

Predicción meteorológica mediante redes neuronales

Castello Facundo, Sikh Lautaro y Yackel Francisco
Facultad de Ingeniería y Ciencias Hídricas – fyackel@gmail.com

Resumen— En la actualidad debido a fenómenos como el calentamiento global, el efecto invernadero, entre muchos otros, se observa una dinámica cambiante con respecto al clima. Por tanto, cobra suma importancia poder establecer un pronóstico del mismo. En este trabajo final de materia nos propusimos como objetivo emplear una red neuronal con el fin de poder llevar a cabo una aproximación a distintas variables que influyen en la conformación del estado climático, como son las precipitaciones, las temperaturas, tanto máxima como mínima, el porcentaje de humedad, la velocidad y el punto de rocío. El método que aplicamos utiliza en una primera parte un mapa auto organizativo para agrupar patrones similares y posteriormente su tratamiento en una red neuronal con entradas desplazadas, utilizando cierta cantidad de patrones de datos anteriores. La red entrenada tiene como fin realizar pronósticos locales a corto plazo. Con los resultados obtenidos en las diferentes pruebas realizadas a lo largo del estudio se puede observar que el modelo propuesto sirve para realizar dicha tarea.

Palabras clave—mapa, red, neuronas, análogos, pronósticos.

I. INTRODUCCIÓN

EN este trabajo proponemos investigar si la realización de pronósticos climáticos mediante herramientas proporcionadas por el cursado, va a brindar un modelo óptimo que arroje resultados aceptables, es decir, realice una buena aproximación de una serie de datos climáticos.

Como datos de estudio para los distintos entrenamientos en las fases del modelo nos servimos de la base de datos pública del Instituto Nacional de Tecnología Agropecuaria (INTA) medidos en la ciudad de Oro Verde, provincia de Entre Ríos. La bondad de los resultados abarcan regiones aledañas a esta localidad por tratarse de pronósticos locales.

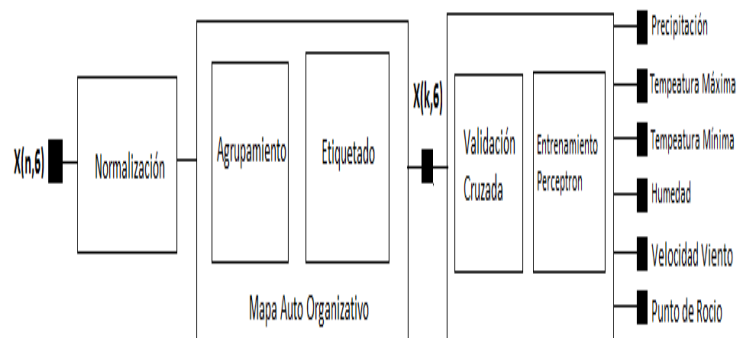
Todos los algoritmos utilizados en el proceso del estudio fueron realizados por los integrantes de este grupo. Implementamos un mapa auto organizativo cuadrado y un perceptrón multicapa con propagación hacia atrás y desplazamiento de sus entradas.

Finalmente comparamos los resultados obtenidos con distintos servicios meteorológicos para comprobar la fidelidad del modelo propuesto.

A grandes rasgos podemos detallar el pipeline del proyecto de la siguiente forma:

- Extracción de los datos
- Normalización de cada dimensión
- Agrupamiento en SOM
- Etiquetado del SOM
- Identificación de días análogos
- Entrenamiento de red neuronal
- Pronóstico

Fig. 1: Modelo completo



II. PROCESAMIENTO DE LOS DATOS

Tanto para alimentar el mapa auto organizativo como la red neuronal, contamos con entradas de seis dimensiones descriptas en el resumen de este informe. Cada patrón de datos consiste en una tupla de seis valores que representa un día del año. Tomamos datos a partir del 1° de Enero de 2010 hasta el 11 de Noviembre del corriente año, por lo que contamos con un lote de 2141 patrones de entrada al modelo.

Previo al análisis de los mismos, debimos normalizar cada dimensión de los datos con el fin de obtener valores numéricos entre -1 y 1, para manipularlo de forma más fácil y correcta, para luego des normalizarlos y realizar comparaciones.

III. AGRUPAMIENTO DE DÍAS ANALOGOS.

Tal como propone José M. Gutiérrez en su trabajo^[1], una vez que normalizamos todos nuestros datos de entrada, procedemos a realizar un agrupamiento de días análogos a partir de un mapa auto organizativo (SOM) cuadrado de 64 neuronas. El mismo nos sirve para identificar días del año que poseen características similares bajo un criterio de organización independiente estipulado por el SOM a partir de un etiquetado de los datos. En el caso de que queden neuronas sin etiquetar, esto significa que no hay ningún conjunto de patrones asociados a ella, se descarta.

Una vez realizado el etiquetado de todos los datos, buscamos aquellos días que sean análogos al día anterior a pronosticar. Esto nos provee de una lista de días con características similares.

IV. ENTRENAMIENTO DEL PERCEPTRÓN MULTICAPA.

A la lista de días análogos provista por el mapa auto organizativo, la utilizaremos para entrenar un perceptrón multicapa, el cual contará como salidas deseadas, para cada día análogo, el día siguiente a sí mismo.

Bajo este método, contamos con la posibilidad de utilizar más información para el entrenamiento de la red con el fin de sumar precisión al pronóstico. Es por esto que decidimos alimentar a la misma no solo con un patrón en un momento dado, sino que también incluimos entradas desplazadas en el tiempo a la red, esto es tomar N días anteriores a cada día análogo provisto por el SOM.

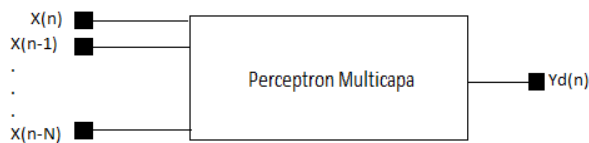


Fig. 2: Perceptron multicapa con entradas desplazadas

La estructura elegida para la red, luego de varias pruebas con distintas configuraciones, quedó determinada por 4 neuronas en la primera capa, 8 neuronas en la capa oculta y 6 neuronas en la capa de salida que corresponden a las seis variables de estudio.

Generamos distintos lotes de entrenamiento y prueba para realizar validaciones cruzadas en función de determinar cuál es la mejor configuración de pesos para entrenar la red que arrojen el menor error posible.

Una cuestión importante es el criterio de error adoptado. Como se trata de un pronóstico, y no de resultados exactos, hemos adoptado un criterio de tolerancia en cuanto a la diferencia entre el valor pronosticado y el deseado independiente a cada variable ya que, por ejemplo, las escalas de magnitudes que se manejan en porcentaje de humedad, no son las mismas que las que se tienen en temperatura. También ponderamos el margen de diferencia tolerado según un criterio que convenimos apropiado para darle más importancia a los errores cometidos en variables que para este grupo son más relevantes que otras, como ser el pronóstico de temperaturas con respecto al punto de rocío. Si la ponderación resultante arroja un valor mayor a cierta tolerancia, contamos esto como un error en el entrenamiento.

V. PRONOSTICO

Una vez entrenados los pesos para un conjunto de días análogos y sus N días anteriores, introducimos las variables del día anterior al día que se quiera pronosticar (y sus N días anteriores si así lo requiere) y obtenemos la salida de la red que deberá ser desnormalizada, es decir, realizar proceso inverso a la normalización que se hizo en un primer momento. El conjunto de datos obtenidos será el resultado final del análisis.

VI. APLICACIÓN

Aquí presentamos algunos de los análisis que realizamos en cuanto a pronósticos de corto plazo y también distintas comparaciones con diversas fuentes de pronósticos meteorológicos.

TABLA 1
Pronóstico para el 8 de Noviembre de 2015

Pronostico para el 8 de Noviembre 2015						
Dias Anteriores	Precipitación (mm)	Temp. Máxima (C°)	Temp. Mínima (C°)	% Humedad	Vel. Viento (km/h)	Punto Rocío (C°)
0 dias	2,91	29,51	15,66	54,79	9,37	12,49
1 dia	3,05	29,98	16,03	58,92	9,98	14,45
2 dias	0,4	30,77	16,68	59,64	7,35	14,83
3 dias	1,47	31,73	15,71	53,18	7,87	14,05
4 dias	0,17	23,84	9,04	44,82	7,66	5
Valores Reales	0.0	29.1	15.4	54	7.2	13.7

En la Tabla 1 se puede observar el pronóstico hecho para un determinado día utilizando hasta 4 días anteriores. Se alerta que los resultados son buenos debido a la poca diferencia que hay entre el valor pronosticado y el valor real medido para ese día.

Se dejan algunos pronósticos extras en las tablas 2 y 3.

TABLA 2
Pronostico para el 9 de Noviembre de 2015

Pronostico para el 9 de Noviembre 2015						
Dias Anteriores	Precipitación (mm)	Temp. Máxima (C°)	Temp. Mínima (C°)	% Humedad	Vel. Viento (km/h)	Punto Rocío (C°)
0 dias	8.85	28.92	16.14	66.28	8.64	15.68
1 dia	73.41	31.02	18.81	61.2	11.57	17.3
2 dias	35.36	31.03	19.45	69.38	10.4	18.98
3 dias	5.68	29.25	17.3	68.21	7.66	17.06
4 dias	70.48	26.58	17.11	76.01	10.15	17.36
Valores Reales	40.5	31.9	18.6	61	11.8	17.6

TABLA 3
Pronostico para el 10 de Noviembre de 2015

Pronostico para el 10 de Noviembre 2015						
Dias Anteriores	Precipitación (mm)	Temp. Máxima (C°)	Temp. Mínima (C°)	% Humedad	Vel. Viento (km/h)	Punto Rocío (C°)
0 dias	1,03	33,19	16,86	58,03	8,69	17,77
1 dia	1,28	31,95	17,03	59,68	6,96	16,56
2 dias	1,79	29,98	18,83	69,86	5,53	18,44
3 dias	2,39	31,14	19,27	67,66	7,26	17,96
4 dias	3,3	31,41	17,58	62,37	6,72	16,87
Valores Reales	0.0	28.2	16.4	74	7.3	18.2

Incluso se realizaron pruebas con otras estructuras neuronales como se puede ver en la Tabla 4, en este caso 6 neuronas en la primera capa, 10 en la capa oculta y 6 neuronas en la capa de salida; donde los resultados denotan una pequeña diferencia de precisión con respecto a la arquitectura descrita anteriormente.

TABLA 4
Pronostico para el 10 de Noviembre de 2015 con un
Red neuronal (6,10,6)

Pronostico para el 10 de Noviembre 2015						
Dias Anteriores	Precipitación (mm)	Temp. Máxima (C°)	Temp. Mínima (C°)	% Humedad	Vel. Viento (km/h)	Punto Rocio (C°)
0 dias	1,39	26,22	17,91	85,93	9,32	19,38
1 día	3,4	27,06	18,99	88,76	10,08	21,07
2 dias	19,59	25,86	18,65	90,46	9,39	20,88
3 dias	3,65	28,23	19,06	84,38	7,45	20,79
4 dias	7,53	24,1	17,03	89,57	6,18	18,97
Valores Reales	0,0	28,2	16,4	74	7,3	18,2

En general, el tiempo de ejecución del algoritmo completo ronda los 14 minutos ya que la misma cuenta de un agrupamiento de más de 2000 patrones en un mapa de 8x8 y 1000 épocas, seguido del entrenamiento de 10 redes neuronales debido a la validación cruzada que realizamos.

Encontramos un error cuadrático medio del orden de 1.3, que es levemente menor al encontrado por Diana Poma Lima en su trabajo “Predicción meteorológica mediante redes neuronales” [2] en donde arroja un error de entrenamiento cercano a 1.9, por el hecho de que no utiliza agrupamiento de días análogos en su estudio.

VII. CONCLUSIONES

Considerando los alcances propuestos al inicio de este documento, que apuntaban a realizar pronósticos locales a corto plazo, nos encontramos con que el modelo propuesto para realizar esto cumple sobradamente con las expectativas.

Cabe destacar que la implementación de un mapa auto organizativo perfecciona notablemente el rendimiento de nuestro proyecto ya que permite trabajar con lotes similares de datos para procesarlos con una red pequeña, en detrimento del tiempo total del proceso ya que el mismo se lleva gran parte del tiempo en los cálculos.

Como se ve en las tablas, contar con N entradas desplazadas es una buena práctica para aumentar la precisión en el pronóstico, pero cuando dicha cantidad supera los tres días, se percibe que los resultados obtenidos no son tan buenos; como sí lo son aquellos en los que tomamos hasta dos días anteriores.

Otro aspecto importante, que es difícil pronosticar la cantidad exacta en milímetros de precipitación caída, pero sí se puede hablar de una tendencia cuando los valores pronosticados son altos, como se ve en la Tabla 2, por esto es conveniente hablar de probabilidad de precipitaciones y no de cantidad.

También destacamos que la dimensión utilizada del mapa auto organizativo es inversamente proporcional al tamaño del perceptron. Esto es así, porque al utilizar un SOM más grande, reducimos la cantidad de días análogos para un determinado momento, y esto reduce la cantidad de patrones de entrenamiento y prueba disponibles. Por lo tanto, existe un compromiso entre la dimensión del mapa, la cantidad de días anteriores y la estructura de la red en función de mantener un balance entre la cantidad de patrones para entrenar y los pesos que se deben calcular.

VIII. PROPUESTAS A FUTURO

Como conclusión final, dados los buenos resultados obtenidos en la mayoría de los casos, existe la posibilidad de expandir el proyecto con el fin de utilizar el día pronosticado para realimentar la red, se podría intentar pronosticar a más de un día debido a que el error que se arrastra es despreciable.

REFERENCIAS

- [1] José M. Gutiérrez, *PROMETEO: Un Sistema Experto para el Pronóstico Meteorológico Local basado en Redes Neuronales y Cálculo de Analogías* Instituto Nacional de Meteorología, 1999.
- [2] Diana Lucía Poma Lima, “Predicción meteorológica mediante redes neuronales”, *Universidad Técnica Particular de Loja*.