値は小さいほど正則化が強いことを示し、デフォルトではpenaltyがl2（L2正則化）、Cが1となっている。

そこで、2022年6月1日・7月1日に収集した投稿についても手動で分類を行い、投稿データの総数とバリエーションを増やした上で再度学習を行った。有害投稿と有害でない投稿それぞれ1,051件ずつ、計2,102件の投稿を学習に利用した。また、グリッドサーチによる正則化のパラメータの最適化も行った。最適化の結果、penaltyはl2（L2正則化）、Cは1.3となった。

モデルの学習後、再度学習曲線を描き確認を行った。描いた学習曲線を図6に示す。

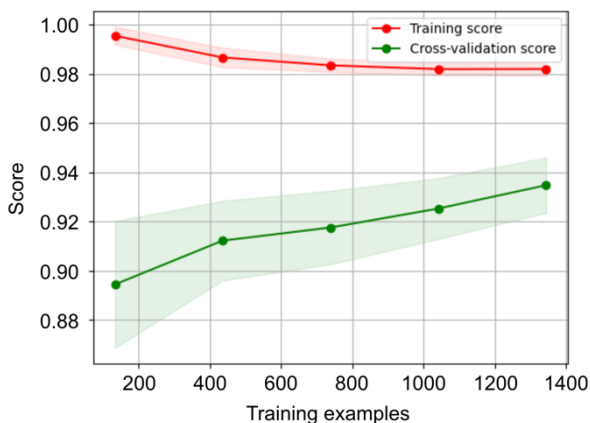


図6 学習曲線（学習データ：2,102件）

図6を図5を比較すると、図6の方がサンプル数（Training examples）の増加に従って訓練データの予測精度（Training score）と評価データの予測精度（Cross-validation score）とがより漸近していることが確認できる。従って、過学習が改善されたといえるため、ノイズ除去機能には、この学習済みモデルを利用することとした。

5.2 閾値の選定

4.2で述べた通り、ノイズ除去機能の要件を満たすためには再現率1の条件下で適合率が高くなるように

閾値を決定する必要がある。そこで、閾値の選定には適合率-再現率曲線を利用することとした。5.1で学習させたモデルに対し、2022年8月1日に収集し手動で分類を行った4,002件の投稿（うち有害投稿311件、有害でない投稿3,691件）を利用して適合率-再現率曲線を描いた。描いた適合率-再現率曲線を図7に示す。なお、閾値による適合率、再現率の変化がわかりやすいように、閾値0から1まで0.05刻みで点をプロットしている。

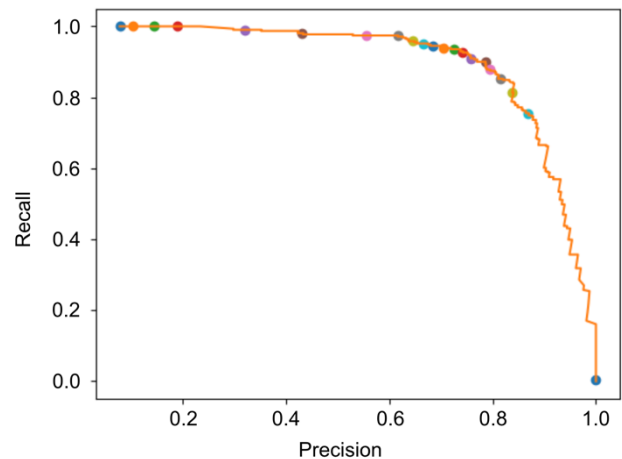


図7 適合率-再現率曲線

図7から、閾値0.15までは再現率が1であること、閾値0.15~0.20の間にも再現率が1である部分が見られることが読み取れる。つまり、適切な閾値はこの範囲内に存在すると考えられる。より細かい確認を行うため、閾値が0.15より大きく0.20より小さい範囲で、0.01刻みで適合率・再現率を算出した。算出した結果を表4に示す。

表4 適切な閾値の探索（0.01刻み）

閾値	適合率 (Precision)	再現率 (Recall)
0.16	0.20	1.00
0.17	0.21	1.00
0.18	0.23	1.00
0.19	0.30	0.99

表4から、再現率が1の条件下で適合率が最も高くなる閾値は0.18であることがわかる。従って、分類の閾値には0.18を利用することとする。

ただし、閾値を0.18とした場合、適合率が0.23と低くなっている点に注意が必要である。つまり、有害と判別された投稿のうち実際に有害な投稿は23%しか含まれず、残り77%はノイズが残ったままとなる。そこで、ロジスティック回帰では各クラスに該当する確率が出力されることを利用し、有害投稿に該当する確率（有害率）による投稿の並べ替えを行い、有害率の高いものから優先的にプッシュ配信するようにすることで対応を行った。

5.3 モデルの効果確認

開発したモデルにより実現したノイズ除去機能をプッシュ配信システムに組み込み、効果確認を行った。プッシュ配信システムの運用期間は2022年7月2日～2022年9月30日であり、ノイズ除去機能の組み込みは2022年8月26日に行ったため、利用者による判別結果をノイズ除去機能の組み込み前後で比較することによりモデルの効果を確認する。なお、対象とする利用者はノイズ除去機能の組み込み前からシステムを利用していた16名とする。また、プッシュ配信システムでは「有害」・「有害でない」の選択肢以外に、投稿が既に削除されていた場合に選択する「削除済み」、有害かどうか判断できなかった場合に選択する「保留」の項目も用意した。ノイズ除去機能の組み込み前後での判別結果の内訳を図8に示す。

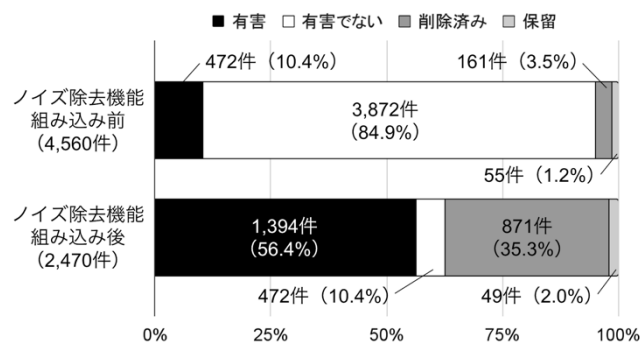


図8 ノイズ除去機能組み込み前後での判別結果の内訳

図8から、「有害」が選択される割合について、ノイズ除去機能組み込み前は10.4%だが、ノイズ除去機能組み込み後には56.4%に増加していることがわかる。すなわち、開発したモデルを利用したノイズ除去機能により、有害投稿の割合が約5倍になったといえる。また、「削除済み」が選択される割合についても、ノイズ除去機能組み込み前は3.5%だが、ノイズ除去機能組み込み後には35.3%まで増加していることがわかる。Twitterにおける有害投稿は、アカウントの凍結などを防ぐために短時間で削除されることが多い。そのため、ノイズ除去機能で有害である可能性が高いとされた投稿が、利用者が判別する頃には既に削除されているというケースが多かったと考えられる。

6. おわりに

本研究では、サイバーパトロール活動の効率化・活性化のために開発したプッシュ型配信を利用した活動促進システムに対して組み込むことを目的とした、機械学習による投稿の有害判定モデルの開発を行った。開発の際には、対象とする投稿の特徴に応じた前処理や、システムの要件を満たす適切な閾値の選定などの工夫を凝らした。開発したモデルを利用したノイズ除去機能により、プッシュ配信される有害投稿の割合が約5倍になることを確認した。

今後の課題として、モデルの定期的な更新が挙げられる。現状、モデルの更新を行う仕組みはシステムに組み込まれていない。新しい有害投稿にも柔軟に対応していくためには、新しい学習データでモデルの再学習、更新を定期的に行う仕組みが必要である。また、今回は児童性被害防止のための通報活動を対象に有害投稿判別モデルの開発を行ったが、サイバー防犯ボランティアでは他にも様々な有害情報に対して検索・通報活動を行っている。そのため、今後は他の活動への機械学習の活用も期待される。

謝辞 本研究は、東北工業大学学内公募研究(2022-14)の援助により行われたものである。ここに記して謝意を表す。

参考文献

- [1] “令和3年における少年非行、児童虐待及び子供の性被害の状況”. <https://www.npa.go.jp/bureau/safetylife/syonen/pdf-r3-syonen-hikoujyokyo.pdf>, (参照 2022-09-29).
- [2] “総合セキュリティ対策会議”. <https://www.npa.go.jp/cyber/csm-eting/index.html>, (参照 2022-09-29).
- [3] “サイバー防犯ボランティア活動のためのマニュアル(モデル)”. <https://www.npa.go.jp/cyber/policy/volunteer/manual.pdf>, (参照 2022-09-29).
- [4] “令和3年末におけるサイバー防犯ボランティア団体数等”. https://www.npa.go.jp/cyber/pdf/r3_volunteer.pdf, (参照 2022-09-29).
- [5] “CSE”. <https://help.twitter.com/ja/forms/safety-and-sensitive-content/cse>, (参照 2022-09-29).
- [6] 横山未有, 角田裕. プッシュ型配信を活用したサイバーパトロール活動促進システムの検討. FIT2022 第21回情報科学技術フォーラム. 2022, O-007.
- [7] 安彦智史, 加藤諒, 北川悦司. 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. 2020, Vol. 61, No. 3, p. 535-543.
- [8] 羽田拓朗, 清雄一, 田原康之, 大須賀昭彦. コーパス間での類似語の差異に着目したマイクロブログにおける隠語検出. 情報処理学会研究報告. 2020, Vol. 2020-ICS-200, No. 2, p. 1-8.
- [9] 池辺正典. サイバー防犯ボランティアのためのサイバーパトロール支援システムについて. 警察学論集. 2018, Vol. 71, No. 5, p. 146-158.
- [10] 池辺正典, 川合康央, 櫻井淳. SNSを対象とした都道府県警察のサイバーパトロール支援システムについて. 情報教育シンポジウム論文集. 2021, Vol. 2021, p. 196-201.
- [11] 千葉翔也, 角田裕. 児童被害防止のためのSNSの有害コメント収集プラットフォームの開発. 情報処理学会第81回全国大会講演論文集. Vol. 2019, No. 1, p. 35-36.
- [12] “SlackをあなたのDigital HQに | Slack”. <https://slack.com/intl/ja-jp/>, (参照 2022-09-29).
- [13] 横山未有, 角田裕. サイバー防犯ボランティアの通報活動を効率化するシンプルなGUIアプリケーション. 2021年度電気関係学会東北支部連合大会. 2021, 3C04.
- [14] “ブラウザー自動化プロジェクト | Selenium”. <https://www.selenium.dev/ja/documentation/>, (参照 2022-09-30).
- [15] “Welcome to janome's documentation! (Japanese) — Janome v0.4 documentation (ja)”. <https://mocabeta.github.io/janome/>, (参照 2022-10-26).
- [16] “scikit-learn: machine learning in Python — scikit-learn 1.1.3 documentation”. <https://scikit-learn.org/stable/#>, (参照 2022-11-1).