 doi:10.20944/preprints202207.0221.v1

*Review*

# Harnessing Artificial Intelligence to Improve Food Assistance: A Scoping Review of Machine Learning Tools

**Nina M. Martin [1*], João Sedoc [2], Lisa Poirier [1], Andrew J. Rosenblum [3], Melissa M. Reznar [4], Daniel J. Barnett [1] and Joel Gittelsohn[1]**

[1] Johns Hopkins Bloomberg School of Public Health
[2] New York University
[3] American University Washington College of Law
[4] Oakland University
* Correspondence: nmarti38@jhu.edu

**Abstract:**

**Background:** Machine learning has revolutionized situational awareness during disaster management by classifying, clustering, and predicting impacted locations and people. Despite its importance, no review has been conducted on machine learning tools for food assistance efforts during emergency or non-emergency situations. The purpose of this scoping review is to address that gap.

**Methods:** Keywords were defined within the concepts of food assistance and machine learning. After the database searches, PRISMA guidelines were followed to perform a partnered, two-round scoping literature review. Text mining and Latent Dirichlet Allocation topic modeling algorithms were used to determine trends.

**Results:** 28 articles met criteria and were included in the analysis. The types of study designs included: model development (42.9%), non-study (i.e., text and opinion) (28.6%), qualitative research (14.3%), case study (10.7%), and meta-analysis (3.6%). There were no quantitative studies. The machine learning tools' main functions were improving SNAP programs (32.1%), detecting needs and resources (25%), predicting food insecurity (21.4%), and situational awareness of current food insecurity issues (21.4%). None of these studies took place during a disaster or explicitly addressed emergency mitigation, preparedness, or recovery. All of the studies were in early phases of development and implementation.

**Conclusion:** Machine learning tools for improving situational awareness, resource allocation, policymaking, and prediction have the potential to improve food assistance, but there is a lack of implementation and evaluation during all disaster phases. Also needed is more formative work on generating food-related queries and defining variables and features of food security.

**Keywords:** Artificial intelligence, machine learning, food security, food assistance

**Abbreviations:** GPS (Global Positioning System), LDA (Latent Dirichlet Allocation), PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) Checklist, SNAP (Supplemental Nutrition Assistance Program), U.S. (United States)

## 1. Introduction

Since 2015, the world has been experiencing an increase in food insecurity and undernourishment, and, uncoincidentally, the expanding impacts of climate change, natural

 doi:10.20944/preprints202207.0221.v1



disasters, and humanitarian crises.1,2 These threats impede the achievement of the United Nations' 2030 Sustainable Development Goal of "zero hunger."3 Reducing the impact of these threats to food security through early warning and response to food security crises is not only important for saving lives and reducing suffering, but early action is also more efficient since it can yield significant cost savings of humanitarian aid.2,4

Detecting and predicting food insecurity can be challenging because of the underlying network of variables from food chain aspects, such as food availability, access, utilization, and stability, to socio-economic factors such as poverty, race, and neighborhood characteristics. Food must be available for people to access it, and target populations must be aware of its availability and able obtain it.

From a disaster's impact through long-term recovery, one of the most basic management tasks is to support nutritional needs by classifying and predicting the food needs and regional availability. This is a key component of situational awareness, or the "ability to identify, process, and comprehend the critical information about an incident."5 During the COVID-19 pandemic, while other fields have matured their real-time situational awareness technologies, these efforts have been notably lacking for food distribution agencies.6,7 A recent scoping review of digital tools for emergency food assistance found that the food security sector is lagging behind others in terms of use and research on modern digital technologies to improve situational awareness during disasters.8

One approach that other fields are using for classifying and predicting needs is machine learning, a sub-field of artificial intelligence (Figure 1). In the simplest sense, machine learning algorithms are models of a system where the parameters are learned via the data. Machine learning algorithms have proven vital for a variety of tasks related to classification, predictive modeling, and analysis of data.9

Machine learning has been heavily implemented already by the agriculture sector to predict and detect food insecurity issues like pest control, weather changes, crop shortages and yields, food safety issues, and food fraud.10–24 Other sectors, such as the restaurant industry, have used it for social listening, which is the process of tracking, analyzing, and iteratively improving services, brands, or products in response to client conversations on social media.25,26 Public health researchers have also used machine learning for social listening to understand and predict trends on social media, recognize image patterns, and identify relationships between diet, personality traits and obesity.27–30 During the COVID-19 pandemic, it's been used to understand public discourse on COVID-19 in general31–33, wearing masks34, vaccine hesitancy35, racism and equity36, as well as document government action37, predict new cases38,39 and categorize symptoms.40 A critical gap in situational awareness and intervention remains in predicting and identifying disaster victims' needs and getting the proper supportive resources to the right place, at the right time.41

While these strategies have been applied to some extent for disasters such as earthquakes, floods and tsunamis to generally detect resource need and availability on social media,42–44 limited attention has been given to how these strategies can be harnessed to track food assistance efforts for individuals in need.

To better assess: the growing availability of machine learning as a tool for public health researchers and practitioners,9,45–48 advancements in capabilities for predicting and detecting food insecure individuals, and review the capabilities of social listening for available resources, we conducted a scoping literature review. This scoping review aims to assess the past decade's knowledge base of machine learning for food assistance that

 doi:10.20944/preprints202207.0221.v1



aims to improve situational awareness of food needs and resources and predict food insecure events and populations. The specific aims were to:

1. Define the scope and methods of current research regarding machine learning tools used for improving food assistance to vulnerable individuals and communities

2. Describe the available tools in terms of key functions, data collected, machine learning models, platform, and food security settings

3. Use text mining and topic modeling to define themes in the corpus of reviewed articles to define gaps and challenges in the field

## 2. Materials and Methods

### *2.1. Literature Search*

To define the current literature concerning the use of machine learning to improve food assistance for vulnerable populations, a scoping review was conducted of peer-reviewed and grey literature in August 2021. PubMed, Embase, PsychInfo, Scopus, and Web of Science databases were used to identify studies published from 2011 to 04 August 2021. Search terms were developed using a combination of controlled vocabulary and keywords to define the concepts of food assistance and machine learning, and terms were adapted for use in each database. The key terms for machine learning were adapted from methods validated by Wang et al.(49) To address the grey literature, similar terms were used to develop searches for Google, Science.gov, Worldwidescience.org, the World Food Programme, United States Agency for International Development, UNICEF, and Feeding America websites. A summary of the search strategy is presented in Table 1 and the full search strategy for the PubMed database is presented in Supplementary Table 1.Using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)[50] Checklist as a guide throughout, all citations were imported into Covidence systematic review software (Veritas Health Innovation, Melbourne, Australia) to facilitate removal of duplicates, screening, and full-text review.[51] Utilizing a blinded, dual review process, each article was first screened by the title/abstract, followed by a full-text review. Conflicts were resolved by a third reviewer. Inclusion and exclusion criteria are presented in Table 2. Machine learning descriptions were included from any country, provided that the article included machine learning that was, or could be, used in improving food security by connecting individuals or groups to food resources and was in the English language. A data extraction tool was developed in Covidence to extract the included the articles' title, first author last name, year of publication, data type, food insecurity setting country and city, food insecurity cause, study aim, study design, type and purpose of intervention/tool, digital platform type, intervention methods, intended users, intended beneficiaries.

### *2.2. Thematic analysis using topic modeling*

Topic modeling is an unsupervised machine learning technique (see Figure 1) that uses statistical modeling to scan a set of documents, detect word and phrase patterns within them, and automatically cluster word groups and similar expressions that best characterize a set of documents.(53) The type of unsupervised learning algorithm used in this study was Latent Dirichlet allocation (LDA). This is a commonly used statistical-based model for exploratory data analysis of previously unseen, unstructured textual data.(32,53–55) LDA topic modeling may be preferred to traditional qualitative approaches because it summarizes textual data at a scale beyond the capacity of human annotation.(54) By deriving a probabilistic clustering, or latent topic distributions, LDA topic models can identify patterns within the texts; it enables efficient classification of large bodies of data based on patterns and features.(32) The output of the LDA model, therefore, is a set of 30 key terms and their likelihood that the given word will be used in conjunction withing a given topic. This technique can be used for literature reviews as an additional, objective method to understand overarching themes per study and per the entire corpus (aka set of documents to be analyzed).

LDA was run via the Python(56) coding language and in the online coding notebook Google Colaboratory(57). The full Google Colaboratory notebooks are available on GitHub.(58–60) The protocol used by Prabhakaran(54) was employed to run and visualize the results of the LDA modeling. Briefly, all pdf articles included in this review were converted to text files, exclusive of meta data, sections headings, and references. This text was further preprocessed to remove punctuation, "na", and convert all letters to lowercase. Stop words, i.e. the/in/or, were removed. Finally, the generated string was converted to a list. Word, bigram, and trigram frequency per document and the entire corpus was calculated and visualized. In order to select the optimal number of topics to analyze from the corpus, the coherence score, alpha, and beta model hyperparameters were calculated on 75% of the corpus as described by Kapadia.[61] An additional model evaluation method was used via the pyLDAvis package to produce a dynamic intertopic distance map (via multidimensional scaling) including a list of the top 30 most salient terms overall and 30 most relevant terms per topic.[62] The topic modeling algorithm, Latent Dirichlet Allocation (LDA), was run via the Python[53] coding language and in the online coding notebook Google Colaboratory[54]. The full Google Colaboratory notebooks are available on GitHub.[55–57] The protocol used by Prabhakaran[58] was employed to run and visualize the results of the LDA modeling. Briefly, all pdf articles included in this review were converted to text files, exclusive of meta data, sections headings, and references. This text was further preprocessed to remove punctuation, "na", and convert all letters to lowercase. Stop words, i.e., the/in/or, were removed. Finally, the generated string was converted to a list. Word, bigram, and trigram frequency per document and the entire corpus was calculated and visualized. In order to select the optimal number of topics to analyze from the corpus, the coherence score, alpha, and beta model hyperparameters were calculated on 75% of the corpus as described by Kapadia.[59] An additional model evaluation method was used via the pyLDAvis package to produce a dynamic intertopic distance map (via multidimensional scaling) including a list of the top 30 most salient terms overall and 30 most relevant terms per topic.[60]

## 3. Results

Of the 3,984 unique articles identified in the final search (which included 72 articles from the grey literature and 3,912 peer-reviewed publications), 3,224 were excluded at the title/abstract screening stage. Of the 131 studies included in the full text review, 39 (66.1%) peer-reviewed studies were excluded for the following reasons: 21 were not food security related, 6 did not include machine learning, 5 were related to agricultural settings, 5 pertained to food safety issues, 1 was outside of the date range, and 1 pertained to food quality. The PRISMA diagram is presented in Figure 2 and the final list of reviewed articles is presented in Table 3.

### *3.1 Descriptive Research Statistics*

The final synthesis included 20 peer reviewed and 8 gray literature studies (Table 3). The types of data presented in the studies were 57.1% qualitative, 28.6% quantitative and qualitative, and 14.3% models/codes. Types of study designs included: model development (42.9%), text and opinion (28.6%), qualitative research (14.3%), case study

(10.7%), and meta-analysis (3.6%). There were no trials or wide scale implementation of the tools.

During the 2011-2021 search period, over half (82%) of the 28 articles were published in the last five years (Figure 3). The year with the highest volume of publications was 2020 (25%), followed by 2018 (21.4%). There were zero studies published in 2011, 2012, 2013, or 2017.

Geographically, the largest portion of the included studies took place in the United States (50%), followed by the United Kingdom (14%), Brazil (4%), Ethiopia (4%), Malawi (4%), India (4%), Pakistan (4%), South Africa (4%), and Uganda (4%). Five studies aimed for a global impact and did not specify one location. Based on these locations, the majority of articles focused on high income countries (61%), followed by low income (11%), upper middle income (7%), and lower middle income (7%) countries.

Articles were also classified by food insecurity cause as well as intended users and beneficiaries. The distribution of food insecurity causes was largely undefined across the articles; half of the articles did not define a cause. Causes mentioned included famine or lack of food (13.2%), poverty (13.2%), economic or financial crisis (7.9%), hurricanes (5.3%), terrorism or political unrest (5.3%), and COVID-19 (2.6%). None of these studies took place during any phase of the disaster lifecycle or during an active disaster. Two studies looked at historical disaster data. Among the wide range of intended users and beneficiaries, the most frequently mentioned are: food banks and pantries, humanitarian relief agencies, policymakers, food donors, food recipients, researchers, SNAP program managers and users.

*3.2 Thematic analyses via topic modeling*

The results of the model evaluation indicated that 9 topics had the highest coherence score (Supplementary Figure 1) and thus were selected for further analysis. Using topics equal to 9, the model had the highest coherence value with the hyperparameters alpha equal to 'asymmetric' and beta equal to .31. To confirm these results, an intertopic distance map was constructed using pyLDAvis algorithm[60] and visualized in Supplementary Figure 2. This map shows an optimal spread of topics, with little to no overlapping circles.

Using the data presented in Supplementary Table 2 along with qualitative analysis of representative text, topics were labeled and are presented in Table 4. The nine topics, with top three associated terms, are: detecting SNAP application fraud (churn, client, benefit), characterizing new and existing SNAP users (snap, household, new), predicting donations to food banks (amount, supermarket, receive), World Food Program initiatives (security, blockchain, level), sentiment [classification] analysis of tweets (tweet, classifier, sentiment), changes in food supply and demand during the pandemic and other disasters (demand, change, country), forecasting food rescue needs and resources (system, rescue, value), predicting household food insecurity (security, household, level), and hurricane preparedness and response on Twitter (mention, hurricane, twitter).

Not all topics in the corpus are equally important. To understand which topics are the most important, topic dominance and weightage were calculated per document and per the corpus were calculated (Figure 4). Seven out of the nine topics were dominant, meaning that they had high contribution rates within many of the documents. These dominant topics (and the top three frequent words) were detecting SNAP application fraud (churn, client, benefit), characterizing new and existing SNAP users (snap, household, new), sentiment analysis of tweets (tweet, classifier, sentiment), 6 (demand, change, country), predicting food rescue needs and resources (system, rescue, value), predicting food insecurity (security, household, level), and hurricane preparedness and response on Twitter (mention, hurricane, twitter). Two topics were considered less important in the corpus: predicting food bank donations (amount, supermarket, receive) and World Food Programme Initiatives (security, blockchain, level). The number of documents per dominant topic and topic weightage are shown in Figure 4. For example, topics 3 and 4 had both zero to few documents by dominant topic and by topic weightage. Conversely, Topic 2 had both a high number of documents by dominant topic and topic weightage. This means that topic 2 is dominant to topic 3, for example.

Finally, the frequency and density of keywords in the topics were calculated, shown in Supplementary Figure 3. These graphs show word count (left axes) versus word density (right axes) for three representative dominant topics (topics 1, 2, and 7) and one representative non-dominant topic (topic 3). Words that had lower word density than word count across multiple topics were removed from the analysis. This included, for example, food, data, and model.

### *3.3. Word frequency patterns in entire corpus*

After removing stop words and low-density words such as food, data, and model, a list of the most to least frequent words, word pairs (bigrams) and trigrams was computed. The top 30 of these words and phrases are shown in Figure 5. The top five words identified were tweet, household, information, different, and level. The top words related to machine learning applications, specified with examples in Figure 5, were classifying, mining text, mapping and clustering, predicting, assisting, and image detection. The top 8 referenced algorithms, defined in Figure 5, were artificial neural networks, support vector machine, sentiment analysis, regression, decision tree, topic modeling, random forest, and k nearest neighbors. The top 10 word frequencies related to disasters and emergencies were food insecurity, hurricane, disaster, pandemic, hunger, drought, natural disaster, earthquake, conflict, and climate change (Supplementary Figure 4).

### *3.4 How Machine Learning is Currently Being Used to Address Food Insecurity*

Using the results from the descriptive statistics and topic modeling, Table 5 summarizes trends from the entire corpus, including categories of machine learning tool types, goals, challenges, frequently used algorithms, and intended users and beneficiaries. The four tool types described in the corpus included: improving SNAP programs (32.1%), detecting needs and available resources (25%), predicting food insecurity (21.4%), and situational awareness & characterization of current food insecurity issues (21.4%).

#### *i. Improving SNAP programs*

The most frequent goal presenting within this corpus was characterizing and improving SNAP programs and its users. This included tools for application pre-screening, understanding churn, predicting fraud, characterizing user needs, and classifying types of users. The motivators for developing these tools were addressing hunger and famine, poverty, and financial crises. Most of these studies (n = 6) did not specify a specific motivator other than food insecurity. The intended users and/or beneficiaries of these studies were food banks, policymakers, and SNAP program manager and users. The most frequently used algorithms for characterizing SNAP programs were natural language processing algorithms for sentiment analysis, topic modeling, and artificial neural networks. The top three reported challenges for this category were (1) poor accuracy of algorithms;

(2) fear of government using machine learning to discriminate against SNAP users; and (3) lack of complete and/or accurate data.

*ii. Detecting needs and available resources*

This tool category pertains to improving resource allocation between food donors and recipients inside and outside of disaster settings. This includes allocation between food banks and food pantries, as well as between supermarkets and food rescue organizations and food banks. During disasters, these tools can also be used to search for disaster-impacted individuals and connect them with relief organizations. The causes of food insecurity to be addressed by these related studies included COVID-19 (n = 1), hurricanes (n = 3), or not specified (n = 4). The intended beneficiaries and/or users of these tools included food banks, food donors, grocery stores, policymakers, producers, humanitarian relief workers, food recipients, and volunteers. The top three challenges to implementing these tools were (1) some countries do not frequently use social media; (2) Internet may be unavailable during disasters; (3) requirement of adequate data collection from multiple sources to return accurate results. Although some of these tools were intended to address response during a disaster, none of these took place in real-time, but were more proof-of-concept studies using historical, for example, hurricane data.

*iii. Predicting food insecurity*

Predicting food insecurity included both nowcasting and forecasting at the individual, household, regional, and global levels. The motivation for these studies was to address poverty, threats from climate change, or lack of food. Two studies did not specify a motivator beyond food insecurity. The intended users and/or beneficiaries of these tools were humanitarian relief workers and agencies; policymakers; researchers; food donors; and food recipients. Most frequently used algorithms were Random Forest,[89] Regression,[90] and XgBoost.[91] The top three reported challenges with predicting food insecurity were: (1) data issues (i.e. lack of data; countries collecting different types of metrics); (2) algorithm and feature selection: complex variables and confounders; and (3) lack of political will to predict food security issues.

*iv. Situational awareness and characterization of current food insecurity issues*

These situational awareness tools were intended for classifying and nowcasting current food security needs for both disaster and non-disaster settings. A major goal of these tools was for decision support for policymakers and relief organizations. This included characterizing and nowcasting food security individuals and their needs at the household level, improving policymaking for current users of food insecurity programs, and understanding opinions of policymakers and media outlets on current food security issues. Most studies in this category (n = 6) did not list a cause for food insecurity; one study mentioned poverty as an important cause. The intended users and beneficiaries of these tools included policymakers, researchers, stakeholders, food recipients, or OLIO (a food rescue application) users. There were many proposed algorithms for this type of tool, including Support Vector Machine, Bayesian belief network, K nearest neighbors, and sentiment analysis. The top challenges with these models were (1) data issues: country-level monitoring is poor & varied; (2) model accuracy is dependent on expert understanding & selection of variables; (3) lack of use in disaster settings.

## 4. Discussion

The application of big data and machine learning techniques to the food security field is a relatively new and growing area of research; eighty-two percent of the articles

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 July 2022 doi:10.20944/preprints202207.0221.v1



in this study are from the last five years. This scoping literature review, the first of its kind, sought to explore machine learning tools for food assistance. This review highlights the types of machine learning methods being implemented and evaluated, while identifying gaps in related publications, use of algorithms, and breadth of coverage of algorithms, topics, disaster types, phases, and locations. Our research revealed several important findings. We will first discuss the major technical findings of this review, such as preferred algorithms, typical set up steps, data considerations and technical challenges. Second, we will discuss the major applications and gaps of machine learning in food assistance strategies.

i. Major technical findings and challenges

The studies included in this review had similar approaches to setting up machine learning analyses. The typical research and development steps performed include research question development, literature review, selection of best or most accurate algorithm, dataset collection and/or selection, pre-processing of data, splitting of data into training and testing sets (for supervised learning algorithms, Figure 1), feature engineering and/or selection, parameter selection, running of training set, running of test set, statistical validation of model, and experimentation with new datasets.

These studies always selected the algorithm based on a literature review and applying an algorithm from another field to food assistance. No study indicated that new algorithms were needed, created or tested. Some studies chose their algorithm from among the most popular or based on past evidence[77] and experience, while other studies chose their algorithm by testing of multiple algorithms within their research[73,78] to determine the best fit and/or accurate model.

A major difficulty in machine learning success is ensuring that the correct variables, features, parameters, and hyperparameters are selected and at the right, if applicable, proportions. This is a challenge in the food security sector as the definition of food insecurity is complex and the variables represent complex socioeconomic variables interacting with each other.

The articles included in this review have tried to overcome this challenge in six ways. First, multiple articles employ a separate, initial algorithm for selecting features and/or modeling many features to select best fit. For example, Westerveld et al. [84] created 50 additional variables using this approach, which led to improved model accuracy. Second, some study teams recruited expert committees to define causal relationships between variables and select real-life variables based on experience before comparing these generated features to data/statistical model generated variables.[77] The mixture of expert committees plus statistical model generated variables has improved model accuracy. Third, multiple studies performed generalization testing of many algorithms at once to verify which model led to the most accurate predictions.[81,82,85,88] Fourth, a portion of studies employed systematic model evaluation methods like f-measures, sensitivity/specificity tests, ROC (receiver operating characteristic curve or a graph showing the performance of a classification model at all classification thresholds) or AUC (Area under the ROC Curve or an aggregate measure of performance across all possible classification thresholds) analysis and using training and test datasets. More machine learning studies in this field need to employ quantitative systematic model evaluation methods. Fifth, some studies initial their feature selection through an unsupervised learning algorithm.[78,81] The results of this model are used as the features in a second, supervised learning algorithms (defined in Figure 5). The combination of the two types of algorithms have been particularly useful in improving classification of needs and resource social media posts. Last, cross-validation, a commonly used statistical method to estimate the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 July 2022 doi:10.20944/preprints202207.0221.v1



performance (or accuracy) of machine learning models, was used in one study to protect against overfitting of a food security prediction model.[67]

Related to the issue of defining parameters is that the model will only be as good as the data collected and/or available. Data was obtained through: household surveys and interviews, national historical records, Google trends, and Twitter scraping. Two studies reported the lack of national data as a major setback in predicting food insecurity at a sub-national level.[73,78] It is unreasonable for single studies to collect the national-level data that is needed for some models to run accurately. They overcame this challenge through historical records from years where food security data was collected at a national level and use of available apps with large in-country user bases (e.g., OLIO and conducting cross-sectional household interviews to approximate the larger population).

An important aspect and challenge for machine learned-based studies are also the proper pre-processing of data before running it through the algorithm. Pre-processing refers to steps taken to transform raw data into appropriately processed data usable by the algorithms. This may include converting all letters to lowercase, or removing punctuation, stop words, or metadata. Different algorithms require differently processed data; and thus, understanding and taking the appropriate steps is critical to effective analysis.

ii. Applications and gaps

Considering 50% of the studies are set in the United States, it is understandable that the most frequently cited purpose for using machine learning is to improve and/or characterize SNAP as well as characterize its users (32.1%).

A set of these studies focused on understanding churn and fraud. Characterizing and predicting churn and fraud was observed in Topic 1, which was also a dominant topic in the corpus. Churn is when customers join a program, quit, then join again in a short period. While there are some legitimate reasons for customers to do so, it is also a generally recognized indicator of possible fraud.[75] The United States government became more fixated on detecting fraud during the Trump presidency, and this matches with the highest volume of publications on this topic (2016-2020).[92]

Another theme regarding SNAP, topic 2, concerns characterizing users past and present and understanding their changing needs. The pool of SNAP users has changed dramatically since the Great Recession.[71] Hamad et al. are considering how new SNAP users may have different needs and backgrounds than those prior to 2008.[71] This information is critical to informing policymakers and SNAP program officers on how best to meet these changing needs. This topic was not considered less important within the corpus according to topic modeling than the detection of fraud.

The next frequently mentioned purpose of the identified tools was using machine learning for the classification of food needs and resources (25%). Using text mining and machine learning to classify needs and resource tweets is not a new feature set of concepts. Specifically, prior studies identified social media posts containing information on the need and availability of resources in disaster situations. These studies used different methods of detecting needs and available resources, for example, natural language processing[93] and neural networks.[42,94] These studies, while not particularly focused on food, concluded that traditional Information Retrieval and Natural Language Processing approaches often do not perform well over short, informally written microblogs. A preferred approach is to use neural network-based information retrieval models to identify need-tweets and availability-tweets. Current work in this field is testing neural network algorithms to additionally match the need-tweets with appropriate availability-tweets. Methods specifically regarding food availability and needs remain sparse. The techniques described by this corpus to detect food resources and needs were based on studies from other fields and/or a general detection of needs during disasters (not necessarily food-related needs and resources).

The largest challenge to success is generating an appropriate query with terms that retrieve all related content and its location in relation to a disaster setting. Purohit et al. (2013) provided a set of 18 regular expressions to identify specific types of need-tweets indicating donation requests for different resources, and tweets indicating the availability of resources to be donated.[95] There have also been attempts to develop tools that tap the capabilities of social media to identify need-requests and resource providers.[95,96] Basu et al.[94] and Khosla et al. [87] organized their searches using an Information Retrieval (search) task method, where appropriate query-terms are selected to form queries corresponding to needs and availabilities. However, the studies included in this current scoping review have only tried limited search terms, such as "food poverty Uganda."[77] Not only does this return a wide range and non-resource or need related tweets, but it is also poor at specifying location. Retrieving location is integral to linking resources with needs. This study by Lukyamuzi et al. did clearly state this limitation and justified their query due to the atypically low numbers of Twitter users in Uganda.[77] Instead, they considered perspectives from anyone mentioning Uganda. This would be a suitable technique for opinion mining on food security issues in Uganda. However, this a poor technique for real-time detection of food needs and resources.

Another gap in this corpus is poor to no evaluation of the algorithms' performance, particularly for studies not solely focused on nutrition issues. For example, Basu et al.[94] uses only F-scoring to evaluate the accuracy of the models used. In contrast to other fields, it is the gold standard to compare accuracy, precision, and F-scores.[94]

In summary, tools for matching food resources to needs are lacking in modern methodologies used outside of the food focused literature. Rigorous search terms, location data, and evaluation methods are critical to being able to understand the data landscape and progressing the field forward. Additionally, all of the studies are in the early development phases. Although the studies use algorithms that are well established, many do so as pilot studies or other proof of concept style research and not widely implemented or evaluated.

Next, 21.4% of the studies were concerned with predicting food insecurity of households and communities. There were multiple reasons for this focus. First, the authors stated that certain countries, i.e., the United Kingdom and Malawi, do not collect information specifically on food security. Policymakers currently make decisions based on local reports which vary in quality, types of data collected, and data frequency. It is difficult for policymakers to understand the scope of food insecurity in their jurisdiction and classically policymakers have tended to underestimate food insecurity when not provided with adequate data. Most of the studies aimed to influence or improve policymaking via use of machine learning technologies for predicting food security. Specific aims included evidence-based resource allocation, intervention design and implementation, call for better or consistent food security data collection, and showcasing the dearth of food insecurity information in specific countries, like the United Kingdom. However, a major criticism of this approach is that for policy makers, causal models are most likely more important, whereas the models used in this corpus are mostly correlative. In addition, the accuracy of the predictive models in this corpus was generally low, making it hard for decision makers to be confident in the predictions. Also, there is an inherent lack of explainability, from model input to output, in neural network models. This trait may pose future challenges when these models are further developed—and this is a place where food specific methods may be important. Therefore, prediction of food security might not be the right metric for policymakers, as robustness of datasets where there's domain or distribution shift is important.

Other reasons for the focus on predicting food insecurity includes countries prefer to compare economic status through users of food insecurity programs over more traditional indicators like GDP.[68] Some stakeholders worry that these traditional measures are corrupt, inaccurate, or are produced too slowly to be representative of current economic status.

Because of the complex interactions among dependent and independent variables of food insecurity, linear regression models (one to one or many to one variables) are not accurate in this context. Instead, artificial neural networks and other types of 'many to many' (i.e., many variables are related to many other variables) algorithms show promise.[67]

A gap in the literature, notably grey literature, was State- and local-government use of machine learning for detecting and predicting food needs and available resources. Scraping, analyzing, and predicting social media for tweets has become an accessible form of machine learning and data science. It is cheap and relatively easy to set up within the last two years with the prevalence of packages.[97–102] A search of state and local governments did not return any results. It is possible these methods are used but not advertised publicly. Future studies should include interviews and other qualitative research with relevant government officials, e.g., emergency operations managers, to gauge the use of machine learning, if any, in these settings.

While food assistance tools were not found through this search, State governments used machine learning for other public health issues. For example, New York City used machine learning algorithms to predict homelessness for more efficient public health interventions and resource allocation.[103] Maryland State's Task Force on Vulnerable Populations uses machine learning to calculate COVID-19 risk analytics to prioritize deployment of manage-in-place field teams and coordinate emergency interventions with county health officers, area health plans, and local non-profit organizations.[104] Massachusetts State reported on the funding of projects to improve data and situational awareness of the opioid epidemic.[105]

### *4.1 Limitations*

This study is subject to several limitations. While we used search terms for machine learning from highly cited sources, there remains the possibility that relevant specific model were not included in the search. Despite this, we have taken care to include the most common machine learning models. Additionally, the complexity of machine learning creates many permutations of article types and publication locations. While we built our search iteratively to attempt to capture the broadest possible subset of the literature, it is possible that our search strategy missed critical studies. Furthermore, while our inclusion criteria was specifically tailored to English language literature and several key governmental and non-governmental search engines and websites, our approach may have systematically excluded otherwise worthy studies in non-English languages or that are not formally published in the major databases used in our research. Additionally, to ensure a specific focus on food security we did not consider machine learning tools that were (1) Agriculture-based machine learning for detecting/predicting pests, weather changes, crop shortages or yields; (2) articles on food supply chain – previously reviewed (3) food recognition systems, food safety detection systems, food image detection (4) customer satisfaction with restaurant or hotel food (non-food insecurity setting), (5) diet or obesity related without food insecurity component, (6) food fraud (7) forecasting food prices, (8) GIS/food mapping that didn't involve machine learning. Lastly, our study addressed the scope of digital food emergency tools and not the quality of the studies themselves.

Unsupervised learning algorithms, such as the LDA algorithm used for topic modeling, are known to have limitations, such as lower accuracy of assigning tweets to the correct topic.[55,61,97] For exploratory data analysis, observations, and preliminary detection of topics in an unfamiliar corpus, it is acceptable to use LDA modeling despite these known limitations.[61,62] In addition, the study team used both AI-based and traditional methods for data extraction analysis to confirm results. The topic modeling results were in line with the traditional qualitative analysis results as shown in Table 4. Despite the limitations, this study highlights several key gaps in emergency preparedness interventions and response while defining the scope of the existing literature with a food assistance lens.

**5. Conclusions**

Based on the findings in this review, there are many opportunities for leveraging innovative machine learning technologies to improve food security globally. These advances remain ever pressing as global food security threats continue to emerge from a multitude of sources, including climate change, pandemics, and political unrest.[2] The studies presented in this review uniformly were in early development and research phases and are mostly established in developed countries. This highlights the timely need for current work focused on developing machine learning technologies to improve a range of tools, including SNAP in the United States, real time detection of food needs and resources, and predicting food insecurity at the individual, household (most common), and global scale.

Avenues for additional research, policy, and practice include re-focusing the research agenda on low- and middle- income countries and under-resourced communities, instead of the current focus on mostly developed countries. In the United States, research is heavily concentrated on machine learning strategies to improve SNAP in terms of resource allocation, fraud detection, and characterizing new and existing users. The focus on fraud detection is not an evidence based best use of this novel approach to improving SNAP, as only 1.5% or less of applicants are deemed fraudulent and there is a high chance for discrimination.[65,75,92] It would be highly beneficial to expand the research agenda in the United States to real-time detection of needs and resources, as is being studied in the United Kingdom and elsewhere. Local and state policymakers can highly benefit from this information to ensure that resources and needs are efficiently allocated during disasters. The machine learning models in this review, although in early research and development phases, show promise in being able to accurately detect real time needs and resources in a target population.

There are two major barriers for the research regarding predicting food security. First, there is a lack of food security data available globally. Even developed countries, such as the United Kingdom, do not consistently collect this data. It would be highly beneficial for countries or international aid organizations to develop and routinely measure a single metric for food security. As Nica-Avram et al. points out, "unless the organizations chiefly responsible for emergency food assistance (e.g., foodbanks, community cafes, food sharing applications) compile shared aggregate figures the prevalence of food insecurity (and its relation to food surplus) will remain obscure."[77] Second, machine learning algorithms are generally better at making observations or correlations than causal relationships. Therefore, the prediction models presented in this review have lower-than-ideal accuracy. The intended end-users of these models, typically policymakers and aid organizations, may not find the results of these models helpful or may not feel confident enough to use the predictions for decision-making and resource allocation.

Considering these data and prediction limitations, research should be conducted in partnership with policymakers, aid organizations, and other stakeholders. This will ensure that the developed machine learning tools and associated dashboards meet the needs—and is used by—the intended end-users. As mentioned before, more useful and reliable are the machine learning tools for real time detection of food needs and resources—and determining how to deliver the resources to those in need quickly.

As reflected in this review, there are many policy and practice implications. The use of machine learning to classify new and existing food assistance program users can help policymakers more efficiently meet the needs of food insecure individuals and families while appropriately allocating government budgets. During disasters, real time situational awareness will allow policymakers to find and allocate resources efficiently to those in need. Predicting households and individuals that may experience food insecurity in the short- and long-term can help policymakers and aid organizations better prepare to meet future needs.

This review supports the expansion of two major practice initiatives. First, complementary dashboards to Fewsnet and the World Food Programme's Hunger Map Live should be built to show real time food needs and resources. As found by this review, machine learning algorithms for real time classification of needs and resources is more accurate and, most likely, helpful for policymakers and aid organizations to allocate resources. Second, the use of machine learning tools can have ethical implications that need to be considered. For example, based on the discussion in this review, it is important for wider engagement on the misconceived perception of fraud in welfare programs like SNAP and whether machine learning tools are appropriate for this context.

**Acknowledgements:** We would like to thank the Bloomberg American Health Initiative for funding support.

**Author contributions:** Thank you to the following co-authors for their contributions to this manuscript. Conceptualization, N.M.M, J.G., D.J.B., J.S.; methodology, N.M., J.S.; formal analysis, N.M., L.P.; data curation, N.M.; writing—original draft preparation, N.M., L.P.; writing—review and editing, N.M.M, D.J.B., J.S., J.G., L.P., A.R.,M.R.; visualization, N.M.M.; supervision, D.J.B, J.S., J.G., N.M.; funding acquisition, N.M. All authors have read and agreed to the published version of the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

1. Kovács G, Spens KM. Humanitarian logistics in disaster relief operations. Jahre M, Persson G, eds. Int J Phys Distrib Logist Manag. 2007;37(2):99-114. doi:10.1108/09600030710734820
2. World disasters report 2020. International Federation of Red Cross and Red Crescent Societies. Accessed January 17, 2021. https://media.ifrc.org/ifrc/world-disaster-report-2020/
3. Hassani H, Huang X, Macfeely S, Entezarian MR. Big data and the united nations sustainable development goals (UN SDGs) at a glance. Big Data Cogn Comput. 2021;5(3). doi:10.3390/bdcc5030028
4. Global Early Warning – Early Action Report on Food Security and Agriculture. Food and Agriculture Organization of the United Nations; 2017:32.
5. PrepTalks | FEMA.gov. Accessed October 27, 2021. https://www.fema.gov/emergency-managers/practitioners/preptalks
6. McLoughlin GM, Fleischhacker S, Hecht AA, et al. Feeding Students During COVID-19—Related School Closures: A Nationwide Assessment of Initial Responses. J Nutr Educ Behav. 2020;52(12):1120-1130. doi:10.1016/j.jneb.2020.09.018
7. Knorr D, Khoo CSH. COVID-19 and Food: Challenges and Research Needs. Front Nutr. 2020;7. doi:10.3389/fnut.2020.598913

8. Martin NM, Rex S, Barnett D, et al. Digital tools for emergency food assistance: a scoping review. Disaster Med Public Health Prep. Published online September 2021. doi: 10.1017/dmp.2021.281
9. Mooney SJ, Pejaver V. Big Data in Public Health: Terminology, Machine Learning, and Privacy. Annu Rev Public Health. 2018;39(1):95-112. doi:10.1146/annurev-publhealth-040617-014208
10. Nosratabadi S, Ardabili S, Lakner Z, Mako C, Mosavi A. Prediction of food production using machine learning algorithms of multilayer perceptron and anfis. Agric Switz. 2021;11(5). doi:10.3390/agriculture11050408
11. Nosratabadi S, Szell K, Beszedes B, Imre F, Ardabili S, Mosavi A. Comparative Analysis of ANN-ICA and ANN-GWO for Crop Yield Prediction. In: Proc. - RIVF Int. Conf. Comput. Commun. Technol., RIVF. Institute of Electrical and Electronics Engineers Inc.; 2020. doi:10.1109/RIVF48685.2020.9140786
12. Aguilar E, Bolanos M, Radeva P. Exploring Food Detection Using CNNs. In: COMPUTER AIDED SYSTEMS THEORY - EUROCAST 2017, PT II. Vol 10672. ; 2018:339-347. doi:10.1007/978-3-319-74727-9_40
13. Jean N, Burke M, Xie M, Davis WM, Lobell DB, Ermon S. Combining satellite imagery and machine learning to predict poverty. Science. 2016;353(6301):790-794. doi:10.1126/science.aaf7894
14. Kollia I, Stevenson J, Kollias S. Ai-enabled efficient and safe food supply chain. Electron Switz. 2021;10(11). doi:10.3390/electronics10111223
15. Deng X, Cao S, Horn A. Emerging Applications of Machine Learning in Food Safety. In: Doyle M, McClements D, eds. ANNUAL REVIEW OF FOOD SCIENCE AND TECHNOLOGY, VOL 12, 2021. Vol 12. ; 2021:513-538. doi:10.1146/annurev-food-071720-024112
16. Glen N, Korol O, André L, Robert F, Tom G. Extracting pest risk information from risk assessment documents. In: Bonn M., Wu D., Downie S.J., Martaus A., eds. Proc. ACM IEEE Joint Conf. Digit. Libr. Vol 2019-June. Institute of Electrical and Electronics Engineers Inc.; 2019:368-369. doi:10.1109/JCDL.2019.00074
17. Loti N, Noor M, Chang S. Integrated analysis of machine learning and deep learning in chili pest and disease identification. J Sci FOOD Agric. 2021;101(9):3582-3594. doi:10.1002/jsfa.10987
18. Wu M, Sun J, Lu B, Ge X, Zhou X, Zou M. Application of deep brief network in transmission spectroscopy detection of pesticide residues in lettuce leaves. J FOOD PROCESS Eng. 2019;42(3). doi:10.1111/jfpe.13005
19. Vinushree N, Hemalatha B, Kaliappan VK. Efficient kernel-based fuzzy C-means clustering for pest detection and classification. In: Proc. - World Congr. Comput. Commun. Technol., WCCCT. IEEE Computer Society; 2014:179-181. doi:10.1109/WCCCT.2014.61
20. Agarwal M, Bohat V, Ansari M, et al. A Convolution Neural Network based approach to detect the disease in Corn Crop. In: PROCEEDINGS OF THE 2019 IEEE 9TH INTERNATIONAL CONFERENCE ON ADVANCED COMPUTING (IACC 2019). ; 2019:176-181.
21. Alibabaei K, Gaspar P, Lima T. Crop Yield Estimation Using Deep Learning Based on Climate Big Data and Irrigation Scheduling. ENERGIES. 2021;14(11). doi:10.3390/en14113004
22. Huang Z, IEEE. Poverty Prediction Through Machine Learning.; 2021:324. doi:10.1109/ECIT52743.2021.00073
23. Yin L, Zhang Y. Village precision poverty alleviation and smart agriculture based on FPGA and machine learning. Microprocess Microsyst. Published online 2020. doi:10.1016/j.micpro.2020.103469
24. Yin J, Qiu Y, Zhang B. Identification of Poverty Areas by Remote Sensing and Machine Learning: A Case Study in Guizhou, Southwest China. ISPRS Int J GEO-Inf. 2021;10(1). doi:10.3390/ijgi10010011
25. Adewopo JB, Solano-Hermosilla G, Colen L, Micale F. Using crowd-sourced data for real-time monitoring of food prices during the COVID-19 pandemic: Insights from a pilot project in northern Nigeria. Glob Food Secur. 2021;29. doi:10.1016/j.gfs.2021.100523
26. Kim J, Cha M, Lee JG. Nowcasting commodity prices using social media. PeerJ Comput Sci. 2017;2017(7). doi:10.7717/peerj-cs.126
27. Cesare N, Dwivedi P, Nguyen Q, Nsoesie E. Use of social media, search queries, and demographic data to assess obesity prevalence in the United States. PALGRAVE Commun. 2019;5. doi:10.1057/s41599-019-0314-x
28. Anguita-Ruiz A, Segura-Delgado A, Alcala R, Aguilera C, Alcala-Fdez J. eXplainable Artificial Intelligence (XAI) for the identification of biologically relevant gene expression patterns in longitudinal human studies, insights from obesity research. PLOS Comput Biol. 2020;16(4). doi:10.1371/journal.pcbi.1007792
29. Bentley RA, Ormerod P, Ruck DJ. Recent origin and evolution of obesity-income correlation across the United States. Palgrave Commun. 2018;4(1). doi:10.1057/s41599-018-0201-x
30. DeJesus JM. Implicit attitudes, eating behavior, and the development of obesity. Behav Brain Sci. 2017;40:e114. doi:10.1017/S0140525X16001394
31. Pascual-Ferrá P, Alperstein N, Barnett DJ. Social Network Analysis of COVID-19 Public Discourse on Twitter: Implications for Risk Communication. Disaster Med Public Health Prep.:1-9. doi:10.1017/dmp.2020.347
32. Xue J, Chen J, Chen C, Zheng C, Li S, Zhu T. Public discourse and sentiment during the COVID 19 pandemic: Using Latent Dirichlet Allocation for topic modeling on Twitter. PLOS ONE. 2020;15(9):e0239441. doi:10.1371/journal.pone.0239441
33. Rauchfleisch A, Vogler D, Eisenegger M. Public Sphere in Crisis Mode: How the COVID-19 Pandemic Influenced Public Discourse and User Behaviour in the Swiss Twitter-sphere. Javn Ljubl Slov. 28(2):129-148. doi:10.1080/13183222.2021.1923622
34. Al-Ramahi M, Elnoshokaty A, El-Gayar O, Nasralah T, Wahbeh A. Public Discourse Against Masks in the COVID-19 Era: Infodemiology Study of Twitter Data. JMIR Public Health Surveill. 2021;7(4):e26780. doi:10.2196/26780

35. Fazel S, Zhang L, Javid B, Brikell I, Chang Z. Harnessing Twitter data to survey public attention and attitudes towards COVID-19 vaccines in the UK. Sci Rep. 2021;11(1):23402. doi:10.1038/s41598-021-02710-4
36. Criss S, Nguyen TT, Norton S, et al. Advocacy, Hesitancy, and Equity: Exploring U.S. Race-Related Discussions of the COVID-19 Vaccine on Twitter. Int J Environ Res Public Health. 2021;18(11):5693. doi:10.3390/ijerph18115693
37. Li L, Aldosery A, Vitiugin F, et al. The Response of Governments and Public Health Agencies to COVID-19 Pandemics on Social Media: A Multi-Country Analysis of Twitter Discourse. Front Public Health. 2021;9:716333. doi:10.3389/fpubh.2021.716333
38. Moulaei K, Shanbehzadeh M, Mohammadi-Taghiabad Z, Kazemi-Arpanahi H. Comparing machine learning algorithms for predicting COVID-19 mortality. BMC Med Inform Decis Mak. 2022;22(1). doi:10.1186/s12911-021-01742-0
39. Margus C, Brown N, Hertelendy AJ, Safferman MR, Hart A, Ciottone GR. Emergency Physician Twitter Use in the COVID-19 Pandemic as a Potential Predictor of Impending Surge: Retrospective Observational Study. J Med Internet Res. 2021;23(7):e28615. doi:10.2196/28615
40. Velichko E, Shariaty F, Orooji M, et al. Development of computer-aided model to differentiate COVID-19 from pulmonary edema in lung CT scan: EDECOVID-net. Comput Biol Med. 2022;141. doi:10.1016/j.compbiomed.2021.105172
41. How artificial intelligence may help the Covid-19 pandemic: Pitfalls and lessons for the future - Malik - 2021 - Reviews in Medical Virology - Wiley Online Library. Accessed January 15, 2022. https://onlinelibrary-wiley-com.proxy1.library.jhu.edu/doi/full/10.1002/rmv.2205
42. Dutt R, Basu M, Ghosh K, Ghosh S. Utilizing microblogs for assisting post-disaster relief operations via matching resource needs and availabilities. Inf Process Manag. 2019;56(5):1680-1697. doi:10.1016/j.ipm.2019.05.010
43. Sindhu S, Nair DS, Maya VS, Thanseeha MT, Vishnu Hari C. Disaster management from social media using machine learning. In: Institute of Electrical and Electronics Engineers Inc.; 2019:246-252. doi:10.1109/ICACC48162.2019.8986198
44. Machine Learning for Social Media Data Analysis in Disaster Management. Presented at: ASEAN Information and Communications Technology; November 27, 2018; Sari Pacific Jakarta, Jakarta, Indonesia. Accessed April 3, 2021. https://www.nict.go.jp/en/asean_ivo/lde9n2000000iqi5-att/23-a1544493279473.pdf
45. Baclic O, Tunis M, Young K, Doan C, Swerdfeger H, Schonfeld J. Challenges and opportunities for public health made possible by advances in natural language processing. Can Commun Dis Rep. 2020;46(6):161-168. doi:10.14745/ccdr.v46i06a02
46. Conway M, Hu M, Chapman WW. Recent Advances in Using Natural Language Processing to Address Public Health Research Questions Using Social Media and ConsumerGenerated Data. Yearb Med Inform. 2019;28(1):208-217. doi:10.1055/s-0039-1677918
47. Morgenstern JD, Rosella LC, Daley MJ, Goel V, Schünemann HJ, Piggott T. “AI’s gonna have an impact on everything in society, so it has to have an impact on public health”: a fundamental qualitative descriptive study of the implications of artificial intelligence for public health. BMC Public Health. 2021;21(1):40. doi:10.1186/s12889-020-10030-x
48. Patel CJ, Rasooly D. Data Science and Machine Learning in Public Health: Promises and Challenges | Blogs | CDC. Genomics and Precision Medicine. Published September 20, 2019. Accessed October 4, 2021. https://blogs.cdc.gov/genomics/2019/09/20/data-science/
49. Wang W, Kiik M, Peek N, et al. A systematic review of machine learning models for predicting outcomes of stroke with structured data. PLOS ONE. 2020;15(6):e0234722. doi:10.1371/journal.pone.0234722
50. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst Rev. 2021;10(1):1-11. doi:10.1186/s13643-021-01626-4
51. Covidence - Better systematic review management. Covidence. Accessed April 1, 2021. https://www.covidence.org/
52. Topic Modeling: An Introduction. MonkeyLearn Blog. Published September 26, 2019. Accessed October 27, 2021. https://monkeylearn.com/blog/introduction-to-topic-modeling/
53. Welcome to Python.org. Accessed September 27, 2021. https://www.python.org/
54. Google Colaboratory. Accessed September 27, 2021. https://colab.research.google.com/notebooks/intro.ipynb?utm_source=scs-index
55. Martin N. Github Repository for LDA Topic Modeling Code.; 2021. Accessed September 27, 2021. https://github.com/drscicomm/food-security/blob/a70b06f68c4e36f6a4a49bab75f765061a226f1e/LDA_viz_C_20210924.ipynb
56. Martin N. Github Repository for Polarity and Subjectivity Analysis via TextBlob.; 2021. Accessed September 27, 2021. https://github.com/drscicomm/food-security/blob/a70b06f68c4e36f6a4a49bab75f765061a226f1e/TextBlob_Sentiment_Analysis_ML_Corpus.ipynb
57. Martin N. Github Repository for Sentiment Analysis via Vader.; 2021. Accessed September 27, 2021. https://github.com/drscicomm/food-security/blob/a70b06f68c4e36f6a4a49bab75f765061a226f1e/Vader_Sentiment_Analysis_ML_corpus.ipynb
58. Topic modeling visualization - How to present results of LDA model? | ML+. Machine Learning Plus. Published December 4, 2018. Accessed October 27, 2021. https://www.machinelearningplus.com/nlp/topic-modeling-visualization-how-to-present-results-lda-models/
59. Kapadia S. Evaluate Topic Models: Latent Dirichlet Allocation (LDA). Medium. Published December 29, 2020. Accessed September 27, 2021. https://towardsdatascience.com/evaluate-topic-model-in-python-latent-dirichlet-allocation-lda-7d57484bb5d0
60. Sharma H. Topic Model Visualization using pyLDAvis. Medium. Published June 5, 2021. Accessed October 4, 2021. https://towardsdatascience.com/topic-model-visualization-using-pyldavis-fecd7c18fbf6
61. Barbosa RM, Nelson DR. The Use of Support Vector Machine to Analyze Food Security in a Region of Brazil. Appl Artif Intell. 2016;30(4):318-330. doi:10.1080/08839514.2016.1169048

62. Brock LG III, Davis LB. Estimating available supermarket commodities for food bank collection in the absence of information. Expert Syst Appl. 2015;42(7):3450-3461. doi:10.1016/j.eswa.2014.11.068
63. Scott D, Oh J, Chappelka M, Walker-Holmes M, DiSalvo C. Food for thought: Analyzing public opinion on the Supplemental Nutrition Assistance Program. J Technol Hum Serv. 2018;36(1):37-47. doi:10.1080/15228835.2017.1416514
64. Chrisinger BW, Kinsey EW, Pavlick E, Callison-Burch C. SNAP judgments into the digital age: Reporting on food stamps varies significantly with time, publication type, and political leaning. PloS One. 2020;15(2):e0229180. doi:10.1371/journal.pone.0229180
65. Gilman M. AI Algorithms Intended to Detect Welfare Fraud Often Punish the Poor. US News & World Report. Published February 14, 2020. Accessed October 1, 2021. //www.usnews.com/news/best-states/articles/2020-02-14/ai-algorithms-intended-to-detect-welfare-fraud-often-punish-the-poor-instead
66. Eskandari F, Lake AA, Weeks G, Butler M. Twitter conversations about food poverty: an analysis supplemented with Google Trends analysis. The Lancet. 2019;394((Eskandari F., F.Eskandari@tees.ac.uk; Lake A.A.; Weeks G.; Butler M.) School of Science, Engineering and Design, Teesside University, Middlesbrough, United Kingdom):S38. doi:10.1016/S0140-6736(19)32835-1
67. Eyre RW, House T, Xavier Gómez-Olivé F, Griffiths FE. Bayesian belief network modelling of household food security in rural South Africa. BMC Public Health. 2021;21(1). doi:10.1186/s12889-021-10938-y
68. Fantazzini D. Nowcasting and forecasting the monthly food stamps data in the us using online search data. PLoS ONE. 2014;9(11). doi:10.1371/journal.pone.0111894
69. Feeding America Annual Report 2020. Feeding America; 2020. Accessed August 18, 2021. https://www.feedingamerica.org/sites/default/files/2020-12/2020%20Annual%20Report_FINAL.pdf
70. Gadzalo I, Sychevskiy M, Kovalenko O, Deineko L, Yashchenko L. Assessment of global food demand in unexpected situations. Innov Mark. 2020;16(4):91-103. doi:10.21511/im.16(4).2020.08
71. Hamad R, Templeton ZS, Schoemaker L, Zhao M, Bhattacharya J. Comparing demographic and health characteristics of new and existing SNAP recipients: application of a machine learning algorithm. Am J Clin Nutr. 2019;109(4):1164-1172. doi:10.1093/ajcn/nqy355
72. Lentz EC, Michelson H, Baylis K, Zhou Y. A data-driven approach improves food insecurity crisis prediction. World Dev. 2019;122:399-409. doi:10.1016/j.worlddev.2019.06.008
73. Lukyamuzi A, Ngubiri J, Okori W. Tracking food insecurity from tweets using data mining techniques. In: Proc Int Conf Software Eng. IEEE Computer Society; 2018:27-34. doi:10.1145/3195528.3195531
74. Madichetty S, M S. A stacked convolutional neural network for detecting the resource tweets during a disaster. Multimed Tools Appl. Published online September 25, 2020:1-23. doi:10.1007/s11042-020-09873-8
75. Mills. Understanding the Rates, Causes, and Costs of Churning in the Supplemental Nutrition Assistance Program (SNAP). Urban Inst. Published online 2015:11.
76. Bharadwaj R. AI in Government - Current AI Projects in the Public Sector. Emerj. Accessed August 27, 2021. https://emerj.com/ai-sector-overviews/ai-government-current-ai-projects-public-sector/
77. Nica-Avram G, Harvey J, Smith G, Smith A, Goulding J. Identifying food insecurity in food sharing networks via machine learning. J Bus Res. 2021;131:469-484. doi:10.1016/j.jbusres.2020.09.028
78. Nica-Avram G, Harvey J, Goulding J, et al. FIMS: Identifying, Predicting and Visualising Food Insecurity. In: Web Conf. - Companion World Wide Web Conf., WWW. Association for Computing Machinery; 2020:190-193. doi:10.1145/3366424.3383538
79. Torres R. Can machine learning help more people access SNAP benefits? Technical.ly Philly. Published October 1, 2018. Accessed August 18, 2021. https://technical.ly/philly/2018/10/01/machine-learning-model-help-more-people-access-benefits/
80. Pugh N, Davis LB. Forecast and analysis of food donations using support vector regression. In: Obradovic Z., Baeza-Yates R., Kepner J., et al., eds. Proc. - IEEE Int. Conf. Big Data, Big Data. Vol 2018-January. Institute of Electrical and Electronics Engineers Inc.; 2018:3261-3267. doi:10.1109/BigData.2017.8258309
81. Razzaq A, Ahmed UI, Hashim S, et al. An Automatic Determining Food Security Status: Machine Learning based Analysis of Household Survey Data. Int J Food Prop. 2021;24(1):726-736. doi:10.1080/10942912.2021.1919703
82. Shi ZR, Lizarondo L, Fang F. A recommender system for crowdsourcing food rescue platforms. In: Web Conf. - Proc. World Wide Web Conf., WWW. Association for Computing Machinery, Inc; 2021:857-865. doi:10.1145/3442381.3449787
83. Turner-McGrievy G, Karami A, Monroe C, Brandt HM. Dietary pattern recognition on Twitter: a case example of before, during, and after four natural disasters. Nat Hazards. 2020;103(1):1035-1049. doi:10.1007/s11069-020-04024-6
84. Westerveld JJL, van den Homberg MJC, Nobre GG, van den Berg DLJ, Teklesadik AD, Stuit SM. Forecasting transitions in the state of food security with machine learning using transferable features. Sci Total Environ. 2021;786:147366. doi:10.1016/j.scitotenv.2021.147366
85. Widener M, Li W. Using geolocated Twitter data to monitor the prevalence of healthy and unhealthy food references across the US. Appl Geogr. 2014;54:189-197. doi:10.1016/j.apgeog.2014.07.017
86. Famine Action Mechanism (FAM). World Bank. Accessed August 18, 2021. https://www.worldbank.org/en/programs/famine-early-action-mechanism
87. WFP and Alibaba unveil next generation of machine learning technology in the fight against hunger | World Food Programme. Accessed August 18, 2021. https://www.wfp.org/news/wfp-and-alibaba-unveil-next-generation-machine-learning-technology-fight-against-hunger
88. HungerMap LIVE. World Food Programme. Published 2018. Accessed August 18, 2021. https://hungermap.wfp.org/

89. Yiu T. Understanding Random Forest. Medium. Published September 29, 2021. Accessed October 16, 2021. https://towardsdatascience.com/understanding-random-forest-58381e0602d2
90. Polzer D. 7 of the Most Used Regression Algorithms and How to Choose the Right One. Medium. Published October 3, 2021. Accessed October 16, 2021. https://towardsdatascience.com/7-of-the-most-commonly-used-regression-algorithms-and-how-to-choose-the-right-one-fc3c8890f9e3
91. Morde V. XGBoost Algorithm: Long May She Reign! Medium. Published April 8, 2019. Accessed October 16, 2021. https://towardsdatascience.com/https-medium-com-vishalmorde-xgboost-algorithm-long-she-may-rein-edd9f99be63d
92. Cohn S. Trump proposal to crack down on food stamp fraud reignites a heated debate. CNBC. Published August 30, 2019. Accessed January 15, 2022. https://www.cnbc.com/2019/08/30/trump-plan-to-crack-down-on-food-stamp-fraud-reignites-heated-debate.html
93. Varga I, Sano M, Torisawa K, et al. Aid is Out There: Looking for Help from Tweets during a Large Scale Disaster. In: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics; 2013:1619-1629. Accessed October 4, 2021. https://aclanthology.org/P13-1159
94. Basu M, Ghosh K, Das S, Dey R, Bandyopadhyay S, Ghosh S. Identifying Post-Disaster Resource Needs and Availabilities from Microblogs. In: Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017. ACM; 2017:427-430. doi:10.1145/3110025.3110036
95. Purohit H, Castillo C, Diaz F, Sheth A, Meier P. Emergency-relief coordination on social media: Automatically matching resource requests and offers. First Monday. Published online 2014. doi:10.5210/fm.v19i1.4848
96. Nazer TH, Xue G, Ji Y, Liu H. Intelligent Disaster Response via Social Media Analysis A Survey. SKDD. Published online 2017. doi:10.1145/3137597.3137602
97. scikit-learn: machine learning in Python — scikit-learn 1.0 documentation. Accessed October 4, 2021. https://scikit-learn.org/stable/
98. NLP Gensim Tutorial - Complete Guide For Beginners - GeeksforGeeks. Accessed October 4, 2021. https://www.geeksforgeeks.org/nlp-gensim-tutorial-complete-guide-for-beginners/
99. Matplotlib: Python plotting — Matplotlib 3.4.3 documentation. Accessed October 4, 2021. https://matplotlib.org/
100. TextBlob: Simplified Text Processing — TextBlob 0.16.0 documentation. Accessed October 4, 2021. https://textblob.readthedocs.io/en/dev/
101. Textblob vs Vader For Sentiment Analysis in Python. Analytics Vidhya. Published January 6, 2021. Accessed October 4, 2021. https://www.analyticsvidhya.com/blog/2021/01/sentiment-analysis-vader-or-textblob/
102. Python | Sentiment Analysis using VADER - GeeksforGeeks. Accessed October 4, 2021. https://www.geeksforgeeks.org/python-sentiment-analysis-using-vader/
103. Vissing Y, Hudson CG, Nilan D. Changing the Paradigm of Homelessness. Taylor and Francis; 2019:233. doi:10.4324/9780429431814
104. How Maryland Optimized Data Analytics to Target COVID-19 Outreach to Vulnerable Populations. AJMC. Accessed August 18, 2021. https://www.ajmc.com/view/how-maryland-optimized-data-analytics-to-target-covid-19-outreach-to-vulnerable-populations
105. Black JC, Margolin ZR, Olson RA, Dart RC. Online Conversation Monitoring to Understand the Opioid Epidemic: Epidemiological Surveillance Study. JMIR Public Health Surveill. 2020;6(2):e17073. doi:10.2196/17073

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 July 2022 doi:10.20944/preprints202207.0221.v1



# TABLES

**Table 1: Databases Included, Search Terms Utilized, and Search Results by Database**

| | |
|---|---|
| Databases | |
| Bibliographic Databases | PubMed, Embase, Scopus, Web of Science, PsychInfo |
| Gray Literature | Google, Science.gov, Worldwidescience.org, World Food Programme Website, U.S. Agency for International Development Website, UNICEF Website, Feeding America Website |
| Key Search Terms* | Machine Learning; Artificial Intelligence; Natural Language Processing; Support Vector Machine; Naive Bayes; bayesian learning; Neural network; random forest boosting; deep learning; sentiment analysis; text mining; opinion mining; social network analysis<br><br>Food Assistance; Food Supply; Food Bank; Food Security |
| Search Results | PubMed: 154<br>Embase: 132<br>Scopus: 1989<br>Web of Science: 1601<br>PsychInfo: 26 |

*The key search terms were edited/formatted for each database.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 July 2022 doi:10.20944/preprints202207.0221.v1



**Table 2: Article Inclusion and Exclusion Criteria**

| Field | Inclusion Criteria | Exclusion Criteria |
|---|---|---|
| Language | English only | Non-English language |
| Publication Date Range | 2011-2021 | Outside of date range |
| Technology Type | Machine Learning | Non-Machine Learning |
| Technology Purpose | Food assistance, Food security | Non-food/nutrition related, agriculture:<br>✗ Detection or prediction of pests, weather changes, crop shortages or yields<br>✗ Food supply chain<br>✗ Food recognition systems<br>✗ Food safety detection systems<br>✗ Food image detection<br>✗ Customer satisfaction with restaurant or hotel food<br>✗ Diet or obesity related<br>✗ Food fraud<br>✗ Forecasting food price |

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 July 2022 doi:10.20944/preprints202207.0221.v1



9 **Table 3: Characteristics of reviewed articles (n = 28)**

| | Study ID | Year Published | Food Insecurity Cause | Policy Focus? | Tool Purpose | Type of Tool | Type of Learning | Algorithm(s) Used |
|---|---|---|---|---|---|---|---|---|
| 1 | Barbosa[65] | 2016 | Poverty | Yes | Classifying households that are food insecure | Classification | Supervised | Support Vector Machine |
| 2 | Brock[66] | 2015 | Not specified | No | Prediction of in-kind donations to food banks | Prediction | Supervised | Multi-layer perceptron neural network |
| 3 | Chappelka[67] | 2018 | Ending hunger by 2025 | Yes | Classification of food security tweets, text mining on SNAP & policymaker opinions. | Opinion mining | Supervised & Unsupervised | Natural language processing sentiment analysis |
| 4 | Chrisinger[68] | 2020 | Not specified | Yes | Assess how, why, and when media discusses SNAP and SNAP users | Topic modelling; qualitative content analysis | Unsupervised | Structural topic modeling |
| 5 | Gilman[69] | 2020 | Not specified | Yes | Detection of SNAP fraud | Classification | Supervised | Not specified |
| 6 | Eskandari[70] | 2019 | Not specified | Yes | Classification of food insecurity trends | Classification | Not specified | Not specified |
| 7 | Eyre[71] | 2021 | Not specified | Yes | Show future impact of interventions on food insecurity in the community | Classification | Supervised and expert network learning | Bayesian belief network |
| 8 | Fantazzini[72] | 2014 | Not specified | Yes | Nowcasting and forecasting SNAP users | Prediction | Supervised | Test 2000 algorithms, e.g., neural network |
| 9 | Feeding America[73] | 2020 | Not specified | No | Increase ability of food banks to reallocate food waste and large donations | Classification | Not specified | Not specified |
| 10 | Gadzalo[74] | 2020 | COVID-19 | Yes | Impact of disaster on consumer buying trends | Food trends during disasters | Not specified | Google Trends |
| 11 | Hamad[75] | 2019 | Great Recession; Poverty | Yes | Classification of SNAP users | Classification | Unsupervised | (1) Lasso regression; (2) Ordinary least squares followed by regression |
| 12 | Lentz[76] | 2019 | Poverty | Yes | Predict food insecurity | Prediction; Clustering | Supervised | Linear and Non-Linear Regression |

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 July 2022 doi:10.20944/preprints202207.0221.v1



| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 13 | Lukyamuzi[77] | 2018 | Not specified | Yes | Classify patterns from tweets | Classification | Supervised | K Nearest Neighbors, Naïve Bayes, Support Vector Machine, Neural Network, Decision Trees |
| 14 | Madichetty[78] | 2020 | Natural disaster (hurricane) | No | Automatic detection of Need and Available Resource Tweets | Classification | Supervised | Stacked convoluted neural network |
| 15 | Mills[79] | 2015 | Not specified | Yes | Detection of churn in SNAP | Classification | Supervised | Not specified |
| 16 | New York City Government[80] | 2019 | Not specified | Yes | Classify applicants to SNAP as eligible or not | Classification | Not specified | Not specified |
| 17 | Nica-Avram[81] | 2021 | Not specified | Yes | Predict food insecurity; Classify users of OLIO app who are food insecure | Classification | Supervised | Random Forest, AdaBoost, Support Vector Machine, k Nearest Neighbors |
| 18 | Nica-Avram[82] | 2020 | Not specified | Yes | Classification of food insecure users of OLIO | Classification | Supervised | Random Forest, Support Vector Machine, Gradient Boosted, k Nearest Neighbors |
| 19 | Philadelphia Government[83] | 2018 | Not specified | Yes | Predict needs of new and existing SNAP users | Prediction | Not specified | Not specified |
| 20 | Pugh[84] | 2018 | Not specified | No | Prediction of in-kind donations to food banks | Prediction | Supervised | Support Vector Regression |
| 21 | Razzaq[85] | 2021 | Multiple | No | Prediction of food insecurity; rural and agricultural villages | Prediction | Supervised | Random Forest, Support Vector Machine, k Nearest Neighbors, Neural Network, Naïve Bayes, Logistic Regression |
| 22 | Shi[86] | 2021 | Not specified | No | Improve volunteer recruitment rate | Recommender | Supervised | Neural Network |
| 23 | Turner-McGrievy[87] | 2020 | Hurricanes | No | Classification | Text mining; Trends analysis | NA | Text mining |
| 24 | Westerveld[88] | 2021 | Poverty, Famine | No | Prediction of food insecurity | Prediction | Supervised | Random Forest, Extreme Gradient Boosting, Xgboost |

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 July 2022 doi:10.20944/preprints202207.0221.v1



| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 25 | Widener[89] | 2014 | Not specified | Yes | Classification of food security tweets | Opinion Mining | Supervised | Sentiment Analysis |
| 26 | World Bank[90] | 2018 | Lack of food | | Prediction and early warning system of food insecurity events & potentially impacted regions | Prediction | Not specified | Not specified |
| 27 | World Food Programme[91] | 2019 | Not specified | No | Classification of food insecure households | Classification; Image Recognition | Supervised | Digital Engine for Emergency Photo-analysis |
| 28 | World Food Programme[92] | 2018 | Not specified | Yes | Prediction of food insecure regions where monitoring data is lacking via HungerMap LIVE | Prediction | Supervised | XgBoost |

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 July 2022 doi:10.20944/preprints202207.0221.v1



**Table 4: Topic, top ten words per topic, and representative text**

| CATEGORY | TOPIC | | TERMS | REPRESENTATIVE TEXT |
|---|---|---|---|---|
| IMPROVING SNAP | 1 | Detecting SNAP application fraud | churn, client, benefit, state, cost, government, recertification, program, agency, fraud | "President Donald Trump recently suggested there is "tremendous fraud" in government welfare programs… For instance, the food stamp program, formally called the Supplemental Nutrition Assistance Program, currently serves about 40 million people monthly at an annual cost of US$68 billion. Despite regular denigration of food stamp recipients, less than 1% of benefits go to ineligible households, according to the federal government." (Daley, 2018) |
| | 2 | Characterizing new and existing SNAP users | snap, household, new, recipient, topic, program, exist, insecurity, include, level | "We sought to identify new and existing SNAP recipients, and to examine differences in sociodemographic characteristics, health, nutritional status, and food purchasing behavior between new and existing recipients of SNAP after the recession." (Hamad, 2019) |
| DETECTING NEEDS AND AVAILABLE RESOURCES | 3 | Predicting donations to food banks | amount, supermarket, receive, bank, warehouse, layer, event, method, approximation, regional | With network members and external partners, we tested a collaborative sourcing prototype that uses machine-learning algorithms that enable food banks to more effectively reallocate large-scale donations of perishable food such as fresh produce. We also co-developed tools to help improve particularly challenging food rescue situations such as identifying a nearby recipient for a grocery store donation. By marrying the algorithms with food sourcing and logistics processes, we hope to accelerate the evolution of supply chain systems." (Feeding America, 2020) |
| | 6 | Changes in food supply and demand (e.g., of staples like flour) during the pandemic | demand, change, country, product, increase, structure, consumer, value, pandemic, flour | "Under the conditions mentioned above, it is very important for the modern food market operators to correctly assess consumer demand, which can help governments solve food problems; food producers, on the other hand, ought to ensure a rational ratio of production and consumer demand, effectively manage their stocks and product range, develop production programs and apply effective methods of product promotion." (Gadzalo, 2020) |
| | 7 | Forecasting food rescue needs and resources (e.g., volunteers) | system, rescue, value, volunteer, network, donation, forecast, give, problem, expert | "Since they rely on external volunteers to pick up and deliver the food, some Food Rescues use web-based mobile applications to reach the right set of volunteers. In this paper, we propose the first machine learning based model to improve volunteer engagement in the food waste and security domain. We (1) develop a recommender system to send push notifications to the most likely volunteers for each given rescue, (2) leverage a mathematical programming based approach to diversify our recommendations, and (3) propose an online algorithm to dynamically select the volunteers to notify |

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 July 2022 doi:10.20944/preprints202207.0221.v1



| | | | | |
|---|---|---|---|---|
| | | | | without the knowledge of future rescues." (Shi, 2021) |
| | 8 | Predicting household food insecurity | security, household, level, measure, class, crisis, insecurity, country, access, prediction | "Here we use a methodology that allows us to consider causal interactions between the multiple, interacting variables involved in food security, applied to data from a low resource, rural South African community. In particular, we construct belief networks between variables using Bayesian reasoning and consider how these compare with the community's own beliefs about determinants of food security. Such belief networks can be interpreted, with policy implications drawn out, by nonspecialists, and as such we believe that the methodology could usefully be applied more frequently in food security research." (Eyre, 2021) |
| CHARACTERIZING CURRENT FOOD SECURITY ISSUES & OPINIONS | 4 | World Food Programme initiatives | security, blockchain, level, measure, household, insecurity, drone, grateful_l, coordinator, relentlessly | "The unmanned aerial vehicle (UAV) project seeks to enhance the ability of WFP and partners to prepare for, and respond to, emergencies in multiple ways – rapid damage analysis, topographical data collection and interpretation through machine learning and Artificial Intelligence, high-resolution imagery for the mapping of vulnerabilities, and connectivity in disasters. Separately, a blockchain pilot in refugee camps in Jordan is making WFP's cash transfers to 100,000 vulnerable Syrians more efficient and transparent. Piloted in collaboration with other agencies, the 'Building Blocks' project is delivering more for less, offering donors better value for money." (WFP, 2019) |
| | 5 | Sentiment [classification] analysis of tweets | tweet, classifier, sentiment, word, feature, trend, propose, dataset, negative, positive | "This project explores public opinion on the Supplemental Nutrition Assistance Program (SNAP) in news and social media outlets, and tracks elected representatives' voting records on issues relating to SNAP and food insecurity. We used machine learning, sentiment analysis, and text mining to analyze national and state level coverage of SNAP in order to gauge perceptions of the program over time across these outlets." (Chappelka, 2018) |
| | 9 | Hurricane preparedness and response on Twitter | mention, hurricane, twitter, healthy, tweet, category, list, day, group, examine | "Little is known about what foods/beverages (F&B) are common during natural disasters. The goal of this study was to track high-frequency F&B mentions during four hurricanes affecting the coast of South Carolina for quantifying dietary patterns in Twitter." (Turner-McGrievy, 2020) |

13

 doi:10.20944/preprints202207.0221.v1



**Table 5: Machine learning tools for food assistance research summary**

| Tool Categories | Description of Tool Category | # of Articles | Food Insecurity Causes | Intended Users and/or Beneficiaries | Top Algorithms | Challenges (Top 3) |
|---|---|---|---|---|---|---|
| Predict Food Insecurity | Includes nowcasting (prediction of the present, the very near future, and the very recent past) and forecasting at the individual, household, regional and global levels | 6 | Lack of food (1), Poverty (2), Not specified (2), Climate Change (1) | Humanitarian relief workers & agencies, policymakers, researchers, food donors, food recipients | Random Forest, Regression, XgBoost | (1) Data issues: Lack of data; countries collecting different types of metrics (2) Algorithm & feature selection: complex variables & confounders (3) Lack of political will to predict food security issues |
| Classify Needs and Available Resources | Improving detection of resources and needs in and out of disaster settings. | 7 | COVID-19 (1), Hurricanes (3), Not specified (4) | Food banks, donors, grocery stores, policymakers, producers, humanitarian relief workers, food recipients, volunteers | Neural Network, Support Vector Regression | (1) Some countries' do not frequently use social media (2) Internet may be unavailable during disasters (3) Requires adequate data collection from multiple sources to return accurate results |
| Characterizing & Improving SNAP | Includes application prescreening, churn, fraud, characterizing needs of users | 9 | Ending hunger by 2025 (1), Not specified (6), Poverty (1), Great Recession (1) | food banks, stakeholders, policymakers, SNAP program officers, SNAP users | Natural language processing sentiment analysis, topic modeling, neural network | (1) Poor accuracy of algorithms (2) Fear of government using machine learning to discriminate against SNAP users (3) Lack of complete and/or accurate data |
| Situational Awareness of Current Food Security Issues | Classifying and nowcasting current food security needs for both disaster and non-disaster settings | 6 | Not specified (5), Poverty (1) | Policymakers, researchers, stakeholders, food recipients, OLIO users | Support Vector Machine, Bayesian belief network, K nearest neighbors, sentiment analysis | (1) Data issues: country-level monitoring is poor & varied (2) Model accuracy is dependent on expert understanding & selection of variables (3) Lack of use in disaster settings |