Detecting False Claims in Low-Resource Regions: A Case Study of Caribbean Islands

Jason S. Lucas **Limeng Cui** **Thai Le** **Dongwon Lee**

The Pennsylvania State University, PA, USA
{jsl5710,lzc334,thaile,dongwon}@psu.edu

Abstract

The COVID-19 pandemic has created severe threats to global health control. In particular, misinformation circulated on social media and news outlets has undermined public trust in government and health agencies. This problem is further exacerbated in developing countries or low-resource regions where the news may not be equipped with abundant English fact-checking information. This poses a question: "*are existing computational solutions toward misinformation also effective in low-resource regions?*" In this paper, to answer this question, we make the first attempt to detect COVID-19 misinformation in English, Spanish, and Haitian French populated in the Caribbean region, using the fact-checked claims in US-English. We started by collecting a dataset of real & false claims in the Caribbean region. Then we trained several classification and language models on COVID-19 from high-resource language regions and transferred this knowledge to the Caribbean claim dataset. The experimental results show the limitations of current false claim detection in low-resource regions and encourage further research toward the detection of multi-lingual false claims in long tail.

1 Introduction

In this work, we refer to *false claim* as assertions that are not supported by facts and are made with the objective of misleading or deceiving the public (Molina et al., 2021). Social media platforms enable people to independently publish and share media content without scrutiny filters for credibility and integrity[1]. Therefore, inaccurate, false, malicious, and propagandistic content have become abundant in social media. Furthermore, when false claims travel across regions and often get translated/modified, it becomes increasingly difficult for machine learning (ML) models to detect such false claims. Online surveillance (i.e., false claim detectors) systems are often primarily pre-trained on high-resource languages (e.g., English, Chinese). Despite significant progress in ML models, however, building and maintaining ML models in low-resource languages (e.g., Tagalog, Haitian Creole) are still challenging due to its scarce data or language lexicon and translation barriers which are indigenous to low-resource language settings.

This poses a natural question: "**how effective are computational ML solutions developed in high-resource regions to detect false claims circulating in low-resource regions?**" In this paper, to answer this question, we propose the first thorough case study on the detection of false claims in the Caribbean Islands.

Fact-checking initiatives are scarce and inept in low-resource settings, especially for the Caribbean Islands due to the cultural and linguistically diverse nature of their languages. The Caribbean region is a developing, heterogeneous, interconnected archipelago that is vulnerable to false claims campaigns. It consists of 35 states and territories bordering the Gulf of Mexico and Caribbean Sea[2]. The Caribbean has six official languages: Spanish, English, French, and Dutch, as well as two indigenous Creoles (Haitian Creole and Papiamento)[3]. Our data curation initiative shows that this region lacks essential technological resources and infrastructure to combat false claim propagation. Few fact-checking organizations exist, and they have limited data covering the Caribbean. Major news outlets such as Loop News make significant efforts to debunk false claims. These initiatives are essential but inadequate to effectively respond to prevailing false claims during crises.

In particular, we studied two research questions:

[1] https://www.who.int/news-room/feature-stories/detail/immunizing-the-public-against-misinformation

[2] https://studyincaribbean.com/about-caribbean.html

[3] https://www.caribbeanandco.com/caribbean-languages/

RQ1: How do ML models trained in high-resource languages perform with current Caribbean false claims?

RQ2: Are more sophisticated ML techniques (e.g., Transfer Learning), useful to detect false claims in the Caribbean?

Note that the focus of our investigation is on the COVID-19 related false claims in the Caribbean islands. ML models trained in high-resource languages are not easily transferable to low-resource languages. One of the main challenges comes from data scarcity (i.e., lack of labeled training data in low-resource languages). This issue is further exacerbated by the application of false claims detection that suffers from imbalance (i.e., where the number of labeled false claims is significantly smaller than that of labeled true claims). Therefore, to thoroughly study false claims in the Caribbean Islands, more sophisticated ML techniques that address indigenous nuances need to be tested.

2 Related Work

Since the onset of the COVID-19 pandemic, misinformation in different languages has been circulating on social media. The COVID-19 misinformation datasets can be roughly divided into two categories: monolingual and multilingual. CoAID (Cui and Lee, 2020), ReCOVery (Zhou et al., 2020), CMU-MisCOV19 (Memon and Carley, 2020), CHECKED (Yang et al., 2021) and COSTRAINT task dataset (Patwa et al., 2020) are monolingual datasets in high-resource languages (English or Chinese). CoAID is a diverse COVID-19 misinformation dataset, including 5,216 news about COVID-19, and ground truth labels. Multilingual datasets contain news pieces in multiple languages. MM-COVID (Li et al., 2020) contains false & real news content in 6 different languages. FakeCovid (Shahi and Nandini, 2020) has 5,182 COVID-19 fact-checking news pieces in 40 languages.

With the urge to combat the infodemic in developing countries or immigrant communities speaking low-resource languages, researchers have been studying how to transfer the pre-trained models on high-resource domains to low resource domains. Du et al. (2021) proposed a cross-lingual false claims detector called "CrossFake", which is trained based on a high-resource language (English) COVID-19 news corpus and used to predict news credibility in a low-resource language (Chinese). Bang et al. (2021) proposed two model generalization methods on COVID-19 fake news for more robust fake news detection in different COVID-19 misinformation datasets. In this paper, we chose the false claim detection in the Caribbean region as a showcase. It is a challenging problem due to the multiculturalism and multilingualism of Caribbean people. We studied how to leverage the pre-trained models from high-resource regions (CoAID) to detect misinformation in a low-resource region (Caribbean false claim data).

3 Main Proposal: Datasets and Research Questions

3.1 Caribbean Claims Dataset

This investigation utilized CoAID, a high-resources language COVID-19 false claims dataset written in English and curated from the United States (Cui and Lee, 2020). CoAID corpus comprises of 260,037 claims and news articles (Cui and Lee, 2020). This study assessed CoAID's pre-trained baseline models ability to accurately detect false claims in Caribbean dataset, given indigenous data challenges such as scarcity and language barrier.

Fact-checking institutions are trustworthy sources for determining the veracity of claims (Shu et al., 2019). They use rigorous methods to investigate the veracity and correctness of assertions, including references and URLs where false claims originate (Shu et al., 2019). Unfortunately, the Caribbean territory lacks these critical technological resources, notably fact-checking institutions with adequate regional data to combat the spread and growth of false claims. Instead, majority of fact-checking is primarily performed by respected Caribbean news outlets such as *Loop News* that do not consistently adhere to stringent fact-checking procedures. As a result, Caribbean fact-checked false claims are primarily assertions rarely linked to original content or the origin of such claims. This is the reason why we study Caribbean false claims detection in this work (Molina et al., 2021).

We manually crawled the accessible fact-checking and news organization websites given the aforementioned status quo. Then, we extract only original assertions, or alternatively extract the annotated claims when the original assertions were inaccessible. See Table 1 for all web sources that are crawled. We further inspect the Caribbean web

Table 1: Web sources and news claim articles curated from each source

Institution	Source Name	# Articles
News Outlet	Loop	188
News Outlet	Diario Libre	35
News Outlet	Aljazeera	25
News Outlet	St. Lucas Times	7
News Outlet	GBN	3
News Outlet	St. Vincent Times	3
News Outlet	Barbados Today	2
News Outlet	Mikey LiVE	1
Fact-checker	Poynter	9

Table 2: The language composition of the curated Caribbean dataset.

Language	Qty.	%
English	171	63%
Spanish	66	24%
French	36	7%

sources and solicited data from 9 institutions' websites detailed in Table 1. The final dataset totaled 273 articles published mostly between 2019 and 2022. All data collected are COVID-19 claims except for two Dominican Republic vaccine-related health claims published in 2010. The corpus consists of 121 annotated news and 152 original news claims. The dataset covers 3 of 6 official languages spoken in the Caribbean: English, Spanish and French (Table 2). The labels are comprised of 54% real claims and 46% false claims (Table 4). See Table 4 for the character length distribution of the two labels. The contents of our Caribbean dataset contains language cues that help ML model distinguish between false and real claims (Cui et al., 2020).

3.2 RQ1: Baseline Model Performance on Caribbean False Claims

To establish a baseline, we used pre-trained models trained on a large amount of English moderated COVID-19 data. Since CoAID contains a large amount of English news claims in the United States (Cui and Lee, 2020), the baseline models were trained on CoAID. We sectionized **RQ1** experiment into three sub tasks to ascertain empirical explainability. Each task uses different test sets to answer **RQ1**.

Task I Get the baseline performance using the CoAID dataset . Test set is CoAID dataset.

Task II Assess CoAID models' ability to predict Caribbean English false claims. Test set is Caribbean English claims.

Task III Assess the baseline model with another English Caribbean claims translated from Spanish and French. Test set is a translated to English Caribbean claims dataset .

3.3 RQ2: Applying Transfer Learning

This experiment adopted a self-supervised BERT-based transformer model, pre-trained on a large corpus of monolingual data. We encode the news using BERT. We adopt the binary cross-entropy loss function in the training. We fine-tuned the BERT model using the CoAID dataset and used it to conduct **RQ2** experiments.

Our hypothesis is that the answer to **RQ1** will not be sufficient to solve the task of detecting false claims accurately in Caribbean languages. Therefore, we propose a more sophisticated method to improve model's performance. Specifically, we studied the performance of transfer learning using a pre-trained BERT model. We break **RQ2** experiment in two tasks to answer this question and maintain empirical consistency with **RQ1** experiments.

Task IV Assess fine-tuned BERT model's ability to predict Caribbean English false claims. Test set is Caribbean English claims.

Task V Assess the fine-tuned BERT model with another English Caribbean claims translated from Spanish and French. Test set is a translated to English Caribbean claims dataset .

Table 3: Caribbean dataset composition of false and real news by **RQs** tasks respectively

RQS Tasks	Claims	False	Real	Total
RQ1: T2 & RQ2: T4	Original-En	95	76	171
RQ1: T3 & RQ2: T5	Translated-En	52	50	102

4 Empirical Evaluation

4.1 Set-Up

This research has three main test sets.

Table 4: Dataset statistics

Corpus	Size	Min_{char}	$\text{Mean}_{\text{char}}$	Max_{char}
Real claims	126	67	1187	3141
false claims	147	26	183	969

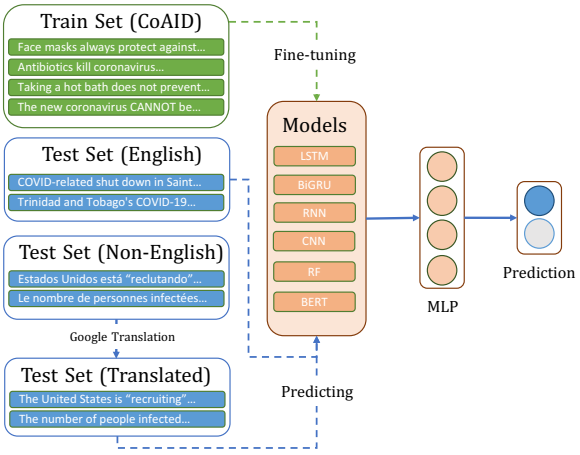

Figure 1: The framework overview of the false claim detector. For **RQ1**, we train the models on CoAID dataset and test on English Caribbean dataset and Translated English Caribbean dataset. For **RQ2**, we fine-tune the BERT model with CoAID dataset, and English Caribbean dataset and Translated English Caribbean dataset.

1. CoAID Test Set: this is only used for **RQ1**.
2. Original Caribbean English Set: this is used for **RQ1**: Task II and **RQ2**: Task IV (Table 3).
3. Translated-English Caribbean Set: this is used for **RQ2**: Task III and **RQ2**: Task V (Table 3).

Given the unique challenges with Caribbean false claims data, this research selected five baseline models:

- Long short-term memory (LSTM)
- Bidirectional Gated Recurrent Unit (BiGRU) (Bahdanau et al., 2015)
- Recurrent Neural Network (RNN)
- Convolutional Neural Network (CNN)
- Random Forest (RF)

The framework overview is shown in Figure 1. For the first task in **RQ1**, we first encode the news using GloVe (Pennington et al., 2014), a language pre-training model, and fit the embeddings into the models. The Glove wordembedding is used for all the baseline models except for Random Forest, which encodes the text with TF-IDF.

The baseline models were evaluated using F1, Kappa and Precision-Recall Area Under the Curve (PR AUC) scores from the models' output:

1. **Area Under the Precision-Recall Curve** (PR-AUC):

$$\text{PR-AUC} = \sum_{k=1}^{n} \text{Prec}(k)\Delta\text{Rec}(k),$$

where k is the k-th precision and recall operating point $(\text{Prec}(k), \text{Rec}(k))$.

2. **F1 Score**: $\text{F1 Score} = 2 \cdot (\text{Prec} \cdot \text{Rec})/(\text{Prec} + \text{Rec})$, where Prec is precision and Rec is recall.

3. **Cohen's Kappa**: $\kappa = (p_o - p_e)/(1 - p_e)$, where p_o is the observed agreement (identical to accuracy), and p_e is the expected agreement, which is probabilities of randomly seeing each category.

One of our primary interests is the precision-recall of the positive class, which is the positive false claim classification in our assessment of the models' performance.

We implement all models with Keras. The train and test sets use the 75:25 ratio, respectively. For all models, we use RMSProp (Hinton et al., 2012) with a mini-batch of 50 and the training epoch is 30. In order to have a fair comparison, we set the hidden dimension as 100 for all models. For the pre-trained BERT model, we use a BERT base model[4] (uncased) pre-trained on a large corpus of English data. All methods are trained on an Ubuntu 20.04 and Nvidia Tesla K80 GPU.

4.2 Results

First, to establish the research baseline performance, we pre-trained machine learning models on CoAID claims in English and tested them on English Caribbean false claims. Table 5 details the performance of the baseline models. LSTM model demonstrated high accuracy with F1 and Kappa evaluation matrices; however, CNN has the highest PR AUC predictive accuracy.

Next, Task II was performed using a total of 171 claims consisting of 95 false and 76 real Caribbean news claims detailed in table 3. Task I results are shown in table 6. Compared to Task I baseline output, Task II shows a general decline with all models' performance. Task II evaluation matrix scores are within a lower range compared to Task I. Task I output shows F1: 0.34 - 0.60, Kappa: 0.33 - 0.57 and PR AUC: 0.61 - 0.76. Task II matrix

[4]https://huggingface.co/bert-base-uncased

scores show: F1: 0. 33 - 0.54, Kappa: -0.64 - 0.02 and PR AUC: 0.51 - 0.56. LSTM outperformed all models with F1 while RNN having the highest Kappa and PR AUC scores.

The Task III assesses CoAID models' ability to classify Caribbean false claims translated from Spanish/French to English using Google Translate API. As shown in table 3, a total of 102 claims were used; 52 were false and 50 were real Caribbean news. Task III results, as shown in table 7, show an overall decrease in all models' predictive power in comparison to the baseline output in Task I. Task III evaluation matrix scores are within a lower ranges compared to Task I. Task I output shows F1: 0.34 - 0.60, Kappa: 0.33 - 0.57 and PR AUC: 0.61 - 0.76. Task III matrix scores shows: F1: 0. 30 - 0.53, Kappa: -0.52 - 0.02 and PR AUC: 0.50 - 0.55. BiGRU outperformed all models with F1 scores whereas RNN has the highest Kappa and PR AUC scores. Overall, all models showed a drop in performance when classifying translated Caribbean news claims in English.

Task IV encompasses running English Caribbean news claims through the refined BERT model and assessing its performance. The result from this experiment shows that transfer-learning with BERT out-performed Task II for **RQ1** models which used the same dataset detailed in table 3. The BERT model's F1 score is 0.55, whereas Task II for **RQ1** top F1 score is 0.54. Also, BERT's PR AUC score is 0.59, whereas Task II for **RQ1** top PR AUC is 0.56. However, BERT Kappa score of -0.16 was less than Task II for **RQ1** score, 0.02. Transfer learning technique using BERT achieved better predictive performance.

Finally, in the Task V, we assessed the pre-trained, fine-tuned BERT model's ability to accurately predict Caribbean false claims translated from French/Spanish to English. The results from this experiment indicate that BERT transfer-learning out-performs Task III for **RQ1** models which basically used the same dataset detailed in table 3. The BERT model's F1 score is 0.55, whereas Task III for **RQ1** top F1 score is 0.52. Also, BERT's PR AUC score is 0.57, whereas Task III for **RQ1** top PR AUC is 0.55. However, BERT Kappa score of -0.17 was less than Task III for **RQ1** score, 0.02.

Table 5: Comparison on Task I for **RQ1**. The false claims classification performance with standard deviation across five runs. The final prediction denotes the average of each evaluation matrix's score from all runs. The results in this table show that LSTM has the best F1 & Kappa scores, while CNN has the highest PR AUC score.

Model	F1	Kappa	PR AUC
LSTM	**0.5991** $_{0.060}$	**0.5721**$_{0.062}$	0.6923$_{0.032}$
BiGRU	0.5708$_{0.062}$	0.5457$_{0.062}$	0.6792$_{0.026}$
RNN	0.4147$_{0.188}$	0.3950$_{0.186}$	0.6651$_{0.074}$
CNN	0.5326$_{0.181}$	0.510$_{0.178}$	**0.7565**$_{0.097}$
RF	0.3439$_{0.121}$	0.3261$_{0.118}$	0.6152$_{0.085}$

Table 6: Comparison on **RQ1** Task II. The false claims classification performance with standard deviation across five runs. The final prediction denotes the average of each evaluation matrix's score from all runs. This experiment shows an overall performance declined observed compared to Task I baseline models output in table 5.

Model	F1	Kappa	PR AUC
LSTM	**0.5405**$_{0.059}$	-0.0704$_{0.099}$	0.5361$_{0.042}$
BiGRU	0.5020$_{0.056}$	-0.3164$_{0.139}$	0.4632$_{0.049}$
RNN	0.2013$_{0.120}$	**0.0213**$_{0.027}$	**0.5603**$_{0.040}$
CNN	0.3574$_{0.134}$	-0.1864$_{0.200}$	0.5151$_{0.045}$
RF	0.3316$_{0.012}$	-0.6427$_{0.015}$	0.5121$_{0.008}$

5 Discussion

5.1 RQ1 Experiments

RQ1: **Task I**. We established our baseline performance. It is clear from Task I results that CoAID baseline models are resilient with classifying claims despite imbalance dataset with majority real claims. The CNN PR AUC score was approximately 76% accurate in predicting the minority false claims regardless of imbalanced binary classification in the dataset. This suggest that CoAID high-resource language models perform fairly well at predicting news claims curated from the US high-resource language settings.

RQ1: **Task II**. assessed CoAID models' ability to accurately detect Caribbean news claims originally written in English. When classifying Caribbean news claims in English, we observed an overall performance decline in all models. Thus, this outcome suggests that pre-trained high-resource detection models perform poorly on low-resource language context data written in En-

Table 7: Comparison on Task III for **RQ1**. The false claims classification performance with standard deviation across five runs. The final prediction denotes the average of each evaluation matrix's score from all runs. This experiment shows an overall performance declined observed compared to Task I baseline models output in table 6.

Model	F1	Kappa	PR AUC
LSTM	$0.4649_{0.168}$	$-0.0735_{0.100}$	$0.4990_{0.089}$
BiGRU	$\mathbf{0.5268}_{0.049}$	$-0.1809_{0.166}$	$0.4954_{0.018}$
RNN	$0.2963_{0.175}$	$\mathbf{0.0226}_{0.114}$	$\mathbf{0.5543}_{0.037}$
CNN	$0.4884_{0.097}$	$-0.0830_{0.175}$	$0.5164_{0.091}$
RF	$0.3923_{0.009}$	$-0.5196_{0.008}$	$0.5384_{0.007}$

Table 8: Comparison on Task IV & V for **RQ2**. The false claims classification performance with standard deviation across five runs. The final prediction denotes the average of each evaluation matrix's score from all runs. A performance increase was observed in these experiments compared to Task II & III models output in table 6 and table 7 respectfully.

Task	F1	Kappa	PR AUC
Bert $_{IV}$	$0.5476_{0.018}$	$-0.1578_{0.306}$	$0.5852_{0.113}$
Bert $_{V}$	$0.5485_{0.047}$	$-0.1656_{0.039}$	$0.5695_{0.117}$

glish.

RQ1: **Task III**. assessed CoAID models' ability to accurately detect Caribbean news claims translated to English. When claims translated to English, pre-trained high-resource detection models under-perform on low-resource language context data. These results suggest a language translation loss. We propose the term language translation loss to encapsulate the phenomena that occur when a model's predictive power decreases due to translation nuances. Examples are politically loaded COVID-19 false claims propaganda and slang hidden in datasets that weaken signals impacting ML models' predictive power.

RQ1 Summary. **RQ1** results show a steady decline in all models' performance when introduced to Caribbean news claims that are originally written or translated to English (see Fig 3 & 2). These findings are clear indicators that high-resource language ML models are substandard with detecting low-resource language false claims such as the Caribbean region news claims data. These findings validated the research hypothesis: high-resource language models are not appropriate for detecting COVID-19 false claims in diverse, low-resources

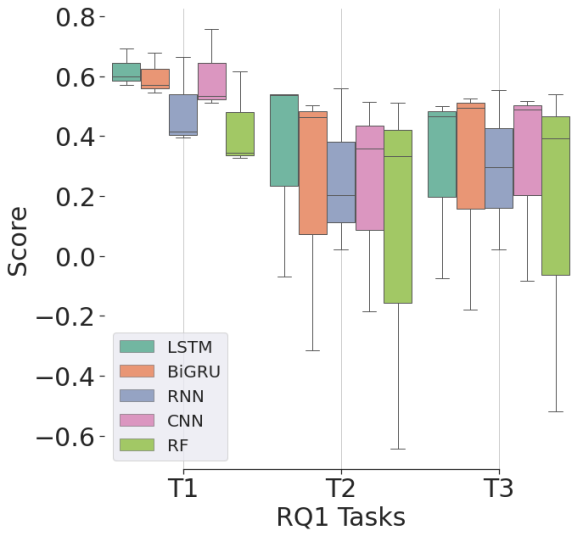

Figure 2: Overview of **RQ1** ML models' performance from Tasks I to III. The box plot shows that decline in ML models' performance on Caribbean data

language settings.

5.2 RQ2 Experiments

The above results prompted the need for more robust, novel, and clever techniques to best address the nuances and false claims phenomena specific to the Caribbean. Thus, we experiment with transfer learning methodology to garner insight on Caribbean false claims detection challenges.

RQ2: **Task IV & V** assessed transfer learning technique on Caribbean false claims detection. Task IV results indicate that the transfer learning technique using BERT achieved better predictive performance than English pre-trained high-resource language models. Similarly, Task V data demonstrate that the transfer learning technique achieves better model performance. Given indigenous Caribbean data challenges, these findings indicate that advance ML techniques have better learning mechanisms to address low-resource language setting detection (see Fig 4 & 5).

RQ2 Summary: results give clear indication that sophisticated, refined ML approaches achieve better performance. Transfer learning is shown to optimize performance with addressing Caribbean data scarcity issues. The linguistic similarity between CoAID and Caribbean false claims leveraged the model's performance through transfer learning.

6 Research Implication

News outlet websites, Factcheckcaribbean.com and Poynter.com are most reputable organizations to

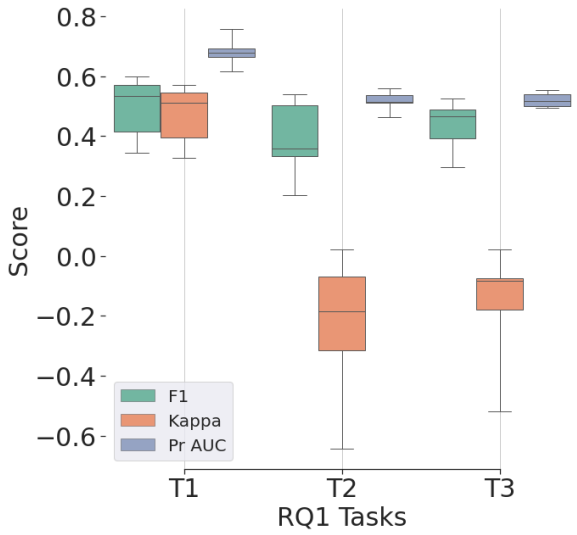

Figure 3: Overview of **RQ1** Models Evaluation Matrices from Tasks I to III. This box plot is shows a decline in performance using F1, Kappa and PR AUC.

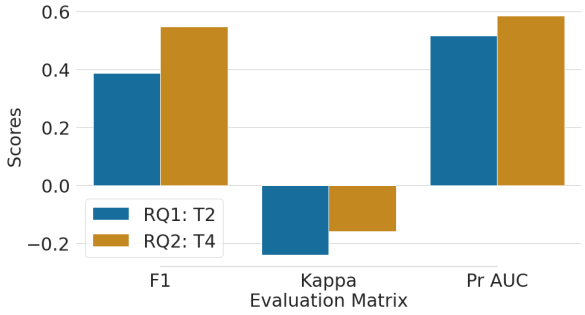

Figure 4: Overview of **RQ1** ML models' performance matrices scores compared to **RQ2** scores. This bar chat compares the performance of CoAID **RQ1**: Task II models performance with **RQ2**: Task IV fine-tuned Bert transformer model. This graph shows that transfer learning achieves better performance.

curate Caribbean false claims data. These institutions have limited data covering only a few islands. Loop news has the largest coverage and quantity of Fact-checked news claims compared to other sources. Although news outlets have more data, fact-checking institutions have better quality data. News outlet organizations do their best to verify and debunk false claims. In the Caribbean region there is need for more rigorous processes for false claims fact checking (Seo et al., 2022). This initiatives can be establish by non-government organization (NGOs) such as the Pan American Health Organization (PAHO) and Caribbean Public Health Agency.

This research did not address data imbalances in Caribbean data, which can be addressed by future work using state-of-the-art techniques. Future studies can focus in developing or utilizing interesting AI techniques such as meta-transfer learning, data augmentation techniques and Multilingual Bert transformer model to address false claims propagation in the Caribbean low-resource setting.

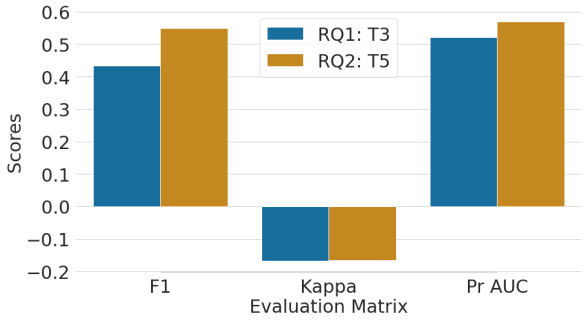

Figure 5: Overview of **RQ1** performance compared to **RQ2**. This bar chat compares the performance of CoAID **RQ1**: Task III models with **RQ2**: Task V fine-tuned Bert transformer model. This graph shows that transfer learning via Bert achieves better performance.

Context is imperative when considering computational solutions to address low-resource language setting false claims phenomena. In the Caribbean region context, numerous barriers implicate false claims detection when using high-resources language ML models. These barriers include: language, data scarcity, and rare full-coverage fact-checking institutions. Such barriers are not researched and thus poorly understood. This suggest the need for more exploratory studies to have in depth understanding of the false claims phenomena in the Caribbean region.

7 Conclusion

High-resource detection models have low accuracy with classifying Caribbean false claims data. Region-specific data challenges have shown to reduce the performance of high-resource ML models. This encourages the use of sophisticated ML techniques and AI methodologies to capture signals that current models are unable to recognize.

Our experiment with transfer learning has shown improvements with ML models' performance. The findings in this research support our hypothesis: high-resource language model performs poorly on low-resource language data. Future studies need to focus efforts on improving false claims detection in the Caribbean. A major challenge is that every island has its unique Creole, which complicates ML models trained in formal settings. Since the

Jamaican languages are a combination of several languages, even the best language translator are ineffective in accurately translating the language to English. This poses another difficulty to the problem of false claims detection.

False claims are the greatest threat to public health in the Caribbean and globally. As we saw with COVID19, if we do not address false claims, epidemic/pandemic diseases will spread exponentially (Brainard and Hunter, 2020).

Acknowledgements

This research was supported in part by NSF awards #1820609, 915801, and #2114824.

References

Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2015. Neural machine translation by jointly learning to align and translate. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.

Yejin Bang, Etsuko Ishii, Samuel Cahyawijaya, Ziwei Ji, and Pascale Fung. 2021. Model generalization on covid-19 fake news detection. In *International Workshop on Combating Online Hostile Posts in Regional Languages during Emergency Situation*, pages 128–140. Springer.

Julii Brainard and Paul R Hunter. 2020. Misinformation making a disease outbreak worse: outcomes compared for influenza, monkeypox, and norovirus. *Simulation*, 96(4):365–374.

Limeng Cui and Dongwon Lee. 2020. Coaid: Covid-19 healthcare misinformation dataset. *arXiv preprint arXiv:2006.00885*.

Limeng Cui, Haeseung Seo, Maryam Tabar, Fenglong Ma, Suhang Wang, and Dongwon Lee. 2020. Deterrent: Knowledge guided graph attention network for detecting healthcare misinformation. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 492–502.

Jiangshu Du, Yingtong Dou, Congying Xia, Limeng Cui, Jing Ma, and S Yu Philip. 2021. Cross-lingual covid-19 fake news detection. In *2021 International Conference on Data Mining Workshops (ICDMW)*, pages 859–862. IEEE.

Geoffrey Hinton, Nitish Srivastava, and Kevin Swersky. 2012. Neural networks for machine learning lecture 6a overview of mini-batch gradient descent. 14:8.

Yichuan Li, Bohan Jiang, Kai Shu, and Huan Liu. 2020. Mm-covid: A multilingual and multimodal data repository for combating covid-19 disinformation. *arXiv preprint arXiv:2011.04088*.

Shahan Ali Memon and Kathleen M Carley. 2020. Characterizing covid-19 misinformation communities using a novel twitter dataset. *arXiv preprint arXiv:2008.00791*.

Maria D Molina, S Shyam Sundar, Thai Le, and Dongwon Lee. 2021. "fake news" is not simply false information: A concept explication and taxonomy of online content. *American behavioral scientist*, 65(2):180–212.

Parth Patwa, Shivam Sharma, Srinivas PYKL, Vineeth Guptha, Gitanjali Kumari, Md Shad Akhtar, Asif Ekbal, Amitava Das, and Tanmoy Chakraborty. 2020. Fighting an infodemic: Covid-19 fake news dataset.

Jeffrey Pennington, Richard Socher, and Christopher D Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543.

Haeseung Seo, Aiping Xiong, Sian Lee, and Dongwon Lee. 2022. If you see a reliable source, say something: Effects of correction comments on covid-19 misinformation. In *Proceedings of the AAAI Conference on Web and Social Media*.

Gautam Kishore Shahi and Durgesh Nandini. 2020. Fakecovid–a multilingual cross-domain fact check news dataset for covid-19. *arXiv preprint arXiv:2006.11343*.

Kai Shu, Limeng Cui, Suhang Wang, Dongwon Lee, and Huan Liu. 2019. defend: Explainable fake news detection. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 395–405.

Chen Yang, Xinyi Zhou, and Reza Zafarani. 2021. Checked: Chinese covid-19 fake news dataset. *Social Network Analysis and Mining*, 11(1):1–8.

Xinyi Zhou, Apurva Mulay, Emilio Ferrara, and Reza Zafarani. 2020. Recovery: A multimodal repository for covid-19 news credibility research. In *Proceedings of the 29th ACM international conference on information & knowledge management*, pages 3205–3212.