A News Recommendation System for Environmental Risk Management

Hamed Aboutorab[1], Ran Yu[2], Alishiba Dsouza[2], Morteza Saberi[3], Omar Khadeer Hussain[1]

[1]University of New South Wales, Canberra, Australia, [2]DSIS Group, University of Bonn, Bonn, Germany
[3]University of Technology Sydney, Sydney, Australia

Second International Workshop on Linked Data-driven Resilience Research
co-located with ESWC2023
28th May 2023

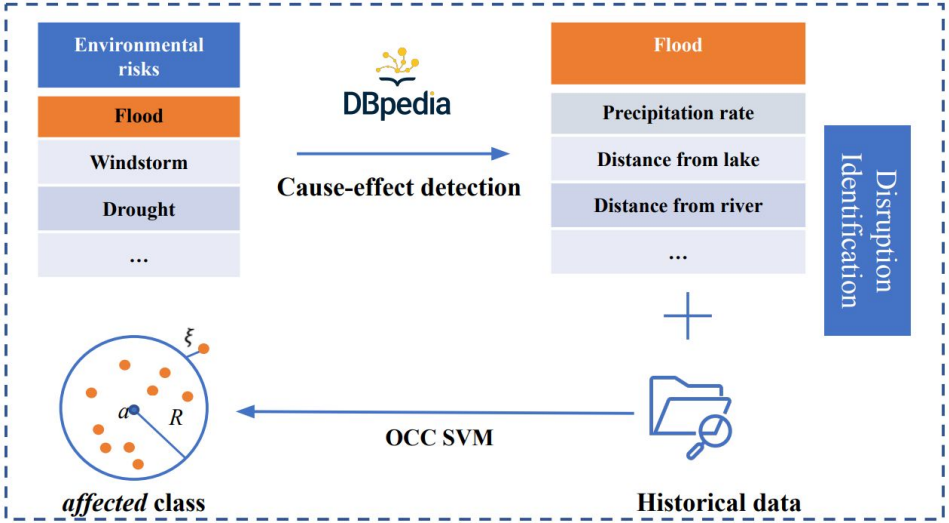
Motivation

- Environmental risk events can disrupt supply chain
  - Example: Floods – Disruption on the freight routes
- Prediction of the Environmental risk events can help businesses prevent losses
- News can help in notifying such risks to users.
- Challenges:
  - Location accuracy and precision
  - News recommendation system can have a lot of unnecessary information
- Additional knowledge from knowledge bases can help in reduction of the irrelevant news

NR-ERIA Approach

Disruption Identification:

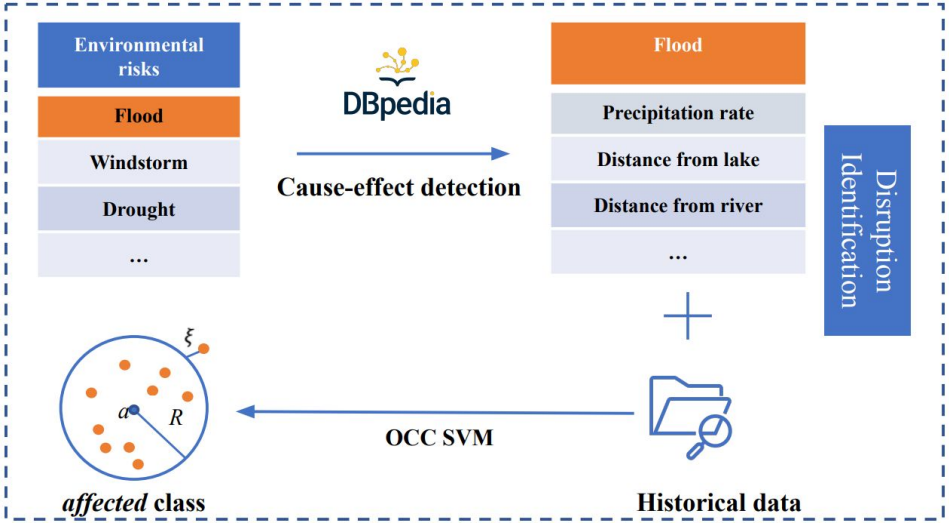
- Extraction of the important factors for the given environmental risk
- Feature extraction from textual descriptions from DBpedia knowledge graph
- BERT based cause-effect detection tool to determine the main causes of the risk



NR-ERIA Approach

Disruption Identification:

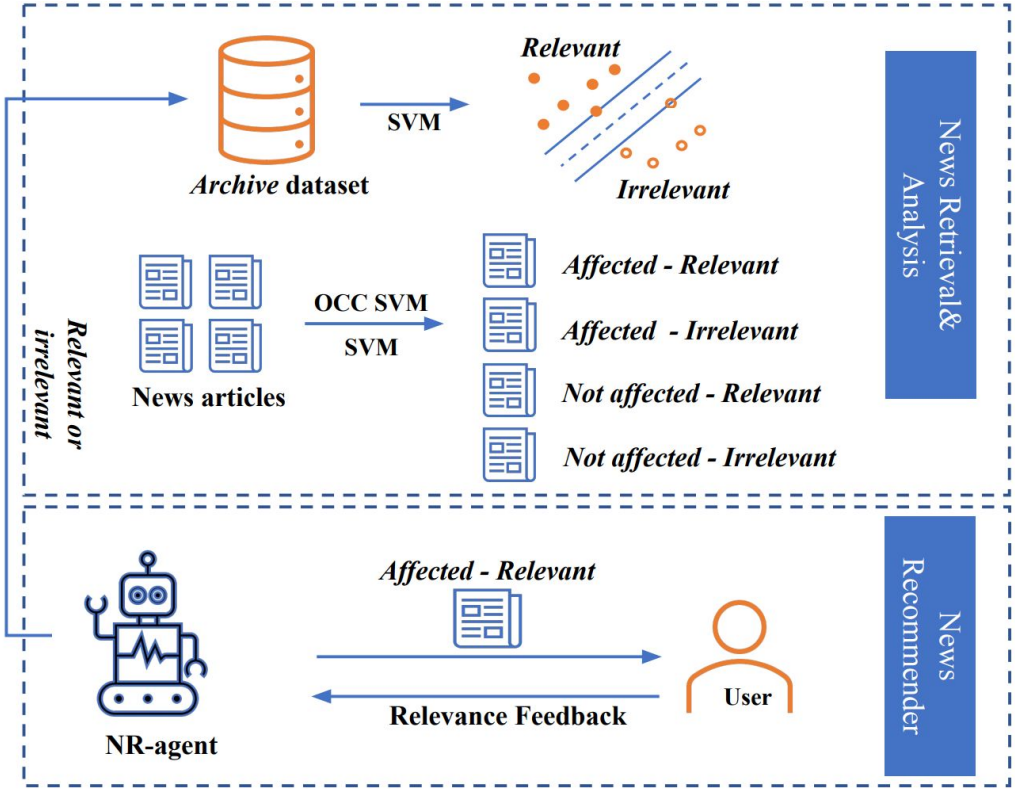
- One-Class Classification Support Vector Machine (OCC SVM) to model the boundaries of the radius R around a given location based on the feature vector
- Radius R tells whether the location falls in the affected areas or not



NR-ERIA Approach

News Retrieval and Analysis and News Recommender:

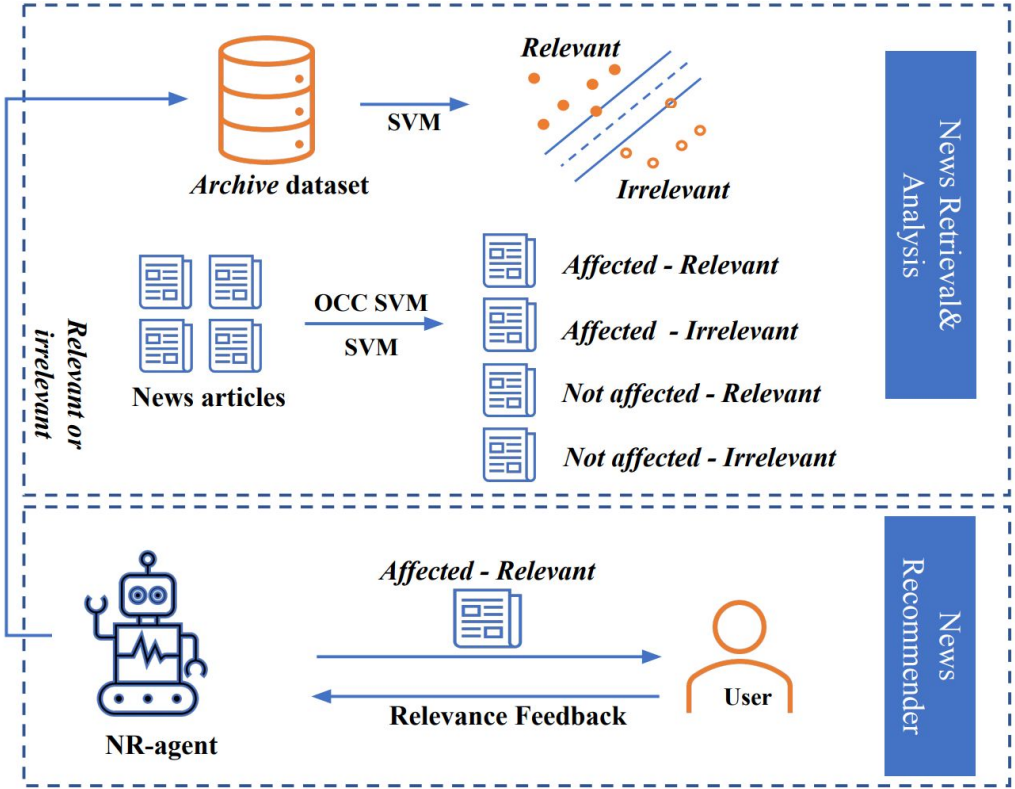
- Analyze the news article against the affected class
- Named entity Extraction for the location extraction
- For each location, check if it in the affected class
- Only the relevant articles are shown to the user



NR-ERIA Approach

News Retrieval and Analysis and News Recommender:

- Align the archived dataset using the user feedback (Based on the news recommender)
- News articles are further classified into relevant irrelevant classes using SVM classifier using TF-IDF vectorization



Data and Setup

- Historical data of road closures due to flooding from the Australian government website
- Features collection using Meteostat and Natural earth.
- For each road closure, collect news 10 days before and after the road closure, using the name of the road as the keyword.
- 10-fold cross validation with 90:10 train-test split.

1 row of *normal* dataset for 1000 road closures due to flood

Date	Location	Coordinate	Distance from (km)			Precipitation (mm)
			River	Lake	Ocean	
12/10/20	Armstrongs Rd	(-38.568, 145.982)	227.322	228.008	62.471	6.428

Performance Analysis

NR-ERIA achieves an accuracy of 0.92 in identifying the location of disruptions caused by flooding.

Comparison to RL-PRI:

- NR-ERIA outperforms RL-PRI in terms of accuracy and F1 score
  - NR-ERIA considers more environmental risk-specific factors important to classify the news

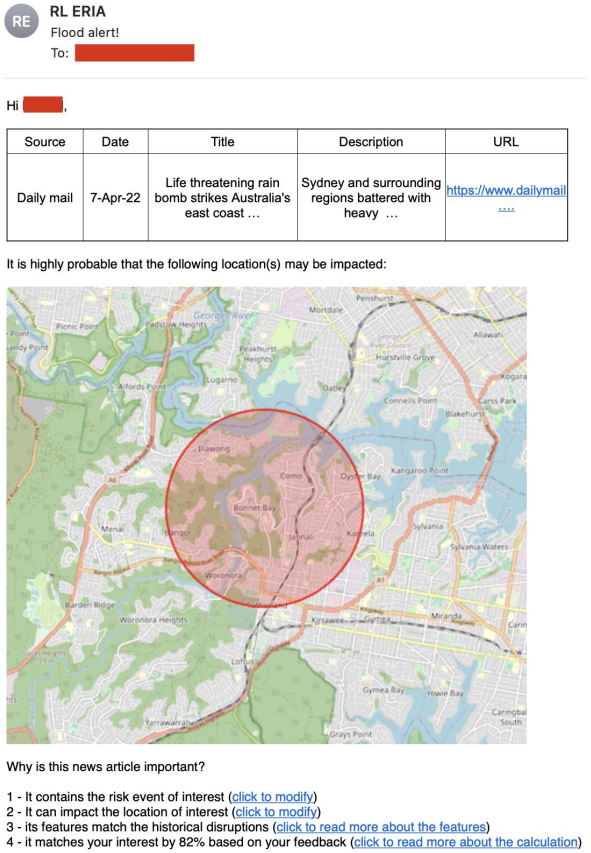
Category	RL-PRI	NR-ERIA
TP	87	78
FP	235	17
FN	6	15
TN	20	238

Metric	RL-PRI	NR-ERIA
accuracy	0.31	**0.91**
precision	0.27	**0.82**
recall	**0.94**	0.84
F1	0.42	**0.83**

News Recommendation Performance

- Simulate the user relevance feedback by considering the news mentioning at least one road closure location as relevant
- Performance test on 500 self-annotated news articles
- 0.74 F1-score performance.

	accuracy	*precision*	*recall*	*F1*
NR-ERIA	0.71	0.77	0.71	0.74



Conclusion

- Framework for new recommendation for environmental risk events
- NR-ERIA - proactively presents users with relevant news
- Outperforms baselines by 41 percentage points in terms of the F1-Score
- Demonstrates good performance in adapting to relevant feedback from users
- Can be applied to any environmental risk.

Thank You! Questions?