FORCE: A Benchmark Dataset for Foodborne Disease Outbreak and Recall Event Extraction from News

Sudeshna Jana
TCS Research
sudeshna.jana@tcs.com

Manjira Sinha
TCS Research
sinha.manjira@tcs.com

Tirthankar Dasgupta
TCS Research
dasgupta.tirthankar@tcs.com

Abstract

The escalating prevalence of food safety incidents within the food supply chain necessitates immediate action to protect consumers. These incidents encompass a spectrum of issues, including food product contamination and deliberate food and feed adulteration for economic gain leading to outbreaks and recalls. Understanding the origins and pathways of contamination is imperative for prevention and mitigation. In this paper, we introduce FORCE (**F**oodborne disease **O**utbreak and **R**e**C**all **E**vent extraction from openweb). Our proposed model leverages a multi-tasking sequence labeling architecture in conjunction with transformer-based document embeddings. We have compiled a substantial annotated corpus comprising relevant articles published between 2011 and 2023 to train and evaluate the model. The dataset will be publicly released with the paper. The event detection model demonstrates fair accuracy in identifying food-related incidents and outbreaks associated with organizations, as assessed through cross-validation techniques.

1 Introduction

The escalating number of food safety concerns remains a source of significant apprehension(Amico et al., 2018; Kase et al., 2017; Boatemaa et al., 2019; Nerín et al., 2016; Kase et al., 2017). Globally, foodborne diseases continue to plague populations and stand as leading contributors to both illness and mortality(Bouzembrak and Marvin, 2016; Potter et al., 2012; Pádua et al., 2019; Lüth et al., 2019). Recent estimates have identified norovirus and Campylobacter as the most common reason behind foodborne illnesses, while fatalities have been associated with non-typhoidal Salmonella enterica, Salmonella Typhi, Taenia solium, hepatitis A virus, and aflatoxin(Djekic et al., 2017; Kleter et al., 2009; Bouzembrak and Marvin, 2019).

One of the repercussions of food safety issues is the necessity for food recalls, which pose substantial economic threats to both businesses and nations alike(Deng et al., 2016). This underscores the imperative of uncovering the root causes behind these incidents and the factors contributing to contamination (Zhou et al., 2020). Cross-contamination in food and beverages is a multifaceted issue that can transpire at various stages of the food processing chain, including external raw food contamination, transportation, cleaning processes, heating, food packaging, and even during food storage. Contamination events resulting in outbreaks can manifest at any point before, during, or after food processing (Scallan and Mahon, 2012; Gupta et al., 2004).

Consequently, the pivotal task of identifying the origins of contamination or the triggers for recalls is of paramount importance (Hall et al., 2013). It is essential to gain insights into potential sources and pathways of contamination leading to foodborne outbreaks and product recalls, and to devise effective measures for prevention(Tao et al., 2020; Zhou et al., 2021; Jin et al., 2020; Marvin et al., 2017). It is worth noting that there is a dearth of substantial work in the development of computational and/or analytical models addressing these concerns(Allard et al., 2016; Moumni Abdou et al., 2019). This scarcity of research can largely be attributed to the limited availability of data concerning the contributory factors associated with food safety incidents. As such, there is a pressing need to develop an automated tool that can mine reported food safety incidents and recalls to bridge this knowledge gap.

Applications of language technologies and data science for food-borne risk assessment are gaining ground(Harris et al., 2017; Altenburger and Ho, 2019; Maharana et al., 2019; Deng et al., 2021). data-intensive systems play important roles in tracking food-borne illness cases and agents(Pujahari and Khan, 2022; Oldroyd et al., 2021; Nychas et al., 2021; Gupta et al., 2004; Scallan and Mahon, 2012; Wang et al., 2021). Examples at the US federal

Proceedings of The 9th Social Media Mining for Health Applications (#SMM4H) Workshop and Shared Tasks, pages 163–169

level include PulseNet (Swaminathan et al., 2001), the National Antimicrobial Resistance Monitoring System (NARMS) (Gupta et al., 2004), FoodNet (Scallan and Mahon, 2012), and the National Outbreak Reporting System (Hall et al., 2013). Implementation of whole-genome sequencing (WGS) in surveillance and outbreak investigation has fueled an explosion of publicly available foodborne pathogen genomes in new systems such as GenomeTrakr (Allard et al., 2016), EnteroBase (Zhou et al., 2020), and the National Center for Biotechnology Information's Pathogen Detection. These works are primarily focused on a) identifying the sentiment polarity of the documents, and b) identifying the occurrences of a set of predefined types of entities and events. However, none of the above works perform deep linguistic analysis of the textual contents to identify food-related incidents based only on the linguistic structure of the text. The model presented in this paper is distinct from the earlier approaches since it is capable of detecting novel and heretofore unknown incidents from reports based only on semantic and linguistic analysis of content.

Keeping in mind the above-mentioned requirements of end users, in this work we propose an event detection model that can identify food safety-related insights, recalls, and outbreaks mentioned in regulatory reports and social media platforms. The proposed model uses a multi-tasking sequence labeling architecture that works with transformer-based document embeddings. We have created a large annotated corpus containing relevant articles published by multiple regulatory agencies over twelve years (2011 - 2023) for training and evaluating the model. The dataset will be publicly released with this paper. The model has been thereafter applied to recent publications. Aggregate analysis of these extracted insights reveals interesting trends.

2 Dataset Creation

The dataset comprises regulatory news articles from two sources namely, a) A corpus of 6000 regulatory articles comprising around 121080 sentences, under Outbreak(O) and Recall(R) categories published between 2011 and 2023 by Food Safety News (FSN)[1] and b) A collection of around 2200 news articles from United States Food and Drug Administration (FDA) recall and outbreak announcements[2]. All together there are 8100 articles.

All the news articles were manually annotated to mark the various target entities. The annotation was done by multiple annotators following a rigorous procedure to ensure acceptable inter-annotator agreement.

The entities and events to be identified by the annotators are as follows:

Target Organization (TO): Of the many organization names that may appear in a document and are detected by the NER earlier, the task during annotation is to identify and tag the organization whose product has been recalled.

Product Name (P): The name of the product that has been recalled or caused the outbreak.

Infection Name (I): The name of the bacterial infection mentioned in the report that causes the outbreak/recall.

Safety Incident (SI) - Annotators have to tag phrases or sequences of words that collectively are indicative of the food safety incident.

Cause of Incident (CI) - Annotators have to tag phrases or sequences of words that indicate the primary cause of the food safety incident.

Number of People Affected (N): Annotators have to tag phrases or sequences of words that collectively are indicative of a number of people affected due to the outbreak.

To help the annotators, each document is first processed using the Stanford NER (Manning et al., 2014) to obtain the organization names, locations, and currency values as named entities. This helps in quick localization of the first four elements, if present in the document. The annotators are domain experts who are knowledgeable about the domain.

16 annotators took part in the annotation, with each expert annotating 700 documents using the Stanford simple manual annotation tool[3]. This included 200 documents, which were sent to all the annotators to compute the inter-annotator agreement later. The average length of a document is around 23 sentences. The experts read each document and performed the following tasks,

Task-1: - Label sentences of each document as **Food Recall** - if the document reported events of a product recall or **Disease Outbreak** - if the document mentions events that report an foodborne disease outbreak or **NEUTRAL**- in case none of the above factors hold.

Task-2: - This task had two components: (a)

[1]https://www.foodsafetynews.com/

[2]https://www.fda.gov/safety

[3]https://nlp.stanford.edu/software/

From among the named entities, the target organization, product name, infection, and locations were marked, if any, and (b) Mark phrases in the text that indicate food safety incidents, and its cause. At the end of the annotation, each word in the document is assigned a label TO, P, I, SI, CI, N, or None. Table 3 in A.2 illustrates the annotation with some example News texts. For the sake of understanding, we have shown labels of only the phrases that belong to any one of the following classes {TO, P, I, SI, CI, L, or N}.

Using the annotations obtained for 200 common documents, we measured the inter-annotator agreement using the Fleiss Kappa (Fleiss et al., 1981) measure (κ). This is computed as: $\kappa = \frac{\bar{P}-\bar{P}_e}{1-\bar{P}_e}$ The factor $1 - \bar{P}_e$ gives the degree of agreement that is attainable above chance, and $\bar{P} - \bar{P}_e$ gives the degree of agreement achieved above chance. It was observed that the inter-annotator score for Task-1 was 0.83, which is appreciably high. For Task 2, it was found to be 0.71. The scores are computed using word-label matches assigned by different annotators. The very high scores indicate that all experts were marking fairly uniformly and therefore, the expert annotated dataset is reliable to be used for training incident detection systems. Out of the 165,080 sentences from 82000 documents, around 27500 sentences were found to contain words belonging to at least one type mentioned in the incident knowledge schema. Altogether, we obtained 13000 safety incidents, 12100 causes, 2223 Target Organizations, 3300 locations, 4695 product names, and 4101 infection names. The entire annotated data will be publicly released with this paper.

Model No.	Model Name	Sequence Classification		
		P	R	F1
I.	Single task CNN-BiLSTM	0.63	0.59	0.62
II.	Single task PreTrained BERT	0.75	0.72	0.73
III.	Single-task-BERT-CNN-BiLSTM	0.72	0.75	0.73
IV.	LLAMA-2	0.75	0.79	0.78
V.	Mistral7B	0.73	0.79	0.77
VI.	Multi-task BERT-CNN-BiLSTM	**0.83**	**0.89**	**0.86**

Table 1: Results reporting accuracy of classifying food safety events as *Food Recall* or *Disease Outbreak*.

3 A Multi-tasking Neural Model for Food Safety Knowledge Extraction

The proposed model works on each sentence at a time to detect elements of interest that are defined in the incident knowledge schema. Multi-task learning utilizes the correlation between related tasks to improve classification by learning tasks in parallel. In the present work, the two related tasks are *task-1:* classifying a sentence into either *Recall* or *Outbreak* classes as discussed earlier and *task-2:* labeling appropriate phrases in the text as per the incident knowledge schema.

A cascaded CNN-BiLSTM layer for the combined tasks of sentence classification and sequence label prediction, using the fine-tuned BERT for creating the sequence embeddings.

To obtain the multi-tasking model for dual tasks of sequence classification and sequence labeling, the $BERT-CNN-BiLSTM$ layers have been trained with two separate loss functions L_1 and L_2. Where, $L_1(\theta) = -\sum_{t=1}^{M}\sum_{k=1}^{K}\bar{y}_t^k log(y_t)$ and $L_2(\theta) = -\sum_{t=1}^{N}\sum_{j=1}^{J}\bar{q}_t^{i,j} log(q_t^i)$.

Here, q_t is the vector representation of the predicted output of the model for the input word w_t^i. K and J are the number of class labels for each task. The model is fine-tuned end-to-end by minimizing the cross-entropy loss.

We define the joint loss function using a linear combination of the loss functions of the two tasks:

$$L_{joint}(\theta) = \lambda * L_1(\theta) + (1-\lambda) * I_{[y_{sentence}==1]} * L_2(\theta) \quad (1)$$

Where λ controls the contribution of losses of the individual tasks in the overall joint loss. $I_{[y_{sentence}==1]}$ is an indicator function that activates the loss only when the corresponding sentence classification label is 1 since we do not want to back-propagate sequence labeling loss when the corresponding sequence classification label is 0.

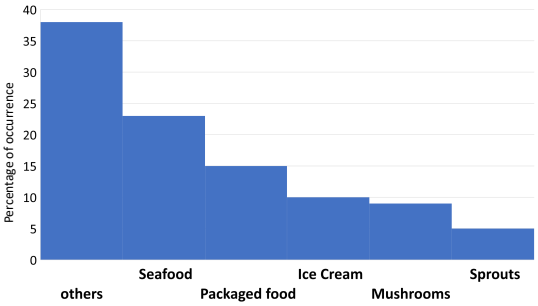

Figure 1: Distribution of occurrences of recalled products.

4 Evaluation and Results

The performance of the proposed model has been compared with a number of baseline models used for single-objective document classification and sequence labeling tasks as well as large language

Model	Sequence labeling task																	
No.	TO			P			I			SI			CI			N		
	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F
I.	0.76	0.78	0.77	0.71	0.67	0.69	0.77	0.78	0.77	0.72	0.77	0.74	0.77	0.8	0.78	0.72	0.78	0.72
II.	0.77	0.78	0.78	0.69	0.74	0.71	0.79	0.77	0.78	0.78	0.75	0.76	0.74	0.75	0.76	0.78	0.80	0.79
III.	0.80	0.81	0.80	0.71	0.75	0.73	0.76	0.86	0.80	0.78	0.79	0.78	0.75	0.76	0.75	0.78	0.77	0.78
IV.	0.82	0.87	0.84	0.75	0.79	0.77	0.82	0.86	0.82	0.78	0.79	0.78	0.78	0.72	0.75	0.84	0.89	0.86
V.	0.82	0.89	**0.85**	0.80	0.82	**0.81**	0.82	0.87	0.83	0.76	0.79	0.78	0.78	0.79	0.79	0.92	0.90	**0.91**
VI.	0.81	0.89	0.84	0.79	0.83	0.80	0.82	0.87	**0.84**	0.82	0.90	**0.84**	0.85	0.91	**0.88**	0.86	0.88	0.88

Table 2: Results reporting the performance of the food safety incident and entity extraction task.

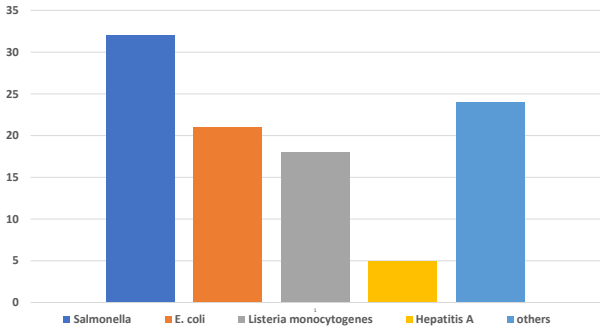

Figure 2: Distribution of germs found in the recalled products.

models like LLAMA-2 7B and fined-tuned Mistral 7B. Table 1 presents the precision, recall, and F1 scores of classifying food safety events as *food recall* or *disease outbreak*. We have obtained the highest F1 score of 0.86 with a high precision of 0.89.

Table 2 presents the accuracy of subsequent labeling of word sequences within a sentence by their respective categories - *TO, P, I, SI, CI, or N*, as described earlier. For both the cases, the performance of the proposed multi-objective architecture has been compared with several baseline state-of-the-art models designed with single objective functions. It was observed that for most of the classes like, S, SI and CI, the Multi-task BERT-CNN-BiLSTM model significantly outperforms the baseline models. On the other hand, mistral-7B model trained over the given dataset performs better recognizing the TO, P and N classes.

The primary reason for the poor performance of LLAMA-2 as well as mistral-7B can be attributed due to two reasons: a) lack of environmental domain knowledge due to which critical domain concepts like, *Salmonella*, *Listeria monocytogenes* and *E.Coli* gets ignored. b) Unable to identify phrase boundaries. We observe that despite in most of the cases LLAMA-2 correctly identified the safety incident and cause phrases, but the span of the phrases are either too long or too short. as a results of which outputs of the model get penalized. Similar observations were made for mistral 7B, however, since the mistral model is fine-tuned over the current dataset, problems related to domain concept mismatch were relatively less. However, the output word span detection for incident and causes still remains a challenge.

Apart from the classification and extraction of food Safety incidents, it is equally important to perform some basic analytics on the dataset. Figure 1 shows the distribution of the top five causes of *food recall* across different locations in the United States. In general, we have observed that infections such as *Salmonella*, *Listeria monocytogenes* and *E.Coli* are the biggest causes of food recall in the United States. Figure 2 depicts the top food products that contain the aforementioned germs.

5 Conclusion

In this paper we present resource creation and extraction of critical information related to food safety and food-borne infection from regulatory reports and social media platforms. The proposed model, founded on a multi-tasking sequence labeling architecture integrated with transformer-based document embeddings, demonstrates its effectiveness in this task. To develop and evaluate our model, we meticulously curated an annotated corpus comprising pertinent articles. Our initial analysis demonstrate the proposed multi-task model surpasses the performance of almost all the baseline models including LLMs such as LLAMA-27B and finetuned mistral-7B.

References

Marc W Allard, Errol Strain, David Melka, Kelly Bunning, Steven M Musser, Eric W Brown, and Ruth Timme. 2016. Practical value of food pathogen traceability through building a whole-genome sequencing network and database. *Journal of clinical microbiology*, 54(8):1975–1983.

Kristen M Altenburger and Daniel E Ho. 2019. Is yelp actually cleaning up the restaurant industry? a reanalysis on the relative usefulness of consumer reviews. In *The World Wide Web Conference*, pages 2543–2550.

Priscilla D' Amico, Daniele Nucera, Lisa Guardone, Martino Mariotti, Roberta Nuvoloni, and Andrea Armani. 2018. Seafood products notifications in the eu rapid alert system for food and feed (rasff) database: Data analysis during the period 2011–2015. *Food Control*, 93:241–250.

Sandra Boatemaa, McKenna Barney, Scott Drimie, Julia Harper, Lise Korsten, and Laura Pereira. 2019. Awakening from the listeriosis crisis: Food safety challenges, practices and governance in the food retail sector in south africa. *Food Control*, 104:333–342.

Yamine Bouzembrak and Hans JP Marvin. 2016. Prediction of food fraud type using data from rapid alert system for food and feed (rasff) and bayesian network modelling. *Food Control*, 61:180–187.

Yamine Bouzembrak and Hans JP Marvin. 2019. Impact of drivers of change, including climatic factors, on the occurrence of chemical food safety hazards in fruits and vegetables: A bayesian network approach. *Food control*, 97:67–76.

Xiangyu Deng, Shuhao Cao, and Abigail L Horn. 2021. Emerging applications of machine learning in food safety. *Annual Review of Food Science and Technology*, 12:513–538.

Xiangyu Deng, Henk C den Bakker, and Rene S Hendriksen. 2016. Genomic epidemiology: whole-genome-sequencing–powered surveillance and outbreak investigation of foodborne bacterial pathogens. *Annual review of food science and technology*, 7:353–374.

Ilija Djekic, Danijela Jankovic, and Andreja Rajkovic. 2017. Analysis of foreign bodies present in european food using data from rapid alert system for food and feed (rasff). *Food control*, 79:143–149.

Joseph L Fleiss, Bruce Levin, Myunghee Cho Paik, et al. 1981. The measurement of interrater agreement. *Statistical methods for rates and proportions*, 2(212-236):22–23.

Amita Gupta, Jennifer M Nelson, Timothy J Barrett, Robert V Tauxe, Shannon P Rossiter, Cindy R Friedman, Kevin W Joyce, Kirk E Smith, Timothy F Jones, Marguerite A Hawkins, et al. 2004. Antimicrobial resistance among campylobacter strains, united states, 1997–2001. *Emerging infectious diseases*, 10(6):1102.

Aron J Hall, Mary E Wikswo, Karunya Manikonda, Virginia A Roberts, Jonathan S Yoder, and L Hannah Gould. 2013. Acute gastroenteritis surveillance through the national outbreak reporting system, united states. *Emerging infectious diseases*, 19(8):1305.

Jenine K Harris, Jared B Hawkins, Leila Nguyen, Elaine O Nsoesie, Gaurav Tuli, Raed Mansour, and John S Brownstein. 2017. Research brief report: Using twitter to identify and respond to food poisoning: The food safety stl project. *Journal of Public Health Management and Practice*, 23(6):577.

Cangyu Jin, Yamine Bouzembrak, Jiehong Zhou, Qiao Liang, Leonieke M Van Den Bulk, Anand Gavai, Ningjing Liu, Lukas J Van Den Heuvel, Wouter Hoenderdaal, and Hans JP Marvin. 2020. Big data in food safety-a review. *Current Opinion in Food Science*, 36:24–32.

Julie Ann Kase, Guodong Zhang, and Yi Chen. 2017. Recent foodborne outbreaks in the united states linked to atypical vehicles—lessons learned. *Current Opinion in Food Science*, 18:56–63.

Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

GA Kleter, ALDO Prandini, LAURA Filippi, and HJP Marvin. 2009. Identification of potentially emerging food safety issues by analysis of reports published by the european community's rapid alert system for food and feed (rasff) during a four-year period. *Food and chemical toxicology*, 47(5):932–950.

Stefanie Lüth, Idesbald Boone, Sylvia Kleta, and Sascha Al Dahouk. 2019. Analysis of rasff notifications on food products contaminated with listeria monocytogenes reveals options for improvement in the rapid alert system for food and feed. *Food Control*, 96:479–487.

Adyasha Maharana, Kunlin Cai, Joseph Hellerstein, Yulin Hswen, Michael Munsell, Valentina Staneva, Miki Verma, Cynthia Vint, Derry Wijaya, and Elaine O Nsoesie. 2019. Detecting reports of unsafe foods in consumer product reviews. *JAMIA open*, 2(3):330–338.

Christopher D. Manning, Bauer Surdeanu, Mihai, Finkel John, Bethard Jenny, Steven J., and David. McClosky. 2014. The stanford corenlp natural language processing toolkit. In *52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 55–60.

Hans JP Marvin, Esmée M Janssen, Yamine Bouzembrak, Peter JM Hendriksen, and Martijn Staats. 2017.

Big data in food safety: An overview. *Critical reviews in food science and nutrition*, 57(11):2286–2295.

Houda Moumni Abdou, Ilham Dahbi, Mohammed Akrim, Fatima Zahra Meski, Yousef Khader, Mohammed Lakranbi, Hind Ezzine, and Asmae Khattabi. 2019. Outbreak investigation of a multipathogen foodborne disease in a training institute in rabat, morocco: case-control study. *JMIR public health and surveillance*, 5(3):e14227.

Cristina Nerín, Margarita Aznar, and Daniel Carrizo. 2016. Food contamination during food process. *Trends in food science & technology*, 48:63–68.

George-John Nychas, Emma Sims, Panagiotis Tsakanikas, and Fady Mohareb. 2021. Data science in the food industry. *Annual Review of Biomedical Data Science*, 4:341–367.

Rachel A Oldroyd, Michelle A Morris, and Mark Birkin. 2021. Predicting food safety compliance for informed food outlet inspections: a machine learning approach. *International Journal of Environmental Research and Public Health*, 18(23):12635.

Inês Pádua, André Moreira, Pedro Moreira, Filipa Melo de Vasconcelos, and Renata Barros. 2019. Impact of the regulation (eu) 1169/2011: Allergen-related recalls in the rapid alert system for food and feed (rasff) portal. *Food control*, 98:389–398.

Antony Potter, Jason Murray, Benn Lawson, and Stephanie Graham. 2012. Trends in product recalls within the agri-food industry: Empirical evidence from the usa, uk and the republic of ireland. *Trends in food science & technology*, 28(2):77–86.

Rakesh Mohan Pujahari and Rijwan Khan. 2022. Applications of machine learning in food safety. In *Artificial Intelligence Applications in Agriculture and Food Quality Improvement*, pages 216–240. IGI Global.

Elaine Scallan and Barbara E Mahon. 2012. Foodborne diseases active surveillance network (foodnet) in 2012: a foundation for food safety in the united states. *Clinical Infectious Diseases*, 54(suppl_5):S381–S384.

Bala Swaminathan, Timothy J Barrett, Susan B Hunter, Robert V Tauxe, and CDC PulseNet Task Force. 2001. Pulsenet: the molecular subtyping network for foodborne bacterial disease surveillance, united states. *Emerging infectious diseases*, 7(3):382.

Dandan Tao, Pengkun Yang, and Hao Feng. 2020. Utilization of text mining as a big data analysis tool for food science and nutrition. *Comprehensive reviews in food science and food safety*, 19(2):875–894.

Hanxue Wang, Wenjuan Cui, Yunchang Guo, Yi Du, Yuanchun Zhou, et al. 2021. Machine learning prediction of foodborne disease pathogens: Algorithm development and validation study. *JMIR medical informatics*, 9(1):e24924.

Qinqin Zhou, Hao Zhang, and Suya Wang. 2021. Artificial intelligence, big data, and blockchain in food safety. *International journal of food engineering*, 18(1):1–14.

Zhemin Zhou, Nabil-Fareed Alikhan, Khaled Mohamed, Yulei Fan, Mark Achtman, Derek Brown, Marie Chattaway, Tim Dallman, Richard Delahay, Christian Kornschober, et al. 2020. The enterobase user's guide, with case studies on salmonella transmissions, yersinia pestis phylogeny, and escherichia core genomic diversity. *Genome research*, 30(1):138–152.

A Appendix

A.1 Fine-tuning the BERT language model

The basic $BERT_{base}$ model was first fine-tuned with a portion of the food safety corpus using oversampling, to create FoodSafety-BERT, referred to as FS-BERT hereafter. A labeled document is broken into multiple smaller chunks, such that each chunk can be fed as a unit to $BERT_{base}$ to create its corresponding vector. Each chunk is associated with a label that is the same as its parent document. A classification task is now defined with these chunks during which the basic BERT model is fine-tuned while training. This model is designed as a fully connected layer over the BERT base model, with softmax as the activation function. Training was done with learning rate set to 2×10^{-5} using the Adam optimizer (Kingma and Ba, 2014). The model is fine-tuned for a few epochs (3-4) only to avoid over-fitting. The chunk representations are saved from the CLS token embeddings created during the process. The fine-tuned BERT model, FS-BERT, is subsequently used for document and incident recognition tasks.

A.2 Example Annotation Sentences

[ORGName]/**TO** of Poughkeepsie, NY, is *[recalling its 2-lb., 5-lb. and 15-lb. boxes]*/**SI** of "*[Abady Highest Quality Maintenance & Growth Formula for Cats]*/**P**" because they have the *[potential to be contaminated with Salmonella]*/**CI**.
[ORGName]/**TO** of Brampton, Ontario, *[recalled 36 pounds of fully cooked pork baby back ribs in recalled 36 pounds of fully cooked pork baby back ribs in]* /**SI** today because they were not presented for inspection at the U.S. border. The problem was discovered when U.S. Department of Agriculture Food Safety and Inspection Service import staff reviewed records and discovered that the independent third-party carrier *[did not present a product for USDA inspection at the U.S.-Canadian border]*/**CI**. According to a Public Health Alert released by FSIS, being recalled are 18-pound cases containing 1.5-pound packages of "*[Cobblestone Farms Fully Cooked Pork Baby Back Ribs in Honey Garlic Barbeque Sauce]*/**P**" bearing package code "Sell By 2015-AL-08" and case code "15201" bearing the Canadian mark of inspection with establishment number "624." The product was distributed to a retailer in *[New York]*/**L**. In its announcement, FSIS stated that it is working on solutions to prevent future failure-to-present episodes from occurring, including outreach to industry, foreign food-safety agencies, and importers.
A Listeria outbreak in the *[Midwest]*/**L** linked to *[one death]*/**N** and a miscarriage likely was caused by *[contamination during the cheese-making process]*/**CI**, according to a new report from the U.S. Centers for Disease Control and Prevention. The Minnesota Department of Agriculture tested samples of the cheese from two retail outlets, revealing the outbreak strain to be *[Listeria monocytogenes]*/**I**.
About *[96,000 pounds]*/**SI** of *[Oscar Mayer Classic Wieners]*/**P** *[were recalled]*/**SI** Sunday by *[ORGName]*/**TO** of Columbia, MO, because of a *[packaging error]*/**CI**.
[ORGName]/**TO** of Detroit, MI, is *[recalling approximately 1.8 million pounds]*/**SI** of *[ground beef products]*/**P** that may be *[contaminated with E. coli O157:H7]*/**CI**, the U.S. Department of Agriculture's Food Safety and Inspection Service (FSIS) announced Monday. At the time that the recall was issued, there were *[11 illnesses]*/**N** linked to the recalled product.

Table 3: Sample Food Safety News texts with the respective annotated entities and events. Note that all the target organization names were intentionally masked by the token [ORGName] to maintain anonymity.