Exploring sensor temperature relationships for located frost prediction

Ana Laura Diedrichs [1,3], Facundo Bromberg[1,2]

[1]Laboratorio Desarrollo de Herramientas de Aprendizaje de Máquinas
Dpto. Ingeniería en Sistemas de Información
Universidad Tecnológica Nacional - Facultad Regional Mendoza
[2]Ph.D. advisor [3]ana.diedrichs at frm.utn.edu.ar

Introduction

The damage caused by the frost takes place when the temperatures are below than a tolerable limit for the plants. Each phenological state, e.g flowering, has a variable cold hardiness [12], so the lethal temperature is also variable. Freezing climatic events are the most dangerous, because they affect a large land surface. Mendoza is not an exception. According to the Instituto Nacional de Vitivinicultura (INV), in 2013 the loss of the vine crop reached up to 27% [2]. Big part of that loss of yield was during the early spring. In order to study the micro-climate phenomenon of frost in Mendoza, sensors should be distributed in the vineyards vertically as well horizontally, because the air temperatures change in both directions, and the plant has also different cold hardiness in the organs like trunk, flowers, shoots. Previous works on frost prediction have worked with data taken from meteorological stations very distant between them [6][4][7] or using wireless sensor networks (WSN) [9]. All of them have used supervised machine learning algorithms, such as artificial neural networks and support vector machines, with an particular configuration. For a better understanding of the phenomenon, we propose a study the sensor relationships in order to improve the frost prediction. We are exploring the variables relationships using the independence approach by learning Markov Network structures [5] from the environmental data for later to corroborate with the opinion of an expert. The analysis of the Markov blanket of particular sensors helps to identify which neighbor sensors could improve the prediction. Hypothesis: It is possible to improve the temperature sensor prediction taking advantage of the sensor neighbors information. Given the temperature prediction of a place S_i, we are asking if it could improve with the information given by a sensor neighbor S_j.

Markov network approach

In order to analyze how related is a sensor respect others, we are going to use the probabilistic independence approach. A common way of probability distribution representations is conditional independence. Two events α and β are independent iff $P(\alpha|\beta) = P(\alpha)$. It means if you know that β occurs, it does not change the probability of occurence of α. Markov networs are undirected graphical models which represent a joint probability distribution over the variables. Each node of the graph is a random variable of the domain and each edge between nodes represents a dependency conditional relationship between both variables. Given a Markov network it is feasible to find the conditional independence over the variables respect others. It is possible to learn Markov network from data using structure learning algorithms. We have chosen independence structure learning approach [11] for knowledge discovering.

Entropy approach

Entropy measures the amount of information (or uncertainty) of a random variable. Given X, a discreted random variable and P(x) its probability distribution, the entropy H(X) is calculated as (1)

$$H(X) = \sum_{x \in X} P(x) \log P(x) \qquad (1)$$

If the log base 2 is used, the unit is the bit. By definition $H(X) \geq 0$. The joint entropy represents the amount of information needed on average to specify the value of two discrete random variables X and Y which is given by (2)

$$H(X,Y) = \sum_{x \in X} \sum_{y \in Y} P(x,y) \log P(x,y) \qquad (2)$$

The conditional entropy indicates how much extra information you still need to supply on average to communicate Y given that the other party knows X. Conditional entropy $H(X|Y)$ measures the amount of uncertainty in X after we know the value of Y (on average): $H(Y|X) = H(X,Y) - H(X)$

$$H(Y|X) = \sum_{x \in X} \sum_{y \in Y} P(x,y) \log P(y|x) \qquad (3)$$

Mutual information measures the information that X and Y share: it measures how much knowing one of these variables reduces uncertainty about the other.

$$I(X,Y) = \sum_{x \in X} \sum_{y \in Y} P(x,y) \frac{\log P(x,y)}{P(x)P(y)} \qquad (4)$$

where p(x,y) is the joint probability distribution function of X and Y, and p(x) and p(y) are the marginal probability distribution functions of X and Y respectively. It is the marginal additional information someone, analyzing X , gains from knowing Y. Moreover, $I(X,Y)$ is symetric and non-negative. $I(X,Y) = I(Y,X)$. If $I(X,Y) = 0$, then X and Y are independent.
We analyze the conditional entropy of the sensor and their relationships in a present time t and a future time T in order to answer:

- How much reduces the entropy of $H(S_{iT}|S_{it})$ which indicates how much information gives S_{iT} given that we have S_{it}
- How much information gives a neighbor S_{jt}, calculating $H(S_{iT}|S_{it}, S_{jt})$

Experimental setup

Six temperature sensor [1] from a research project [8] where placed at different heights, as we can see on Table 1. The sampling interval was set to a minute. They represent six temperature variables located in the same (x, y) position varing their height. The data was acquiered between 07/21/2012 and 07/26/2012, a week with frost events. For each sensor there are 7330 datapoints.
For analyzing the relationship between the variables S_{it} in $t = t_0$, where $i = 1..N$ and N is the total number of sensors, versus the prediction of S_{iT} in $t = T$, a dataset was build whose columns concatenate the variables as follow: one column per each S_{it} and then one column per each S_{iT}.

Temperature sensor	Sensor at time t	Sensor at t + T
ground level	S_1	S_7
40 cm height	S_2	S_8
75cm height	S_3	S_9
1.5m height	S_4	S_{10}
2m height	S_5	S_{11}
3m height	S_6	S_{12}

Table 1: Variables of interest

The setup of structure learning is listed below on Table 2. The first column is the algorithm used ("SL slgorithm"), the second the test used to calculated if two variables are independent, and the third is the threshold used for the test.

SL algorithm	Test	Threshold
HHC-MN [10]	G^2	0.001, 0.01, 0.05
GSMN [3]	G^2	0.001, 0.01, 0.05
PC [13]	G^2	0.001, 0.01, 0.05
IBMAP-HC [10]	bayesian test	0.5

Table 2: Setup for structure learning

Result and discussions

Table 1 presents how the entropy is decreced thanks to have the information of the neighbor. The output are nats (no bits), because base e was used to compute entropy. Each cell (i,j) of the matrix stores $H(S_{iT}|S_{it}, S_{jt})$. For example the row (S_2) and column S_3 stores the result of calculate $H(S_{2T}|S_{2t}, S_{3t})$. Mutual information results $I(X, Z|Y)$, with $X = Si_T, Y = Si_t, Z = Sj_t$, on Table 1 indicate how much knowing one of these variables reduces uncertainty about the other. We can see the mutual dependence because $H(S_{iT}|S_{it}, S_{jt}) \neq 0$.

	S_1	S_2	S_3	S_4	S_5	S_6
S_1	1.65	1.50	1.45	1.49	1.46	1.48
S_2	1.41	1.56	1.50	1.48	1.43	1.49
S_3	1.34	1.47	1.52	1.47	1.43	1.46
S_4	1.33	1.43	1.45	1.50	1.43	1.43
S_5	1.32	1.39	1.42	1.44	1.49	1.40
S_6	1.32	1.42	1.42	1.41	1.38	1.48

Table 3

	S_1	S_2	S_3	S_4	S_5	S_6
S_1	0.00	0.15	0.19	0.16	0.18	0.17
S_2	0.16	0.00	0.06	0.08	0.14	0.07
S_3	0.18	0.05	0.00	0.05	0.09	0.05
S_4	0.17	0.07	0.05	0.00	0.07	0.07
S_5	0.17	0.11	0.07	0.05	0.00	0.09
S_6	0.16	0.06	0.06	0.07	0.10	0.00

Table 4

The Figure 1 shows the result of counting all edges between all the adjacency matrix obtained from the structured learning algorithms with the following setup: threshold 0.001, except IBMAP-HC with 0.5, and T = 3 (12 hs). The columns and rows show the variable number, for example 1 is S_1, which are listed on Table 1. We can see that still with a big confidence interval, dependencies remains between the variables. The relationships between one sensor with its next-door neighbor, which is located over or below itself, reflect the typical thermodinamic behavior because one layer of air is related with other. Edges between S_1 and S_5 explain where the inversion layer is located. The inversion layer is the division between the cool air and heat air.

	1	2	3	4	5	6	7	8	9	10	11	12
1	0	2	1	1	4	1	1	2	0	1	0	1
2	2	0	4	1	1	2	1	0	0	0	0	2
3	1	4	0	4	1	1	1	0	1	0	0	0
4	1	1	4	0	4	1	0	0	0	1	0	1
5	4	1	1	4	0	4	0	0	0	0	0	1
6	1	2	1	1	4	0	2	1	2	0	1	2
7	1	1	1	0	0	2	0	3	1	0	3	2
8	2	0	0	0	0	1	3	0	3	2	0	2
9	0	0	1	0	0	2	1	3	0	3	0	1
10	1	0	0	1	0	0	0	2	3	0	3	1
11	0	0	0	0	0	1	3	0	0	3	0	3
12	1	2	0	1	1	2	2	2	1	1	3	0

Issues

- Complete interpretation with an expert of the markov networks: could they help to understand the frost phenomenon?
- Setup of the prediction model: input and output variables
- Discretization of continuous features: sensor values and time
- How to infer the best sensor location in order to optimize the prediction and the number of sensor needed.

References

[1] Logtag: humidity and temperature recorder. datasheet. Online. http://logtagrecorders.com/products/pdf/LogTag_HAXO-8_Product_Specification.pdf.

[2] El INV presentó el ajuste de la Estimación de Cosecha 2014. online news, Instituto Nacional de la Vitivinicultura, Mendoza, Argentina, February 2014.

[3] Facundo Bromberg, Dimitris Margaritis, Vasant Honavar, et al. Efficient markov network structure discovery using independence tests. *Journal of Artificial Intelligence Research*, 35(2):449, 2009.

[4] Luca Ghielmi and Emanuele Eccel. Descriptive models and artificial neural networks for spring frost prediction in an agricultural mountain area. *Computers and electronics in agriculture*, 54(2):101–114, 2006.

[5] Daphne Koller and Nir Friedman. *Probabilistic graphical models: principles and techniques.* MIT press, 2009.

[6] Brenda B Lin. Agroforestry management as an adaptive strategy against potential microclimate extremes in coffee agriculture. *Agricultural and Forest Meteorology*, 144(1):85–94, 2007.

[7] Imran Maqsood, Muhammad Riaz Khan, and Ajith Abraham. An ensemble of neural networks for weather forecasting. *Neural Computing & Applications*, 13(2):112–122, 2004.

[8] Remugnán Daniel (Co-director) Rojas Daniel (Director). Defensa activa contra heladas. Utilización de energía geotérmica de baja temperatura. Proyecto de investigación, Facultad de Ciencias Agrarias, Universidad Nacional de Cuyo.

[9] P Sallis, M Jarur, M Trujillo, and A Ghobakhlou. Frost prediction using a combinational model of supervised and unsupervised neural networks for crop management in vineyards. In *18th IMACS World Congress-MODSIM09 International Congress on Modelling and Simulation*, pages 13–17, 2009.

[10] F. Schlüter, F. Bromberg, and A. Edera. The IBMAP approach for Markov network structure learning. *Annals of Mathematics and Artificial Intelligence*, pages 1–27, 2014.

[11] Federico Schlüter. A survey on independence-based markov networks learning. *Artificial Intelligence Review*, pages 1–25, 2012.

[12] Richard L Snyder and J Paulo Melo-Abreu. *Frost protection: fundamentals, practice and economics. Volume 1.* FAO, 2005.

[13] Peter Spirtes, Clark Glymour, and Richard Scheines. Causation, prediction, and search. adaptive computation and machine learning, 2000.