Different Bayesian Network Models in the Classification of Remote Sensing Images

Cristina Solares * and Ana Maria Sanz

University of Castilla-La Mancha, Spain
Cristina.Solares@uclm.es
Ana.Sanz@uclm.es

Abstract. In this paper we study the application of Bayesian network models to classify multispectral and hyperspectral remote sensing images. Different models of Bayesian networks as: Naive Bayes (NB), Tree Augmented Naive Bayes (TAN) and General Bayesian Networks (GBN), are applied to the classification of hyperspectral data. In addition, several Bayesian multi-net models: TAN multi-net, GBN multi-net and the model developed by Gurwicz and Lerner, TAN-Based Bayesian Class-Matched multi-net (tBCM^2) (see [1]) are applied to the classification of multispectral data. A comparison of the results obtained with the different classifiers is done.

1 Introduction

Classification problems (see [2]) occur in a wide range of situations in real life such as disease diagnosis, image recognition, fault diagnosis, etc.

Probabilistic models, especially those associated with Bayesian networks, are very popular as a formalism for handling uncertainty. The increasing number of applications developed these last years show that this formalism has practical value also.

In this paper we apply different models of Bayesian networks to the classification of remote sensing images, considering multispectral and hyperspectral data sets. In a multispectral image the number of spectral bands for each pixel is less than 20, otherwise the image is called hyperspectral.

The paper is organized as follows. Section 2 introduces the Bayesian networks and the Bayesian networks as classifiers. Six models of Bayesian networks are introduced: General Bayesian network(GBN), Naive Bayes (NB), Tree Augmented Naive Bayes (TAN), TAN Bayesian multi-net, GBN Bayesian multi-net and the TAN-Based Bayesian Class-Matched multi-net (tBCM^2). Section 3 presents the application of the above models to the classification of remote sensing images. In Sect. 4 some conclusions are given.

* The author is indebted to the Spanish Ministry of Science and Technology (Project MTM2006-15671) and to the Junta de Comunidades de Castilla-La Mancha (Project PAI-05-044) for partial support.

2 Bayesian Networks

A Bayesian Network (BN) (see [3] for further details) over $\boldsymbol{X} = (X_1, \ldots, X_n)$ is a pair (D, P), where D is a directed acyclic graph with one node for each variable in $\boldsymbol{X}$ and $P = \{p_1(x_1|\pi_1), \ldots, p_n(x_n|\pi_n)\}$ is a set of n conditional probability distributions, one for each variable, given the values of the variables on its parent set Π_i (CP table). Each node in D represents a domain variable (eg, a dataset attribute) and each arc in D represents a probabilistic dependence between two variables quantified using the above CP table.

Here x_i and π_i denote realizations (instantiations) of X_i and Π_i, respectively. The joint probability distribution (JPD) of $\boldsymbol{X}$ can then be written as

$$p(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} p_i(x_i|\pi_i) \ . \tag{1}$$

2.1 Bayesian Network Classifiers

The application of Bayesian network models to classification involves two subtasks: Learning the BN structure (the graphical structure D) and the BN parameters (CP table). It is trivial to learn the parameters for a given structure, simply use the empirical conditional frequencies from the data (see [4]). Constructing the BN structure can be performed using expert knowledge or directly from the data. There are different methods of learning a BN structure, as the score-based methods (see [4]) and the methods that learn the structure by identifying the conditional independence relation-ships among the nodes (CI-based methods). The score-based methods incorporate a search procedure to find a network structure and a score is employed to evaluate each structure in the search space. The K2 algorithm, introduced in [4], is a search algorithm for finding a high quality Bayesian network in a reasonable time. An example of CI-based method is the algorithm described in Cheng et al. [5]. Cheng et al. in [6] show that the CI-based learning algorithms are very efficient and the learned BN classifiers can give very good prediction accuracy. Next we describe the different models of Bayesian network classifiers used in this paper.

General Bayesian Network (GBN). A GBN with JPD $p(a_1, a_2, \ldots, a_n, c)$ defined as in (1), can be constructed to solve a classification problem (see Fig. 1). The variables $\boldsymbol{A} = (A_1, \ldots, A_n)$ are the attributes of the problem and C is the class variable having k different states. The resulting model can be used to classify a given set of attribute values $\boldsymbol{a} = (a_1, \ldots, a_n)$ (see [7]). The vector $\mathbf{a}$ belongs to the class $c \in C$ that maximizes the posterior probability $p(c|\mathbf{a})$. The structure of the GBN can be learned using a score-based method as the K2 algorithm (see [4]) or a CI based method as the algorithm introduced in [5]. In this paper, we use the K2 search algorithm.

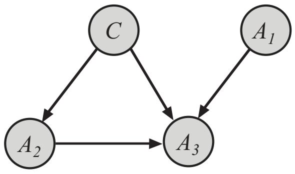

Fig. 1. Example of General Bayesian Network (GBN). C is the class variable and A_1, A_2, A_3 are the attribute variables.

Naive Bayes (NB). A NB is a simple structure of Bayesian network, the class node C is a parent of all other nodes (attributes) and there are not other connections between the nodes (see [7]).

Tree Augmented Naive Bayes (TAN). The very strong assumption of independence of all the attributes given its parents set in the Naive Bayes, is relaxed in this type of network. The TAN algorithm constructs a tree structure between the attribute nodes and after that adds a link from the class node C to the attribute nodes $A_i, i = 1, \ldots, n$ (see [7]). This model is based in the algorithm described by Chow et al. in [8], for learning tree-like Bayesian networks.

GBN Bayesian Multi-net. A GBN Bayesian multi-net is a generalization of the GBN, a different GBN is built for each class value and a set of networks is used as a classifier (see Fig. 2). For that, we partition the training data set by classes and for each class value we construct a GBN for the attribute variables.

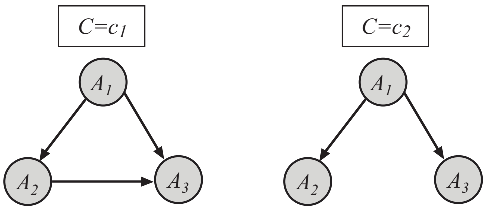

Fig. 2. Example of Bayesian Multi-net. C is the class variable that takes two values c_1 and c_2, and A_1, A_2, A_3 are the attribute variables.

TAN Bayesian Multi-net. In the TAN model, the relations between the attributes are the same for all the different instances of the class variable C. A Bayesian TAN multi-net is a generalization, a different TAN is built for each class value and a set of networks is used as a classifier (see [7]). This model allows the relations among the attributes to be different for the different values of the class. For that, we partition the training data set by classes and for each class value we construct a TAN for the attribute variables.

TAN-Based Bayesian Class-Matched Multi-net (tBCM^2). The tBCM^2 is a multi-net classifier that learns each local network (BN associated to each class value) using a detection-rejection measure (see [1]). The algorithm searches for the structure maximizing a discrimination-driven score that is calculated using training data for all the classes. The structure of each local network in tBCM^2 is based on the TAN model and it is learned using the SuperParent algorithm (see Keogh et al. in [9]). In Gurwicz et al.[1] the average superiority of the tBCM model in comparison with other classifiers, as the TAN multi-net, is shown.

In the next section the above models of Bayesian network classifiers are applied to the classification of remote sensing images.

3 Remote Sensing Image Classification

The models of Bayesian networks introduced in Sect. 2.1 can be applied to classify remote sensing spectral images. For the implementation of the proposed models, we use the Bayes Net toolbox in matlab (see [10]) and the BNT Structure Learning Package (see [11]).

A remote sensing spectral image consists of an array of multidimensional vectors assigned to particular spatial regions (pixel locations), reflecting the response of a spectral sensor at various wavelengths. Formally these images can be described as a matrix $\boldsymbol{V} \equiv (\boldsymbol{v}_{11}(x^1, y^1), \ldots, \boldsymbol{v}_{nm}(x^n, y^m))$ where $\boldsymbol{v}_{ij}(x^i, y^j) \in R^l, i = 1, \ldots, n, j = 1, \ldots, m$ is the vector of spectral information associated with pixel location (x^i, y^j) and the vector components $v_{ijk}(x^i, y^j), k = 1, \ldots, l$ reflects the responses of a spectral sensor at various wavelengths.

In this application all variables (class variable and attributes of the problem) are assumed to be discrete, that is, each variable has a finite set of possible values.

3.1 An Example of Multispectral Data Set Analysis

In the present contribution we consider a LANDSAT TM image from Sierra de Gredos (Spain). This image has been obtained from the GIS IDRISI 32 tutorial (http://www.clarklabs.org/). LANDSAT TM satellite-based sensors produce images of the Earth in different spectral bands. In this work six bands (bands 1-5 and band 7) are strategically determined for optimal detection and discrimination of water, soil and four different forest type, these are the class values for the classification problem. Band 6 is often dropped from analysis because of the lower spatial resolution. The spectral information, associated with each pixel of a LANDSAT scene is represented by a vector $\boldsymbol{v}(x, y) \in R^6$, these vectors are the attribute values of the problem. This is a classification problem with six attributes and six class values.

The GBN, NB and TAN models, have been previously applied by the authors, to the analysis of a multispectral data (see [12]). A GBN multi-net model also has been previously applied to this problem (see Ouyang et al. in [13]). In this paper, we apply the GBN multi-net, TAN multi-net and the tBCM^2 multi-net models to the classification of multispectral remote sensing images.

The above classification problem is analyzed using 5-fold cross-validation (CV5). We apply the different models of Bayesian multi-net classifiers to classify the multispectral image, the training and test accuracy obtained are shown in Table 1. A comparison shows negligible differences between the TAN and the GBN Bayesian multi-net models. All the models obtain almost 85% of accuracy with a slight advantage of the tBCM^2 multi-net.

Table 1. Training and test accuracy (mean (±std) in %) obtained with each classifier in Sect. 3.1

Classifier	Training	Test
TAN Multi-net	84.38(±0.11)	84.29(±0.40)
GBN Multi-net	84.84(±0.19)	84.47(±0.68)
tBCM^2 Multi-net	85.50(±0.12)	85.20(±0.45)

3.2 An Example of Hyperspectral Data Set Analysis

For some years, the above application has been limited to data of low dimensionality, less than 10 bands (multispectral data). Recent advances in sensor technology make possible to work with several hundred bands (hyperspectral data). In this paper, we do the novel application of the NB, GBN and TAN models to the classification of hyperspectral data. The hyperspectral data used in our experiments is a section of a scene taken over northwest Indiana's Pines by the AVIRIS sensor in 1992 (ftp://ftp.ecn.purdue.edu/biehl/MultiSpec/). The AVIRIS sensor collects 224 bands of data but four of these bands contain only zeros and consequently they are eliminated. The initial 220 bands are reduced to 200 because the bands covering the region of water absorption: $[104 - 108], [150 - 163], 220$ are removed. In this work, 200 bands are considered for optimal detection and discrimination of 9 different classes: Corn-no till, Corn-min till, Grass/Pasture, Grass/Trees, Hay-windrowed, Soybean-no till, Soybean-min till, Soybean-clean till and Woods. From the initial 16 land-cover classes, seven were eliminated, since only few training samples were available for them. The above is a classification problem with 200 attributes and 9 class values.

We analyze the effectiveness of Bayesian networks in classifying hyperspectral images directly in the original hyperdimensional attribute space. The problem is studied using 5-fold cross-validation (CV5). We apply the different models of Bayesian network classifiers (NB, TAN and GBN) to the above classification problem, the training and test accuracy obtained are shown in Table 2. A comparison shows slight differences between the TAN and GBN Bayesian network models, both are superior on accuracy to the NB model. The very strong assumption of independence of all the attributes given its parents set in the NB model is not realistic in the case of study.

Table 2. Training and test accuracy (mean (±std) in %) obtained with each classifier in Sect. 3.2

Classifier	Training	Test
NB	58(±0.24)	58(±0.20)
TAN	88(±0.23)	80(±0.58)
GBN	84(±0.34)	80(±0.80)

4 Conclusions

Bayesian networks appear as powerful tools in hyperspectral remote sensing image classification. Different models of Bayesian networks as: Naive Bayes (NB), Tree Augmented Naive Bayes (TAN) and General Bayesian Network (GBN), have been applied to the classification of an hyperspectral image. In addition, several Bayesian multi-net models: TAN multi-net, GBN multi-net and the model developed by Gurwicz and Lerner, TAN-Based Class-Matched multi-net (tBCM^2) are applied to the classification of multispectral data. Feature (attribute) selection is an important task in remote sensing data processing, particularly in case of hyperspectral images. Actually, we are studying the application of Bayesian network models to the classification of hyperspectral data, combined with a band selection method to reduce the dimensionality of the feature space.

References

1. Gurwicz, Y., Lerner, B.: Bayesian class-matched multinet classifier. Lecture Notes in Computer Science **4109** (2006) 145–153
2. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification. John Wiley & Sons, New York (2001)
3. Castillo, C., Gutiérrez, J.M., Hadi, A.S.: Expert Systems and Probabilistic Network Models. Springer-Verlag, New York (1997)
4. Cooper, G.F., Herskovitz, E.: A Bayesian method for the induction of probabilistic networks from data. Machine Learning **9** (1992) 309–347
5. Cheng, J., Bell, D.A., Liu, W.: An algorithm for Bayesian belief network construction from data. In Proc. AI & STAT'97 (1997) 83–90
6. Cheng, J., Greiner, R.: Learning Bayesian belief network classifiers: Algorithms and system. In Proc. 14th Canadian Conf. on Artificial Intelligence (2001) 141–151

7. Friedman, N., Geiger, D., Goldszmidt, M.: Bayesian network classifiers. Machine Learning **29** (1997) 131–163
8. Chow, C., Liu, C.: Approximating discrete probability distributions with dependence trees. IEEE Trans. Inf. Theory **14**(3) (1968) 462–467
9. Keogh, E.J., Pazzani, M.J.: Learning the structure of augmented Bayesian classifiers. Int. J. on Artificial Ingelligence Tools **11**(4) (2002) 587–601
10. Murphy, K.P.: The Bayes Net Toolbox for matlab. Computing Science and Statistics **33** (2001)
11. Leray, P., Francois, O.: BNT, Structure learning package: documentation and experiments. Technical Report. Laboratoire PSI-INSA Rouen-FRE CNRS 2645 (2004)
12. Solares, C., Sanz, A.M.: Bayesian network classifiers. Some engineering applications. In Proc. 9th IASTED Int. Conf. Artificial Intelligence and Soft Computing (2005) 331–335
13. Ouyang, Y., Ma, J., Dai, Q.: Bayesian multinet classifier for classification of remote sensing data. Int. J. of Remote Sensing **27** (2006) 4943–4961