APPLIED ARTIFICIAL INTELLIGENCE
2016, VOL. 30, NO. 4, 318–330
http://dx.doi.org/10.1080/08839514.2016.1169048

The Use of Support Vector Machine to Analyze Food Security in a Region of Brazil

Rommel Melgaço Barbosa[a] and Donald R. Nelson[b]

[a]Instituto de Informática, Universidade Federal de Goiás, Goiânia, GO, Brazil; [b]Department of Anthropology, University of Georgia, Athens, GA, USA

ABSTRACT

The application of Support Vector Machine (SVM) to classify food security in a northeast region in Brazil is explored. This type of application represents a novel use of the SVM in addressing contemporary social science questions. The results demonstrate an accuracy > 75% and a recall of 84% for classifying households that are food insecure. The variables identified by the model are consistent with contemporary theories of food security and vulnerability. The successful application of SVM in this instance and the growing availability of large-scale social science datasets suggest that data mining techniques will have a larger role to play in answering critical social science questions in the future.

Introduction

Currently, 850 million people around the globe are chronically hungry (FAO 2012) and over 2 billion suffer from stunting, wasting, and other effects of inadequate nutrition. Although the rate of food insecurity has dropped over the last 50 years, the rate of reduction has slowed recently; this is a trend that is influenced by a population that is increasing at a rate higher than production increases, high levels of poverty, and a changing climate, which contributes to greater risk in agricultural subsistence activities (Wheeler and Von Braun 2013). National governments as well as international and civil organizations invest billions of dollars per year in programs designed to reduce the incidence of food insecurity, and the United Nations recently launched an initiative entitled Zero Hunger Challenge (United Nations 2012) The quantification of food security is a first step in the design of these programs. The greater challenge is to identify the suite of interacting social, economic, and material factors that contribute to food insecurity in order to design effective policies that target the most food insecure populations. This article presents an application of Support Vector Machine modeling utilizing a dataset collected from farming households in northeast Brazil. The objective

CONTACT Rommel Melgaço Barbosa rommel@inf.ufg.br Instituto de Informática, Universidade Federal de Goiás, 74690-900 - Goiânia, GO, Brazil.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/UAAI.

was to assess the utility of the method for identifying the household-level factors most strongly associated with food insecurity. To the authors' knowledge, this is a novel application of the analytical technique.

The World Health Organization defines food security as existing when "all people at all times have access to sufficient, safe, nutritious food to maintain a healthy and active life" (FAO 1996). The three aspects of food security include availability, access, and utilization. Availability, which requires that sufficient quantities of food are available on a regular basis, is a necessary but insufficient condition to guarantee food security. Access to food refers to the ability of an individual or household to mobilize resources, either productive or financial, to gain access to a sufficient quantity and quality of food for their needs. Utilization refers to appropriate use of foods as well as adequate water and sanitation. A household is considered food insecure if it fails to meet the requirements for one or more of these conditions.

Brazil is a major food exporter and has fully functioning markets, but rates of food insecurity remain high. Therefore, in the present analysis we focus on the measure of food access. Even when food is readily available in the market, for example, many households do not have the resources to produce or purchase the quantity or quality that they require. Various factors, such as having sufficient land for production, formal and informal employment, and the number of workers in the household, among others, all contribute to the ability of a household to access sufficient food. Food insecurity might be a chronic state, in which households spend years without being able to meet their needs. It might also represent a short-term or temporary state that occurs as a result of a shock, such as increased food prices, a health emergency for the primary breadwinner, or a natural disaster.

Increases in food security require multiple types of investments, which include human, productive, economic, and natural resources (Rosegrant and Cline 2003). Although there is strong empirical evidence that demonstrates the range of needed investments, the individuals responsible for developing food security programming need to know how this set of variables plays out within a particular context. Because of the complexity of factors influencing food security, there is no universal approach that serves as a panacea. Although the suite of investments might be similar, the way investments interact impact food security in local contexts could be very different.

Data for this study was collected from farming households in the state of Ceará, northeast Brazil, and a brief description of the study region characterizes the complexity of food insecurity and informs the interpretation of the results of the analysis. Although Brazil is one of the largest economies in the world, portions of the country, including the northeast, still have much in common with developing countries around the world. For example, the level of food insecurity for the country is 35%, but rises to 54% in the northeast (IBGE 2010). This disparity in poverty rates follows a similar pattern. The northeast represents 28% of the

Brazilian population, yet contains over 70% of the poor population (de Souza and Osorio 2013).

Historically, agriculture was the economic mainstay of the Brazilian northeast. In the semiarid regions, such as Ceará, cotton was the primary cash export that provided a living for farmers during the early to mid-20th century. In the mid-1980s, with the arrival of the boll weevil and depressed cotton prices on the international market, cotton production all but disappeared from the region. Following the crash, the region's economy began to shift to an industrial and services base, but many people remained in the agricultural sector. Recent figures show that whereas 28% of the population's primary activity is in agriculture, agriculture accounts for only 4% of the state's economic activity (IPECE 2012). The lack of economic vigor within the agricultural sector is compounded by severe periodic droughts. The many families who produce most of the food that they consume are highly exposed to droughts due to the lack of irrigation infrastructure, which is unviable in most locations due to inaccessibility of subsurface water or the high saline content.

Drought impacts food access, and consequently food security, in several ways. First, food prices increase. Much of the region's agricultural production is for local markets, and, during a drought, staple crops must be imported, which leads to prices that are sometimes five times higher than normal. Second, agricultural labor demand disappears. For many, agricultural labor is a principal source of cash. When drought hits, the agricultural sector shrinks and many are left without cash income. Finally, almost all of the rural families are dependent to some degree on production to meet their food needs. Because the margin between production and consumption demands is so slim, drought losses can have significant impacts on food security.

Chronic food insecurity is a significant challenge in the region, a situation that was exacerbated in 2012 by the worst drought in over 50 years. By August 2012, nearly 85% of the urban water systems were near collapse and agricultural production losses reached 100% in some areas. At the time of the 475 interviews, only 38% of the households were food secure and 12% were severely food insecure. Individuals in these severely insecure households reported that there were sometimes nothing to eat in the house or that they went more than 24 hours without eating. With the application of the analysis in this study, we seek to determine the portfolio of factors that discriminate the level of food security among households.

Material and Methods

Data collection

Data were collected through household surveys during June–July 2012 using a two-stage stratified random sample. Six municípios were selected from among the 184 municípios in Ceará; each represented one of the six agro-ecological

Table 1. Description of capitals in a livelihoods framework.

Type of capital	Examples
Financial	Savings, access to credit, sources and value of income
Natural	Land size and quality, water, access to other natural resources
Material	Productive assets such as agricultural implements, basic infrastructure such as housing, energy, and communications
Social	Social resources such as networks, safety nets, group memberships
Human	Level of education, health, family size, skills, and knowledge

zones in the state. Approximately 80 households were interviewed in each of the selected municípios. The participant households were randomly selected from a list of agriculturalists maintained by the local Rural Workers Unions. Union membership is used by farmers to prove their vocation in order to be eligible for social security, and therefore, the lists are close to exhaustive. Data were entered and cleaned in SPSS (IBM Corp. 2012).

The purpose of the survey was to understand agricultural vulnerability to drought and the questionnaire was based on a livelihoods-vulnerability framework. Livelihoods refer to the suite of endowments and strategies that a household employs to meet its daily and future needs and goals. For analytical purposes, endowments are divided into five different categories of capitals: financial, natural, physical, social, and human (see Table 1). The questionnaire thus queried information related to demographics, economics, as well as production and migration data in order to understand the context and status of each of the households. It also included questions about risk-management practices including crop diversification, dependence on climate-sensitive income (such as agricultural-wage labor), and diversification of income source among others (see Table 2). In the results and discussion section, for each of the variables used in the model, we expanded on the relationship with livelihoods and food security.

A food security module was included to measure the impact of drought on consumption at a household level. The module was based on the Household Food Insecurity Access Scale (HFIAS) (Coates, Swindale, and Bilinsky 2007). This scale was designed to measure food insecurity in multiple cultural contexts and has been previously tested and calibrated for use in Brazil (Segall Corrêa et al. 2003). The module comprises questions related to household uncertainty of food availability, food quality, and food quantity. For each question, respondents are asked about the occurrence of a response (e.g., going a day without eating) and then frequency of a variable (e.g., rarely, sometimes, often). Four indicators can be calculated using the HFSAS module. For the current analysis, we calculated the indicator that measures insufficient quality of food, which was based on affirmative responses to any of the three food-quality-related questions (in order of increasing severity):

Table 2. Description of features in model.

Variable description	Relation to food security	Capital type or risk strategy	Variable type
Ability to access needed health care	Health is correlated with ability to work and earn income	Human	Binary
Extent of land cultivated	Correlated with volume of agricultural production	Natural	Continuous
Quality of soils	Quality of soil relates to the ability to avoid drought losses	Natural	Categorical
Asset index	Wealth indicator for household	Material	Continuous
Quality of irrigation source	Robustness of irrigation sources relates to ability to avoid drought losses	Material	Categorical
Per capita income	Correlated with the ability to purchase food	Financial	Continuous
Savings account	Wealth indicator and provides consumption buffer	Financial	Binary
Value of climate sensitive income	Correlates with drought sensitivity of household income	Financial	Continuous
Receive government cash transfers	Indicates a nonclimate sensitive source of household income	Financial	Binary
Crop diversity	Higher crop diversity spreads production risk	Risk management	Continuous (1–5)
County	Captures political and other structural differences where individuals reside	NA	Categorical

(1) In the past four weeks, were you or any household member not able to eat the kinds of foods you preferred because of a lack of resources?
(2) In the past four weeks, did you or any household member have to eat a limited variety of foods due to a lack of resources?
(3) In the past four weeks, did you or any household member have to eat some foods that you really did not want to eat because of a lack of resources to obtain other types of food?

The types of dietary changes of families that answered affirmatively to one or more of these questions included eliminating all meat from their diet; reducing dietary variety to beans and rice; just rice; or just cassava flour, which is made from a locally grown tuber.

Classification algorithms

Given an unlabeled sample, a classification problem concerns determining to which category it belongs, based on a training set of data having variables whose category is already known. The set of variables associated with a sample is represented as a vector. Machine learning entails the use of mathematics, statistics, and computational methods with the goal of finding efficient and accurate algorithms for classification.

Machine learning algorithms for classification have been successfully used in many different applications, such as: food science (Barbosa et al. 2014a, 2014b) animal science (Aguiar et al. 2012), crime hot-spot location prediction (Kianmehr and Alhajj 2008), pipeline defect prediction (Isa and Rajkumar 2009), seismic assessment of bridge diagnostic (Cheng, Wu, and Syu 2014), and the detection and location of leaks in water pipe networks (Mashford et al. 2012), to name a few. The learning steps of our classification problem start with an existing collection of labeled examples from which we divide the data into a training set and a test set. We use k-fold cross-validation for model selection and for performance evaluation. In our problem, the original set is partitioned into k equal-size subsets. In k subsets, one subset is set apart for testing, and the others, consisting of k–1 sets, are used for training. This process is repeated k times, with each one of the k subsets being used exactly one time as the test (Witten, Frank, and Hall 2011). In our problem, we used 10-fold cross-validation (Figure 1). After applying the classification algorithm and analyzing its accuracy using all original given features, we applied the same methods to a reduced number of selected features. This reduced number of features was obtained by a feature selection algorithm. The accuracy and computational complexity of the algorithm can be influenced by the total number of variables in a problem. By using a variable selection algorithm, we presuppose that a dataset might have variables that do not provide additional information beyond the selected features. Here, in our study, we used the correlation feature selection (CFS) subset algorithm (Hall 1998).

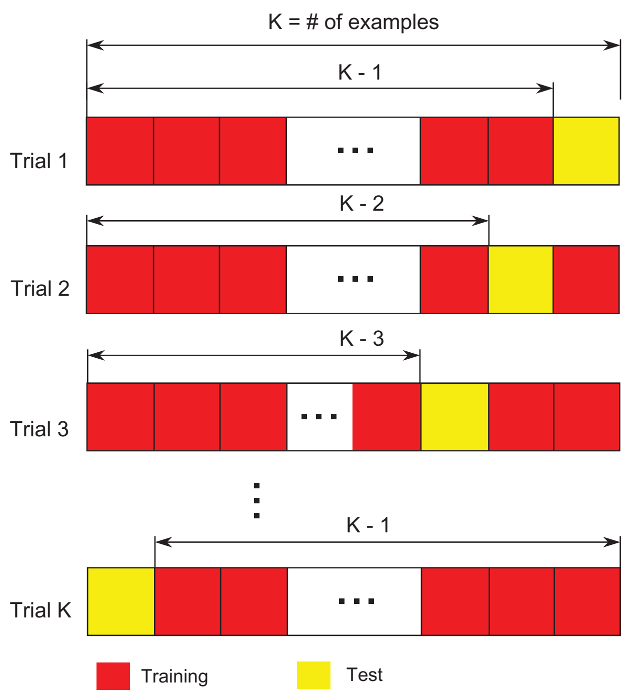

Figure 1. *K*-fold cross validation.

The present study employed the well-known Support Vector Machine classification algorithm. The Weka (Waikato Environment for Knowledge Analysis) software was used to perform the algorithm. Weka, which is available free under the GNU General Public License, contains a collection of visualization tools and algorithms for data analysis and classification (Witten, Frank, and Hall 2011).

Here, we provide an intuitive explanation of the algorithms. A more in-depth discussion can be found in Hamel (2011).

Support vector machines

Support Vector Machines (SVM) was introduced by Vapnik in 1963, and it is one of the most used classification algorithms. SVM is based on a method that finds a special type of linear model called the *maximum-margin hyperplane*. To picture a maximum-margin hyperplane, consider a two-class dataset whose classes are linearly separable, meaning that a hyperplane in the input space classifies all training instances correctly. Another way to define the maximum-margin hyperplane is that it is the one that gives the greatest separation between the classes. The instances that are closest to the maximum-margin hyperplane are called support vectors. Each class has at least one support vector, and often there are more. What is critical is that the set of support vectors singularly defines the maximum-margin hyperplane for the learning problem (Witten, Frank, and Hall 2011).

The classifier should choose a hyperplane that minimizes errors when classifying new samples. Decision boundaries with greater margins tend to generate fewer errors than those with smaller margins (Tan, Steinbach, and Kumar 2006). That is, a linear SVM is a classifier that looks for the hyperplane with the widest margin.

In our binary classification problem, each sample can be represented by one n-tuple (x_i, y_i), $i = 1, \ldots, N$, in which $(x_{i1}, x_{i2}, \ldots x_{id})^t$ corresponds to social features obtained from the ith example and $y_i \in \{-1, 1\}$.

The decision boundary of the classifier can be written as:

$$\mathbf{w} \cdot \mathbf{x} + b = 0,$$

in which w and b are parameters of the model (Tan, Steinbach, and Kumar 2006).

Thus, for a new sample z data, we can classify this example with the following:

$$y = 1, \mathit{if}\ \mathbf{w} \cdot \mathbf{z} + b > 0, \mathit{or}$$

$$y = -1, \mathit{if}\ \mathbf{w} \cdot \mathbf{z} + b < 0$$

The margin of the decision boundary is given by the distance of the hyperplanes that contain the support vectors of each class. This margin is given as

$$d = \|w\|^2/2$$

In the training phase of SVM, the values of w and b are obtained. Then, we should solve the following optimization problem:

$\min_w \|w\|^2/2$

Subject $y_i(\boldsymbol{w} \,.\, \boldsymbol{x_i} + b) \geq 1,\ i = 1, 2, \ldots, \text{N}.$

In some instances, two classes within a dataset cannot be divisible on a two-dimensional plane. When this occurs, a model that draws a decision boundary can be used. This process is known as *soft margin* (Tan, Steinbach, and Kumar 2006). In this case, slack variables (ζ_i) should be added to the constraints of the optimization problem.

$$\mathbf{w}.\ \mathbf{x_i} + b \geq 1 - \zeta_i, \mathit{if}\ y_i = 1$$

$$\mathbf{w}.\,\mathbf{x_i} + b \leq\ -1 + \zeta_i, \mathit{if}\ y_i = -1$$

with $\zeta_i > 0, \forall i$.

$\min_w \|w\|^2/2 + c\left(\sum_{i=1}^{N} \zeta_i\right)$

subject to $y_i\ .\ (\boldsymbol{w} \,.\, \boldsymbol{x_i} + b) \geq 1 - \zeta_i$.

When linear decision boundaries cannot be obtained, we convert the original dataset x into a new space $\varphi(x)$, so that in this new space a linear decision boundary separates the examples. The linear decision boundary in the newly constructed space fulfills the equation

$$\mathbf{w}.\boldsymbol{\varphi}(\mathbf{x}) + b = 0$$

We now have to resolve the optimization problem:

$\min_w \|w\|^2/2$

Subject to $y_i(\boldsymbol{w} \,.\, \boldsymbol{\varphi}(\boldsymbol{x_i}) + b \geq 1,\ \text{I} = 1, 2, \ldots, \text{N}.$

After solving the optimization problem, a new sample z can be classified using the function

$f(z) = \text{sign}(\boldsymbol{w} \,.\, \boldsymbol{\varphi}(\boldsymbol{z}) + b) = \text{sign}\left(\sum_{i=1}^{m} x_i\, y_i\, \boldsymbol{\varphi}(\mathbf{x_i}) \cdot \boldsymbol{\varphi}(\mathbf{z}) + b\right).$

The dot product in the transformed space can be expressed in terms of a similar function in the original space:

$$k(\mathbf{u}, \mathbf{v}) = \boldsymbol{\varphi}(\mathbf{u}) \cdot \boldsymbol{\varphi}(\mathbf{v}) = (\mathbf{u} \cdot \mathbf{v} + 1)^2,$$

This function k, which is computed in the original space, is referred to as the *kernel* function.

Analysis criteria

In order to evaluate the models, three performance criteria were used:

(1) Accuracy = (True Positives + True Negatives)/(True Positives+ True Negatives + False Positives + False Negatives)
(2) Recall(Sensitivity) = True Positives/(True Positives + False Negatives)
(3) Specificity = True Negatives/(True Negatives + False Positives)

Such that, true positive means X forecasted to belong to class C and is in fact in it; false positive means X forecasted to belong to C but is not actually in it; true negative means X not forecasted to belong to C and is not in fact in it; false negative X not forecasted to belong to C but is in fact in it.

Results and discussion

A number of simulations were run with different subsets of the original variables to achieve the highest possible level of accuracy in classifying food insecure households. Results for the original and final model are presented in Tables 3–5. The final 11 features were obtained through a feature selection algorithm and define a model with significant improvement in relation to the accuracy levels derived from the 75 original features (Table 3). The final model demonstrates 77% accuracy and a recall precision of 84%. In the current application, in which the objective is to identify food insecure households, the recall precision is the most relevant outcome. Policy makers, for example, are most interested in targeting the food insecure. In this case, the model demonstrates a 16% "leakage" in which food secure households are incorrectly identified as food secure.

The key features identified through the SVM are consistent with the expectations that emerge from a livelihoods framework. All of the capitals, with the exception of social capital, are relevant and included in the classification of food insecure households. Financial capital represents the highest number of features in the model. Per capita income is an indicator of the ability of a household to purchase required food, and the existence of savings provides a buffer against household income shocks. The value of climate-sensitive income represents the exposure of household income to drought events. The variable represents the value of agricultural-wage labor for a household. In the event of a drought, the availability of this type of labor disappears, and thus, households with high values of climate-sensitive income are particularly vulnerable to drought events. Household income from cash transfers refers to either social security or to the federal poverty reduction program. The monetary value of social security is higher than the poverty reduction payments but both provide stable, climate-neutral sources of income that are reliable even during drought events.

The asset index was developed based on an inventory of household consumer and productive assets. A pairwise comparison was made between each asset to identify which asset represents greater wealth, and a relative weight was derived for each type of asset. The index was calculated for each household using the weights along with the presence or absence of each asset in the household. Households with a higher asset index demonstrate greater purchasing power, but a higher index value also suggests that households are more food secure over long time periods. Drought shocks reduce available cash for asset purchase and households often respond to severe shocks through the sale of assets. Irrigation serves as a critical drought buffer and reduces agricultural production losses during drought. The variable captures differences in the robustness of irrigation sources. Households are classified as having irrigation, less dependable irrigation—which includes access to shallow reservoirs that dry up during severe droughts—and dependable irrigation, which refers to deep wells.

Two types of natural capital are included in the final model. Quality of soils is a self-reported value that represents the ability of soils to retain moisture. Farmers were asked whether their fields are able to produce well, moderately, or not at all in the event of a drought. This is a rough proxy that provides information on relative agricultural yields (kg/ha) of individuals in the sample. Extent of land cultivated refers to the size (ha) of the holdings that are planted by each household, based on the consideration that all things being equal, households that plant a larger area have a larger harvest. Both of these variables provide information on the level of drought risk that is experienced by a household. Related to this is a measure of risk management that is captured in the crop diversity indicator. Each crop has a particular demand for water and a tolerance for variation in rainfall volume and timing. Thus, through crop diversification a farmer mitigates against the chance of complete production failure by planting crops with a variety of water demands and tolerances. The indicator is the sum of the number of different types of crops for a household. These include basic subsistence crops (corn and beans), other subsistence, horticulture, orchards, and other cash crops such as cotton.

Several human capital indicators were used in the original model, including education levels and dependency ratios (ratio of workers to nonworkers in a household). In the final model, however, the only relevant indicator is whether a household has access to health care. Health care is critical for maintaining livelihoods. When a working household member becomes ill, household income suffers either because of reduced hours or wages or due to the inability to work at all.

Although the results are consistent with a livelihoods approach to food security, there is one aspect of the findings that deserves particular attention: the climate and drought vulnerability literature increasingly focus on the relationship between poverty and vulnerability, which is proxied here with a food insecurity indicator. Although there is recognition that poverty is not synonymous with vulnerability to food insecurity, the relationship is not well understood (Eriksen and O'Brien 2007; Patt 2012). The results for this cross-sectional study confirm the role of financial capitals in assuring food security, but they also underscore that agricultural related features continue to be critical determinants. In terms of food security policy, they suggest that poverty reduction efforts are necessary but insufficient to reduce drought vulnerability and increase food security in the region.

Conclusions

The objective of this research was to evaluate the application of SVM algorithm to classify the food security of agricultural households. The model provided accuracy of 77% and a recall of 84% (Tables 3, 4, and 5). The features identified by the model are consistent with the theoretical framework for understanding food insecurity, which suggests that the SVM is an appropriate approach to utilize toward understanding the complexity of household-level

Table 3. Classification method using the 75 attributes.

			Confusion matrix		Accuracy	Recall
			prediction		73%	79%
SVM			−1	1		
	Actual	−1	131	73		
		1	55	216		

Table 4. Classification using the 14 most important attributes: Municipio, F_1, F_3, F_6, H_1, N_1, N_2, N_4, N_7, M_2, LS_2C, AS_2, AS_3, R_1.

			Confusion matrix		Accuracy	Recall
			prediction		75%	80%
SVM			−1	1		
	Actual	−1	141	63		
		1	54	217		

Table 5. Classification using the 11 most important attributes: Municipio, F_1, F_3, F_6, H_1, N_2, N_4, M_2, AS_2, AS_3, R_1.

			Confusion matrix		Accuracy	Recall
			prediction		77%	84%
SVM			−1	1		
	Actual	−1	137	67		
		1	43	228		

food insecurity. The type of information provided by SVM can contribute to the development of policies that consider the livelihoods of a population and are sensitive to local context.

Currently, SVM and other data mining techniques are underutilized in applications that make use of household livelihood data. There is an increasing availability of large datasets, some of which capture time series, which are collected to answer food security and other livelihoods and development questions. The use of data mining techniques makes it possible to gain new insights into the understanding of the nuances of food insecurity and other critical social questions. The increased use of these techniques, with questionnaires and other data collection instruments developed specifically for them, will provide refined models with higher accuracy and recall.

Funding

This work was done with the support of the National Science Foundation – Decision, Risk and Management Sciences – SES 1061966.

References

Aguiar, G. F. M., B. L. Batista, J. L. Rodrigues, L. R. S. Silva, A. D. Campiglia, R. M. Barbosa, and F. Barbosa Jr. 2012. Determination of trace elements in bovine semen samples by inductively coupled plasma mass spectrometry and data mining techniques for identification of bovine class. *Journal of Dairy Science* 95:7066–73.

Barbosa, R. M., B. L. Batista, R. M. Varrique, V. A. Coelho, A. D. Campiglia, and F. Barbosa. 2014a. The use of advanced chemometric techniques and trace element levels for controlling the authenticity of organic coffee. *Food Research International* 61:246–51.

Barbosa, R. M., L. R. Nacano, R. Freitas, B. L. Batista, and F. Barbosa. 2014b. The use of decision trees and naïve Bayes algorithms and trace element patterns for controlling the authenticity of free-range-pastured hens eggs. *Journal of Food Science* 79 (9):C1672–C1677.

Cheng, M., Y. Wu, and R. Syu. 2014. Seismic assessment of bridge diagnostics in Taiwan using the evolutionary support vector machine inferende model (ESIM). *Applied Artificial Intelligence* 28:449–69.

Coates, J., A. Swindale, and P. Bilinsky. 2007. *Household food insecurity access scale for measurement of food access: Indicator guide*. Washington, DC: Food and Nutrition Technical Assistance Project (FANTA).

de Souza, P. H. G. F., and R. G. Osorio. 2013. O Perfil da Pobreza no Brasil e suas Mudanças Entre 2003 e 2011. In *Programa Bolsa Família. Uma década de inclusão e cidadania*, ed. T. Campello and M. C. Neri, 139–56. Brasília, Brazil: IPEA.

Eriksen, S. H., and K. O'Brien. 2007. Vulnerability, poverty and the need for sustainable adaptation measures. *Climate Policy* 7 (4):337–52. doi:10.1080/14693062.2007.9685660.

FAO. 1996. *Rome declaration on world food security and world food summit plan of action*. Rome, Italy: World Food Summit.

FAO, World Food Program, & International Fund for Agricultural Development. 2012. *The state of food insecurity in the world 2012: Economic growth is necessary but not sufficient to accelerate reduction of hunger and malnutrition*. Rome, Italy: FAO.

Hall, M. A. 1998. Correlation-based feature subset selection for machine learning. PhD Thesis, University of Waikato.

Hamel, L. H. 2011. *Knowledge discovery with support vector machines*. Hoboken, NJ: Wiley-Interscience.

IBGE. 2010. *Segurança Alimentar 2004/2009*. Rio de Janeiro, Brazil: Programa das Nações Unidas para o Desenvolvimento.

IBM Corp. 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.

IPECE. 2012. *Anuário Estatístico do Ceará*. Fortaleza, Brazil: Governo do Ceará.

Isa, D., and R. Rajkumar. 2009. Pipeline defect prediction using support vector machines. *Applied Artificial Intelligence* 23:758–71. doi:10.1080/08839510903210589.

Kianmehr, K., and R. Alhajj. 2008. Effectiveness of support vector machine for crime hot-spots prediction. *Applied Artificial Intelligence* 22:433–58. doi:10.1080/08839510802028405.

Mashford, J., D. De Silva, S. Burn, and D. Marney. 2012. Leak detection in simulated water pipe networks using SVM. *Applied Artificial Intelligence* 26:429–44. doi:10.1080/08839514.2012.670974.

Patt, A. 2012. Multi-level climate adaptation policy and causation narratives. *Geografisk Tidsskrift-Danish Journal of Geography* 112 (2):174–82. doi:10.1080/00167223.2012.742967.

Rosegrant, M. W., and S. A. Cline. 2003. Global food security: Challenges and policies. *Science Communication* 302 (5652):1917–19. doi:10.1126/science.1092958.

Segall Corrêa, A. M., R. P. Escamilla, L. K. Maranha, and M. F. A. Sampaio. 2003. *(In) Seguraça Alimentar no Brasil: Validação de Metodologia para Acompanhamento e Avaliação*. Campinas, Brazil: UNICAMP.

Tan, P., M. Steinbach, and V. Kumar. 2006. *Introduction to data mining*. Boston, MA: Addison-Wesley.

United Nations. 2012. UN secretary general challenges all nations to achieve zero hunger. http://un-foodsecurity.org/sites/default/files/Zero_Hunger_Challenge_Press%20release.pdf.

Wheeler, T., and J. von Braun. 2013. Climate change impacts on global food security. *Science* 341 (6145):508–13. doi:10.1126/science.1239402.

Witten, I. H., E. Frank, and M. A. Hall. 2011. *Data mining: Practical machine learning tools and techniques*. Burlington, MA: Morgan Kaufmann.