

Prediction of frost location using machine learning and wireless sensor networks: exploring sensor relationships

Ana Laura Diedrichs

Laboratorio DHARMa, Dpto Sistemas,
Facultad Regional Mendoza, Universidad Tecnológica Nacional,
Rodríguez 273, Ciudad de Mendoza, Argentina, 5500
{ana.diedrichs,facundo.bromberg}@frm.utn.edu.ar
<http://dharma.frm.utn.edu.ar>

Abstract. Las heladas son un evento meteorológico que ocasiona grandes pérdidas a la producción. Se caracterizan por descenso de la temperatura a niveles que pueden dañar los cultivos. La topografía del terreno es un factor importante que ocasiona que el fenómeno no afecte de la misma manera a toda la finca. Dado que el aire frío es más denso, el mismo se estratifica a las zonas más bajas. Por ello es importante evaluar la variación vertical de las temperaturas. En el presente trabajo realizaremos el análisis de seis sensores posicionados verticalmente para evaluar la variación de temperaturas en época de heladas. Los resultados nos muestran que existen correlaciones entre los sensores vecinos y es posible predecir el comportamiento de uno a partir del otro.

1 Introducción a la problemática de las heladas

El daño ocasionado por heladas toma lugar cuando las temperaturas se encuentran bajo un límite tolerable por las plantas, ya que las mismas presentan distinto grado de resistencia al frío según el estado fenológico en el que se encuentren [?]; por lo que la temperatura de daño es variable. Los eventos de heladas son muy dañinos afectando a grandes superficies. Mendoza no es una excepción. De acuerdo al Instituto Nacional de Vitivinicultura (INV), en 2013 la pérdida de viñedo alcanzó un 27%[?], ocurriendo gran parte de la pérdida durante la primavera temprana. Con el objeto de estudiar el fenómeno microclimático de las heladas, los sensores de temperatura deben ser distribuidos verticalmente y horizontalmente porque la temperatura del aire varía en ambas direcciones y la planta tiene diferentes umbrales de resistencia al frío en sus órganos (tronco, flores, yemas)

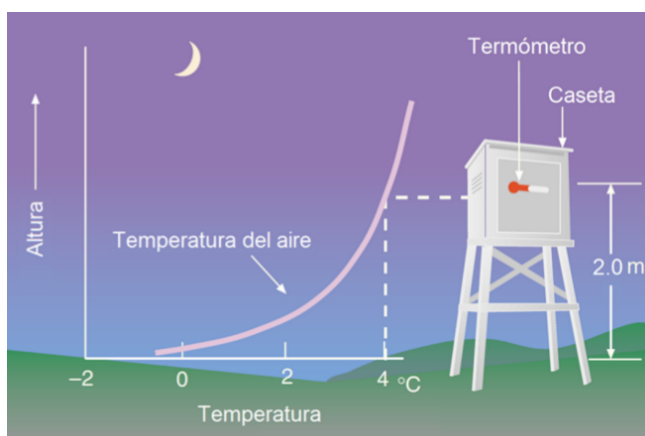
2 Hipótesis del Trabajo

Existe la posibilidad de desarrollar mecanismos más efectivos respecto al estado del arte actual para la predicción localizada de heladas en los cultivos, innovando

28 en la técnica predictiva mediante un estudio de la distribución espacio-temporal
29 de las variables involucradas y el análisis de sus relaciones *colaborativas entre*
30 *sensores vecinos* para: anticipar la predicción de la helada y *localizar la zona*
31 *de ocurrencia* haciendo uso de tecnología de sensado inalámbrica y aprendizaje
32 automático (machine learning).

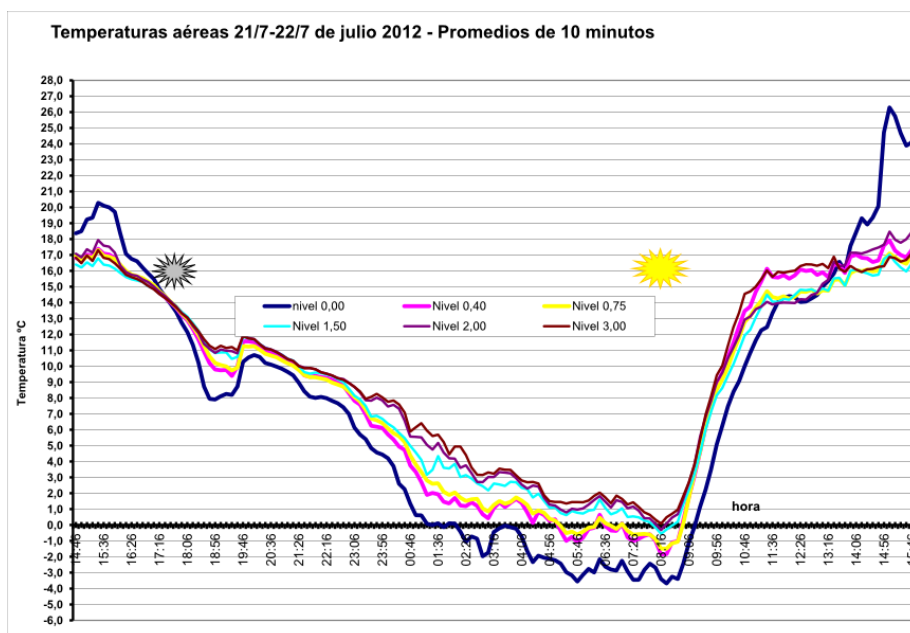
33 3 Problemática de las Heladas

34 Las heladas constituyen uno de los accidentes de tiempo que causan grandes
35 pérdidas económicas, a la agricultura en la Argentina y gran parte del mundo,
36 e impacto social al verse afectado los cultivos de los productores, debido a que
37 no son fenómenos locales sino extensivos. Existen varias definiciones de helada
38 como considerar helada a las temperaturas mínimas menores a 0°C . La más
39 apropiada desde el punto de vista agronómico es *considerar a la helada como*
40 *el evento meteorológico que ocurre cuando los cultivos y otras plantas experi-*
41 *mentan daño por congelación*. El daño causado por heladas ocurre cuando las
42 temperaturas están debajo de un límite tolerable para los cultivos. El umbral de
43 resistencia de las plantas al frío varía de acuerdo al estado fenológico en el que se
44 encuentren (floración, frutos o yemas presentes, etc) Según el Instituto Nacional
45 de Vitivinicultura (INV) en el 2013 la pérdida de viñedo por helada llegó a un
46 27% [?]. Las heladas tardías en Mendoza suceden entre septiembre y noviembre
47 siendo muy peligrosas porque empieza la floración y ya hay yemas brotadas.



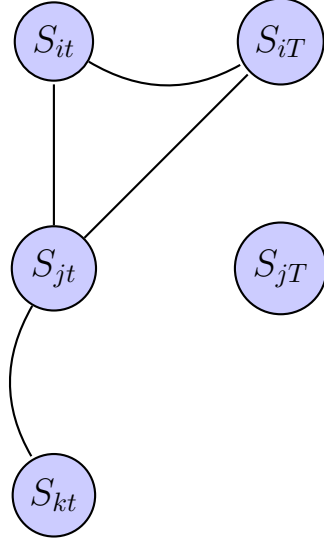
48 Existen dos tipos de heladas. Las heladas advectivas se caracterizan por los al-
49 tos niveles de humedad, escarcha y cielo nublados. Las heladas radiativas suceden
50 bajo cielo despejado, escaso viento y muy baja humedad. Estas últimas son con-
51 sideradas muy peligrosas porque el balance calórico disminuye drásticamente en
52 la noche despejada al perderse el calor recibido durante el día. Para la medición

de variables ambientales comúnmente se utilizan estaciones meteorológicas. Actualmente las mismas están instaladas muy distanciadas unas de las otras, por lo que no son suficientes para caracterizar un fenómeno micro-climático. A esto se suma que la casilla meteorológica está a una altura promedio entre los 1.5m y 2m, como lo ilustra la figura[?] anterior; impidiendo caracterizar el fenómeno de la inversión térmica. En el siguiente gráfico [?] se muestran las temperaturas durante una noche de heladas de seis sensores posicionados verticalmente (a nivel del suelo, 40cm, 75 cm, 1.5m, 2 m, 3 m) donde podemos visualizar la variabilidad de la amplitud térmica a distintas alturas. Esto sucede porque el aire frío es más denso y fluye hacia las capas más bajas, en consecuencia drena hacia la parte más baja del terreno; mientras que el aire más caliente queda estratificado en la altura. Por esto es importante medir la variable de interés (temperatura) in situ y a distintas alturas. La predicción de heladas es importante porque permite activar con tiempo los mecanismos de prevención activa. Durante las noches de heladas suelen utilizarse quemadores, molinos de viento, aspersores, entre otras técnicas que permiten generar o mejorar la circulación del calor donde se encuentran los cultivos. Por otra parte, la predicción localizada de heladas permitiría saber no sólo si helará o no ese día, sino también las zonas de una finca o región que se verían afectadas.



72 4 Hyphotesis

73 Hyphotesis: It is possible to improve the temperature sensor prediction taking
74 advantage of the sensor neighbors information. Given the temperature predic-
75 tion of a place S_i , we are asking if it could improve with the information given
76 by a sensor neighbor S_j .



77

78 5 Markov networks approach

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

91 6 Mutual Information approach

92 Entropy measures the amount of information (or uncertainty) of a random vari-
93 able. Given X , a discretized random variable and $P(x)$ its probability distribution,
94 the entropy $H(X)$ is calculated as (1)

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

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

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

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

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

102 Mutual information measures the information that X and Y share: it mea-
 103 sures how much knowing one of these variables reduces uncertainty about the
 104 other.

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

105 where $p(x, y)$ is the joint probability distribution function of X and Y, and
 106 $p(x)$ and $p(y)$ are the marginal probability distribution functions of X and
 107 Y respectively. It is the marginal additional information someone, analyzing
 108 X, gains from knowing Y. Moreover, $I(X, Y)$ is symmetric and non-negative.
 109 $I(X, Y) = I(Y, X)$. If $I(X, Y) = 0$, then X and Y are independent.

110 We analyze the conditional entropy of the sensor and their relationships in a
 111 present time t and a future time T in order to answer:

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

115 7 Experimental setup

116 Se procedera a explicar las características del origen de los datos en 7.1 y su
 117 posterior procesamiento en 7.2 (discretización, consideración temporal) para dar
 118 origen a otros datasets, que sirvan como entrada para el aprendizaje de estruc-
 119 tura de markov y para el análisis de mutual info. En 7.2 se presenta la config-
 120 uración utilizada para el aprendizaje de estructuras. Por último se desglosa los
 121 resultados de mutual info.

7.1 Data source

Six temperature sensor[?] from a research project [?] where placed at different heights, as we can see on table 7.1. The sampling interval was set to a minute. They represent six temperature variables located in the same (x, y) position varying their height. The data was acquired between 07/21/2012 and 07/26/2012, a week with frost events. For each sensor there are 7330 datapoints.

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

7.2 Pre-processing

The continuous temperature data is ranging from a minimum -9.8°C to 26.8°C . It was discretized with eight intervals with the same width.

For the temporal discretization, three options of time displacement interval were chosen: 6, 8 and 12 hours; which were labeled with T_1, T_2 y T_3 respectively.

Dataset design Para tomar en cuenta el valor de la variable S_i en un tiempo T posterior, se diseño dos tipos de datasets que pasaré a explicar.

ESTE NO VA A L POSTER Para analizar la relación de todas las variables S_{it} en $t = t_0$, versus la predicción de alguna de ellas S_{kT} , $i \neq k$, en un $t = T$, donde $i = 1..N$ y N el número total de sensores. Se construyó un dataset cuyas columnas concatenan las variables de la siguiente manera: Una columna para cada S_{it} (total N columnas) y una columna S_{kT} .

- 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 built whose columns concatenate the variables as follow: one column per each S_{it} and then one column per each S_{iT} .

El número total de datasets resultantes está definido por $(N + 1) * T$. Dado que tomaremos 3 períodos futuros, $T \in (6, 8, 12)$, y $N = 6$ para nuestro caso, contamos con 21 datasets generados.

148 **Aprendizaje de estructuras: setup** The setup of structured learning is listed
 149 below on Table 7.2. The first column is the algorithm used ("SL algorithm"),
 150 the second the test used to calculated if two variables are independent, and the
 151 third is the threshold used for the test.

152 **Agregar una intro a cada uno de ellos**

Table 2. Setup para aprendizaje de estructuras

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

153 7.3 Resultados Markov networks

154 Las estructuras resultantes para los algoritmos pc, gsmn, hhc se encuentran en
 155 scripts/structured-learning/output/structures, y para ibmaph en scripts/structured-
 156 learning/output-ibmaph/structures. Los archivos que guardan las estructuras
 157 son extensiones .ugraph. La nomenclatura a utilizar es la indicada en la Tabla
 158 7.1

159 The Figure 7.3 shows the result of apply OR-operand (counting all edges)
 160 between all the Adjacency matrix obtained from the structured learning algo-
 161 rithms with the following setup: threshold 0.001, except IBMAPH with 0.5, and
 162 $T = 3$ (12 hs). The columns and rows show the variable number, for example 1
 163 is S_1 , which are listed on Table 7.1. We can see that still with a big confidence
 164 interval remains dependencies between the variables. The relationships between
 165 one sensor with its next-door neighbor, which is located over or below itself, re-
 166 flects the typical thermodynamic behavior because one layer of air is related with
 167 other. Edges between S_1 and S_5 explain where the inversion layer is located. The
 168 inversion layer is the division between the cool air and heat air.

169 En el siguiente enlace de gdrive contiene las tablas como la de arriba https://docs.google.com/spreadsheets/d/16cIB4ZQ2kGtz6V0BqyA9paCYtCYshQuPt7_hV0eqDxs/edit?usp=sharing
 170
 171

172 8 Resultados mutual information

173 Los cálculos de mutual information se encuentran implementados en el archivo
 174 mutual.information.R, método run.experiment.cond.entropy. Se procedió a cal-
 175 cular los siguiente:

	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

Fig. 1. —

1) La entropía de S_{iT} dado S_{it} se refleja como $H(S_{iT}|S_{it})$, que significa cuanta información extra requiere S_{iT} dado que tiene/sabe S_{it} . ¿En cuánto me ayuda a disminuir la incertidumbre de S_{iT} saber S_{it} ?

2) Por otra parte, nos interesa saber cuanta información me otorga agregar a un vecino S_{jt} , y se refleja en la entropía $H(S_{iT}|S_{it}, S_{jt})$

Los cálculos son en base e, en consecuencia el resultado son nats (no bits).

En la vector de resultados de $H(S_{iT}|S_{it})$ cada posición i guarda el valor de uno de los cálculos. En la matriz de resultados de $H(S_{iT}|S_{it}, S_{jt})$, un resultado en la posición (i,j) siendo i la fila y j la columna; por ej, la fila de s_2 y columna s_{3t} guarda el resultado de calcular $H(S_{2T}|S_{2t}, S_{3t})$.

Each cell (i,j) , where i is the row and j the column, of the matrix stores one result of $H(S_{iT}|S_{it}, S_{jt})$. For example the row 2 (S_2) and column S_3 stores the result of calculate $H(S_{2T}|S_{2t}, S_{3t})$.

```

> run.experiment.cond.entropy()
[1] "/home/usuario/phd/heladas/frost/heladas/scripts/datasets/T_1_b_8.csv"
CALCULO DE H(X)  X=S_T
value
s1  8.69
s2  8.74
s3  8.76
s4  8.77
s5  8.77

```



```

198 s6 8.78
199 CALCULO DE H(X | Y) donde: X=S_T Y=S_t
200 value
201 s1 1.65
202 s2 1.56
203 s3 1.52
204 s4 1.50
205 s5 1.49
206 s6 1.48
207 CALCULO DE H(X | Y,Z) donde: X=Si_T Y=Si_t Z=Sj_t
208 s1 s2 s3 s4 s5 s6
209 s1 1.65 1.50 1.45 1.49 1.46 1.48
210 s2 1.41 1.56 1.50 1.48 1.43 1.49
211 s3 1.34 1.47 1.52 1.47 1.43 1.46
212 s4 1.33 1.43 1.45 1.50 1.43 1.43
213 s5 1.32 1.39 1.42 1.44 1.49 1.40
214 s6 1.32 1.42 1.42 1.41 1.38 1.48
215 MUTUAL INFO I(X,Z|Y) donde: X=Si_T Y=Si_t Z=Sj_t
216 s1 s2 s3 s4 s5 s6
217 s1 0.00 0.15 0.19 0.16 0.18 0.17
218 s2 0.16 0.00 0.06 0.08 0.14 0.07
219 s3 0.18 0.05 0.00 0.05 0.09 0.05
220 s4 0.17 0.07 0.05 0.00 0.07 0.07
221 s5 0.17 0.11 0.07 0.05 0.00 0.09
222 s6 0.16 0.06 0.06 0.07 0.10 0.00
223 METRICA razon o division del 3ro con el 2do
224 s1 s2 s3 s4 s5 s6
225 s1 1.00 0.91 0.88 0.90 0.89 0.90
226 s2 0.90 1.00 0.96 0.95 0.91 0.95
227 s3 0.88 0.97 1.00 0.97 0.94 0.96
228 s4 0.88 0.95 0.97 1.00 0.95 0.95
229 s5 0.88 0.93 0.95 0.96 1.00 0.94
230 s6 0.89 0.96 0.96 0.96 0.93 1.00
231 porcentaje de reducci3n de la entropía
232 s1 s2 s3 s4 s5 s6
233 s1 0.00 0.09 0.12 0.10 0.11 0.10
234 s2 0.10 0.00 0.04 0.05 0.09 0.05
235 s3 0.12 0.03 0.00 0.03 0.06 0.04
236 s4 0.12 0.05 0.03 0.00 0.05 0.05
237 s5 0.12 0.07 0.05 0.04 0.00 0.06
238 s6 0.11 0.04 0.04 0.04 0.07 0.00
239 [1] "/home/usuario/phd/heladas/frost/heladas/scripts/datasets/T_2_b_8.csv"
240 CALCULO DE H(X) X=S_T
241 value
242 s1 8.67

```

```

243 s2 8.72
244 s3 8.74
245 s4 8.75
246 s5 8.75
247 s6 8.76
248 CALCULO DE H(X | Y) donde: X=S_T Y=S_t
249 value
250 s1 1.58
251 s2 1.45
252 s3 1.43
253 s4 1.35
254 s5 1.40
255 s6 1.36
256 CALCULO DE H(X | Y,Z) donde: X=Si_T Y=Si_t Z=Sj_t
257 s1 s2 s3 s4 s5 s6
258 s1 1.58 1.41 1.38 1.40 1.39 1.39
259 s2 1.36 1.45 1.39 1.38 1.37 1.39
260 s3 1.30 1.37 1.43 1.34 1.34 1.34
261 s4 1.25 1.29 1.29 1.35 1.30 1.29
262 s5 1.26 1.29 1.30 1.33 1.40 1.33
263 s6 1.23 1.29 1.28 1.29 1.29 1.36
264 MUTUAL INFO I(X,Z|Y) donde: X=Si_T Y=Si_t Z=Sj_t
265 s1 s2 s3 s4 s5 s6
266 s1 0.00 0.17 0.20 0.18 0.19 0.19
267 s2 0.09 0.00 0.06 0.08 0.08 0.06
268 s3 0.12 0.06 0.00 0.08 0.09 0.08
269 s4 0.10 0.06 0.06 0.00 0.05 0.07
270 s5 0.14 0.11 0.10 0.07 0.00 0.07
271 s6 0.13 0.07 0.07 0.06 0.07 0.00
272 METRICA razon o division del 3ro con el 2do
273 s1 s2 s3 s4 s5 s6
274 s1 1.00 0.89 0.87 0.89 0.88 0.88
275 s2 0.94 1.00 0.96 0.95 0.94 0.96
276 s3 0.91 0.96 1.00 0.94 0.94 0.94
277 s4 0.92 0.95 0.96 1.00 0.96 0.95
278 s5 0.90 0.92 0.93 0.95 1.00 0.95
279 s6 0.91 0.95 0.95 0.95 0.95 1.00
280 porcentaje de reducci3n de la entropía
281 s1 s2 s3 s4 s5 s6
282 s1 0.00 0.11 0.13 0.11 0.12 0.12
283 s2 0.06 0.00 0.04 0.05 0.06 0.04
284 s3 0.09 0.04 0.00 0.06 0.06 0.06
285 s4 0.08 0.05 0.04 0.00 0.04 0.05
286 s5 0.10 0.08 0.07 0.05 0.00 0.05
287 s6 0.09 0.05 0.05 0.05 0.05 0.00

```

```

288 [1] "/home/usuario/phd/heladas/frost/heladas/scripts/datasets/T_3_b_8.csv"
289 CALCULO DE H(X) X=S_T
290     value
291 s1  8.63
292 s2  8.69
293 s3  8.70
294 s4  8.71
295 s5  8.72
296 s6  8.73
297 CALCULO DE H(X | Y) donde: X=S_T Y=S_t
298     value
299 s1  1.40
300 s2  1.37
301 s3  1.31
302 s4  1.33
303 s5  1.32
304 s6  1.24
305 CALCULO DE H(X | Y,Z) donde: X=Si_T Y=Si_t Z=Sj_t
306     s1 s2 s3 s4 s5 s6
307 s1 1.40 1.31 1.29 1.29 1.29 1.29
308 s2 1.28 1.37 1.33 1.32 1.31 1.31
309 s3 1.19 1.27 1.31 1.28 1.25 1.24
310 s4 1.21 1.26 1.27 1.33 1.28 1.26
311 s5 1.22 1.25 1.26 1.28 1.32 1.26
312 s6 1.15 1.19 1.17 1.18 1.19 1.24
313 MUTUAL INFO I(X,Z|Y) donde: X=Si_T Y=Si_t Z=Sj_t
314     s1 s2 s3 s4 s5 s6
315 s1 0.00 0.09 0.11 0.11 0.12 0.12
316 s2 0.09 0.00 0.04 0.06 0.06 0.06
317 s3 0.12 0.04 0.00 0.03 0.06 0.07
318 s4 0.11 0.07 0.05 0.00 0.05 0.07
319 s5 0.09 0.07 0.06 0.04 0.00 0.05
320 s6 0.09 0.05 0.07 0.06 0.05 0.00
321 METRICA razon o division del 3ro con el 2do
322     s1 s2 s3 s4 s5 s6
323 s1 1.00 0.93 0.92 0.92 0.92 0.92
324 s2 0.94 1.00 0.97 0.96 0.96 0.96
325 s3 0.91 0.97 1.00 0.98 0.95 0.95
326 s4 0.91 0.95 0.96 1.00 0.96 0.95
327 s5 0.93 0.95 0.96 0.97 1.00 0.96
328 s6 0.92 0.96 0.94 0.95 0.96 1.00
329 porcentaje de reducci3n de la entropía
330     s1 s2 s3 s4 s5 s6
331 s1 0.00 0.07 0.08 0.08 0.08 0.08
332 s2 0.06 0.00 0.03 0.04 0.04 0.04

```

```

333 s3 0.09 0.03 0.00 0.02 0.05 0.05
334 s4 0.09 0.05 0.04 0.00 0.04 0.05
335 s5 0.07 0.05 0.04 0.03 0.00 0.04
336 s6 0.08 0.04 0.06 0.05 0.04 0.00

```

337 9 Open issues

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

344 Appendix: Otros issues

- 345 1. Se puede optimizar la ubicación de los sensores en una finca para maximizar
- 346 el poder de los algoritmos predictivos (Kriging, Conditional Random Fields).
- 347 2. *Discretización*: Existen otros criterios para discretizar como el criterio de
- 348 Sturge que computa $k = \log_2 n + 1$, donde n es el número total de datapoints
- 349 (instancias de dato, hechos, filas). Hasta la versión (commit)

350 NOMBRAR LA VERSION

351 no investigué adecuadamente como discretizar las variables y elegí probar con

352 distintos números de intervalos k. Un método de discretización interesante

353 que encontré fue Ameva, whose objective is to maximize the dependence

354 between the intervals that divide the values of an attribute and the classes

355 to which they belong, providing at the same time the minimum number of

356 intervals. En R, la librería discretizacion lo trae implementado. Para usarlo,

357 llamar al método `disc.Topdown(data,method=3)`. La ventaja de AMEVA es

358 que uno puede desligarse de indicarle el valor de k o h al algoritmo, lo calcula

359 solo. mas info: unas slides introductorias pero mencionan todos los aspectos

360 vinculados a discretización (SLIDES <http://stat.skku.ac.kr/myhuh/homepage/specialLectures/SDULE>)

361 Recopilación de info interesante sobre mutual information y su implementación

362 en R.

363

364 RECOMENDABLE!!! --> Entropy-Based Inference using R and the np Package:

365 A Primer http://socserv.mcmaster.ca/racine/entropy_np.pdf

366

367 OTRO DOC más info en "Inequality, Entropy and Goodness of Fit"

368 <http://personal.lse.ac.uk/BANDYOPS/GoF1.pdf>

369 UNAS SLIDES ACLARATORIAS http://www.unc.edu/~maguilar/metrics/Myslides_trans2.pdf