

Sistema de inteligencia artificial para la predicción temprana de heladas meteorológicas

Artificial intelligence system for early prediction of weather frost

Franklin Riabani Mercado¹, Willman García Fernández², Johnny A. Herrera Acebey³

¹ Universidad Católica Boliviana, Maestría en inteligencia de negocios, Cochabamba, Bolivia

² Universidad Católica Boliviana, Carrera de Ingeniería Ambiental, Cochabamba, Bolivia

³ Universidad Católica Boliviana, Carrera de Telecomunicaciones, Cochabamba, Bolivia, c Márquez s/n, Zona Tupuraya, Cochabamba Bolivia.

franklin.riabani@gmail.com

Resumen: La helada es un factor meteorológico importante para la hidrología, climatología y agricultura. Este estudio propone crear una máquina de aprendizaje extremo (ELM) utilizando el algoritmo propuesto por Huang sobre una red neuronal monocapa con propagación hacia adelante, como base para la predicción temprana de heladas meteorológicas. El estudio fue desarrollado en el Valle Alto del departamento de Cochabamba-Bolivia, los datos fueron colectados en 6 estaciones meteorológicas llegando a un total de 178450 mediciones, se para el entrenar la red neuronal. Para la verificación se utilizaron los datos obtenidos de las estaciones meteorológicas de Tiquipaya y Arque del año 2016. Con márgenes de confianza superiores al 90%. Como resultados de la investigación se demostró que una red neuronal entrenada con el algoritmo propuesto por Huang, es un buen predictor de heladas meteorológicas tanto en niveles de confianza y tiempos de respuesta.

Palabras claves: Inteligencia computacional, inteligencia Artificial, Machine Learning, SciPy, Raspberry Pi, Agricultura inteligente, Heladas

Abstract: Summary: Frost is an important weather factor for hydrology, climatology and agriculture. This study proposes to create a learning machine end (ELM) using the method proposed by Huang on a layer neural network monolayer as a basis for early frost weather prediction algorithm. The study was developed in the Valle Alto of Cochabamba-Bolivia, data were collected at 6 meteorological stations reaching a total of 178450 measurements, for training the neural network. For verification, meteorological data were used stations from Arque and Tiquipaya 2016. With confidence margins above 90% they were used. As research results showed that a neural network trained with the proposed Huang algorithm, it is a good predictor of weather frost from both, confidence levels and times response.

Keywords: Computational Intelligence, Artificial Intelligence, Machine Learning, SciPy, Raspberry Pi, intelligent Agriculture, Frosts

1 Introducción

1.1 Heladas

Las heladas son un fenómeno atmosférico que se presenta cuando la temperatura del aire en cercanías del suelo, aproximadamente 1,5 a 2 metros, desciende por debajo de cero grados centígrados. Técnicamente, la palabra “helada” se refiere a la formación de cristales de hielo sobre las superficies, tanto por congelación del rocío como por un cambio de fase de vapor de agua a hielo [16].

La helada se define en dos ámbitos: “advectiva” y “radiativa”. Las heladas advectivas están asociadas con incursiones a gran escala de aire frío con una atmósfera con viento y bien mezclada y una temperatura que a menudo está por debajo de cero, incluso durante el día. Las heladas de radiación están asociadas con el enfriamiento debido a la pérdida de energía por el intercambio radiante durante las noches despejadas y en calma, y con inversiones de temperatura [17]. Otro evento que favorece la ocurrencia de heladas es el desplazamiento de masas de aire frío provenientes del polo sur, las heladas se clasifican tanto por su tipo como por el daño que produce la helada blanca cuando deja una capa de hielo sobre la superficie y helada negra cuando produce quemadura o mata las células vegetales.

El daño que provocan las heladas a los cultivos no se debe a las frías temperaturas sino principalmente a la formación de hielo extracelular (i.e. fuera de las células) dentro del tejido de la planta, provocando la salida de agua y dañando las células por deshidratación. Las pérdidas en los cultivos como consecuencia de las heladas varían entre 20 a 100% dependiendo de la intensidad de la helada, estado de desarrollo y tipo de cultivo.

En la región de los Valles de Bolivia, tanto la agricultura de subsistencia como la agricultura comercial, enfrentan problemas climáticos diversos, entre estos las heladas que provocan disminución o pérdida en las cosechas afectando la economía familiar campesina.

La predicción precisa de heladas puede potencialmente reducir el daño por heladas, proporcionando una oportunidad a los agricultores de prepararse contra ellas. Es importante poder predecir cuándo cae la temperatura hasta un valor crítico para poner en marcha los métodos de contingencia activos de protección contra las heladas.

1.2 Aprendizaje Automático

El aprendizaje automático es una rama de la Inteligencia Artificial que busca que las computadoras sean capaces de aprender e inducir conocimiento de datos. El

aprendizaje automático se clasifica según el tipo de aprendizaje Supervisado y no supervisado. El aprendizaje supervisado está compuesto por dos fases o etapas la primera conocida como *Entrenamiento* donde se observa las características más relevantes de un evento que actúan como datos de entrada y la segunda conocida como *Prueba* en la que compara estas características con nuevos datos de prueba y se obtiene al final de estas dos etapas un modelo entrenado y probado que intenta explicar los datos de entrada.

En base a los antecedentes, el objetivo de la presente investigación es demostrar la aplicabilidad del algoritmo propuesto por Huang denominada “*Extreme Learning Machine*” sobre una red neuronal monocapa con propagación hacia adelante (Single Layer Feedforward neural network - SLNF), para la predicción temprana de heladas en los valles y zona andina del departamento de Cochabamba.

2 Teoría del Algoritmo Extreme Learning Machine (ELM)

Las Máquinas de Aprendizaje Extremo o *Extreme Learning Machine* (ELM) propuesto por Huang [1] es un algoritmo que está atrayendo la atención de miles de investigadores como una técnica de inteligencia artificial que resuelve muchos de los problemas a los que se enfrentan otras técnicas, por ejemplo su lenta velocidad de aprendizaje [2]. ELM inicialmente trabaja en redes neuronales monocapa con propagación hacia adelante (SLFNs) [3] (Figura 1).

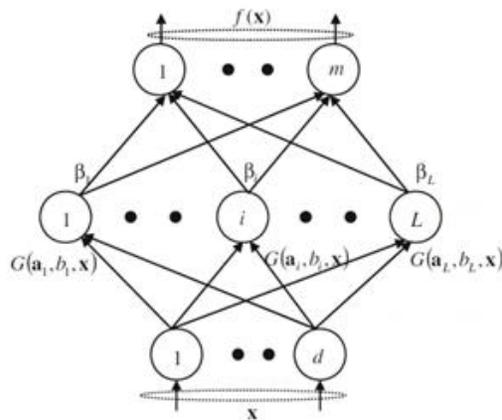


Figura 1: Arquitectura Single Layer Feedforward neural network (SLFN)

El algoritmo del ELM está basado en el hecho de que si los pesos de entrada y los sesgos de la capa oculta de una SLFN, son asignados al azar, la SLFN puede ser

considerada como un sistema lineal y los pesos de salida se pueden determinar analíticamente mediante la inversa generalizada [4]. Esta idea ya fue estudiada por Huang et al [5]; y hay un estudio profundo acerca de las diferencias entre el ELM y trabajos previos. Sin embargo, Huang demuestra la capacidad que tiene el ELM para ser un aproximador universal donde los pesos de entrada elegidos al azar pueden ser generados de acuerdo a cualquier distribución de probabilidad continua, sin ningún conocimiento previo [6]. Esta asignación aleatoria se puede hacer si las funciones de activación en la capa oculta son infinitamente diferenciables [3, 4], la más popular es la una función signoide.

En general, las SLFNs estándar con L neuronas en la capa oculta y función de activación $f(x)$, pueden aproximar las N muestras con error cero.

$$\sum_{i=1}^L \|O_i - t_i\| = 0 \quad (\text{ec. 1})$$

2.1 Teoría de las máquinas de aprendizaje extremo (ELM)

Según Huang [5], la función de salida de una red SLFN con L nodos ocultos puede ser representada por:

$$f_L(x) = \sum_{i=1}^L \beta_i g_i(x)$$

$$f_L(x) = \sum_{i=1}^L \beta_i G(a_i, b_i, x), x \in R^n, a_i \in R^n \quad (\text{ec. 2})$$

Dónde: g_i denota la función de salida $G(a_i, b_i, x)$ del nodo oculto i . Para nodos aditivos con función de activación g , g_i es definida como:

$$g_i = G(a_i, b_i, x) = g(a_i \cdot x + b_i), a_i \in R^d, b_i \in R \quad (\text{ec. 3})$$

2.2 Teoremas de interpolación

Para [1] las N muestras distintas $(x_i, t_i) \in R^d \times R^m$, las redes neuronales monocapa de tipo “feedforward” con L nodos ocultos son matemáticamente modeladas como:

$$\sum_{i=1}^L \beta_i g_i(x_j) = \sum_{i=1}^L \beta_i G(a_i, b_i, x_j) = o_j, j = 1, \dots, N. \quad (\text{ec. 4})$$

Estas SLFNs pueden aproximar estas N muestras con un error medio nulo $\sum_{j=1}^L \|o_j - t_j\| = 0$ es decir, existen (a_i, b_i) y β_i tal que:

$$\sum_{i=1}^L \beta_i G(a_i, b_i, x_j) = t_j, j = 1, \dots, N. \quad (\text{ec. 5})$$

Las ecuaciones anteriores pueden ser escritas de forma vectorial como:

$$H\beta = T \quad (\text{ec. 6})$$

Dónde:

$$H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_N) \end{bmatrix} = \begin{bmatrix} G(a_1, b_1, x_1) & \cdots & G(a_L, b_L, x_1) \\ \vdots & \cdots & \vdots \\ G(a_1, b_1, x_N) & \cdots & G(a_L, b_L, x_N) \end{bmatrix}_{N \times L}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m} \quad y^T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times m}$$

2.3 Pasos para la implementación del algoritmo Extreme Learning Machine

Para [3]. Dado un conjunto de entrenamiento $N = \{(x_i, t_i) \mid x_i \in R^d, t_i \in R^m, i = 1, \dots, N\}$, una función de salida de los nodos ocultos $G(a_i, b_i, x)$, y un número de nodos ocultos L .

Paso 1: Generar de forma aleatoria los pesos de los nodos de la capa oculta (a_i, b_i)

Paso 2: Calcular la matriz de salida de la capa oculta H .

Paso 3: Calcular los pesos de la capa de salida β .

El algoritmo Extreme Learning Machine puede trabajar con varios tipos de funciones de activación.

3 Metodología

3.1 Dataset de entrenamiento

La investigación se concentra en el Valle Alto del departamento de Cochabamba. Los datos fueron proporcionados por el Servicio Nacional de Meteorología e Hidrología (SENAMHI), de 6 estaciones meteorológicas automatizadas que se encuentran en las poblaciones de Tarata, Arbieto, Punata, Cliza y San Benito (Figura 2).

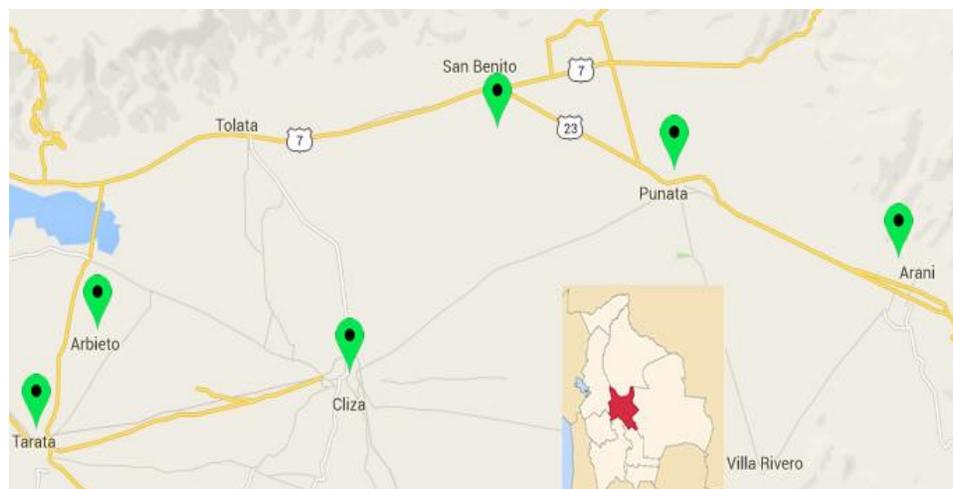


Figura 2: Zona de intervención de la investigación

El periodo de Colecta de datos corresponde a diciembre de 2011 hasta agosto de 2015, se tienen 1.360 días observados. El tiempo entre mediciones que realizan las estaciones meteorológicas es de 15 min, como resultado se obtienen 96 mediciones por día, lo que nos proporciona entre todos los días observados y las 6 estaciones un dataset de 179.320 mediciones.

3.2 Dataset de validación

Para validar el modelo y demostrar la alta capacidad de adaptación del ELM se utilizaran dos nuevas estaciones meteorológicas la primera ubicada en la población de Tiquipaya (valle) y la segunda en la población de Arque (zona andina) del departamento de Cochabamba (Figura 3).



Figura 3: Figura 3: Zona de intervención de la investigación – Validación

Las variables de estudio en la investigación son:

- Temperatura relativa
- Radiación Solar
- Viento
- Humedad
- Precipitación

El motivo para el manejo de todas estas variables para entrenar la red neuronal con el algoritmo de ELM y no solo utilizar la temperatura, se debe a que la radiación y el viento son factores que favorecen la ocurrencia de una helada, debido al balance regional de radiación [7] como se ve en las figuras 4 y 5.



Figura 4: Las nubes reintegran calor a la tierra y disminuyen la ocurrencia de heladas



Figura 5: Cielo despejado aumenta la radiación de calor de suelo hacia el espacio

La lluvia y la humedad determinan el tipo de helada que se producirá [7].

3.3 Scipy

Es una librería para la programación científica que comenzó su desarrollo en 2001 está escrita en Python. SciPy es una expansión de NumPy, que integra nuevos paquetes para el tratamiento científico de datos. Integra gran cantidad de funciones de procesamiento de imágenes, procesado de la señal, estadística, integración numérica.

4 Resultados

Diseño y entrenamiento de un SLFN con el algoritmo Extreme Learning Machine (ELM)

En el proceso de preparación e integración de los datos meteorológicos Conforme con las convenciones de inteligencia artificial los datos fueron sometidos a un proceso de preparación e integración, en conformidad a las convenciones de Machine Learning. El resultado final del proceso es un dataset de entrenamiento con 178.450 mediciones meteorológicas sin datos faltantes ni aberrantes principalmente ocasionados por fallas en las estaciones meteorológicas del SENAMHI, también se preparó un segundo dataset de validación con 14.535 mediciones meteorológicas.

En un análisis exploratorio de los dos Dataset se obtuvieron los siguientes resultados:

Tabla 1. Análisis exploratorio de mediciones meteorológicas

Dataset	Parámetro	Temperatura °C	Radiación W	Humedad relativa %	Viento km h ⁻¹
Entrenamiento	Mínima	-3	0,78	13,90	0,30
	Máxima	30	29,78	99,50	4,88
	Media	18	17,33	66,47	2,4
	Desv. estándar	6,28	0,94	12,64	0,94
Validación	Mínima	-2	1,86	7,10	0,35
	Máxima	30	26,40	87,90	4,35
	Media	19	18,04	63,66	2,38
	Desv. estándar	5,89	4,77	12,58	0,91

Como se mencionó en la introducción del trabajo, en este trabajo de investigación se demuestra que con una estructura de red neuronal mono capa con propagación hacia adelante, utilizando el algoritmo de ELM (Figura 6), se consiguen altos niveles de generalización similares o superiores a otros algoritmos y estructuras de redes neuronales.

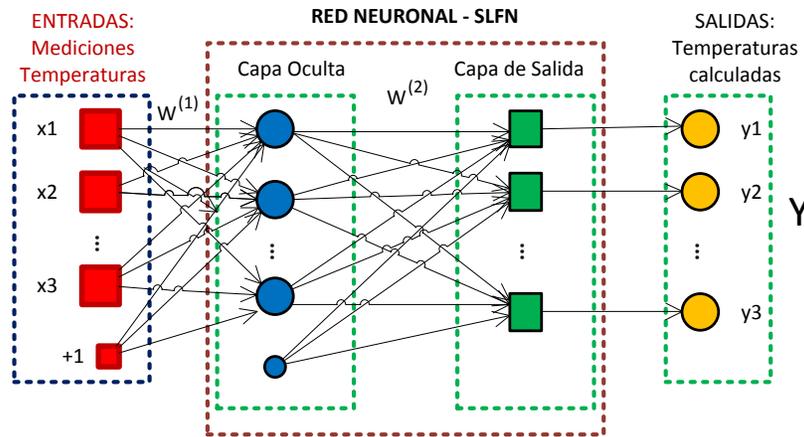


Figura 6: Esquema red de trabajo.

Los resultados obtenidos en la fase de entrenamiento del modelo se pueden apreciar en la Figura 7, donde, la línea negra representa el valor real de la temperatura y los puntos rojos representan el valor de temperatura calculado por la red neuronal entrenada con el algoritmo ELM.

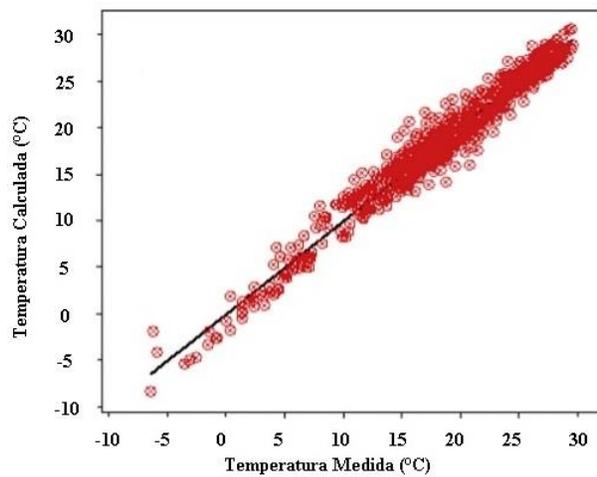


Figura 7: Resultados Dataset de entrenamiento

Podemos notar que los puntos se encuentran dispersos alrededor de la línea negra central que corresponde a la intersección del valor medido y el valor calculado por la red neuronal, con un factor de correlación de Pearson $r = 0.9741$

En la figura 8 se puede apreciar los resultados del dataset de validación para la red neuronal SLFN que ya estaba entrenado con el algoritmo ELM.

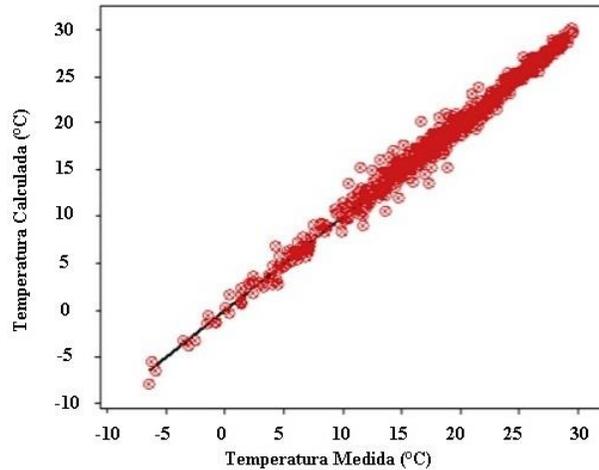


Figura 8: Resultados Dataset de validación.

De la figura 8 se puede concluir que los resultados son mucho más densos alrededor del valor ideal, y un factor de correlación de Pearson $r=0.9894$.

5 Conclusiones

Como una técnica de aprendizaje, el algoritmo de ELM ha demostrado un buen potencial para resolver problemas de clasificación. Así, las principales conclusiones que se han podido sacar son las siguientes:

- La velocidad de aprendizaje de la red neuronal del tipo SLFN con el algoritmo ELM es extremadamente alta.
- Tiene un bajo consumo de recursos de hardware.
- Destaca por su simplicidad, se ha demostrado cómo una red puede ser entrenada con pocos pasos.
- El algoritmo permite una gran capacidad de generalización a la hora de afrontar problemas de clasificación. Se ha podido observar que aumentando el número de datos de entrenamiento se obtienen mejores resultados de forma progresiva porque se crea una red más potente y a su vez el aumento con lleva una disminución de la norma de los pesos de la capa de salida, lo que deriva en un aumento de la capacidad de generalización, suavizando la frontera de decisión y por lo tanto evitando el sobreajuste.

6 Recomendaciones

Se propone como líneas futuras de investigación, continuar explorando las capacidades de otras arquitecturas de redes neuronales, con el objetivo de explotar su capacidad de representación, compresión y clasificación de datos. Una buena aplicación podría ser la de entrenar esta estructura para que sea capaz de mejorar la frontera de predicción minimizando los falsos positivos.

Implementar el ELM sobre otros tipos de redes neuronales nonocapas y multicapas con propagación hacia adelante y atrás para conocer las ventajas que ofrece este algoritmo sobre esas redes neuronales.

Referencias bibliográficas

- [1] G. Nan-Ying Liang, "A Fast and Accurate On-line Sequential Learning Algorithm for Feedforward Networks," *IEEE Transactions on Neural Networks*, vol. 17, no. 6, pp. 1411-1426, 2006.
- [2] David A. Elizondo and Stephen G. Matthews, "Recent Patents on Computational Intelligence, Centre for Computational Intelligence," Leicester, De Montfort University, 2009.
- [3] A. B. Crespo, *Aprendizaje Máquina Multitarea mediante Edición de Datos y Algoritmos de Aprendizaje Extremo*, Cartagena - Colombia: Universidad Politécnica de Cartagena, 2013.
- [4] E. Cambria, "Extreme Learning Machines," *IEEE Intelligent Systems*, pp. 30-59, 2013.
- [5] Q. Huang, "Extreme Learning Machine: Theory and Applications," *Neurocomputing*, no. 70(1-3), pp. 489-501, 2006.
- [6] S. Rajesh R., "Extreme Learning Machines - A Review and State-of the art.," *International Journal of wisdom based Computing*, vol. 1, pp. 35-48, 2011.
- [7] P. Q. Quispe, *Redes neuronales evolutivas para la predicción de heladas*, La Paz - Bolivia: UMSA, 2009.
- [8] Google, "Buscador Google," 5 10 2015. [Online]. Available: www.google.com. [Accessed 5 Octubre 2015].

-
- [9] J. m. m. Olmo, Desarrollo software de técnicas para el diseño automático de Redes Neuronales Artificiales con bajo coste computacional, Cartagena - Colombia: Universidad Politecnica de Cartagena, 2013.
- [10] E. D. a. S. G. Matthews, Recent Patents on Computational Intelligence, Centre for Computational Intelligence, Leicester: University De Montfort, 2009.
- [11] A. P. Engelbrech, Computational Intelligence An Introduction, Wiley: University of Pretoria, 2007.
- [12] F. Fundacion Española para la Ciencia y la tecnología, Meteorología y Climatología, España: Fundacion Española para la Ciencia y la tecnología, FECYT, 2004.
- [13] Z. Huang Guang-Bin, "Extreme Learning Machine: Theory and Applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489-501, 2006.
- [14] L. G. Matías Ramíre, Heladas, CENTRO NACIONAL DE PREVENCIÓN DE DESASTRES - CENAPRE, 2001.
- [15] S. B. Rampal Singh, "Application of Extreme Learning Machine Method for Time Series Analysis," *International Journal of Intelligent Technology*, vol. 2, no. 4, pp. 256-262, 2007.