Natural Language Processing and Machine Learning to assist radiation oncology incident learning

Felix Mathew[a], Hui Wang[b], Logan Montgomery[a], John Kildea[a]

[a]Medical Physics Unit, McGill University, Montreal, QC, H4A3J1, Canada

[b]Unaffiliated, Montreal, QC, Canada

Incident learning

Incident learning is defined as an organization’s ability to identify, report and investigate incidents and near misses, and to take corrective actions that improve the patient care system and reduce the risk of recurrence.

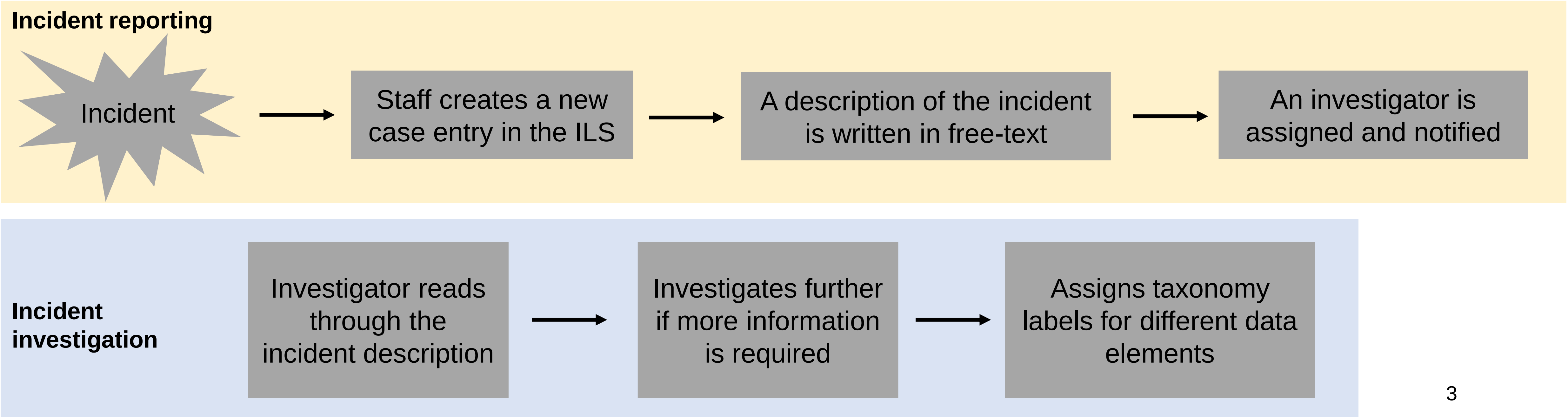
Incident: An unwanted or unexpected change from a normal system behaviour which causes or has the potential to cause an adverse effect to persons or equipment.

Near miss: An event or situation that could have resulted in an accident, injury, or illness but did not either by chance or through timely intervention. Also known as a close call, good catch or near hit.

Incident learning system (ILS)

- Software tool that enables incident learning.
- There are many international, national and local Incident learning systems.

  National System for Incident Reporting – Radiation Treatment (NSIR-RT)
  Safety and Incident Learning System (SaILS)

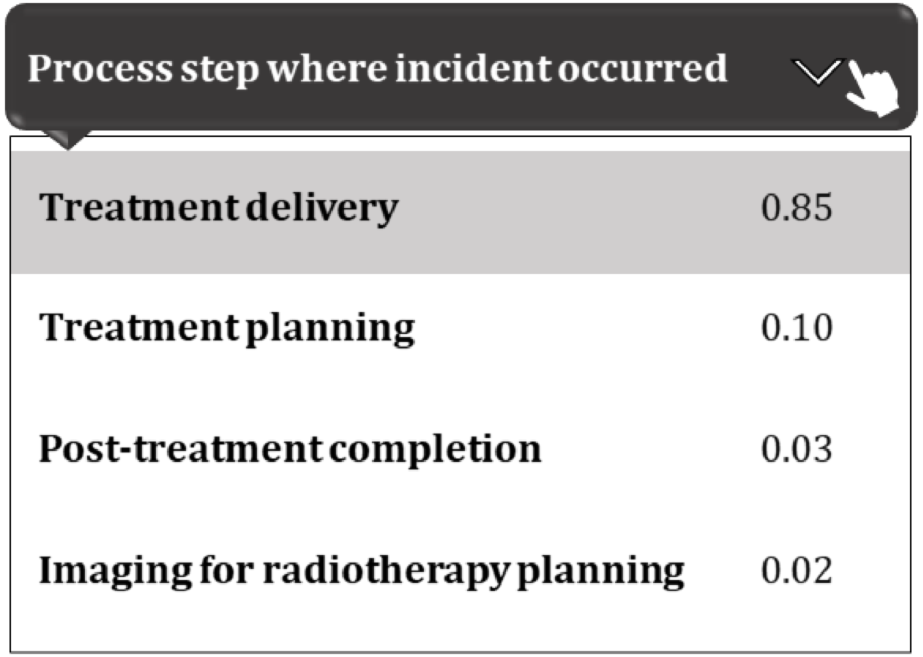


Objective

To use natural language processing and machine learning tools to assist incident investigators with incident classification.

Data elements of interest:

1. Process step where incident occurred (*8 label options*)
2. Problem type of the incident (*16 label options*)
3. Contributing factors of the incident (*25 label options*)



A mock-up of the ranked drop-down list of labels for the process step data element of SaILS

Supervised learning

- Supervised ML methods aim to predict the classification labels (class labels) of new incident reports by learning from expert-labelled training data that are already available.
- We gathered more than 6500 incident reports from CIHI and MUHC SaILS databases.
- Extracted incident descriptions and labels

<u>Incident descriptions</u>

Example:

Plan not ready. Pt was scheduld for 8:45 for plan 2, plan was not ready . Pt was called at 8:00 to come for 11:00. Plan ready @ 12:15.

Training on the data is difficult because:

- Abbreviations
- Shorthand
- Spelling mistakes
- Grammatical errors
- Improper sentence structure

NLP of the incident descriptions

1. Line-break removal

   Replaces newlines or line breaks (‘\n’ characters) with period.

2. Translation

   Bonjour → Hello
   using Google translate API

3. Punctuation and whitespace removal

   Removes all punctuations and unnecessary white spaces between words

4. Lowercase normalization

   All letters are transformed to their lowercase equivalent to ensure that instances of the same word that were written differently, were identified as a single object.
   Eg. Radiation, RADIATION and radiation.

5. Autocorrection

   Corrects spelling mistakes using the PyEnchant dictionary

6. Entity replacement

   Identifies words or ordinals that describe time, date, quantity, and percentage in our text and replaces them with a generic label that describes the entity.

7. Stopword removal

   Words that are used frequently in English such as ‘the’, ‘an’, ‘and’, ‘with’ and ‘but’ are removed. These words have no classification value in ML

8. Lemmatization

   Lemmatization refers to the process of replacing words with their basic dictionary form (lemma)

   Eg: ‘working’, ‘works’ and ‘worked’ → “work”

Ex. No.	Original incident description	Processed text
1	INCCURATE TARGET . CTV WAS BIGGER THAN PTV, WAS NOTICED ONLY AT THE END OF PLANNING PROCESS, TARGET HAD TO BE CORRECTED AND PLAN REDONE.	inaccurate target ctv big ptv notice end planning process target correct plan redone
2	no review status . Pt started rt on May 22, films were never reviewed until May 27th. Reached 7/9 and MD never verified films	review status pt start rt date film review date reach md verify film
3	No anti-emetic . 1 shot spine plus 25/5 rectum. Anti-emetic never prescribed for spine treatment. Danger of patient being sick during rectum iso and treat.	anti emetic shoot spine plus quantity rectum anti emetic prescribe spine treatment danger patient sick rectum iso treat
4	Not enough time to do the work, risk for mistakes. Waiting time for patient . PLAN 2 received last minute to do the plan in dosi 2 hrs before the patient apppointment	time work risk mistake wait time patient plan receive time plan dosi hrs patient appointment
5	Patient orientation . During treatment set up, we noticed that the documented patient orientation was wrong. Patient was scanned HEAD FIRST, but the ct-sim set up sheet indicates FEET FIRST.	patient orientation treatment set notice document patient orientation wrong patient scan head ordinal ct sim set sheet indicate foot ordinal

ML Algorithms

- ML algorithms (Estimators) were obtained from Scikit-learn library.
- Not easy to determine which model is best suited for our dataset.

So we decided to test them all

- Mostly designed for binary classification.
- But we have a multi-label classification problem.

So we extended them to support multi-label compatibility using two techniques: MultiOutputRegressor and RegressorChain

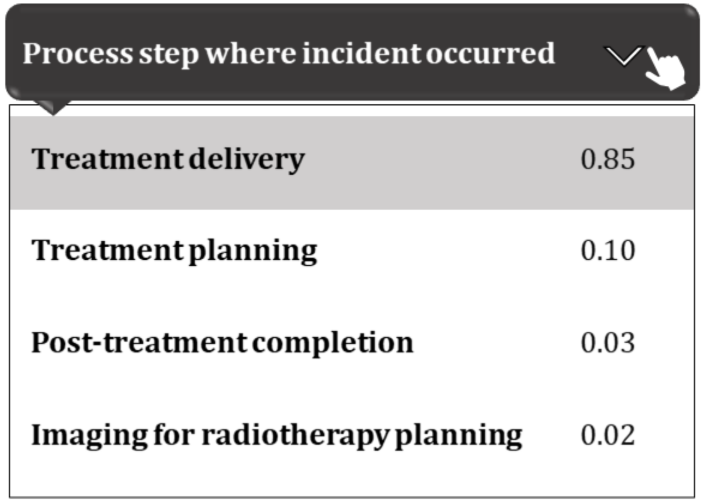
TrueLabelIndex score

Multi-label models generate a ranked list of possible labels

Model prediction: [Label 5, Label 3, Label 1, Label 4, Label 2]

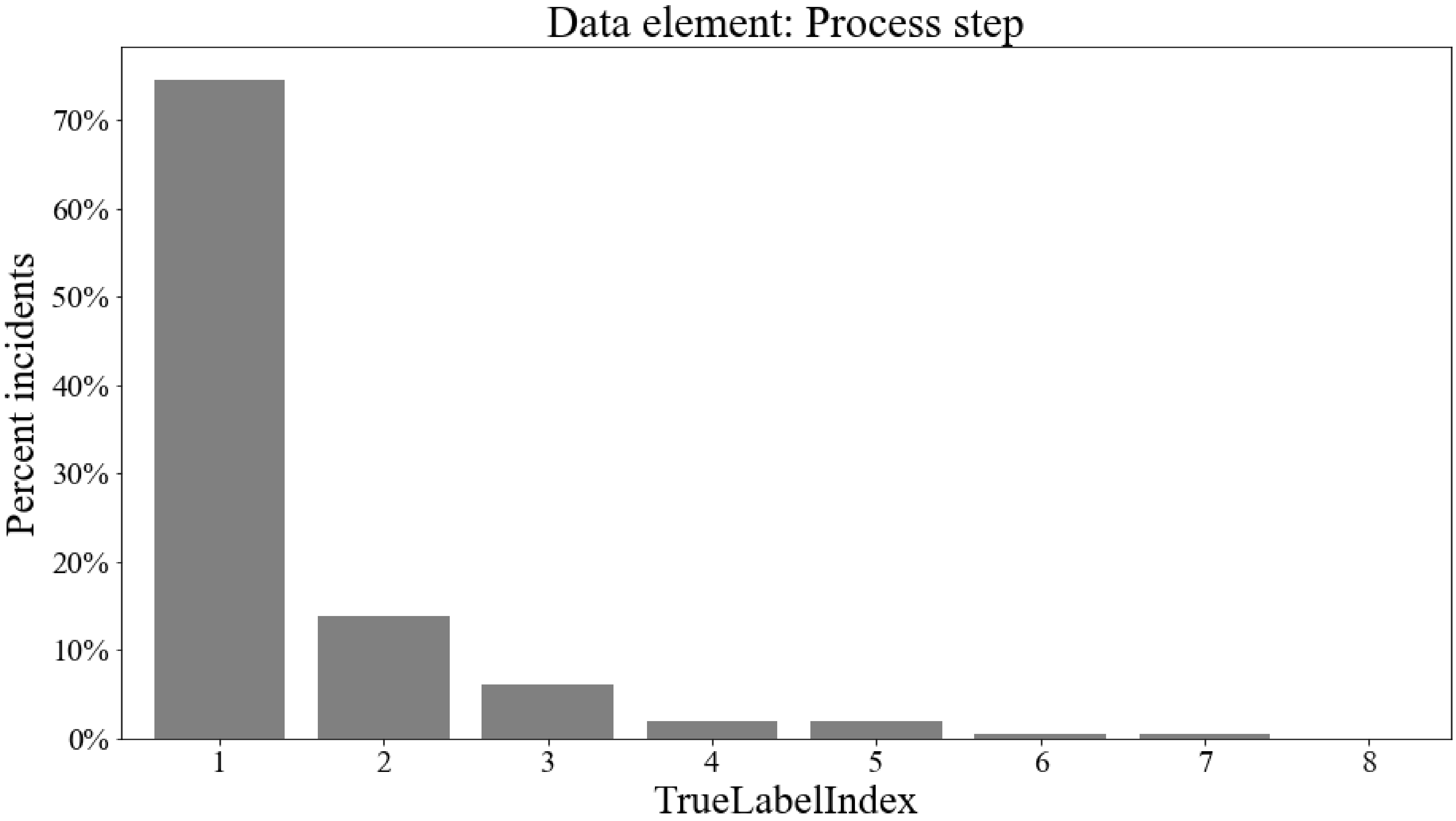
If True label (expert labelled value) = Label 3,

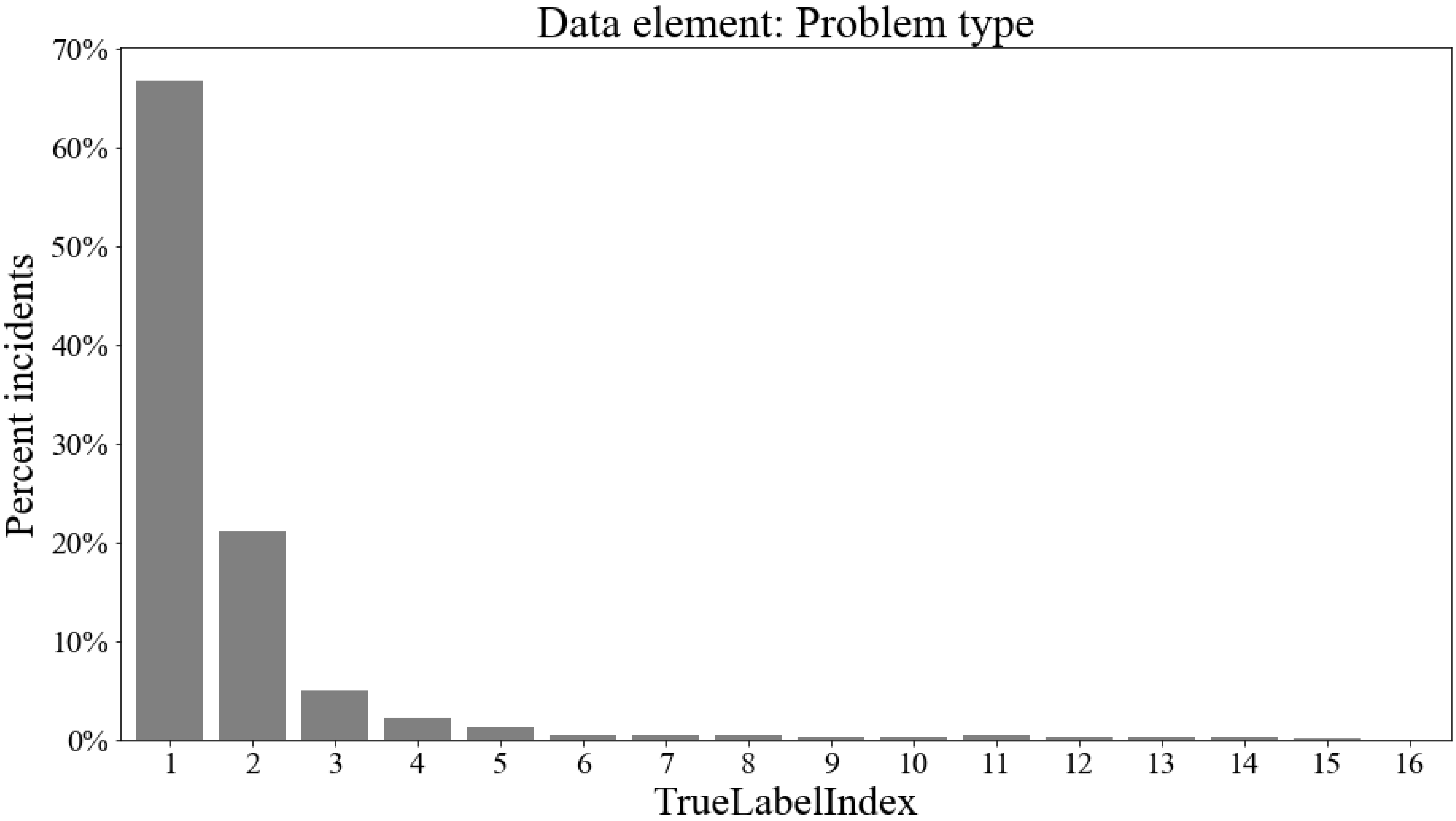
Then, TrueLabelIndex score = 2

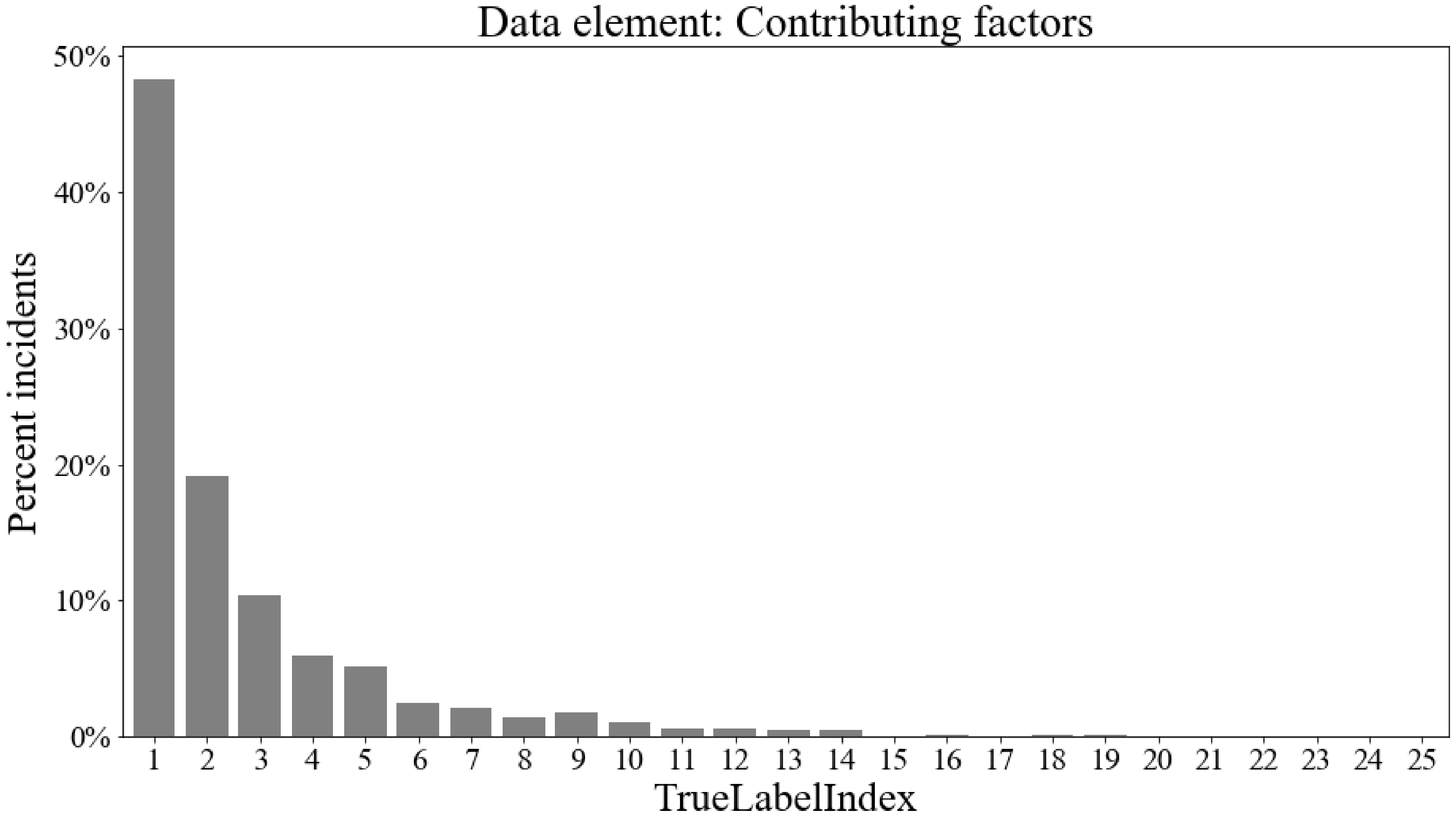


Model performance on test set

Data element	**Best performed model**	**Final test - TrueLabelIndex score** (Best score =1)
Process step	MultiOutputRegressor + Linear SVR	1.47
Problem type	MultiOutputRegressor + Linear SVR	1.72
Contributing factors	MultiOutputRegressor + Linear SVR	2.65


Data element: Process step
Percent incidents
0%
10%
20%
30%
40%
50%
60%
70%
1
2
3
4
5
6
7
8
TrueLabelIndex


Data element: Problem type
Percent incidents
0%
10%
20%
30%
40%
50%
60%
70%
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
TrueLabelIndex


Data element: Contributing factors
Percent incidents
50%
40%
30%
20%
10%
0%
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
TrueLabelIndex

Summary

- Radiotherapy incident investigation is a resource intensive process.
- This work aimed to help the investigator by semi-automating incident classification.
- We used supervised ML technique to develop models that can predict a ranked list of label recommendations for every incident real-time.
- The best model was able to show the appropriate label within the top three label suggestions

Acknowledgments

Dr. John Kildea

Hui Wang

Logan Montgomery

Hossein Naseri

Kelly Agnew

CIHI ICIS

Canadian Institute for Health Information

Institut canadien d'information sur la santé

CPQR

Canadian Partnership for Quality Radiotherapy

PCQR

Partenariat canadien pour la qualité en radiothérapie

McGill

Centre universitaire de santé McGill

McGill University Health Centre