Natural Language Processing for Safer Radiotherapy

Hui Wang[1], Erick Velazquez[2], Logan Montgomery[2], John Kildea[2]

[1]School of Computer Science, McGill University, [2]Medical Physics Unit, McGill University

Background

- Reporting and investigating unintended incidents and accidents, as shown in Figure 1, can help prevent them from recurring.
- Reporting an incident involves describing it in free-form text, which is difficult to analyze in large amounts.

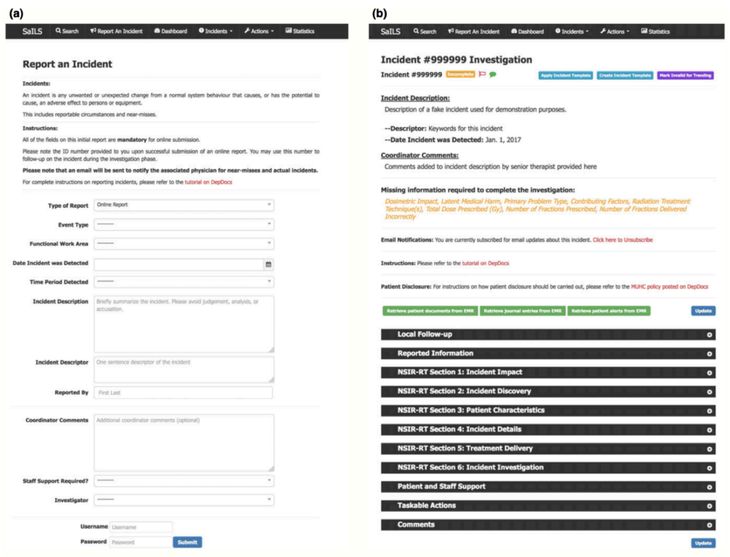

Figure 1: Reporting and investigation interface of the Safety and Incident Learning System[1] as implemented at the McGill University Health Centre in Montreal

Objectives

1 Help label incidents by their process step, problem type, contributing factors, and overall severity using supervised learning.
- Determine the probability of each label being correct.
- Order the labels by probability in a dropdown menu.

2 Discover new problem types and contributing factors using unsupervised learning.
- Group together similar incident descriptions labelled as "other."
- Interpret the groups as possibly new problem types and contributing factors.

Methods

- Natural Language Preprocessing
  - Preprocessed the incident descriptions through machine translation, abbreviation expansion, auto-correction, and linguistic normalization.
  - Generated descriptive statistics that helped to specify the parameters of the word embedding models used in subsequent steps.

1 Supervised Learning
- Assembled over 200 multi-output ensemble regressors by bagging, boosting, and stacking over 60 regressors from Sklearn[2], XGBoost[3], and Keras[4].
- Tuned the ensemble regressors using grid-search, and cross-validated them by how close to the top they placed the correct labels in the dropdown menu.

2 Unsupervised Learning
- Tuned a latent Dirichlet allocation model using ldatuning[5]'s multi-metric[6-9] grid-search over the number of groups to best fit the incident descriptions labelled as "other," as shown in Figure 2.
- Interpreted the groups of similar incident descriptions as possibly new problem types and contributing factors in an interactive visualization by LDAvis[10].

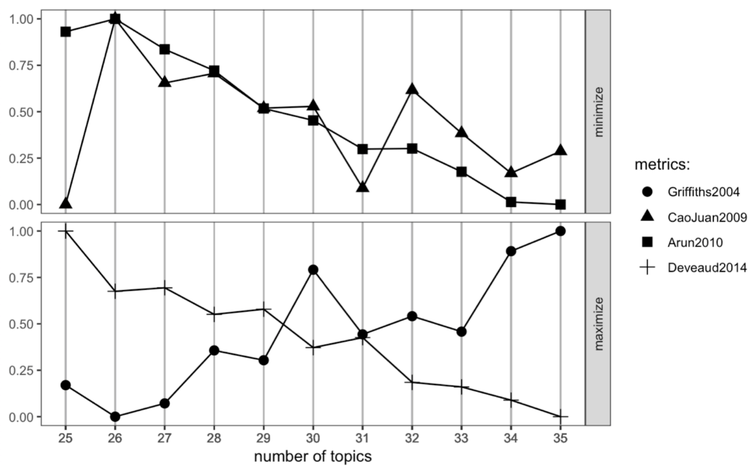

Figure 2: Multi-metric grid-search reveals 30 possibly new problem types or contributing factors in the incident descriptions labelled as "other."

Results

- Natural Language Preprocessing
  - The average incident description is 3.01±3.05 sentences long.
  - 20 incident descriptions are in French.
  - 34.82%, 40.32%, and 7.18% of the incidents have N/A or "other" for process step, problem type, and contributing factors, respectively.

1 Supervised Learning
- The best ensemble regressor was assembled by bagging and stacking a ridge regressor, a support vector regressor, a passive aggressive regressor, a random forest regressor, a gradient boosting regressor, and a long short-term memory recurrent neural network.
- The best ensemble regressor placed the correct labels on average 1.6th, 2.4th, 5.3rd, and 1.1th among the 9 process step, 24 problem type, 21 contributing factors, and 4 overall severity labels, respectively.

2 Unsupervised Learning
- 30 groups best fitted the incident descriptions labelled as "other."
- The interactive visualization of the groups showed which incidents recurred the most, and how they are related to other incidents, as shown in Figures 3-4.

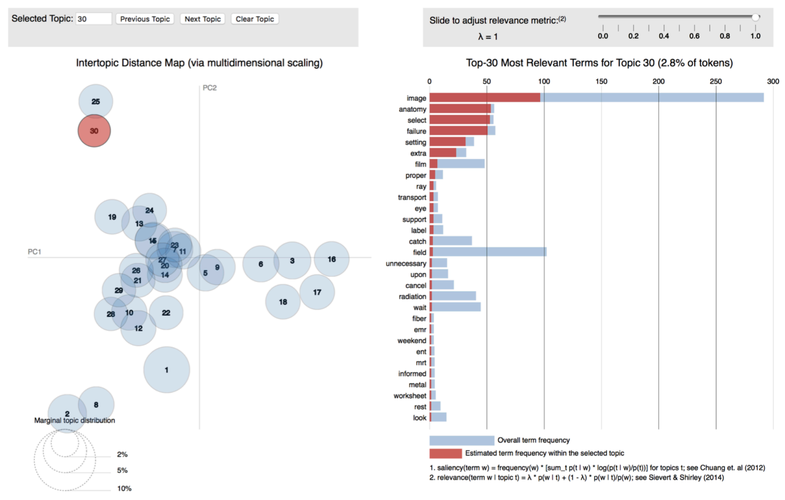

Figure 3: Group 30 reveals that extra images were taken due to failure to select anatomy settings.

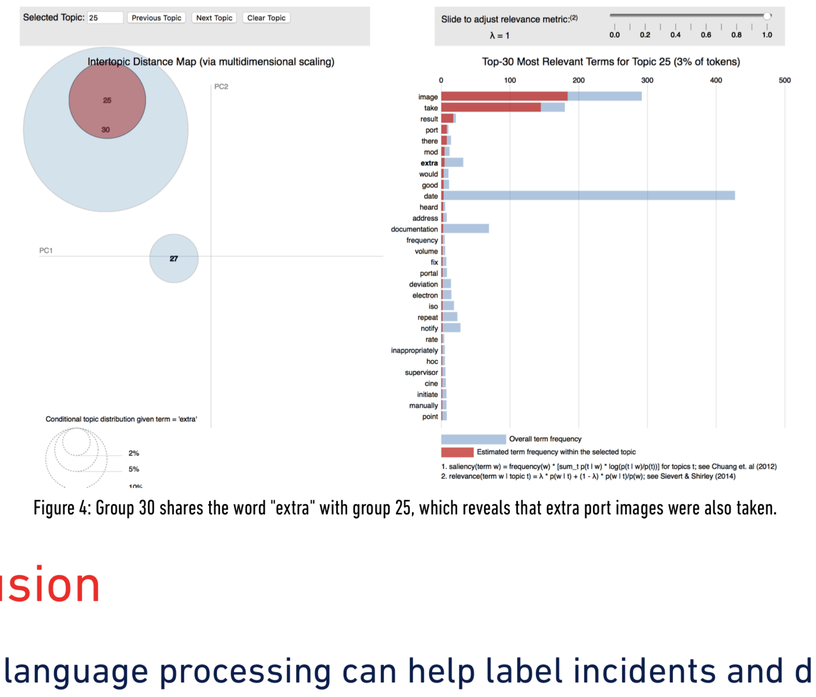

Figure 4: Group 30 shares the word "extra" with group 25, which reveals that extra port images were also taken.

Conclusion

- Natural language processing can help label incidents and discover new problem types and contributing factors.
- More generally, it enables the comprehension of large numbers of incidents without investigating each.

References

1 Montgomery, Logan, et al. "Development and implementation of a radiation therapy incident learning system compatible with local workflow and a national taxonomy." Journal of applied clinical medical physics 19.1 (2018): 259-270.

2 Pedregosa, Fabian, et al. "Scikit-learn: Machine learning in Python." Journal of machine learning research 12.Oct (2011): 2825-2830.

3 Chen, Tianqi, and Carlos Guestrin. "Xgboost: A scalable tree boosting system." Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. ACM, 2016.

4 Chollet, François. "Keras: Deep learning library for theano and tensorflow." URL: https://keras. io/k 7.8 (2015).

5 Murzintcev, Nikita. "Ldatuning: Tuning of the latent dirichlet allocation models parameters." R package version 1 (2014).

6 Arun, Rajkumar, et al. "On finding the natural number of topics with latent dirichlet allocation: Some observations." Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer, Berlin, Heidelberg, 2010.

7 Cao, Juan, et al. "A density-based method for adaptive LDA model selection." Neurocomputing 72.7-9 (2009): 1775-1781.

8 Deveaud, Romain, Ludovic Bonnefoy, and Patrice Bellot. "Quantification et identification des concepts implicites d'une requête." CORIA. 2013.

9 Griffiths, Thomas L., and Mark Steyvers. "Finding scientific topics." Proceedings of the National academy of Sciences 101.suppl 1 (2004): 5228-5235.

10 Sievert, Carson, and Kenneth Shirley. "LDAvis: A method for visualizing and interpreting topics." Proceedings of the workshop on interactive language learning, visualization, and interfaces. 2014.