

クラウドソーシングサービスにおける課題改善への取組みと今後の展望

Approach to Problems Improvement in the Crowdsourcing Service and its Future Prospects

弓山 彬 福島 明 沢田 正 塚本 鋭
Akira Yumiyama Akira Fukushima Tadashi Sawada Ei Tsukamoto

株式会社クラウドワークス
CrowdWorks Inc.

This paper introduces approaches to detect violated job offer posts, and obstacles to increase matching accuracy in the crowdsourcing service. So far, regular expression matching and monitoring by operators are used to detect violated job offer posts. In crowdsourcing service, there are a lot of job offer posts which contain violations. Also, there are a lot of violation reports by end-users. The authors made some prototype softwares to detect violating job offer posts with supervised learning and existing job offer posts as training data set. This method is utilized for picking up violating posts which are not detected for now, and reduces cost for the administrators.

1. はじめに

近年、クラウドソーシングサービスを通じて、インターネット上でクラウドワーカーに対して多様な業務を委託することで、多くの仕事を実際に遂行されるようになってきている。株式会社クラウドワークスでは、「CrowdWorks」という日本最大級のクラウドソーシングサービスを運営している。

CrowdWorksでは流通する「仕事の依頼」のことを「案件」と呼び、表題と本文を中心とするテキスト情報と、仕事カテゴリなどのタグ情報で構成されている。発注者は不特定多数のクラウドワーカーに対して案件への応募を呼びかけ、クラウドワーカーはそれらの公開されている案件に対して応募した後に契約条件を交渉する。発注者とクラウドワーカーの双方が合意した条件にもとづいて、業務委託契約を結ぶことで実際の仕事を開始される。

しかし、身元の確認なしに、不特定多数の人の間で仕事が依頼されることから、法律に違反する案件や、同じような案件を大量に発注するスパム案件の投稿が後を絶たない。CrowdWorksではこのような、法律に違反する案件や、同じ案件の大量投稿といった行為は、利用規約により禁じている。そして、このような違反案件やスパム行為を検出し、適切に公開停止等の手続きを踏むことはサービス運営において非常に重要な機能となっている。

これまで、違反案件やスパム案件については、どういった規約・法律に違反するかの判断も含め、すべて人手で検出・対応を行ってきた。しかし、日々多数の仕事が依頼され続けているサービス上で、常に仕事を監視するのは非常にコストがかかるため、ユーザからの報告機能を用意している。代表的なものとして、違反報告機能や案件評価機能といった機能がある。違反報告機能は利用規約違反の報告に特化しているが、案件評価機能については単価の低さや、依頼内容が曖昧である、といった広範なフィードバックが行えるようになってきている。これらのユーザからのフィードバックを参考に、個別に違反か否かの判断を行っている。

本稿では、この違反案件の検出について、複数の手法により自動検出を試みた事例を紹介する。CrowdWorksでは、今まで前述のようにして違反と判定してきた案件情報や、ユーザからの違反報告データが全て蓄積されている。今回、これらの違

反報告や実際に違反となった案件を教師データとして学習させることで、サービス上で得られる情報から違反案件やスパム案件を自動的に検出していくことを目指す。それにより、今まで人手による検出でかかっていたコストの削減、また今までは検出できていなかった違反案件の検出を実現することができる。

2. 実験

2.1 実験の概要

本実験では2016年にCrowdWorksに投稿された案件を対象に違反案件の検出を試みた。教師データセットとしてランダム抽出した違反案件1,000件と非違反案件1,000件を用意し、検証データセットとして教師データ以外から2,000件の案件情報を用意した。

これらデータセットに対して以下の処理を適用し、各手法では同一の教師データ、検証データを用いて評価を行った。

- 案件の表題と本文を連結して一つのテキストにする
- 生成されたテキストに含まれるHTMLタグなどの不要な要素を削除
- テキスト全体をUnicode正規化

2.2 従来手法の性能評価

現状ですでにサービスに組み込まれている違反案件の仕組みとして、正規表現チェック機構(以下、「正規表現チェック」と呼称)が存在する。この正規表現チェックでは、予め「禁止ワード」を設定しておき、新たな案件が投稿された際に、その禁止ワードのいずれかに一致する文言が含まれていれば自動的に違反の可能性ありとしてマークするという処理を行っている。

正規表現チェックを、上述の検証データセット2,000件に対して適用した。結果を表1に示す。適合率0.765、再現率0.326、F値0.457となった。

適合率より、正規表現チェックで違反とされた案件のうち1/4程度が誤検知であったことが確認できた。また再現率が低い理由は、違反案件が正規表現チェック以外の方法、たとえば目視によるチェックやユーザからの指摘などで違反と判定されたものを含むためである。正規表現チェックだけでは違反案件を網羅的には拾いきれておらず、他の手段に頼っていることが確認された。

連絡先: 弓山彬, akira.yumiyama@crowdworks.co.jp

表 1: 正規表現チェックを用いた手法の評価結果

		正解	
		違反	非違反
推定	違反	326	100
	非違反	674	900

2.3 ベイジアンフィルタの評価

次に、案件のテキストをベイジアンフィルタを用いて違反案件の判定を行った。

上述のデータセットに対して、MeCab を用いた形態素解析を適用し、得られた要素の中から名詞・動詞・形容詞のみを抽出した。これら特定品詞だけの単語を抜き出してベイジアンフィルタを学習させ、上述の検証データセットに対して判定を行った。結果を表 2 に示す。

表 2: ベイジアンフィルタを用いた手法の評価結果

		正解	
		違反	非違反
推定	違反	807	77
	非違反	193	923

適合率 0.913, 再現率 0.807, F 値 0.857 であり、正規表現チェックと比較して、適合率・再現率ともに改善が見られた。

2.4 サポートベクターマシンと Word2Vec を用いた手法の評価

違反判定のアルゴリズムにサポートベクターマシン (SVM) を用い、そのベクトル生成手法として Word2Vec を用いた場合の評価を行った。

上述のデータセットに対して、MeCab を用いた形態素解析を適用し、得られた要素について「日本語 Wikipedia エンティティベクトル」[jawiki] を利用して 200 次元の単位ベクトルを作成した。単語ごとの単位ベクトルの算術平均を案件のベクトルとして、サポートベクターマシンを用いて学習・判定を行った。

結果を表 3 に示す。適合率 0.534, 再現率 0.982, F 値 0.692 という結果になった。

表 3: SVM と Word2Vec を用いた手法の評価結果

		正解	
		違反	非違反
推定	違反	982	858
	非違反	18	142

3. 考察

これらの実験の結果を表 4 に示す。最も高い F 値を示したのはベイジアンフィルタであった。メールのスパム判定に用いられることが多い同手法であるが、メール同様に表題と本文のテキストから構成されている案件データとも親和性が高いと考えられる。

従来手法である正規表現チェックとベイジアンフィルタの差異について検討する。まず、適合率については、正規表現チェッ

表 4: 各手法の評価結果

手法	適合率	再現率	F 値
正規表現チェック	0.765	0.326	0.457
ベイジアンフィルタ	0.913	0.807	0.857
SVM と Word2Vec	0.534	0.982	0.692

クが禁止ワードのいずれか 1 つにマッチした時点で違反と判定していた事に対して、ベイジアンフィルタを用いた場合は複数の語の出現率を元に判定することから、性能向上したことが分かる。再現率については、現状の違反案件判定プロセスがスタッフによる目視チェックなど複数の手法を組み合わせた複合的なものであるため、正規表現チェックで検知できない違反案件をベイジアンフィルタが検出できたことを示している。

次に、正規表現チェックと Word2Vec を用いた手法の差異について検討する。適合率については、正規表現チェックに比べて Word2Vec を用いた手法の方が低下している。これは検証データに含まれる違反案件・非違反案件の大部分を違反と判定したことに起因している。サポートベクターマシンで正しく境界面を見つけられるようなベクトルを作成できなかったためと考える。今回の評価では「日本語 Wikipedia エンティティベクトル」を用いて案件のベクトルを作成したが、CrowdWorks に投稿される案件を区別しやすい空間になっていないことが想定される。この手法を継続する場合、案件の文面の特徴を適切に表現できるようなベクトル空間を予め定義しておく必要があるだろう。

今回の実験では、既存手法 (正規表現チェック) とベイジアンフィルタを用いた方法、Word2Vec を用いた方法について、違反案件の検出性能の評価を行った。結果として、ベイジアンフィルタを用いることによって、既存手法に比べて、適合率・再現率ともに改善することができた。

4. 今後の展望

今回の実験では、案件の依頼文面の情報のみを用いて、違反案件の検出を試みた。細かな違反理由・違反箇所の特徴や、依頼者の属性や行動情報を加えて学習させることによって、より正確な違反案件の検出が可能になると考えられる。

CrowdWorks に登録しているユーザは 120 万人を超え、日々登録される案件の数も数千件に達している。このような状況においては、違反案件を検出し除外するだけでなく、クラウドワーカーの経験・スキルに合致する案件を適切にレコメンドすることが重要となってくる。さらに、仕事内容に対する契約当事者相互の評価内容を活用することによって、個人に対する信用情報を充実させ、新たな価値創造につながるものと考えている。このような領域に対する機械学習・人工知能の適用について継続して取り組んでいきたい。

参考文献

[jawiki] 日本語 Wikipedia エンティティベクトル http://www.cl.ecei.tohoku.ac.jp/~m-suzuki/jawiki_vector/