

# Searching Sexual Predators in Social Networks

Yuridiana Alemán, Darnes Vilariño, and David Pinto

Facultad de Ciencias de la Computación  
Benemérita Universidad Autónoma de Puebla, México  
yuridiana.aleman@gmail.com, darnes@solarium.cs.buap.mx, dpinto@cs.buap.mx

**Abstract.** In this paper we propose a two-step technique for detecting sexual predators from social network dialogues. One step for detecting dialogues in which a sexual predators participates, and the second step is for detecting, from the whole dialogue users, the one that is the sexual predator. From the three different supervised classifier employed, Random Forests obtained the best results in the first step, whereas Neural Networks performed best in the second step.

**Keywords:** Search, Supervised classification, Sexual predators.

## 1 Introduction

Sexual predators have found a new manner of select victims through the use of social networks. It is relatively easy for these predators to pretend to be a child or teenager with the aim of obtain the confidence of their victim. Due to this fact, there exist diverse attempts for detecting these kind of behaviors by analyzing conversations in chat rooms.

Major research works take Pendar [1] as a reference point for this topic. Here, the authors use a dataset gathered from a website named “Perverted Justice”<sup>1</sup> for conducting a study using automatic text categorization techniques for identifying online sexual predators. More recently, Villatoro[2] performed conversation filtering by removing the shortest conversations with unintelligible characters or those conversations in which the chat participants have a very low number of interventions. With this pre-processing step, it is possible drastically to reduce the amount of texts contained in the training set.

## 2 Methodology

The methodology proposed for searching/identifying sexual predators is shown in Figure 1. This proposal is made up of two steps: 1) A classification process that allows to discriminate those conversations in which a sexual predator participate; and, 2) A classification process that allows to discriminate the predator’s dialogues with respect to other participants. In both steps, we use the following classification algorithms: neural networks, random forests and decision trees.

---

<sup>1</sup> <http://perverted-justice.com>

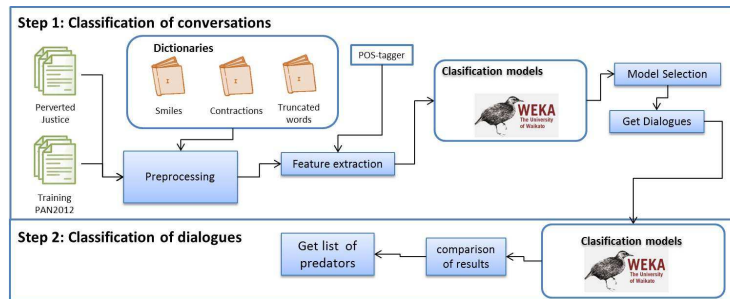


Fig. 1. Methodology for detecting sexual predator conversations

We built and used three lexical resources (dictionaries) for pre-processing step: emoticons (“:-)”) is normalized as “happy”), contractions (“isn’t” is normalized as “is not”), and SMS vocabulary (“10q” is normalized as “thank you”). Afterwards, we extracted the features using POS-tagger[3], we used every morphological feature as attribute<sup>2</sup>. For the experiments carried out, we used the union of two different conversation sets: *Perverted Justice* used in [1] and the *PAN 2012 Training set*<sup>3</sup> that contain conversations provided by the PAN 2012 conference committee, which are structured in XML format.

### 3 Experimental results

We have carried out experiments using several classification algorithms implemented in the *Weka*[4] tool. We selected the following classification algorithms that obtained the best results: Decision trees[5], Random forests[6], and Neural networks (BackPropagation algorithm).

For the evaluation of results, we use the Weka option “Use training set” for step 1 and “Cross-validation” with 10 folds for step 2. Subsequently, we obtain the Precision ( $P$ ), Recall ( $R$ ) and F-Score ( $F$ ), and we used the best  $F$ -score model for using in the second step.

Table 1 shows the results obtained for every classifier in both steps.  $TRC$  is the total of retrieved instances, and  $TCP$  is the total of positive instances. Actually, from 2,353 positive conversations of the dataset, the models identified around 2,000 conversations. The best performance was obtained by *Random Forest* with, 0.983 of precision, 0.882 recall and 0.930 of  $F$ -measure, thus, identifying very well those dialogues in which a sexual predator participates.

Using the conversations detected by the *Random Forest* classifier, we have re-constructed the dialogues in which one sexual predator participates. Thus, the second corpus contains 480 conversations of sexual predators and 442 conversations of non-sexual predators (922 conversations). In this step, *Random*

<sup>2</sup> <http://bit.ly/WHsvBN>

<sup>3</sup> <http://pan.webis.de/>

**Table 1.** Rankings obtained for classification conversations

Classifier	TRC	PCR	Precision	Recall	F-Score
Step 1: Conversations					
Decision Trees	1,146	1,110	0.968	0.471	0.634
Neural Networks	905	766	0.846	0.325	0.470
Random Forests	2,111	2,076	<b>0.983</b>	<b>0.882</b>	<b>0.930</b>
Step 2: Users					
Decision Trees	472	347	0.735	0.723	0.729
Neural Networks	516	399	<b>0.773</b>	0.831	<b>0.801</b>
Random Forests	538	402	0.747	<b>0.838</b>	0.790

*Forests* retrieved more predators than *Neural Networks* did, but it also retrieved false positives. Considering the *F*-score measure, then the best result is obtained using *Neural Networks* (399 of 480 predators retrieved).

## 4 Conclusions and future work

We presented a two step system for detecting sexual predators on-line. The conversation representation using PoS tags allowed to identify terminology employed by sexual predators, as shown by the values obtained in the experiments.

The normalization of texts have had a high impact in the results obtained and, need to be further investigated. Additionally, we are interested on analyzing new features that allow us to detect “all” the conversations in which a sexual predators participates.

## References

1. Pendar, N.: Toward spotting the pedophile telling victim from predator in text chats. In: Proceedings of the International Conference on Semantic Computing. ICSC '07, Washington, DC, USA, IEEE Computer Society (2007) 235–241
2. Villatoro-Tello, E., Juárez-González, A., Escalante, H.J., y Gómez, M.M., Pineda, L.V.: A two-step approach for effective detection of misbehaving users in chats. In: CLEF (Online Working Notes/Labs/Workshop). (2012)
3. Toutanova, K., Klein, D., Manning, C.D., Singer, Y.: Feature-rich part-of-speech tagging with a cyclic dependency network. In: Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology - Volume 1. NAACL '03, Stroudsburg, PA, USA, Association for Computational Linguistics (2003) 173–180
4. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The weka data mining software: an update. SIGKDD Explor. Newsl. **11**(1) (November 2009) 10–18
5. Quinlan, J.R.: C4.5: Programs for Machine Learning (Morgan Kaufmann Series in Machine Learning). 1 edn. Morgan Kaufmann (October 1992)
6. Breiman, L.: Random forests. Mach. Learn. **45**(1) (October 2001) 5–32