

株式会社サイバーエージェント様

自然文解析エンジンが不適切な投稿を高速・高精度で検知 さらに安心・安全なメディアサービスを実現

商品名 PRESS/EX V10 フィルタリングエンジン r2

課題

- 従来のフィルタリングエンジンでは監視の難しい不適切投稿への対処
- 膨大な投稿件数に対処する解析スピード
- 有人逐次監視におけるコストの増加

効果

- 自然文解析エンジンによって検知精度が向上
- 0.01 秒台の高速処理を実現（※実証実験時）
- 監視人員を減らすことによるコスト低減に期待

インターネット業界屈指のサイバーエージェントは、更なる安心・安全なメディアサービスの運営を図るため、自社開発した既存の総合監視基盤システムに Zinrai を活用した自然文解析エンジンを採用。富士通独自の「文の構造や文脈を学習した解析技術」によって、従来のエンジンでは検知が難しかった不適切な投稿も検知可能となり、今後の監視精度向上に期待を寄せている。高速の解析処理も実現し、さらに有人逐次監視を減らすことによる人的コストの削減も見込まれる。

背景

独自の総合監視基盤システムを開発し 安心・安全なメディアサービスの運営に尽力

いまやサービス開始から 14 年を迎え、会員数が六千万人以上となった「Ameba」やインターネットテレビ局「AbemaTV」など、1998 年の創業以来、インターネット産業の変化にあわせて様々なメディアサービスを提供し続けているサイバーエージェント。

「これまで時代に先駆けた様々なサービスを生み出してきましたが、安全に誰にでも気持ちよくサービスを利用してもらうことも、私たちに課せられた使命の一つです」と話すのは、サイバーエージェント秋葉原ラボ研究室長 福田一郎氏。

インターネットの幅広い世代への普及により、インターネットサービスを悪用する犯罪も発生している。特に未成年や高齢者をターゲットに

した犯罪は悪質なものもある。その他にも違法な詐欺広告や特定の個人を攻撃する誹謗中傷など、対策が求められる問題も多い。こうした状況にサービス提供側は日々対策を講じているが、同社はこれまでインターネット事業で蓄積してきた先進のノウハウを活用することでメディアサービスのサイト健全化に尽力している。その中心として機能しているのが秋葉原ラボとカスタマーサポート主導で開発された独自の総合監視基盤システム「Orion（オライオン）」だ。

秋葉原ラボはメディア事業の大規模データを集約し、研究開発を行う組織として 2011 年に開設され、データの処理や分析、機械学習などを専門とするエンジニアが所属している。Orion はモバイルコンテンツ審査・運用監視機構「EMA*」が規定する自社サービスの常時監視をクリアすることの他に、「ブログのコメント監視、投稿画像の監視といった目的別に 5 つあった監視システムを一つに統合し、内製化することでコスト低減につなげるために開発されました」と福田氏は話す。

監視システムが統合・内製化されたことで、投稿の監視は以前に比べ

お客様プロフィール
株式会社サイバーエージェント
本社所在地 東京都渋谷区道玄坂一丁目 12 番 1 号

設立 1998 年 3 月 18 日

ホームページ <https://www.cyberagent.co.jp/>
代表者 代表取締役社長 藤田 晋

事業概要 「21 世紀を代表する会社を創る」というビジョンを掲げ、インターネット広告事業、メディア事業、ゲーム事業を柱にインターネット領域にて環境変化に柔軟に対応しながら展開。


株式会社サイバーエージェント
技術本部 チーフアーキテクト
秋葉原ラボ 研究室長
福田 一郎 氏



株式会社サイバーエージェント
技術本部 秋葉原ラボ
エンジニア
藤坂 祐介 氏

て効率良く行えるようになった。しかし「それですべての課題が解決したわけではありません」と福田氏。

*EMA：青少年の保護と健全な育成を目的に Web サイト及びアプリケーションの運用管理体制の審査・認定および啓発・教育活動を行う

経緯

検知できない不適切投稿の対処と人による逐次監視のコスト低減

現在の監視基盤システムにおいては、自動的な投稿抽出と有人監視（人の目による監視）の二重のチェックを通した上で高い監視精度を維持している。このとき、有人監視に回す投稿を抽出するための方法が、システム全体の監視精度や監視効率を大きく左右する。投稿抽出のためにいくつかの手法を提供しているが、広く使われているものは投稿されたデータから特定のワード（「禁止ワード」）が使われていた場合に、自動的に監視対象を抽出した上で、最終的に人の目によって非表示にするかを判断する方法だ。

「監視ツールに登録する禁止ワードは時代とともに変化するため、常にブラッシュアップが必要です」と話すのは秋葉原ラボ エンジニア 藤坂祐介氏。

また、監視精度の課題として「禁止ワードは含まないが内容的には問題のある」不適切なニュアンスの投稿への対処も挙げられる。

「例えば事件・犯罪を発生させる可能性のある内容でも、あからさまな禁止ワードを使っていない投稿や、アクセス数を稼ぐためのワードサラダ（文法は間違っていないが、文章として意味が破綻している）やスパム（無差別に大量にばらまかれるメッセージ）などです」（藤坂氏）

もちろん、サービスを悪用する側の事例や傾向を掴むことで対策も採られているが、あからさまな禁止ワードを持たない性質上、スルーされる可能性は高い。結果、高い監視精度を保つためには、発見される度に禁止ワードリストを更新していく他なかった。

「フィルタリングの精度を高めていくこと以外にも、解析スピードを上げることで監視作業の効率化を図る、有人監視の負担を減らすことで全体のコストを削減するといった課題もありました」と福田氏。「有人監視は新サービス開始時など、多い時には100人態勢で作業にあたります。新たなサービスは次々と生み出されますし、もちろん24時間体制なので、人的コストの低減は大きな課題でした。そうした中、富士通さんからZinraiを活用した自然文解析エンジンのご提案をいただきました」（福田氏）

ポイント

自然文解析エンジンが文章の構造や文脈を考慮して不適切な投稿を検知

富士通は現在、営業やSEが最先端の技術開発を行う富士通研究所と共に、自社の最新技術とノウハウをお客様に提案する新しいビジネススキームを積極的に推進している。今回、同社に提案した「Zinraiを活用した自然文解析エンジン」は、この活動の成果でもある。

まず学習用の膨大なデータをエンジンに学習させ、解析モデルを生成。そのモデルを用いて検証用データを解析・評価し、エンジンの精度チェックを行うテストを繰り返した。学習用データと検証用データに秋葉原ラボから提供されたブログの生データを使用することで、より実業務に近い形でのテストが可能になった。また、富士通と秋葉原ラボの両環境で検証を実施することで、技術者たちが顔を突き合わせて活発な議論や意

見交換を行うことができ、結果として、高い解析精度を得ることができた。ところが、同社からの回答は不採用。

「問題は解析スピードでした。通常私たちが行っている監視作業は、投稿された時点で逐次的に行うため、一件単位の処理速度が大事です。また、取り扱う件数も膨大なので、一秒遅いだけでも全体では大きなロスになります」（藤坂氏）

その後、逐次的な処理にもスムーズに対応できるようプログラムの実行手法を変更し、処理速度を改善。2018年3月、ついに採用にこぎつけた。「Zinraiを活用した文字解析エンジンの一番の特長は、自然文をチェックできる点です。機械学習によって文章の構造や文脈を学習して、そこから獲得したルールを基に判別を行います」（藤坂氏）

そのため従来のエンジンとは異なり、あらかじめ禁止ワードのリストを設定しなくても文脈を考慮して解析を行い、不適切なニュアンスの投稿を検知することが可能だ。

従来のエンジン	文字解析エンジン
<p>登録した禁止ワードを検知</p> <p>禁 「殺害」「薬物」「犯罪者」...など</p> <p>投稿内容</p> <p>〇〇ちゃんを殺害したい。ちゃんと守っておかないと大変なことになりますよ。。。</p>	<p>文の構造を考慮したルールを基に判別</p> <p>ルールファイル ニュアンスにて不適切判定</p> <p>投稿内容</p> <p>我慢なりません。たぶん明日〇〇ちゃんを刺しにいきます。覚悟しておいてね。KHより</p>

効果と今後の展望

高速、高検知率を記録 実証実験で得た十分な手応え

実証実験で得た結果は、今後に大きな期待を抱かせる。「実際に不適切な投稿が占める割合は、全投稿数の1%もありません。ことわざで“藁の中から針を探す”と言いますが、今回、実証で得た検知率を見ると、富士通の解析エンジンが非常に優秀な監視スキルを持っていることが立証されたと思います」と福田氏。

実証実験では、既存のエンジンで検知できなかった不適切な投稿に対して、94%という高い検知率を記録。またその際に要した時間はわずか0.01秒台。監視精度の向上とともに高速解析処理も実現した。「機械学習のために必要なデータを大量に保持していることが我々の強みですが、今後も日々データを収集、それをシステムが学習するというループを富士通さんと協力して繰り返していくことで、さらなる監視精度の向上が見込まれます」（福田氏）

また、「解析エンジンが不適切な投稿を検知してくれる割合が増すことで、有人監視の負担も減ります。コスト低減が期待されるばかりでなく、監視に割いていた人員が他の仕事に従事できるため会社全体の仕事の効率化にも繋がります」（藤坂氏）

「今後は、このフィルタリングエンジンを当社グループ全体に展開し、さらに快適に利用いただけるメディアサービスの提供を目指します。これからも富士通さんと協力して、様々なアイデアを出し合ってビジネスを拡大していきたいと思います」（福田氏）

お問い合わせ先

富士通株式会社 デジタルソリューション事業本部 フロントビジネス事業部 デジタルソリューション部

03-6424-8044

issdept-camoji@ml.css.fujitsu.com