Assessing Community Needs in Disasters: Transfer Learning for Fusing Limited Georeferenced Data from Crowdsourced Applications on the Community Level

Christin Salley[1]; Neda Mohammadi, Ph.D., A.M.ASCE[2]; Jiajia Xie[3]; Iris Tien, Ph.D., A.M.ASCE[4]; and John E. Taylor, Ph.D., M.ASCE[5]

Abstract: The effectiveness of infrastructure resilience relies on the seamless extraction of information, timely acquisition of critical knowledge, and heightened situational awareness. The ongoing utilization of digital citizen communication through social media with response organizations during disasters remains a valuable avenue for disseminating information, ensuring the effective utilization of public resources in emergency response to crisis events. Public agencies can use this information to examine community sentiments and discussions to assess, determine, and prioritize critical areas in need of assistance. However, there are limitations on harnessing precise geolocation information from social media, as well as a lack of mitigating bias of machine learning models used during such events. These limitations can restrict emergency management personnel's ability to locate and promptly delineate actionable insights. Here, we propose a semisupervised machine learning model that utilizes approaches such as transfer learning, topic modeling (i.e., Latent Dirichlet Allocation), and natural language processing to augment data from historical and current social media posts (i.e., Twitter) with community-driven application alerts (i.e., Waze) to achieve further evidence on the location and context of emergency events. The model is designed to also mitigate machine learning bias using the Wells–Du Bois protocol. A framework was developed for this process and is illustrated through a case study on Hurricane Ian and three previous hurricanes that occurred in Florida. This fusion provides increased situational awareness and may enhance the speed of emergency response. This study establishes a foundation for equitable, real-time crisis event detection, expanding organizations' response capacity in allocating resources and reducing harmful effects of disaster, particularly within public infrastructure systems. **DOI: 10.1061/JMENEA.MEENG-6208.**

Introduction

In emergency response situations, the significance of public infrastructure systems is paramount, serving as a fundamental resource for the efficient and effective management of disasters. This includes response and communication systems for disseminating information and locating those in need. Enhancements to emergency management systems are imperative to improve response execution and better serve society. The record-breaking 2004 and 2005 hurricane seasons (e.g., Hurricanes Ivan, Katrina, Rita, etc.) exposed shortcomings in emergency management, especially in federal response capabilities (Schmidtlein et al. 2008). When a natural disaster event is deemed so severe that it exceeds the ability of both state and local governments to respond, the Federal Emergency Management Agency (FEMA) issues it as a major disaster declaration; however, there is no set definition of what "beyond the combined capabilities of state and local governments to respond" means in order to receive assistance (FEMA 2023c). Thus, subjective judgments have the potential to shape the outcome of declarations and resource allocation. In the majority of cases, before requesting a disaster declaration to receive aid, state and local officials must conduct a damage assessment. With this, emergency management responders can face challenges in providing immediate intervention and relief for ongoing disasters as they await the assessment and declaration of a crisis event. Many areas are underserved by this process, resulting in inequities with distribution of aid (Schmidtlein et al. 2008). Before necessary federal assistance is given, state and local emergency management personnel need to make decisions on potential resources needed to mitigate the effects of disasters, especially when there is little time to decide or they must wait for a drawn-out damage assessment. There needs to be alternative systems in place that can adequately and quickly assess community needs when hazardous events occur that can pose a significant threat to communities and hinder relief endeavors. There is also a need for emergency management personnel to have more effective communication with citizens during a disaster through a tool or interface such as social media (Lovari and Bowen 2019). This is

[1]Ph.D. Candidate, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332. Email: csalley3@gatech.edu

[2]City Infrastructure Analytics Director, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332. Email: nedam@gatech.edu

[3]Ph.D. Student, School of Computational Science and Engineering, Georgia Institute of Technology, Atlanta, GA 30332. Email: jxie@gatech.edu

[4]Williams Family Associate Professor, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332. ORCID: https://orcid.org/0000-0002-1410-632X. Email: itien@ce.gatech.edu

[5]Frederick Law Olmsted Professor, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332 (corresponding author). ORCID: https://orcid.org/0000-0002-8949-3248. Email: jet@gatech.edu

Note. This manuscript was submitted on February 1, 2024; approved on May 28, 2024; published online on August 23, 2024. Discussion period open until January 23, 2025; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Management in Engineering***, © ASCE, ISSN 0742-597X.

where both social media and community-driven applications can further assist with identifying major disasters as they occur and potentially speeding up the process of receiving assistance through enhanced context of community needs.

Utilizing social media for natural disaster assessment continues to trend in research studies since social media platforms, such as Twitter, first emerged in 2006 and gained increased popularity (Wu and Cui 2018). Twitter is one of the world's largest social media platforms, having more than 368 million active users as of December 2022 (Tankovska 2022). Social media can be used for a multitude of activities and initiatives. As it pertains to disaster risk reduction, social media can be used in crisis response to serve as a listening function, to track events, for emergency planning and management, to foster connectedness and volunteering, to promote causes to raise donations or funds for those affected by disasters, and for academic research (Alexander 2014). Social media users can express their worry, relief, and other sentiments on such platforms during a disaster or interact with various community members and stakeholders to share information. It is common during natural disaster events that affected citizens turn to social media for relevant updates, along with seeking help from other individuals or professional organizations during all phases of the disaster cycle (Roy et al. 2020), as social media is faster than "traditional" news outlets for the dissemination of information (Wu and Cui 2018). Social media platforms have the ability to extract pertinent information through crowdsourcing, benefiting emergency management agencies' protocols and practices when this knowledge is analyzed and modeled for detection, prediction, and other aspects of emergency management. Social media can be integrated into the emergency management process, particularly when it comes to decision-making and assessing damages for major disaster declarations.

Platforms such as Twitter are social applications, whereas a community-driven application, as we define it, is a platform that seeks input from users for a particular situation or circumstance. Both types of platforms engage members of society, but social applications allow more creativity with content to build a unique network while community-driven applications result in more of a targeted network to share information. Waze, a popular community-driven navigation application, is used by more than 151 million monthly active users across the globe as of December 2022 (Smith 2023) and employed in operations by US State Departments of Transportation (DOTs), collecting hundreds of thousands of high frequency data pertaining to traffic and incident events a day for a given state. Some US State DOTs are using Waze data to enhance current communication systems already in place (Eastern Transportation Coalition 2017). However, there is a need for augmented systems to be developed and deployed, as currently Waze data is typically input either in parallel with other systems or by itself into DOT feeds. The lack of integration of Waze and other communication systems reveals a gap where Waze data can be merged with an outside source, such as social media data, which is already popularly used in disaster research, to increase situational awareness and aid decision-making. Community-driven applications such as Waze can address select shortcomings that most social media platforms currently possess, regarding lacking interactive features where users can send reports and update other community members specifically on certain harmful events through preestablished categories. Waze also has more precise location data and interactive geographical visualizations.

Social media, however, adds the community individual voice and sentiment of users that Waze lacks. It has also been revealed that despite platforms to aid disaster management using social media data, few are designed for citizen connectedness or use both social media and another platform (i.e., different type of input data such as community-driven applications) (Chair et al. 2019). In essence, Waze enhances Twitter with the higher volume of more precise coordinates related to events beyond the bounding box Twitter provides with its current API, while Twitter enhances Waze with adding more context to the categorized alert types (e.g., a Waze alert deemed "accident" when paired with a tweet in the same area could potentially show how many cars are involved, if someone was seriously injured or needs help, and possibly images related to the event). Thus, the two data streams complement one another. When information provided by active users on public platforms during crises is tagged with geolocation, it aids emergency responders in determining where people are located, evaluating community needs, and providing alerts and warnings to both citizens and first responders in regard to changing environments (Lindsay 2011). Georeferenced posts can strengthen situational awareness and aid in the four phases of emergency management (mitigation, preparedness, response, and recovery) by allowing agency officials to gauge and track community reactions and opinions in real time related to a disaster.

As social media continues to play a significant role in disaster studies and machine learning methods become more integrated into various societal operations, the imperative to develop or deploy strategies that effectively serve populations accessing and utilizing social media during crisis events is crucial. With this comes the growing concern over machine learning bias inherent in computational models intended for societal implementation. A transformative shift is underway in research, urging the normalization of equity-centered tools and approaches (e.g., bias mitigation strategies) within respective scientific domains when employing machine learning techniques (Bozeman 2024). While the effects of existing data analytics approaches and the fairness these techniques have on vulnerable and underserved populations during disasters remain relatively understudied (Yang et al. 2020), there are emerging approaches, such as qualitative measures to mindfully construct machine learning models (Monroe-White and Lecy 2022), that are promising. Incorporating these measures into study design and execution is paramount. Adhering to scholarly-based bias mitigating protocols can help bridge the knowledge gap concerning equitable measures in disaster informatics, thereby advancing our understanding in this domain. Therefore, the consideration of bias issues in social media data and computational models is relevant to our study, as we strive to align with this paradigm shift by designing and executing our research with these considerations at the forefront. To achieve augmented emergency management capacities with platforms such as Twitter and Waze, it is pertinent to understand past research that has been conducted on the use of historical data to enhance data sets, social media as a social sensor, the fusion of different data sets for natural hazards and disasters, and mitigating machine learning biases.

Related Research

Use of Historical Data

Historical data can be used to provide additional content or background knowledge on a particular problem or generate more robust models when trained on a previous event for tasks such as simulations of current or future events. The machine learning concept of transfer learning makes the use of past data easily capable of being integrated into prediction models, typically in situations where data is scarce or limited, as it is the ability of a system to provide the knowledge of the domain it is trained on (i.e., the source) to another domain (i.e., the target) (Neyshabur et al. 2020). The use of a pretrained model on historical data for transfer learning is seen across various infrastructure research areas such as in the energy sector for models that have limited energy related data (e.g., wind power production) (Hooshmand and Sharma 2019) and to infer energy consumption and demand for buildings (Peirelinck et al. 2022; Ribeiro et al. 2018). Other infrastructure areas where this approach is used include smart city applications such as activity recognition and building dynamics (Pinto et al. 2022) and transportation for GPS record estimation on speeding (Yu 2019).

As it pertains to emergency events, Halse et al. (2019) generated a simulator system that emulates real-time tweets from previous tweets based on their temporality with a crisis event. This was designed to replace collecting tweets directly from Twitter. The authors showed that historical tweets can be used for predicting current events and noted a recommendation that custom filtering should be used for training purposes (Halse et al. 2019). Other studies have made use of historical data for natural disasters through scenarios such as remote sensing for flooding (Pollard et al. 2018; Qi et al. 2009), predicting earthquakes (Yuen et al. 2005), and emergency management for validation of emergency vehicle travel times (Henchey et al. 2014) and decision-making (Romanowski et al. 2015). Nevertheless, there is still a need to investigate further how historical data collection can enhance detection and assistance for emergency management, especially for systems tailored for real-time events.

Twitter as a Social Sensor

Infrastructure (e.g., bridges, power systems, etc.) can have physical sensors to monitor or detect damage, but social sensors (e.g., Twitter) have been shown to detect events where physical sensors are lacking, such as providing detailed information about the failure (Tien et al. 2016). Twitter posts (i.e., tweets) can include information such as images and text descriptions, replies, retweets, favorites, and geographical metadata about where the user posted. Currently, about 1%–2% of tweets are geotagged, and location information can be either a precise location or a Twitter "place" (e.g., bounding box) (Twitter Developer Platform 2023). While this can be made useful in crisis situations, it is challenging to extract relevant information to assess and gain actionable insights with precise coordinates. Other typical challenges when dealing with social media pertain to trust, privacy, volume of data, availability of geotagged posts, and "rumors" or fake news that spread when people misuse social media (Rossi et al. 2018).

Social media platforms, like Twitter, have been used in a wide range of ways in the field of civil engineering. Social media analysis has been used to enhance traffic conditions (Athuraliya et al. 2015; Sujon and Dai 2021), detect emergency events via natural language processing (NLP) (Verma et al. 2011; Wang and Taylor 2019), analyze its use in construction operations (Tang et al. 2017; Azhar et al. 2019), and determine disaster-related impact assessments on the built environment (Chen et al. 2020; Fan et al. 2020; Chen and Ji 2021). Also, social media has been used to study human mobility by identifying city-scale patterns (Wang and Taylor 2016), user polarity of sentiments (Wang and Taylor 2018), and urban-level spatiotemporal energy demand prediction (Mohammadi and Taylor 2017). The use of social media integrated into other systems can improve situational awareness through augmenting communications and informing decision makers on resources and aid needed in affected areas (Yin et al. 2012).

Additionally, Twitter with community-driven applications has been used in research involving the Waze navigation application to examine real-time traffic flow data from Waze in comparison to Twitter data congestion (Sidauruk 2018). Twitter has been shown to be less reliable in comparison to other crowdsourced data feeds (Amin-Naseri et al. 2018) in terms of fewer tweets being made at night versus during the day, most being during peak traffic hours, and, while covering arterials well, most tweets coming from the center of a city (i.e., providing less coverage from outside areas) (Gu et al. 2016). Twitter data will continue to be used in various fields for analysis and detection within communities; however, the number of tweets during a disaster can fluctuate depending on the disaster and how engaged community members are on the platform. There are also cases where tweets relevant to a disaster are smaller in volume than expected, necessitating more data points to be ingested into a model for further community perspective (Salley et al. 2022). This requires augmenting social media data sets with applications that are more equipped for real-time event detection (e.g., Waze), which social media largely lacks. While social media still holds potential for enhancing actions taken in emergency management phases to better protect people, property, and the environment during crisis events, it underscores the need for further research to explore the interdependencies of different systems and address gaps in data integration and analysis to further harness its capabilities.

Fusing Data for Natural Disasters

Data integration is critical for timely and effective crisis information collection and communication, data analysis, and emergency personnel decision-making for disasters; however, data integration can be a challenging task (Peng et al. 2011). Within the field of disaster informatics, established research has highlighted significant challenges pertaining to data integration (Ogie and Verstaevel 2020). Purohit et al. (2019) have identified three specific challenges associated with the integration of open-source data for disaster management. These challenges include the heterogeneity of data sources, where the diverse formats of multiple data sources can make merging difficult; the inconsistency of data sources, which results from different words or semantics used across data sources, making the establishment of an interpretable structure challenging; and the incompleteness of data sources, characterized by the scarcity of data or the lack of relevant information (Purohit et al. 2019).

Researchers have initiated efforts to tackle these challenges by devising data fusion methodologies. Some approaches aim to merge data from various sources to assess earthquake impacts, incorporating damage data from forecasts and remote sensing with field measurements (Loos et al. 2020, 2022). Additionally, they have been applied in situations such as the assessment of damage caused by Hurricane Matthew, where unmanned aerial vehicles (UAVs) and social media data, such as tweets, were integrated (Yuan and Liu 2018). Moreover, these fusion techniques have been employed in urban analytics by combining sensor data and social data (Psyllidis et al. 2015). With research emphasizing the intricate nature of data integration in disaster management, there is a continuous need for thoughtful solutions to address them effectively. Research also highlights the importance of approaching data integration responsibly by collecting "good data" (e.g., data that has quality content, truthful, etc.) that is unbiased (Nargesian et al. 2022). While integrating different data sets can help alleviate potential biases, it remains essential to mitigate bias through the implementation of some set of standards or well-defined parameters to ensure reliable computations (Albahri et al. 2023).

Mitigating Machine Learning Bias

Studies have shown how race, social class, and/or placement play a role in populations experiencing social and environmental injustices related to hazards and disasters (Adeola and Picou 2017; Bodenreider et al. 2019; Griego et al. 2020; Hamideh and Rongerude 2018; Nejat et al. 2022; Wright 2011). With the growing integration of machine learning into social decision-making and everyday routines, such as emergency management, there has been a call to control and/or assess fairness in computational efforts to avoid the risk of exacerbating bias. There is no consensus or widespread agreed upon definition of "fairness" as it pertains to bias and equity in machine learning; how fairness is determined depends on the research question and situation it is applied to. This paper defines fairness as the act of addressing bias with the objective of diminishing the potential adverse consequences upon societal integration. Research has established three ways to quantitatively perform bias mitigation before, during, and after model execution: in the training data, while training machine learning models, and on trained machine learning models (Hort et al. 2022). Previous research has investigated fairness through approaches such as fairness testing algorithms (i.e., inconsistencies between existing and mandated fairness requirements of a software); these are typically binary and divide the population into privileged and unprivileged based on a sensitive attribute that protects against unfairness such as age, race, gender, etc. (Chen et al. 2022). Issues with this type of quantitative testing include that it relies on sensitive attributes when in practice that information may not be available in a data set (Awasthi et al. 2021). For instance, Twitter does not provide such demographic information from its users to researchers. Additionally, studies report that current fairness algorithms and metrics cannot handle multiclass and nonbinary problems (Hort et al. 2022). Therefore, if your data set does not have sensitive attribute data or has more than two labels, current models that assess fairness would not be adequate.

Critiques have surfaced asserting that quantitative research undervalues equity, and, when confronted with equity shortcomings, statistical measures are employed to defend the validity of such an analysis (Gillborn et al. 2018). However, with fairness testing there is no guarantee or empirical evidence demonstrating its applicability or effectiveness in real-life scenarios (Chen et al. 2022). Researchers further expose that the report of low bias scores using such quantitative approaches does not automatically equate to actual fair application of models (Hort et al. 2022). Social scientists strongly argue for the imperative of combining machine learning models with a qualitative approach to thoroughly assess bias mitigation efforts (Monroe-White and Lecy 2022). Protocols such as the Wells–Du Bois protocol for machine learning biases could be deployed to overcome systemic inequities ingrained in data sets which historically sought to oppress marginalized communities (Monroe-White and Lecy 2022). Use of intentionally building machine learning models with qualitative protocols is a promising alternative for the limitations and discrepancies within current algorithms for bias control.

This study addresses three research gaps: (1) for social media analysis methods, integration with community-driven applications that may improve capture of incidents relating to emergency preparation and response (Chair et al. 2019) with historical data; (2) creating a method to effectively augment location-specific social media data with community data to address the shortcomings that exist in the ability to more rapidly, and effectively, communicate and respond to crisis events (Lovari and Bowen 2019; Purohit et al. 2019); and (3) incorporating equity-based practices to mitigate machine learning bias (Yang et al. 2020; Monroe-White and Lecy 2022). Our overarching research objective is to integrate these streams while remaining vigilant about potential biases inherent in machine learning within the disaster management context. We will delve deeper into this topic in the subsequent sections, particularly in the articulation of our research question and the remaining discussion in the paper. While the high-level contribution of this study lies in its ability to enhance situational awareness and facilitate the optimized allocation of resources for emergency management efforts, its impact extends to both methodological and engineering management realms. Methodologically, this study advances technical approaches for integrating and harmonizing disaster-related big geo-data and extracting critical semantic information essential for communities or emergency responders. This progress is achieved through a novel and reproducible framework curated with bias mitigation in mind, leveraging multiple machine learning techniques such as transfer learning and semisupervised learning to merge complementary data sets. From an engineering management standpoint, the framework serves as a valuable tool for enhancing community safety and resilience during disasters. Specifically, for DOTs, it offers support for management operations by providing a user-friendly visual interface that aids in prioritizing resources and addressing issues on the state highway system more efficiently. Unlike current setups, which can require operators to navigate through multiple tabs or channels, this streamlined approach consolidates relevant data, advancing the knowledge boundary and offering a comprehensive solution to improve overall safety and resilience.

Intervening and alleviating disasters as they occur in real-time poses an issue for many emergency responders. Again, before necessary federal assistance is given, state and local emergency management personnel need to make decisions to prepare and respond to disasters to mitigate their effects with available resources. This can be facilitated through a more community focused, equitable approach to better understand local needs of citizens and engaging with community discussions that are occurring. To address the aforementioned gaps and research objective, we investigated the following research question:

What is the impact of integrating social media with community-driven applications for the capture of incidents related to emergency management, mitigating machine learning bias, and validating its respective effectiveness (e.g., accuracy)?

In the following sections, we introduce a framework that assesses community needs and provides context for emergency responders using machine learning techniques to train the model on previous events and fuse data from the social media platform Twitter and the community-driven application Waze. We also mitigate machine learning bias of the framework using an equity-based protocol to show how our methodology integrated equity measures. We anticipate that this machine learning-enabled framework can enhance event detection, provide further situational awareness about an emergency event, and thus improve emergency response.

Machine Learning-Enabled Framework

The scope of this framework is twofold: (1) use historical data to develop a robust model and incorporate more community insights; and (2) perform data integration across social media and community-driven platforms at the community scale. The reason this study is at the community scale (i.e., neighborhood to city scale) is to correspond to the bounding box locations of Twitter, which will be explained in more detail later. To achieve these aims, we fused Twitter and Waze data and propose machine learning approaches and spatiotemporal data fusion that utilizes labeling from transfer

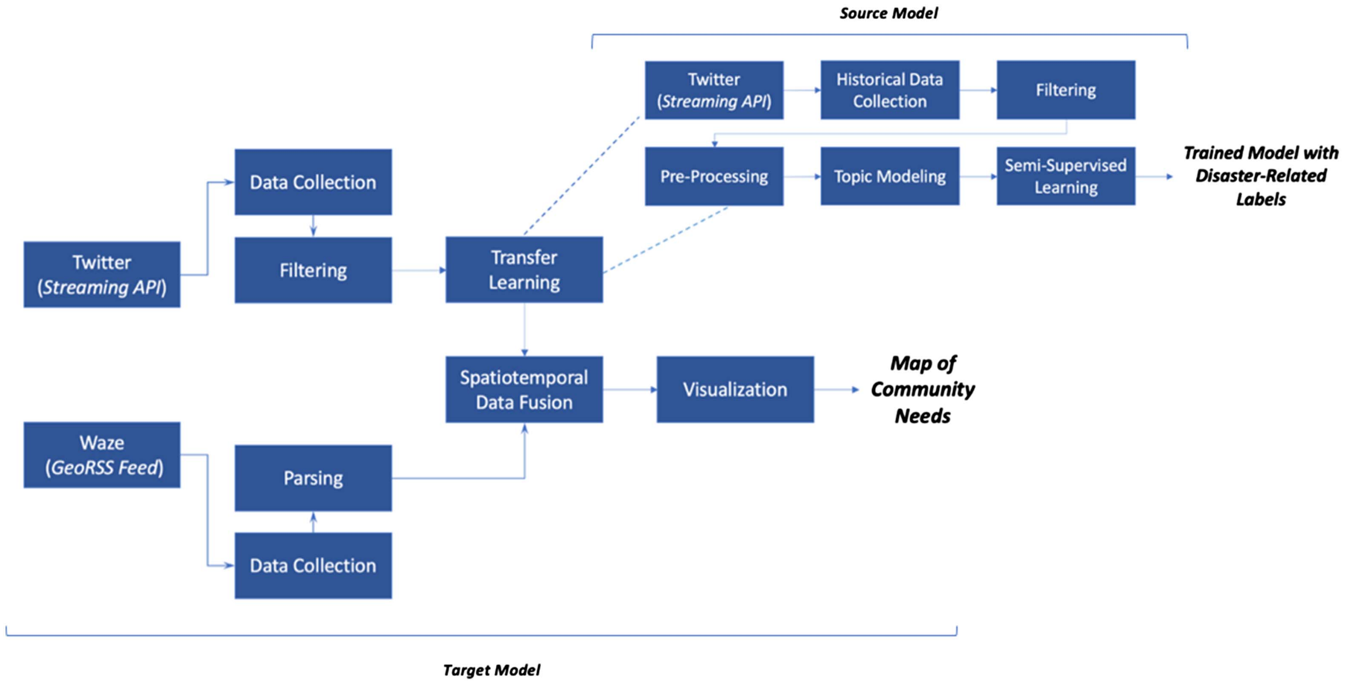

Fig. 1. High-level framework for georeferenced data fusion (Twitter and Waze) workflow, including the process for transfer learning. The transfer learning process leverages pre-existing knowledge, which in this case is derived from historical tweets, to create the source model. Subsequently, the source model is trained and integrated into another domain, referred to as the target model. Here, the domain knowledge from the source model is effectively incorporated to amplify performance and understanding.

learning for Twitter and Waze data sets related to natural "disasters." Fig. 1 illustrates the overall framework developed for the integrated approach with the goal of augmenting georeferenced social media data (i.e., Twitter) with corresponding data from a community-driven application (i.e., Waze). The framework overall utilizes the techniques of transfer learning, NLP, Latent Dirichlet Allocation (LDA), semisupervised learning (SSL), and spatial fusion to produce the output of an augmented data set that classifies each Twitter and Waze pairing to elucidate community conversations and issues. In Fig. 1, the source domain model is the component projected up and to the right from the transfer learning box, which produces the output of labels. The rest of the process occurs in the target domain model, which produces the output of a map of community needs. The following sections will explain in further detail the workflow of the framework outlined in Fig. 1.

The evident biases of social media data should not discourage efforts to mitigate biases in models that utilize this data. Even if acceptable metrics in terms of precision, recall, and F1 score are achieved, it remains essential to assess the potential impacts of this work in practice through recognizing biases. The Wells–Du Bois protocol is a tool designed to assess whether research qualitatively achieves a baseline level of bias mitigation in social scientific research for neutral data collection and machine learning execution. It consists of three dimensions and seven items: bad data—(1) inadequate data and (2) tendentious data; algorithmic bias—(3) harms of identity proxy, (4) harms of subpopulation difference, and (5) harms of misfit models; and human intent—(6) do no harm and (7) harms of ignorance (Monroe-White and Lecy 2022). In this study, these items were viewed through the domain of utilizing social media in emergency management. Detailed information regarding each item and the corresponding steps undertaken in this study to assess the fulfillment of the specified standard is provided in a later section. In the following sections, we outline the methods employed with simultaneous detail provided on how we integrated example data from a case study to illustrate the framework.

Source Domain Model

Historical Data Collection

To enhance the presence of the community's perspective, we incorporated historical data into our framework. We analyzed different major disasters of the same type (i.e., hurricanes) to track sentiments over time and to capture different communities who may have been engaged for one disaster but not another. The assumption posits that within the historical events under examination, varying geographical regions or demographic groups will be represented, as each catastrophic event impacts distinct audiences. Historical data were collected in the form of 43,416 tweets from three hurricanes that occurred in Florida in 2020: Hurricane Eta (November 7, 2020, through November 12, 2020), Hurricane Isaias (July 31, 2020, through August 4, 2020), and Hurricane Sally (September 14, 2020, through September 28, 2020) (FEMA 2023a).

Filtering

In this process, tweets in the state of Florida were extracted and filtered based on location and keywords in the form of a disaster-based glossary we developed (see Table 1). Past studies have shown that the use of hashtags can limit the number of irrelevant tweets (Brunila et al. 2021). However, in this case the quality of data with hashtags was not sufficient; therefore restricted keywords were determined to be used after several tests were run and analyzed using one or the other (or both). Hashtags are also constantly changing and evolving. Therefore, for the model to be more generalizable the decision was made to use only keywords. Hence, we created a disaster-based glossary of common words related to natural disasters that could indicate a crisis event. The disaster-based word

Table 1. List of keywords used to create disaster-related word corpus and their source

Keywords	Source
Affected, Airburst, Avalanche, Chemical, Climate, Coastal, Collapse, Damage, Death, Derecho, Disaster, Disease, Drought, Earthquake, Epicenter, Epidemic, Explosion, Famine, Fire, Flood, Flow, Fog, Food, Freeze, Frost, Hazard, Homeless, Hurricane, Ice, Impact, Injured, Injury, Lahar, Lake, Landslide, Lava, Lightening, Liquefaction, Loss, Missing, Niño, Poisoning, Power, Rain, Risk, Seiche, Shake, Sinkhole, Soil, Storm, Subsidence, Surge, Tornado, Transport, Tsunami, Typhoon, Volcanic, Vulnerability, Wave, Wind, Winter	EM-DAT (CRED 2009)
ARC, CDC, CERT, Community, Crisis, DHS, Drill, Emergency, EMS, EOC, EPA, Evacuate, Evacuation, FEMA, HAZMAT, IMT, Incident, JIC, JIS, NGO, NIMS, Procedure, Protection, Rescue, Responder, Response, Shelter, Structural, Threat, Tree, Warning, Watch, Water	FEMA (2023b)
Building, Critical	UNDRR (2016)
Accident, Alert, Construction, Jam, Road, Traffic, Weather	Waze (2017)

glossary with over 100 words was developed based on the Emergency Events Database (EM-DAT), FEMA, United Nations Office for Disaster Risk Reduction (UNDRR), and Waze. Web scraping was performed to collect the keywords from the respective sites, and manual inspection was done to ensure there were no duplicate terms among the sources and words that were fully applicable or used commonly were represented from longer phrases (e.g., used the word "damage" instead of "estimated damage" in the EM-DAT database). The tactic was designed to maximize the number of relevant tweets that could be collected.

Preprocessing

Textual data can be informal and not structured in a way to enable classification processes. Text from social media can be noisy containing special characters (i.e., emojis and symbols), slang and misspelled words, hashtags, URLs, and more (Salas et al. 2017). Text mining approaches ease the difficulties associated with the time consuming and inconsistent process of manually cleaning data and have been proven to have higher accuracy than no preprocessing techniques being performed at all (Mhatre et al. 2017). In order to prepare all text for the classifiers, we removed these additional elements (e.g., extra URLs, white spaces, special characters, uppercase words, and unnecessary words) using FastText (FastText 2023). This is done through standard techniques such as tokenization (i.e., breaking a sentence into words), stop words removal (i.e., simplifying text and removing words that add no meaning such as "a" and "the"), stemming (i.e., finding the root/stem of the word), and lemmatization (i.e., generating the base or dictionary form of a word) (Mhatre et al. 2017). After these preprocessing techniques, we had a clean corpus of words, and the fused textual information was converted to vectors to be utilized in the LDA model to generate labels based on all text.

Topic Modeling

To obtain the labels for classifying the data set, LDA topic-based modeling was performed. We utilized different LDA-related Python packages to model our preprocessed tweets, running the model with different parameters (altering the number of topics and words within each topic), and using a standard deviation test to determine the number of topics. From the standard deviation test, four to six topics was identified as the preferable range, and running the model on these three different options, five topics was deemed as the most optimal. Also, to further refine the model, we modified the parameters further to only include words that were nouns, adjectives, and verbs in the preprocessing portion. After running several tests, the number of topics was set to five (with 10 words in each topic). This process of identifying topics and refining the model is a conventional practice in the application of topic modeling. The topics were subsequently interpreted through internal validation within the team to ascertain their practical coherence and relevance. Interpreting each topic, the topics were "0": broadcast/news (e.g., anything to do with the news, the government, alerts, etc.); "1": power (e.g., anything to do with power outages, power lines, power systems, lights, Wi-Fi, Internet, etc.); "2": traffic incidents (anything to do with car crashes, congestion on the roads, evacuation, etc.); "3": forecast/weather (anything to do with the climate, flooding, etc.); and "4": miscellaneous (anything that does not fit into these categories and/or has nothing to do with a disaster). The last topic also acts as an additional filter to catch tweets that made it into the text corpus that may have a different interpretation of a word in the disaster-based glossary. Throughout the rest of this paper, the labels will primarily be referred to by their corresponding numerical identifiers as mentioned in the previous sentence. The top eight most frequent words identified in the LDA model were watch (appearing 7,772 times), broadcast (appearing 4,910 times), storm (appearing 4,553 times), chance (appearing 3,293 times), tonight (appearing 2,960 times), live (appearing 2,391 times), forecast (appearing 2,323 times), and alert (appearing 1,960 times). These top words indicate discussion around a time-sensitive storm and that needs pertain most frequently to the topics connected to weather and what is being outlined in news reports. This is beneficial to operators as it can help them with tasks immediately after a disaster such as crafting public safety messaging relevant to what people may or may not already know about the disaster or emerging risks responders will face when dispatched.

Semisupervised Learning

The topics from the LDA model described in the previous section (i.e., "0": broadcast/news; "1": power; "2": traffic incidents; "3": forecast/weather; and "4": miscellaneous) were used as labels in this SSL approach. The model was generated using a label spreading package (Zhou et al. 2004). Roughly 1% of the historical storms data set was manually labeled, leaving 99% unlabeled. To determine the 1% of the fused data that would be manually labeled, the data set was randomized using a function in Python, and then labeled with equal distribution of each topic classification. Labeling 1% of the data was determined to be the guideline for how much of the data should be labeled, as the aim is to limit the manual training of the data, and labeling 1% has been found to achieve high accuracy (Chen and Wang 2017). The annotators consisted of two members from our research team. Annotators divided the 1% of the data set that required labeling according to a well-defined and mutually agreed-upon set of label definitions. After each designated annotator completed their assigned portion, they collaboratively reviewed and discussed the labeling. In the rare event of any disagreement, a third team member was available to arbitrate. This internal validation protocol was implemented throughout the labeling process (Chowdhury and Zhu 2023).

The data were split into 70% being the training set and 30% being the testing set. This was fed into the model, generating

J. Manage. Eng., 2024, 40(6): 04024055

pseudo-labels for the entire data set based on the model's prediction. A validation set of 20% of the data was extracted from the training set prior to this analysis to provide an unbiased evaluation of the model fit on the training set and to tune the hyperparameters. Once the model was completely trained, we ran the testing set to see if the model could predict labels on this data set with adequate accuracy through our evaluation metrics discussed in the next section, and labels were successfully generated from the model for our testing set. With this task completed, we now had a trained model that was ready to be used for the transfer learning process.

Table 2. Classification report for label spread for historical data

Topic	Precision	Recall	F1-score	Support
0	0.95	0.90	0.93	2,270
1	0.82	0.59	0.68	2,221
2	0.93	0.69	0.79	666
3	0.86	0.87	0.87	2,476
4	0.81	0.95	0.88	5,392
Accuracy	—	—	0.85	13,025
Macro average	0.87	0.80	0.83	13,025
Weighted average	0.85	0.85	0.85	13,025

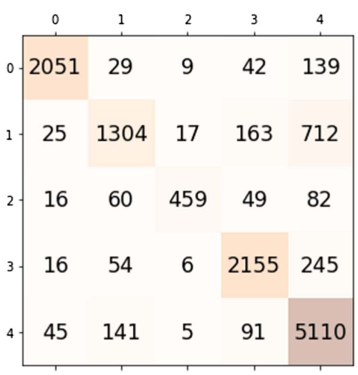

Fig. 2. Confusion matrix for label spread for historical data (no cross validation).

Evaluation Metrics for the Model

To assess the validity of the model, precision (i.e., true positives over all that was predicted as positive), recall (i.e., true positives over all that should have been predicted as positive), and F1-score (i.e., combination of precision and recall, the overall accuracy) were calculated, along with creating confusion matrices. Table 2 shows the classification reports for the historical data set. Higher precision indicates fewer false positive predictions, higher recall indicates capturing most positive cases with few false negatives, and a higher F1 score demonstrates robust performance by achieving both high precision and high recall simultaneously, effectively classifying positive cases while minimizing false positives and false negatives. "Support" outlined in the last column of Table 2 and other classification report tables is the count of occurrences experienced in each class. A confusion matrix was generated (Fig. 2) based on these initial scores and a single-fold analysis. A 10-fold cross validation was then done again on the precision, recall, and F1-score metrics to generate a final accuracy, and a confusion matrix was also produced for this cross validation based on the label spread performance of the model (see Fig. 3). These were completed for the five topics (i.e., the five classes in the classification reports).

Target Domain Model

Case Study

According to FEMA, the state of Florida has experienced more than a dozen major disaster declarations in the last decade alone, ranging from tornadoes to hurricanes, with one of the most recent major disasters being Hurricane Ian (FEMA 2023a). Despite Florida being a coastal state that experiences numerous natural disasters, historically it has just under a 70% success rate with being granted major disaster status for aid disbursement (Schmidtlein et al. 2008). Hurricane Ian is tied as the fifth strongest hurricane to hit the United States and began on September 23, 2022, in the central Caribbean as a tropical storm and three days later on September 26, 2022, became a hurricane (NOAA US Department of Commerce 2022). When Hurricane Ian approached southern Florida on September 28, 2022 it was a Category 4 storm, and it left Florida the next day, with intense winds and rainfall, as a tropical storm again heading to South Carolina (NOAA US

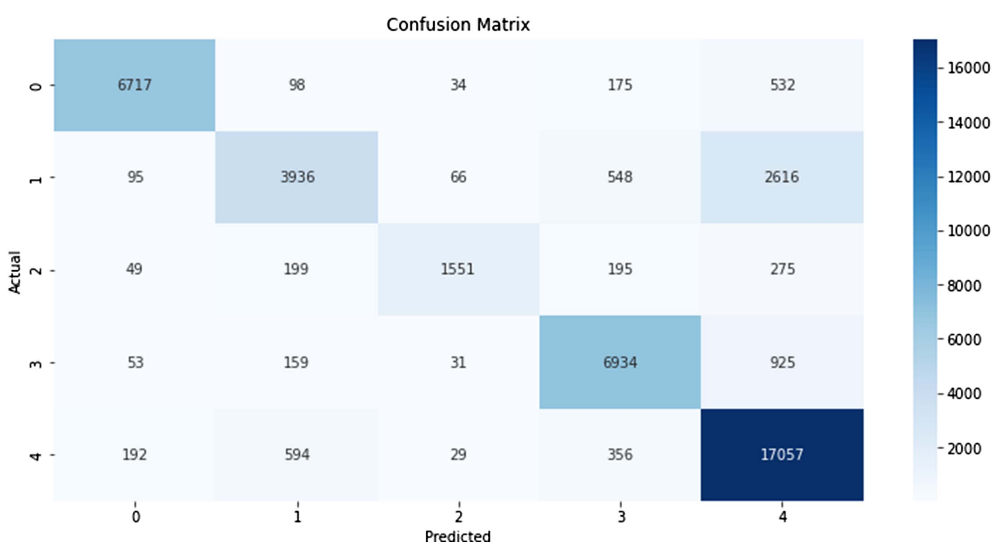

Fig. 3. Confusion matrix for label spread for historical data (cross validation). With average accuracy across folds: 0.833.

Department of Commerce 2022). This case study, which we conducted to demonstrate our framework, focuses on the "immediately after" part of the disaster cycle (i.e., meaning right after the disaster has left an area) to see community conversations based on the impacts of the hazard. This also aligns with when damage assessments would typically take place. Other studies have investigated two-week periods beginning at the landfall or origin of when the storm begins (Samuels and Taylor 2020) and showed that at two weeks the discussion gradually decreases (Zou et al. 2018). Since this study is focused on immediately after the storm exits an area, and investigates when people could be most engaged, a weeklong period was studied for Hurricane Ian, making the "postdisaster period" September 29, 2022, to October 6, 2022. Hurricane Ian also was declared a major disaster on September 29, 2022 (FEMA 2023a), emphasizing the importance of promptly understanding the ongoing situation with the expeditious declaration.

Data Collection

Data were collected from Twitter and Waze during this period for Hurricane Ian with 10,209 filtered tweets and 15,913 Waze alerts. Twitter data were collected through Twitter's public Application Programming Interface (API). Data were retrieved for Hurricane Ian from a live data collection stream developed in Python within our Lab. The data were collected by year, month, day, and hour and stored in JavaScript Object Notation (JSON) format. The Waze data used for this study were Waze alerts, which were initially collected through the Waze GeoRSS Feed that is shared with Connected Citizens Program (CCP) partners, such as the Georgia Department of Transportation (GDOT), for further configuration. The Waze data were collected in Extensible Markup Language (XML) format, showing pertinent information such as the date and time of an incident, precise coordinates, type and subtype of an alert, street name where the alert occurred, country, road type, report rating, confidence, and reliability of incident feeds within the bounding box of the state of Florida.

Parsing (Waze) and Filtering (Twitter)

For Waze alerts, the provided GeoRSS feed collected data needed to be transformed into a readable format for the model. The same filtering process in the source domain model for tweets was executed here to maximize the number of relevant tweets on Hurricane Ian in Florida.

Transfer Learning

The transfer learning process is outlined in Fig. 1 and the "Source Domain Model" section. The model built in the source domain model is already trained and ready to be used at this point in the target domain model. There is no more training or manual processes. It is fully automated since the source domain model was saved and applied here. When the data for Hurricane Ian was run through the saved model, just as in previous evaluations, the predicted labels were assessed with precision, recall, and F1 scores (see Table 3) along with confusion matrices (see Figs. 4 and 5). The outputs demonstrated that the model is a reliable model, even having a higher accuracy score than the source domain model.

Spatiotemporal Data Fusion

The advantage of data integration lies in the ability to enhance and enrich data sets by leveraging their complementary nature. Despite the inherent differences between the Twitter and Waze data sets, such as variations in data collection formats, accuracy, reliability, completeness, and contextual nuances, this study navigated these heterogeneities through careful parsing and selection of the relevant components of each data set, with the aim to fuse them for improved accuracy. Both Twitter and Waze data sets have date, time, location, and textual information pertaining to event detection for a natural disaster. With Twitter, the textual data is the tweet itself (i.e., what the user has posted), and the location is in the form of a precise location or bounding box with the current API (most are the latter). For Waze, the textual data is the alert type and subtype given to the report that the user assigns to the incident, and it provides a single coordinate pair. Waze alerts are classified with the following types: accident, jam, weather hazard/hazard, miscellaneous, construction, and road closed. The subtypes provide more detail for each alert type such as weather hazard/hazard displaying subtypes pertaining to fog, hail, rain, snow, hurricanes, etc.

To fuse these data sets, we identified and paired the tweets and Waze alerts within minimum spatial proximity. This was achieved using the Haversine distance between locations [see Eq. (1)], which can be used to calculate distance between latitude/longitude pairs for real-time classification (Zubiaga et al. 2017). With the current API's bounding boxes for tweets described as being able to be as large as 25 miles in width and height (Twitter Developer Platform 2023), in order to refine the spatial scale of the tweets collected, they were further filtered to identify neighborhood or city information (i.e., shrinking the size of the bounding box). To obtain coordinates for each neighborhood or city bounding box the center of each bounding box was found, which has been done in previous work on a larger scale (Zubiaga et al. 2017). The crowdsourced data produced by Waze reportedly experience about a 30 s delay in reporting, which can cause an incident to be recorded 0.8 km (i.e., approximately half a mile) away (Amin-Naseri et al. 2018). Despite potential reporting delays, Waze data still offer greater location accuracy than Twitter data, providing more precise and reliable coordinates, particularly crucial in scenarios such as transportation planning and management, where precise location

Table 3. Classification report for label spread for Hurricane Ian (with transfer learning from historical data)

Topic	Precision	Recall	F1-score	Support
0	0.93	0.73	0.82	205
1	0.85	0.49	0.62	89
2	0.92	0.87	0.89	146
3	0.77	0.92	0.84	968
4	0.91	0.85	0.88	1,655
Accuracy	—	—	0.86	3,063
Macro average	0.87	0.77	0.81	3,063
Weighted average	0.86	0.86	0.85	3,063

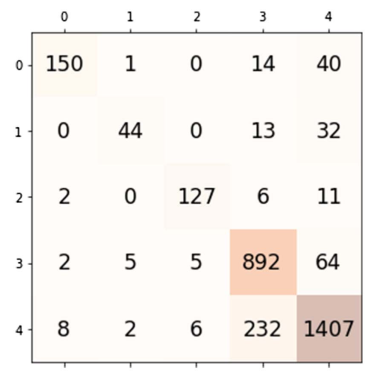

Fig. 4. Confusion matrix for label spread for Hurricane Ian (no cross validation).

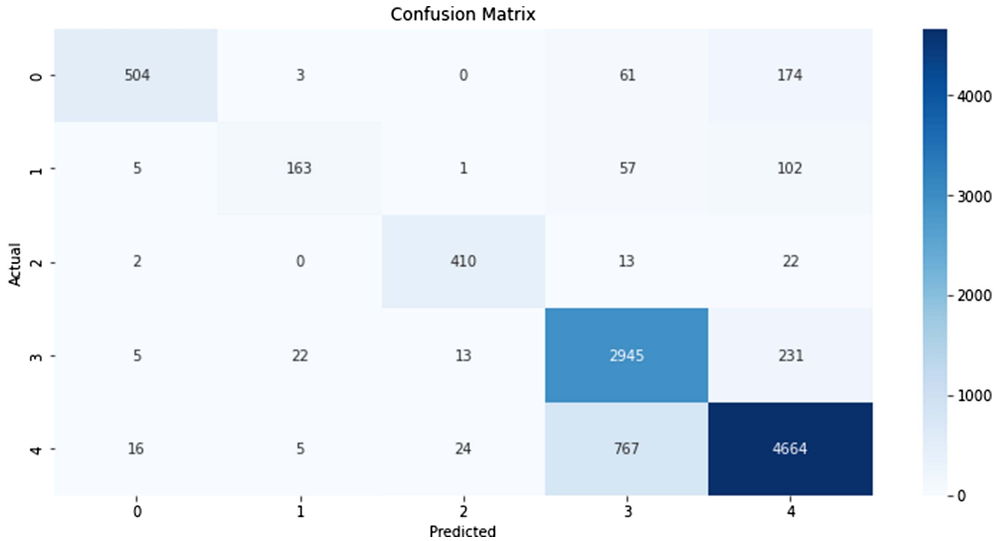

Fig. 5. Confusion matrix for label spread for Hurricane Ian (cross validation). With average accuracy across folds: 0.834.

information is essential. Given this information, the parameters for selecting the tweets closest to Waze reports were set within 1.61 km (i.e., 1 mile) of location to one another to account for delays of up to 60 s in reporting an incident. The merge is based on location and date, displaying all attributes of both feature layers in one data set. The output is a fused data set, matching a tweet with the nearest Waze alert with each data point showing the paired data sets' information along with a classification label and distance from one another. Upon completion, there were 2,566 Twitter and Waze pairings. It is important to note that the same tweet can be paired with multiple Waze alerts depending on proximity. This fusion yields a validation achieved by combining the now reported, dependable Twitter data from the model with the already reliable Waze data, serving as a confirmation for the convergence of these two data sets. This validation is further reinforced by the illustrative case study, emphasizing the efficacy of merging these data sets for enhanced decision-making:

$$d = 2r\arcsin\left(\sqrt{\sin^2\left(\frac{\varphi_2-\varphi_1}{2}\right)+\cos(\varphi_1)\cos(\varphi_2)\sin^2\left(\frac{\lambda_2-\lambda_1}{2}\right)}\right) \quad (1)$$

where φ_1 and φ_2 = latitude coordinates of two points; and λ_1 and λ_2 = longitude coordinates of two points.

Generalizability

To test generalizability of the proposed framework, the model was run again on data in a different state, for a different storm. Tropical Storm Zeta in Georgia was the storm used to test the generalizability of the model. The disaster began on October 21, 2020, in the western Caribbean as troubled weather, being slow to develop, and then on October 23, 2020, forecasts reported Zeta disturbance being brought into the southern Gulf of Mexico along with greater odds of this weather being an actual storm (NOAA US Department of Commerce 2020a). On October 25, 2020, the tropical storm formed and strengthened over the next few days as it began approaching the US, with its highest strength being reported as a category 3 hurricane (NOAA US Department of Commerce 2020a). Georgia news outlets and other mediums (e.g., the National Weather Service) reported on Tropical Storm Zeta with the storm striking and leaving the state of Georgia on October 29, 2020 (FEMA 2021; NOAA US Department of Commerce 2020b). This test focused on dates pertaining to after this incident, the same "right after" timing with a weeklong period done for Hurricane Ian in the case study above. The "postdisaster period" was October 30, 2020, to November 6, 2020. The results of the study (Table 4) show that the model is not only acceptable for the state of Florida but can also be transferred and used effectively for Georgia. The results suggest the possibility of using this framework for other events in other locations as well, as it has an accuracy score that is within a 3% margin of the model outputs from Hurricane Ian. Confusion matrices with and without cross validation were also produced for the Georgia event (see Figs. 6 and 7).

Table 4. Classification report for label spread for Tropical Storm Zeta in Georgia

Topic	Precision	Recall	F1-score	Support
0	0.94	0.48	0.64	127
1	0.95	0.28	0.43	76
2	0.93	0.92	0.93	74
3	0.85	0.65	0.74	343
4	0.81	0.98	0.89	968
Accuracy	—	—	0.83	1,588
Macro average	0.90	0.66	0.72	1,588
Weighted average	0.84	0.83	0.81	1,588

Case Study Visualization

Fig. 8 shows tweets fused with Waze alerts being spatially mapped (displayed with the labels for each pairing), with context embedded in each icon on the map for emergency personnel to have access to in a visual interface. An example pairing is also pictured in Fig. 8 in the inset table, showing how a tweet can add further context to a Waze alert beyond its original classification. As the miscellaneous label was noted as an additional filter, it is not displayed in the final visualization. The visualization shows a substantial amount of fused data points related to forecast/weather and traffic. The discussion of these topics in particular aids emergency operators and responders with actions such as feasibility of potential infrastructure repair, cleanup, or evacuation planning. Engaging in weather-related

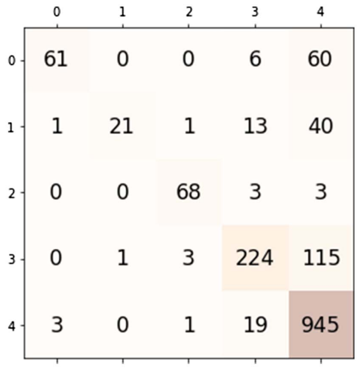

Fig. 6. Confusion matrix for label spread for Tropical Storm Zeta (no cross validation).

discussions enables responders to obtain crucial information, such as the extent of severe flooding in a building or instances of lightning striking trees. Traffic discussion allows them to know what major route or highways are jammed, using the augmented data from Waze. They can then select the proper evacuation routes that avoid congestion, with context of how long it might take for traffic to clear up (e.g., accident, debris on highway, how many lanes are closed, etc.). Furthermore, with the location information, local governments will also be able to see exactly how a portion of a neighborhood might be affected by a disaster, which can help guide what preparedness plans or mitigation tactics can be deployed. All this knowledge assists them in comprehending the specific impacts of the disaster and to tailor their response efforts accordingly.

Mitigating Bias of the Model

As mentioned previously, the model was designed toward equitable practices utilizing the Wells–Du Bois protocol. Again, it comprises three dimensions and seven items, most effectively explained by posing the corresponding questions: bad data—(1) inadequate data (i.e., *Does the data exhibit systematic omissions or misclassifications of certain subpopulations?*) and (2) tendentious data (i.e., *Does the model reflect subjective decisions?*); algorithmic bias—(3) harms of identity proxy (i.e., *Is there a potential for the model to exhibit systematic biases toward specific races, genders, or social classes?*), (4) harms of subpopulation difference (i.e., *Does the algorithm demonstrate varying performance outcomes among different subgroups?*), and (5) harms of misfit models (i.e., *How does the model assess error? What are the broader public and social implications of this research?*); and Human Intent—(6) do no harm (i.e., *Are you ensuring transparency regarding the objectives and aims of your research?*) and (7) harms of ignorance (i.e., *Have you carefully examined the potential unintended consequences of your research?*) (Monroe-White and Lecy 2022). For each dimension, we discuss the following actions taken to mitigate bias in our study:

1. Bad Data
   - Inadequate data: Reporting data sizes and metrics is needed to overcome this. Interpreting this to social media in disaster management, applicable descriptive statistics are given in the training data sets and are separated by group identities (i.e., the classification of labels).
   - Tendentious data: Disclosure of human judgement is needed. We disclose in our model that 1% of our model is manually labeled; however, the labels themselves that were generated do not pose such bias as they were constructed with an LDA model based on the textual information provided by the people and not influenced by the researchers.
2. Algorithmic Bias
   - Harms of identity proxy: The model did not consider race, gender, or social class as factors for the desired outcome. The way Twitter and Waze are both designed, it does not provide such demographic data per post, and only the text and location were needed in this study. This is because in the context of immediate emergency response, it is challenging to prioritize one life over another, as natural disasters strike without regard for such distinctions. While indicators such as community vulnerability would be crucial for assessing measures such as these, the focus of this study was to identify and classify imminent needs of social media users.
   - Harms of subpopulation difference: This study caters to the population that relies on these platforms for communication during a crisis, thereby considering them as a distinct demographic subset in itself. As mentioned previously, only

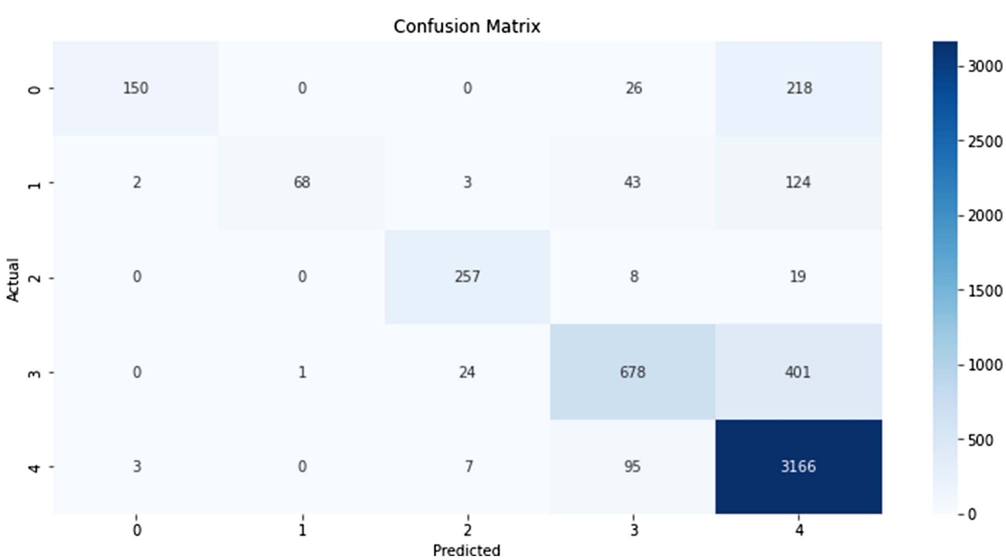

Fig. 7. Confusion matrix for label spread for Tropical Storm Zeta (cross validation). With average accuracy across folds: 0.812.

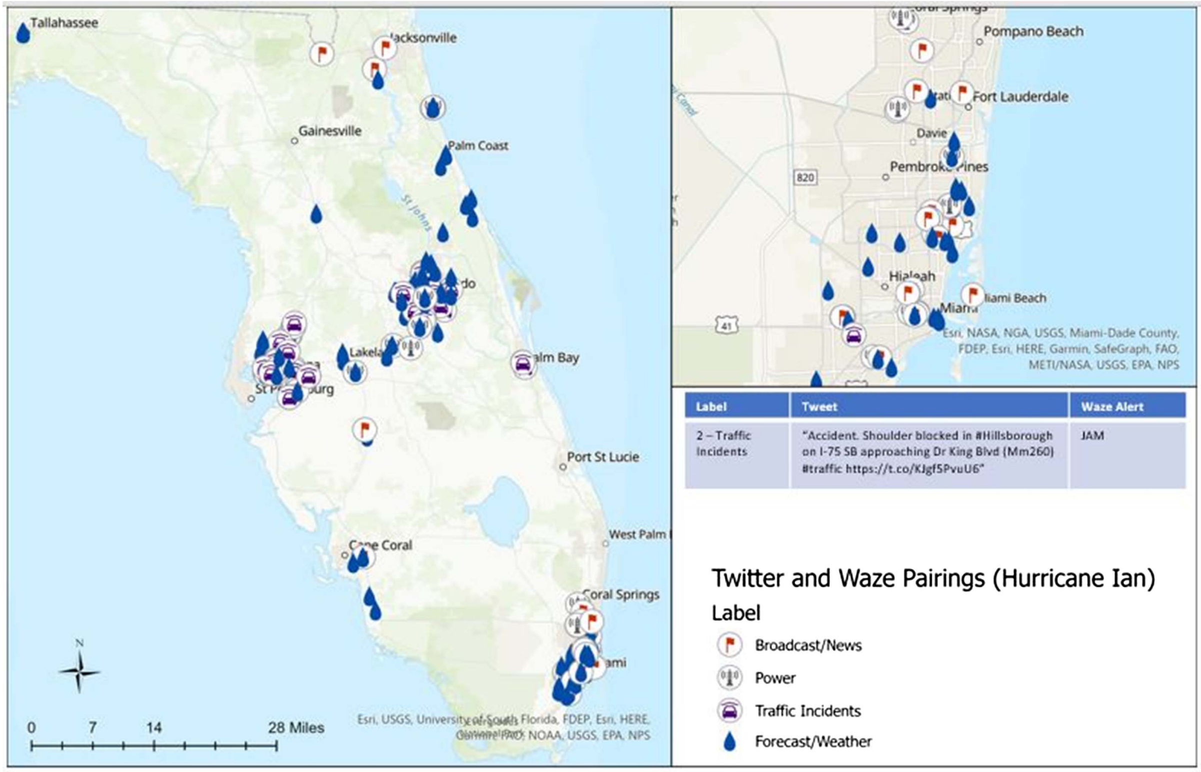

Fig. 8. Example of a classified Twitter and Waze pairing output for Hurricane Ian in Florida, zooming in on an area displaying multiple topics from the study. (Sources: Esri, USGS, University of South Florida, FDEP, Esri, HERE, NOAA, USGS, EPA, NPS; Esri, NASA, NGA, USGS, Miami-Dade county, FDEP, Esri, HERE, Garmin, SafeGraph, FAO, METI/NASA, USGS, EPA, NPS.)

the text and location of each data source were used; the final output does not note who the user was but solely what they said and where they are for enhanced context of the disaster (i.e., maintaining consistency).
- Harms of misfit models: The model undergoes cross validation to avoid overfitting. The goal of this research is serving the population who uses such platforms to communicate during a crisis to aid and enhance the decision-making process for emergency management personnel. The impact of this work can improve allocation of resources for emergency events.

3. Human Intent
   - Do no harm: The goal is for this work to be implemented into agencies such as state DOTs; we make sure to document and share this work with both applicable stakeholders and the research community.
   - Harms of ignorance: We have examined the unintended consequences of our research. For instance, if a cyberattack were to occur on this model as it was deployed into society, adversaries would have knowledge of communities that are currently at risk and what they proclaim to need. Adversaries could send phishing emails, tweets, etc., to try to take advantage of those impacted populations. Inclusion of protective measures should be done for such a system when employed.

It is important to acknowledge that this process does not guarantee the model has no problems in terms of potential bias, but rather serves as a means to implement mitigation strategies and strive toward achieving a threshold for reducing biased research practices; this protocol emphasizes that mitigation efforts are not primarily aimed at solving the issue at hand but rather at acknowledging and addressing the issue prior to the implementation of a model (Monroe-White and Lecy 2022). Implementing the Wells–Du Bois protocol played a significant role in bias mitigation efforts for our model. The protocol guided us in systematically considering harmful impacts, enhancing transparency, and minimizing bias, thereby fostering fairness in our approach. Our confidence in the protocol's efficacy stems from extensive deliberation and discussions with various stakeholders, particularly operators, who provided valuable insights into the challenges and potential biases inherent in our framework, leading us to our final workflow. A notable example that highlights the effectiveness of the protocol is the recognition of the "harms of ignorance" and the critical reflection on the implications of our framework's output. This realization prompted us to engage in further discussions with key personnel, revealing the crucial need for cybersecurity recommendations to protect vulnerable populations from further harm. As a result, new studies were initiated to develop preventive measures for systems similar to ours (Salley et al. 2024). Overall, by directing our attention toward the population we aim to serve, those with access to and who use social media and community-driven applications, we meticulously evaluated our data collection and model implementation and adhered to the seven items outlined in the Wells–Du Bois protocol to actively work toward mitigating potential biases.

Discussion

The framework developed in this study contributes to the larger discussion of enhancing community perspectives in disaster informatics. As discussed relative to FEMA's disaster declaration process, it is crucial for emergency management agencies to receive information, from models such as the one from this study, that can represent and assess what communities need in near real time from the people themselves. Communities, such as areas in Puerto Rico after Hurricane Maria, have been documented as being failed by federal agencies due to these organizations not being fully prepared to respond to disasters or being able to anticipate locals' needs (Sullivan and Schwartz 2018). Knowing what an area needs while a crisis occurs can prevent missteps such as this. It has also been shown that federal disaster relief falls short of equitable measures, leaving disenfranchised and historically marginalized communities at a disadvantage, with FEMA itself stating, "For disaster preparedness, mitigation, response and recovery to drastically improve in 2045, emergency management must understand equity and become equitable in every approach and in all outcomes" (National Advisory Council 2020). This is why some of their goals in their 2022–2026 FEMA Strategic Plan are to have more of a "people first approach" and "meet current and emergent threats" (FEMA 2023d). To address these needs, our framework is centered around community perspective and constructing a system that keeps equity at the forefront and acknowledges current disparities and potential impacts of machine learning efforts. Studies such as this add to the growing body of knowledge of determining ways to more accurately, and effectively, recognize community needs during or after a disaster to better serve society with a community centered approach.

The framework developed in this study also guides decision-making toward equitable response to disasters. It is important that computational models work toward fairness as most are currently unfair due to training data that can disproportionately affect marginalized populations and not thinking of the harmful effects a model can have when integrated in the real world (Monroe-White and Lecy 2022). Disparities such as wage gaps, mortality, and access to care can be seen in all areas of life and the built environment, and when exposed to natural disasters such disparities can be exacerbated when not accounted for properly. Research indicates that there are still few studies on infrastructure and social equity (Dhakal et al. 2021). Social equity systems research in emergency management and disaster research has employed analyses of social vulnerability to expose how disenfranchised populations recover at a slower rate back to their predisaster state (Kim and Sutley 2021). Often the occurrence of natural disasters is viewed as "equal opportunity" in the sense that storms, tornadoes, etc., do not intentionally target a certain population; they just occur haphazardly and can damage everyone just the same (Lieberknecht et al. 2021). While it is true the damage done by major disasters on the surface can be the same (e.g., power outages, extensive flooding, etc.), the postdisaster and recovery phase is not an "equal opportunity" when it comes to the dissemination of resources and the time it takes to rebuild a community depending on its pre-existing conditions. This phenomenon may arise as a result of a limited conceptual framework that fails to account for the disparities inherent in contemporary machine learning techniques employed to assist communities, wherein the incorporation of equity benchmarks or the pursuit of fundamental bias reduction may be overlooked. In some cases, without adequate support a disenfranchised population that is met with an emergency event may never fully recover because they already began at a deficit. Addressing such disparities in the physical, economic, and social environments could improve infrastructure systems and approach equity to establish a culture that provides just assets, funds, policies, and education to communities that need it most.

Limitations and Future Work

Although Twitter is the world's largest microblogging social media network and a popular platform used to extract information for research purposes, latitude and longitude pairs (i.e., precise coordinates) of tweets are no longer automatically attached to all tweets, reducing the number of precisely located posts since about 2016 (Maurer 2020). It is optional for users to share their location; thus most tweets collected through Twitter's streaming API are not georeferenced with exact location but with bounding boxes from place information instead (Maurer 2020). For the framework established in this paper, both precise coordinates (when provided) and bounding box coordinates were utilized for tweet location information on the neighborhood and city level.

As highlighted earlier as a challenge with social media, the volume of data is an ongoing and probable obstacle when dealing with Twitter data. Using social media or community-driven data in disaster research is heavily reliant on citizens participating and providing useful information on such platforms. While this information can be useful to measure other metrics or relationships, in real-time tracking when trying to assess the needs of a community, an extensive community voice is needed. Additional efforts can be made by relevant agencies and stakeholders to educate community members about leveraging these platforms as a means of meaningful interaction, fostering actionable outcomes. Alternatively, they can also prioritize the promotion of their existing systems to ensure greater engagement and effectiveness. However, we discovered, through a disaster-based glossary for filtration and the use of transfer learning, more relevant tweets can be found than previous work (Salley et al. 2022). This framework accounts for scarcity of data and allows for a faster, more automated process when evaluating social media data.

Lastly, as mentioned in the "Mitigating Bias of the Model" section, Twitter and Waze do not provide specific demographic information such as race, gender, or social class on a per-post basis. Consequently, the focus of this study was not on sociodemographic vulnerability but rather on the needs of populations affected by crises that rely on these platforms for communication. These populations can be considered a distinct demographic subset in themselves, highlighting Twitter's role as a social sensor. While it is important to note that these platforms do not represent the entire population, and recent, comprehensive demographic information has not been readily available since around 2013 (Wang and Taylor in 2016), it is worth mentioning that recent data suggest certain trends. For instance, among its multimillion users, approximately 37% of users are female, while 63% are male on Twitter; furthermore, users between the ages of 25 and 34 exhibit high activity, representing around 38% of users worldwide (Dixon 2023). The pursuit of representativeness of the data needs to be continuously asked and answered to further analyze any limitations of the research or further generalizability of its results (Kumar and Ukkusuri 2020) as no data set suits every single task and all can have some sort of limited scope (Nargesian et al. 2022). The most beneficial utilization of social media is achieved when it is used in conjunction with existing emergency management systems at local and government agencies, such as a Department of Transportation (DOT), as it does not holistically represent an entire community and other measures should be used to further contribute to decision-making.

Future work related to this framework should adapt this framework to completely online machine learning labeling, negating any manual process. Future work should also further examine historical data in relation to the typical engagement that communities have with emergency management (e.g., good or bad relationships, levels of engagement on social media, etc.) on different spatial and temporal scales. This study explored neighborhood and city levels day by day, but exploration of county and census level data on an hourly or minute basis may reveal other insights. This could also reveal how a community already utilizes local agencies in these spaces and can provide a baseline for how useful social media networks may be for real-time tracking in a particular area. How citizens currently use social media should also be continually reevaluated, as new platforms are emerging and old ones are obsolescing and updates to current platforms occur often. Additionally, it is essential to undertake extensive quantitative and qualitative investigations when dealing with complex issues like these to effectively counter computational biases during model construction and deployment. Given the nature of these challenges, which rely on data and computational solutions, it becomes imperative to continue to investigate a range of bias mitigation methods. More approaches to determine the most appropriate strategy tailored to the unique demands of the research problem should be investigated.

This study can be taken further in the future through the development of a process that works toward fairness more and establishing measures for proper allocation of resources. Presently, there exists a paucity of scholarly investigation concerning the integration of equity metrics or protocols in the utilization of social media within the scope of emergency management. Further exploration is called for to thoroughly examine ongoing constraints as it relates to equity measures in this domain, such as with specific population subsets, through leveraging quantitative analyses, qualitative insights, and stakeholder feedback to offer a comprehensive evaluation of protocol effectiveness and implications for meeting community needs during and after disaster events. Finally, this study suggests avenues for future exploration regarding the generalizability and transferability of the pretrained model. While our initial assessment showcased transfer learning across two areas within the same region and for similar meteorological phenomena, additional research is needed. Specifically, further investigation is warranted to assess the sensitivity to the features of the training data set and the approach/procedure of model training on other scenarios. Additionally, it is crucial to determine whether the model maintains its accuracy when applied to different areas and types of disasters, based on relevant training data. Methodological considerations in future works must be thoroughly examined to optimize the efficacy of transfer learning in such diverse contexts.

Conclusion

This study was able to identify areas in Florida that were impacted by a disaster with augmented context of specific needs based on classification of a paired data set employing machine learning techniques. The final output for the historical data identified pertinent topics that could be transferred and applied for use in future hurricanes. The final output for the postdisaster period of Hurricane Ian data showed extensive discussion related to the forecast and weather issues related to the storm, as well as the traffic occurring within communities due to the disaster. This research addresses the postdisaster period of a natural disaster, focusing on disasters classified as hurricanes and tropical storms for emergency responders, to be able to aid civilians and distribute the necessary resources to specific areas more quickly and efficiently. The model addressed the following research question: "What is the impact of integrating social media with community-driven applications on improving the capture of incidents related to emergency management, mitigating machine learning bias, and validating their respective effectiveness?" The investigation demonstrated the integration of social media data with community-driven applications, thereby amplifying the efficacy in detecting and documenting incidents from communities relevant to emergency management. Additionally, our model was capable of accurately representing pertinent community needs while concurrently adhering to a baseline standard for equity through mitigation of machine learning bias. This was evidenced through an illustrative case study using a machine-learning based framework which fused Twitter and Waze through transfer learning, NLP, and spatiotemporal analytics on the georeferenced data streams pertaining to emergency events to accurately detect the location and type (i.e., flooding, road closure, etc.) of an event. To the best of our knowledge, this study represents one of the initial endeavors to integrate the Wells–Du Bois protocol in order to achieve a baseline of bias mitigation.

The practical contributions of our study include aiding emergency management decision-making and situational awareness for disasters as well as improving allocation of resources to reduce the harmful effects of disasters. This paper adds to the growing body of knowledge on this topic addressing the shortcomings of Twitter and Waze applications for disaster detection and effective augmentation of platforms such as these. It establishes a foundation for (1) an integrative approach between social media and community-driven applications for crisis event detection toward further expansion of response capacity for real-time decision-making; and (2) including an equity appraisal through incorporating equity protocols into the research process. Recognizing and addressing potential disparities is essential for developing equitable strategies to ease the recovery process for individuals lacking necessary resources, thereby enhancing community resilience. Our approach holds promise in informing the design of future infrastructure systems that can better withstand the impacts of emergencies and efficiently allocate resources. Ultimately, this can positively influence the overall health and well-being of communities by improving systems that facilitate access to emergency healthcare and essential lifeline infrastructure, ensuring more effective and accessible support during critical situations.

Data Availability Statement

Some or all data, models, or code generated or used during the study are proprietary or confidential in nature and may only be provided with restrictions. The Waze and Twitter data are restricted due to the privacy of the API and feed access of the data streams.

Acknowledgments

This study is supported by the Georgia Department of Transportation under Grant No. 20–13. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Georgia Department of Transportation.

References

Adeola, F. O., and J. S. Picou. 2017. "Hurricane Katrina-linked environmental injustice: Race, class, and place differentials in attitudes." *Disasters* 41 (2): 228–257. https://doi.org/10.1111/disa.12204.

Albahri, A. S., et al. 2023. "A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion." *Inf. Fusion* 96 (Jun): 156–191. https://doi.org/10.1016/j.inffus.2023.03.008.

Alexander, D. E. 2014. “Social media in disaster risk reduction and crisis management.” *Sci. Eng. Ethics* 20 (3): 717–733. https://doi.org/10.1007/s11948-013-9502-z.

Amin-Naseri, M., P. Chakraborty, A. Sharma, S. B. Gilbert, and M. Hong. 2018. “Evaluating the reliability, coverage, and added value of crowdsourced traffic incident reports from Waze.” *Transp. Res. Rec.* 2672 (43): 34–43. https://doi.org/10.1177/0361198118790619.

Athuraliya, C. D., M. K. H. Gunasekara, S. Perera, and S. Suhothayan. 2015. “Real-time natural language processing for crowdsourced road traffic alerts.” In *Proc., 2015 15th Int. Conf. on Advances in ICT for Emerging Regions (ICTer)*, 58–62. New York: IEEE.

Awasthi, P., A. Beutel, M. Kleindessner, J. Morgenstern, and X. Wang. 2021. “Evaluating fairness of machine learning models under uncertain and incomplete information.” In *Proc., FAccT ’21: Proc., 2021 ACM Conf. on Fairness, Accountability, and Transparency*, 206–214. New York: IEEE.

Azhar, S., Z. Riaz, and D. Robinson. 2019. “Integration of social media in day-to-day operations of construction firms.” *J. Manage. Eng.* 35 (1): 06018003. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000660.

Bodenreider, C., L. Wright, O. Barr, K. Xu, and S. Wilson. 2019. “Assessment of social, economic, and geographic vulnerability pre- and post-Hurricane Harvey in Houston, Texas.” *Environ. Justice* 12 (4): 182–193. https://doi.org/10.1089/env.2019.0001.

Bozeman, J. F. 2024. “Bolstering integrity in environmental data science and machine learning requires understanding socioecological inequity.” *Front. Environ. Sci. Eng.* 18 (5): 1–9. https://doi.org/10.1007/s11783-024-1825-2.

Brunila, M., R. Zhao, A. Mircea, S. Lumley, and R. Sieber. 2021. “Bridging the gap between supervised classification and unsupervised topic modelling for social-media assisted crisis management.” Preprint, submitted March 22, 2021. http://arxiv.org/abs/2103.11835.

Chair, S., M. Charrad, and N. B. B. Saoud. 2019. “Towards a social media-based framework for disaster communication.” *Procedia Comput. Sci.* 164 (Dec): 271–278. https://doi.org/10.1016/j.procs.2019.12.183.

Chen, X., and T. Wang. 2017. “Combining active learning and semi-supervised learning by using selective label spreading.” In *Proc., 2017 IEEE Int. Conf. on Data Mining Workshops (ICDMW)*, 850–857. New York: IEEE.

Chen, Y., and W. Ji. 2021. “Enhancing situational assessment of critical infrastructure following disasters using social media.” *J. Manage. Eng.* 37 (6): 04021058. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000955.

Chen, Y., Q. Wang, and W. Ji. 2020. “Rapid assessment of disaster impacts on highways using social media.” *J. Manage. Eng.* 36 (5): 04020068. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000836.

Chen, Z., J. M. Zhang, M. Hort, F. Sarro, and M. Harman. 2022. “Fairness testing: A comprehensive survey and analysis of trends.” Preprint, submitted April 23, 2018. http://arxiv.org/abs/2207.10223.

Chowdhury, S., and J. Zhu. 2023. “Investigation of critical factors for future-proofed transportation infrastructure planning using topic modeling and association rule mining.” *J. Comput. Civ. Eng.* 37 (1): 04022044. https://doi.org/10.1061/(ASCE)CP.1943-5487.0001059.

CRED (Centre for Research on the Epidemiology of Disasters). 2009. “EM-DAT glossary.” EM-DAT The International Disaster Database Centre for Research on the Epidemiology of Disasters–CRED. Accessed March 9, 2023. https://www.emdat.be/Glossary.

Dhakal, S., L. Zhang, and X. Lv. 2021. “Understanding infrastructure resilience, social equity, and their interrelationships: Exploratory study using social media data in Hurricane Michael.” *Nat. Hazards Rev.* 22 (4): 04021045. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000512.

Dixon, S. J. 2023. “Distribution of X (formerly Twitter) users worldwide as of January 2023, by gender.” Statista. Accessed October 6, 2023. https://www.statista.com/statistics/828092/distribution-of-users-on-twitter-worldwide-gender/.

ETC (Eastern Transportation Coalition). 2017. *Going my WAZE to closing real time data gaps: Crowdsourcing summit summary brief*, 1–19. Philadelphia, PA: ETC.

Fan, C., Y. Jiang, and A. Mostafavi. 2020. “Social sensing in disaster city digital twin: Integrated textual–visual–geo framework for situational awareness during built environment disruptions.” *J. Manage. Eng.* 36 (3): 04020002. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000745.

FastText. 2023. “Fasttext.” PyPI. Accessed April 12, 2023. https://pypi.org/project/fasttext/.

FEMA. 2021. “FEMA issues public notice for 21 Georgia counties.” FEMA. Accessed March 9, 2023. https://www.fema.gov/press-release/20210211/fema-issues-public-notice-21-georgia-counties.

FEMA. 2023a. “Declared disasters.” FEMA. Accessed March 9, 2023. https://www.fema.gov/disaster/declarations.

FEMA. 2023b. “Glossary.” FEMA. Accessed March 9, 2023. https://training.fema.gov/programs/emischool/el361toolkit/glossary.htm.

FEMA. 2023c. “How a disaster gets declared.” FEMA. Accessed March 9, 2023. https://www.fema.gov/disasters/how-declared.

FEMA. 2023d. “2022–2026 FEMA strategic plan: Building the FEMA our nation needs and deserves.” FEMA. Accessed March 9, 2023. https://www.fema.gov/about/strategic-plan.

Gillborn, D., P. Warmington, and S. Demack. 2018. “Quantcrit: Education, policy, ‘big data’ and principles for a critical race theory of statistics.” *Race Ethnicity Educ.* 21 (2): 158–179. https://doi.org/10.1080/13613324.2017.1377417.

Griego, A. L., A. B. Flores, T. W. Collins, and S. E. Grineski. 2020. “Social vulnerability, disaster assistance, and recovery: A population-based study of Hurricane Harvey in greater Houston, Texas.” *Int. J. Disaster Risk Reduct.* 51 (Mar): 101766. https://doi.org/10.1016/j.ijdrr.2020.101766.

Gu, Y., Z. Qian, and F. Chen. 2016. “From Twitter to detector: Real-time traffic incident detection using social media data.” *Transp. Res. Part C Emerging Technol.* 67 (Jun): 321–342. https://doi.org/10.1016/j.trc.2016.02.011.

Halse, S. E., R. Grace, J. Kropczynski, and A. Tapia. 2019. “Simulating real-time twitter data from historical datasets.” In *Proc., 16th ISCRAM Conf.*, edited by Z. Franco, J. J. González, and J. H. Canós, 780–787. València, Spain: Information Systems for Crisis Response and Management.

Hamideh, S., and J. Rongerude. 2018. “Social vulnerability and participation in disaster recovery decisions: Public housing in galveston after Hurricane Ike.” *Nat. Hazards* 93 (Dec): 1629–1648. https://doi.org/10.1007/s11069-018-3371-3.

Henchey, M. J., R. Batta, A. Blatt, M. Flanigan, and K. Majka. 2014. “A simulation approach to study emergency response.” *J. Simul.* 8 (2): 115–128. https://doi.org/10.1057/jos.2013.20.

Hooshmand, A., and R. Sharma. 2019. “Energy predictive models with limited data using transfer learning.” In *Proc., e-Energy ’19: Proc., 10th ACM Int. Conf. on Future Energy Systems*, 1–5. New York: Association for Computing Machinery.

Hort, M., Z. Chen, J. M. Zhang, F. Sarro, and M. Harman. 2022. “Bias mitigation for machine learning classifiers: A comprehensive survey.” Preprint, submitted April 23, 2018. http://arxiv.org/abs/2207.07068.

Kim, J. H., and E. J. Sutley. 2021. “Implementation of social equity metrics in an engineering-based framework for distributing disaster resources.” *Int. J. Disaster Risk Reduct.* 64 (Dec): 102485. https://doi.org/10.1016/j.ijdrr.2021.102485.

Kumar, D., and S. V. Ukkusuri. 2020. “Enhancing demographic coverage of Hurricane evacuation behavior modeling using social media.” *J. Comput. Sci.* 45 (Dec): 101184. https://doi.org/10.1016/j.jocs.2020.101184.

Lieberknecht, K., D. Zoll, J. Jiao, and K. Castles. 2021. “Hurricane harvey: Equal opportunity storm or disparate disaster?” *Local Environ.* 26 (2): 216–238. https://doi.org/10.1080/13549839.2021.1886063.

Lindsay, B. R. 2011. *Social media and disasters: Current uses, future options, and policy considerations*. Washington, DC: Congressional Research Service.

Loos, S., D. Lallemant, J. Baker, J. McCaughey, S. H. Yun, N. Budhathoki, F. Khan, and R. Singh. 2020. “G-DIF: A geospatial data integration framework to rapidly estimate post-earthquake damage.” *Earthquake Spectra* 36 (4): 1695–1718. https://doi.org/10.1177/8755293020926190.

Loos, S., J. Levitt, K. Tomozawa, J. Baker, and D. Lallemant. 2022. “Efficacy of damage data integration: A comparative analysis of four major earthquakes." *Nat. Hazards Rev.* 23 (4): 04022026. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000581.

Lovari, A., and S. A. Bowen. 2019. "Social media in disaster communication: A case study of strategies, barriers, and ethical implications." *J. Public Aff.* 20 (1): 1–9. https://doi.org/10.1002/pa.1967.

Maurer, S. M. 2020. "Evolving approaches to place tagging in social media." *J. Urban Plan. Dev.* 146 (3): 04020023. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000583.

Mhatre, M., D. Phondekar, P. Kadam, A. Chawathe, and K. Ghag. 2017. "Dimensionality reduction for sentiment analysis using pre-processing techniques." In *Proc., IEEE Int. Conf. on Computing Methodologies and Communication (ICCMC)*, 16–21. New York: IEEE.

Mohammadi, N., and J. E. Taylor. 2017. "Urban energy flux: Spatiotemporal fluctuations of building energy consumption and human mobility-driven prediction." *Appl. Energy* 195 (Mar): 810–818. https://doi.org/10.1016/j.apenergy.2017.03.044.

Monroe-White, T., and J. Lecy. 2022. "The wells-du bois protocol for machine learning bias: Building critical quantitative foundations for third sector scholarship." *VOLUNTAS* 34 (1): 170–184. https://doi.org/10.1007/s11266-022-00479-2.

Nargesian, F., A. Asudeh, and H. V. Jagadish. 2022. "Responsible dataintegration: Next-generation challenges." In *Proc., SIGMOD '22: Proc., 2022 Int. Conf. on Management of Data*, 1–7. New York: Association for Computing Machinery.

National Advisory Council. 2020. *National advisory council report to the FEMA administrator FEMA*. Washington, DC: FEMA.

Nejat, A., L. Solitare, E. Pettitt, and H. Mohsenian-Rad. 2022. "Equitable community resilience: The case of winter storm Uri in Texas." *Int. J. Disaster Risk Reduct.* 77 (Mar): 103070. https://doi.org/10.1016/j.ijdrr.2022.103070.

Neyshabur, B., H. Sedghi, and C. Zhang. 2020. "What is being transferred in transfer learning?" In *Proc., 34th Conf. on Neural Information Processing Systems (NeurIPS 2020)*, 1–12. Vancouver, Canada: Neural Information Processing Systems.

NOAA US Department of Commerce. 2020a. "Hurricane zeta 2020." NOAA's National Weather Service. Accessed May 5, 2023. https://www.weather.gov/tae/zeta2020#:~:text=George%20Island%2031%20MPH%200139,28%2029.79N%2F84.88W%20.

NOAA US Department of Commerce. 2020b. "Tropical storm zeta affects north and central Georgia." Accessed May 5, 2023. https://www.weather.gov/ffc/TropicalStormZeta.

NOAA US Department of Commerce. 2022. "Hurricane Ian's path of destruction." National Environmental Satellite Data and Information Service. Accessed March 8, 2023. https://www.nesdis.noaa.gov/news/hurricane-ians-path-of-destruction.

Ogie, R. I., and N. Verstaevel. 2020. "Disaster informatics: An overview." *Prog. Disaster Sci.* 7 (Mar): 100111. https://doi.org/10.1016/j.pdisas.2020.100111.

Peirelinck, T., H. Kazmi, B. V. Mbuwir, C. Hermans, F. Spiessens, J. Suykens, and G. Deconinck. 2022. "Transfer learning in demand response: A review of algorithms for data-efficient modelling and control." *Energy AI* 7 (Jan): 100126. https://doi.org/10.1016/j.egyai.2021.100126.

Peng, Y., Y. Zhang, Y. Tang, and S. Li. 2011. "An incident information management framework based on data integration, data mining, and multi-criteria decision making." *Decis. Support Syst.* 51 (2): 316–327. https://doi.org/10.1016/j.dss.2010.11.025.

Pinto, G., Z. Wang, A. Roy, T. Hong, and A. Capozzoli. 2022. "Transfer learning for smart buildings: A critical review of algorithms, applications, and future perspectives." *Adv. Appl. Energy* 5 (Jun): 100084. https://doi.org/10.1016/j.adapen.2022.100084.

Pollard, J. A., T. Spencer, and S. Jude. 2018. "Big data approaches for coastal flood risk assessment and emergency response." *WIREs Clim. Change* 9 (5): 1–14. https://doi.org/10.1002/wcc.543.

Psyllidis, A., A. Bozzon, S. Bocconi, and C. Titos Bolivar. 2015. "A platform for urban analytics and semantic data integration in city planning." In *Proc., CAAD Futures 2015: Computer-Aided Architectural Design Futures. The Next City-New Technologies and the Future of the Built Environment*, 21–36. New York: Springer.

Purohit, H., R. Kanagasabai, and N. Deshpande. 2019. "Towards next generation knowledge graphs for disaster management." In *Proc., 2019 IEEE 13th Int. Conf. on Semantic Computing (ICSC)*, 474–477. New York: IEEE.

Qi, S., D. G. Brown, Q. Tian, L. Jiang, T. Zhao, and K. M. Bergen. 2009. "Inundation extent and flood frequency mapping using landsat imagery and digital elevation models." *GISci. Remote Sens.* 46 (1): 101–127. https://doi.org/10.2747/1548-1603.46.1.101.

Ribeiro, M., K. Grolinger, H. F. ElYamany, W. A. Higashino, and M. A. M. Capretz. 2018. "Transfer learning with seasonal and trend adjustment for cross-building energy forecasting." *Energy Build.* 165 (Mar): 352–363. https://doi.org/10.1016/j.enbuild.2018.01.034.

Romanowski, C., R. Raj, J. Schneider, S. Mishra, V. Shivshankar, S. Ayengar, and F. Cueva. 2015. "Regional response to large-scale emergency events: Building on historical data." *Int. J. Crit. Infrastruct. Protect.* 11 (Dec): 12–21. https://doi.org/10.1016/j.ijcip.2015.07.003.

Rossi, C., F. S. Acerbo, K. Ylinen, I. Juga, P. Nurmi, A. Bosca, F. Tarasconi, M. Cristoforetti, and A. Alikadic. 2018. "Early detection and information extraction for weather-induced floods using social media streams." *Int. J. Disaster Risk Reduct.* 30 (Apr): 145–157. https://doi.org/10.1016/j.ijdrr.2018.03.002.

Roy, K. C., S. Hasan, A. M. Mohaimin, and M. Cebrian. 2020. "Understanding the efficiency of social media based crisis communication during hurricane sandy." *Int. J. Inf. Manage.* 52 (Jun): 102060. https://doi.org/10.1016/j.ijinfomgt.2019.102060.

Salas, A., P. Georgakis, and Y. Petalas. 2017. "Incident detection using data from social media." In *Proc., IEEE Conf. on Intelligent Transportation Systems (ITSC)*, 751–755. New York: IEEE.

Salley, C., N. Mohammadi, and J. E. Taylor. 2022. "Semi-supervised machine learning framework for fusing georeferenced data from social media and community-driven applications." In *Proc., Computing in Civil Engineering 2021*, 114–122. Reston, VA: ASCE.

Salley, C., N. Mohammadi, and J. E. Taylor. 2024. "Protecting critical infrastructure: NLP-based automated information retrieval to generate hypothetical cyberattack scenarios." *J. Infrastruct. Syst.* 30 (3): 04024008. https://doi.org/10.1061/JITSE4.ISENG-2407.

Samuels, R., and J. E. Taylor. 2020. "Deepening the divide: Crises disproportionately silence vulnerable populations on social media." *J. Manage. Eng.* 36 (6): 04020083. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000848.

Schmidtlein, M. C., C. Finch, and S. L. Cutter. 2008. "Disaster declarations and major hazard occurrences in the United States." *Prof. Geogr.* 60 (1): 1–14. https://doi.org/10.1080/00330120701715143.

Sidauruk, A. 2018. "Congestion correlation and classification from Twitter and Waze Map using artificial neural network." In *Proc., 3rd Int. Conf. on Information Technology, Information Systems and Electrical Engineering*, 224–229. New York: IEEE.

Smith, C. 2023. "Waze statistics and user count (2023)." DMR. Accessed April 12, 2023. https://expandedramblings.com/index.php/waze-statistics-facts/.

Sujon, M., and F. Dai. 2021. "Social media mining for understanding traffic safety culture in Washington State using Twitter data." *J. Comput. Civ. Eng.* 35 (1): 04020059. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000943.

Sullivan, L., and E. Schwartz. 2018. "FEMA report acknowledges failures in Puerto Rico disaster response." NPR. Accessed April 16, 2023. https://www.npr.org/2018/07/13/628861808/fema-report-acknowledges-failures-in-puerto-rico-disaster-response.

Tang, L., Y. Zhang, F. Dai, Y. Yoon, Y. Song, and R. S. Sharma. 2017. "Social media data analytics for the U.S. construction industry: Preliminary study on Twitter." *J. Manage. Eng.* 33 (6): 04017038. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000554.

Tankovska, H. 2022. "Twitter: Number of users worldwide 2019–2024." Statista. Accessed April 12, 2023. https://www.statista.com/statistics/303681/twitter-users-worldwide/.

Tien, I., A. Musaev, D. Benas, A. Ghadi, S. Goodman, and C. Pu. 2016. "Detection of damage and failure events of critical public infrastructure using social sensor big data." In *Proc., Int. Conf. on Internet of Things and Big Data (IoTBD 2016)*, 435–440. Rome: IoTBD.

Twitter Developer Platform. 2023. "Advanced filtering with Geo Data. Developer Platform, Twitter." Accessed March 7, 2023. https://developer.twitter.com/en/docs/tutorials/advanced-filtering-for-geo-data#:~:text=You%20may%20find%20bounding%20boxes,sqft%20(25%20x%2025).
UNDRR (United Nations Office for Disaster Risk Reduction). 2016. "Terminology." UNDRR. Accessed March 9, 2023. https://www.undrr.org/terminology.
Verma, S., S. Vieweg, W. Corvey, L. Palen, J. Martin, M. Palmer, A. Schram, and K. Anderson. 2011. "Natural language processing to the rescue? Extracting 'situational awareness' tweets during mass emergency." In *Proc., Int. AAAI Conf. on Web and Social Media*, 385–392. Washington, DC: Association for the Advancement of Artificial Intelligence.
Wang, Q., and J. E. Taylor. 2016. "Process map for urban-human mobility and civil infrastructure data collection using geosocial networking platforms." *J. Comput. Civ. Eng.* 30 (2): 04015004. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000469.
Wang, Y., and J. E. Taylor. 2018. "Coupling sentiment and human mobility in natural disasters: A twitter-based study of the 2014 South Napa Earthquake." *Nat. Hazards* 92 (Mar): 907–925. https://doi.org/10.1007/s11069-018-3231-1.
Wang, Y., and J. E. Taylor. 2019. "DUET: Data-driven approach based on latent Dirichlet allocation topic modeling." *J. Comput. Civ. Eng.* 33 (3): 04019023. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000819.
Waze. 2017. *Waze traffic-data specification document*, 1–21. Mountain View, CA: Waze.
Wright, B. 2011. "Race, place, and the environment in the Aftermath of Katrina." *Anthropol. Work Rev.* 32 (1): 4–8. https://doi.org/10.1111/j.1548-1417.2011.01052.x.
Wu, D., and Y. Cui. 2018. "Disaster early warning and damage assessment analysis using social media data and geo-location information." *Decis. Support Syst.* 111 (Dec): 48–59. https://doi.org/10.1016/j.dss.2018.04.005.
Yang, Y., C. Zhang, C. Fan, A. Mostafavi, and X. Hu. 2020. "Towards fairness-aware disaster informatics: An interdisciplinary perspective." *IEEE Access* 8 (Dec): 1–15. https://doi.org/10.1109/ACCESS.2020.3035714.
Yin, J., A. Lampert, M. Cameron, B. Robinson, and R. Power. 2012. "Using social media to enhance emergency situation awareness." *IEEE Intell. Syst.* 27 (6): 52–59. https://doi.org/10.1109/MIS.2012.6.
Yu, J. J. Q. 2019. "Online traffic speed estimation for urban road networks with few data: A transfer learning approach." In *Proc., 2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 4024–4029. New York: IEEE.
Yuan, F., and R. Liu. 2018. "Integration of social media and unmanned aerial vehicles (UAVs) for rapid damage assessment in Hurricane Matthew." In *Proc., Construction Research Congress 2018*, 513–523. Reston, VA: ASCE.
Yuen, D. A., B. J. Kadlec, E. F. Bollig, W. Dzwinel, Z. A. Garbow, and C. R. Silva. 2005. "Clustering and visualization of earthquake data in a grid environment." *Visual Geosci.* 10 (1): 1–12. https://doi.org/10.1007/s10069-005-0023-z.
Zhou, D., O. Bousquet, T. Navin Lal, J. Weston, and B. Schölkopf. 2004. "Learning with local and global consistency." In Vol. 16 of *Proc., Advances in Neural Information Processing Systems*, 1–8. Cambridge, MA: MIT Press.
Zou, L., N. S. Lam, H. Cai, and Y. Qiang. 2018. "Mining Twitter data for improved understanding of disaster resilience." *Ann. Am. Assoc. Geogr.* 108 (5): 1422–1441. https://doi.org/10.1080/24694452.2017.1421897.
Zubiaga, A., A. Voss, R. Procter, M. Liakata, B. Wang, and A. Tsakalidis. 2017. "Towards real-time, country-level location classification of worldwide tweets." *IEEE Trans. Knowl. Data Eng.* 29 (9): 2053–2066. https://doi.org/10.1109/TKDE.2017.2698463.