

Estimación del consumo en rumiantes en pastoreo utilizando redes neuronales artificiales

Mariela N. Uhrig^{1,2}, Julio R. Galli⁴, Hugo L. Rufiner^{2,3}, Diego H. Milone²

¹Centro de Investigaciones Científicas y Transferencia de Tecnología a la Producción – (CICYTTP) CONICET– CP EW3105BWA – Diamante – Entre Ríos – Argentina

²Research Institute for Signals, Systems and Computational Intelligence (sinc(i)), FICH-UNL/CONICET– Santa Fe – Argentina.

³Laboratorio de Investigación de Cibernética, Facultad de Ingeniería, Universidad Nacional de Entre Ríos (LC-FI-UNER) – Oro Verde – Entre Ríos – Argentina

⁴Instituto de Investigación em Ciencias Agrarias de Rosario, Facultad de Ciencias Agrarias, Universidad Nacional de Rosario (IICAR-UNR) – Rosario – Santa Fe – Argentina

marielauhrig@conicet.gov.ar jgalli@lidernet.com.ar

lrufiner@sinc.unl.edu.ar dmilone@sinc.unl.edu.ar

Abstract. *Accurate and rapid measurement of forage intake in ruminants is important for efficient management livestock and forage resources, as well as for animal health and welfare in production systems. The use of intelligent signal processing algorithms to extract relevant information from the sound emitted by ruminants is a promising method to predict the intake of ruminants in grazing conditions. In this work, multilayer perceptrons and extreme learning machines, are used as non-linear multivariate regression models to predict intake. The results show that these non-linear regression techniques can significantly reduce the error in the estimation of forage intake in ruminants.*

Keywords: *Artificial neural network, ingestive behavior in ruminants, non-linear regression, extreme learning machines.*

Resumen. Medir con precisión y rapidez el consumo de forraje en rumiantes es importante para la gestión eficiente del rodeo y los recursos alimenticios, así como para la salud y bienestar animal en los sistemas de producción. El uso de algoritmos de procesamiento inteligente de señales que extraigan información relevante de los sonidos que emiten los rumiantes es una opción prometedora para predecir el consumo de rumiantes en condiciones de pastoreo. En este trabajo se aplican perceptrones multicapa y máquinas de aprendizaje extremo como modelos de regresión no lineales multivariada para la estimación del consumo. Los resultados muestran que éstas técnicas de regresión no lineal para pueden reducir significativamente el error de estimación de la cantidad de materia seca consumida por rumiantes.

Palabras clave: Redes neuronales artificiales, comportamiento ingestivo en rumiantes, regresión no lineal, Máquinas de aprendizaje extremo.

1. Introducción

El sonido que producen los animales al comer es una fuente de información que permite caracterizar y cuantificar las actividades de alimentación en vacas y ovejas (arranques, masticaciones y movimientos combinados), el tipo de pasturas que ingieren y la cantidad total ingerida [Milone et al., 2009; Galli et al., 2011]. Disponer de una metodología precisa y económica para la medición del consumo de rumiantes es importante para mejorar el desempeño de los sistemas de producción basados en el aprovechamiento directo de las pasturas. Este aspecto es esencial para lograr una gestión eficiente del rodeo y de los recursos alimenticios, así como para la salud y el bienestar animal. Además, es necesario que el método sea sencillo desde el punto de vista tecnológico y que no dañe ni perturbe el comportamiento normal de los animales.

La predicción del consumo de materia seca a través del uso de algoritmos de procesamiento inteligente de señales que puedan extraer la información relevante de registros sonoros de rumiantes es todo un desafío.

El análisis acústico de la masticación permite obtener información valiosa para cuantificar el comportamiento ingestivo de novillos [WallisDeVries et al., 1998] y vacas lecheras [Galli et al., 2006b] en pastoreo, y de novillos estabulados [Galli et al., 2006^a]. El uso de este tipo de análisis permitió estimaciones precisas del consumo de forrajes en ovejas en pastoreo, mediante técnicas de regresión lineal, [Galli et al., 2011]. La riqueza de la información contenida en el sonido de las masticaciones de las ovejas al ingerir el forraje permitió predecir de manera confiable el consumo de materia seca en estos animales.

En el presente trabajo se propone utilizar redes neuronales artificiales como regresores no lineales para la estimación del consumo de materia seca en rumiantes. En particular se utilizan dos tipos de redes neuronales artificiales: perceptrón multicapa (MLP, del inglés Multi-Layer Perceptron) [Haykin, S., 1994] y máquinas de aprendizaje extremo (ELM, del inglés Extreme Learning Machines) [Huang et al., 2004]. El MLP se considera un método robusto que tiene buenos resultados predictivos en gran variedad de aplicaciones [Paliwal et al., 2009; Vellido et al., 1999]. Las ELM se han utilizado tanto en problemas de clasificación como de regresión [Huang et al., 2006; Vergara et al., 2015, Huang et al., 2012]. La principal ventaja de las ELM es la ausencia del ajuste en la capa oculta del modelo neuronal, por lo que su tiempo de aprendizaje es mucho menor que el MLP clásico. La capacidad de las ELM para la predicción ha sido demostrada en diferentes problemas actuales [Huang et al., 2006].

El objetivo de este trabajo es desarrollar un método novedoso de estimación del consumo en rumiantes mediante un modelo de regresión no lineal multivariada utilizando redes neuronales artificiales. El resto de este trabajo se organiza de la siguiente forma: en la Sección 2 se describen los datos utilizados y métodos que se proponen; en la Sección 3 se describen los experimentos llevados a cabo y los resultados obtenidos y en la última sección se presentan las conclusiones y los trabajos futuros.

2. Materiales y métodos

2.1. Datos

Se utilizaron los datos obtenidos por Galli et al. (2011), que corresponden a los registros sonoros de tres ovejas mestizas de 2 a 4 años de edad, con un peso promedio de 85 ± 6.0 kg, registradas en el Departamento de Ciencia Animal de la Universidad de California en Davis (UCDavis). Las ovejas realizaron sesiones cortas de pastoreo en cuatro tipos de pasturas puras (alfalfa alta, alfalfa corto, pasto ovillo alto y pasto ovillo corto) ofrecidas en microparcels. Para asegurar el muestreo, los animales fueron ayunados durante una hora previa al inicio de las sesiones de pastoreo. Cada sesión de pastoreo fue grabada usando una cámara digital estándar (Sony DCR-PC100 digital camcorder) y un sistema de micrófono inalámbrico (Nady Systems 151 VR) colocado contra la frente del animal. La ingesta de materia seca se estimó como la diferencia entre la biomasa forrajera de la sesión previa y posterior al pastoreo. Los potes se pesaron individualmente con precisión de 0,1 g usando una balanza digital Setra 140 CP. Se pesaron dos potes por ensayo antes y después de cada sesión de pastoreo para estimar las pérdidas por evapotranspiración. Cada día se seleccionó al azar un subconjunto de potes de cada especie y altura para medir la altura del forraje en cinco hojas extendidas (en pasto ovillo) o tallos (en alfalfa).

Las dos primeras variables consideradas para la estimación del consumo son categóricas: *Especie forrajera* (pasto ovillo, alfalfa) y *Altura* (alto, corto); y las ocho variables restantes son cuantitativas: *Intensidad de masticación* (VC), *Energía total de densidad de flujo de masticación* (EC), *Energía de densidad de flujo de masticación por masticación* (ECc), *Energía de densidad de flujo masticación por unidad de tiempo de comida* (Et), *Intensidad de arranque* (VB), *Energía de densidad de flujo de arranque por arranque* (EBb) y *Energía total de densidad de flujo de masticación por arranque* (ECb). Por último, la variable *Consumo de materia seca* (DMI) es la que deseamos estimar con el modelo de regresión no lineal multivariada.

En la Figura 1 se puede observar un segmento de señal acústica típica, conteniendo los tres eventos mandibulares de los rumiantes (arranque, masticación y movimiento compuesto de masticación-arranque). Este es el tipo de señales registradas y procesadas en los experimentos de Galli et al. (2011) para obtener las 10 variables que luego se utilizaron en este trabajo como entradas a las redes neuronales artificiales.

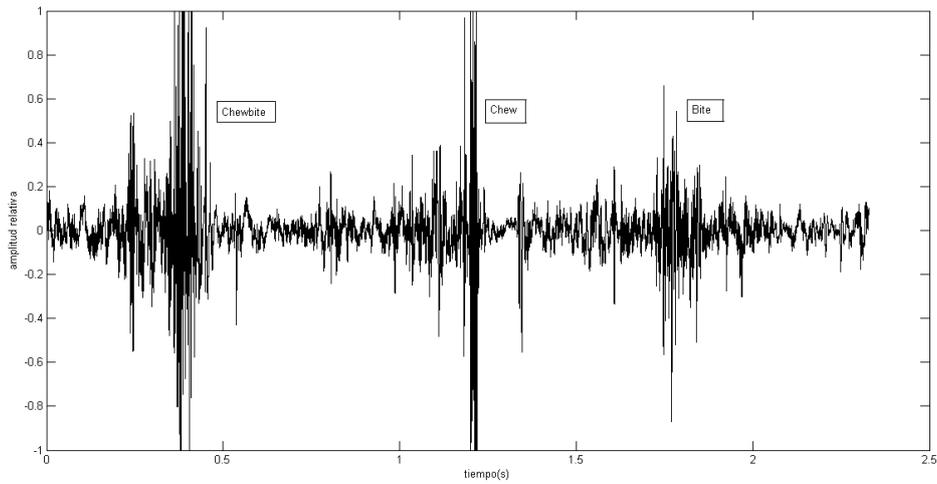


Figura 1. Señal acústica típica con los eventos mandibulares: arranque (bite), masticación (chew), y masticación-arranque (chewbite)

2.2. Redes Neuronales Artificiales

Perceptrón Multicapa

Esta es una de las arquitecturas de red utilizada en este trabajo, con diez neuronas entrada (una por cada variable de entrada) y una de salida (correspondiente a la variable a estimar). En la Figura 1 se muestra esquemáticamente su estructura básica

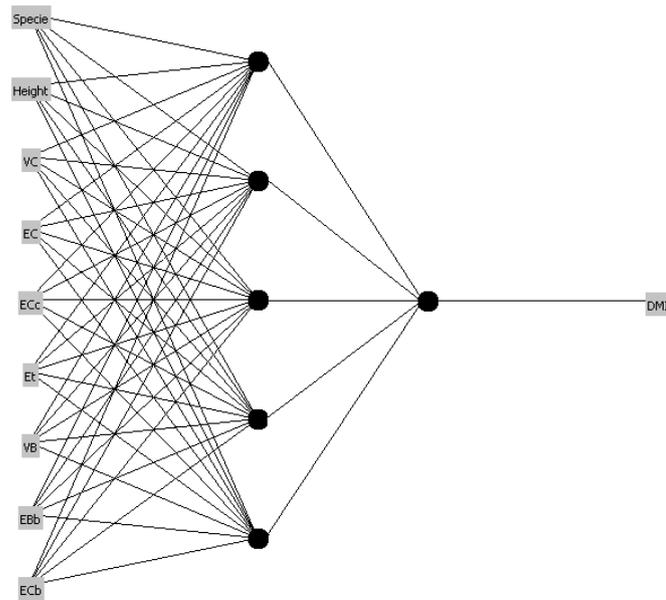


Figura 2. Esquema del perceptrón con las diez entradas en la capa de entrada, 5 neuronas en la capa oculta y la salida, en nuestro caso el DMI

Para la activación de las neuronas se empleó la función sigmoidea:

$$f(x) = 1/(1 + e^{-\beta x})$$

El algoritmo de aprendizaje utilizado fue el de retropropagación¹. [Rumelhart et al., 1986], que durante el entrenamiento cambia iterativamente los pesos de las neuronas con el fin de minimizar el error. Sus pasos principales son: inicialización aleatoria de los pesos, propagación hacia delante de las entradas, propagación hacia atrás del error, adaptación de los pesos y este proceso se repite desde el paso dos hasta su finalización.

Para ajustar los pesos de conexión se utiliza el error cuadrático medio entre la salida deseada y y la salida real:

$$E = \frac{1}{2} (y - f(x))^2$$

Máquinas de aprendizaje extremo

El algoritmo ELM fue desarrollado por Huang *et al.* (2006), tomando como base la estructura del perceptrón de una capa oculta con N_H neuronas y pesos son fijados aleatoriamente. Una de las características más destacadas de ELM es que la capa oculta no necesita ser entrenada.

Los pesos de salida se pueden obtener analíticamente resolviendo:

$$H\beta = Y$$

donde H es la matriz de salida de la capa oculta, β la matriz de pesos de salida y Y el vector de salidas deseadas o correctas. El entrenamiento consiste en resolver un problema de mínimos cuadrados, $\beta' = H^+ Y$, donde $H^+ = (H^T H)^{-1} H^T$ es la pseudoinversa de Moore-Penrose [D. Serre, 2002; C.R. Rao *et al.* 1971].

3. Experimentos y resultados

3.1. Experimentos

Se generaron particiones para el entrenamiento y prueba con los 44 registros utilizados en Galli et al. (2011). Se puede observar que la cantidad de datos disponibles es relativamente pequeña en relación a la cantidad de parámetros ajustables de las redes neuronales. Para generar mayor cantidad de datos, tanto de entrenamiento como de prueba, se realizó remuestreo, dentro del esquema experimental de validación cruzada. En una primera etapa se generaron 20 particiones madre, sólo con datos reales (esto es, sin datos generados por remuestreo). Se definió un tamaño de partición 70/30, lo que resulta en 31 registros para entrenamiento y 13 para prueba. En una segunda etapa se utilizó un método de remuestreo [Good, Philip I., 2006] que agregó a los datos reales,

¹ Se utilizó la versión implementada en Weka v 3.6.12 <http://www.cs.waikato.ac.nz/ml/weka/>

nuevos datos sintetizados a partir del promedio de algunos patrones reales tomados al azar. Con este método se generaron, en cada partición de entrenamiento madre, 40 particiones de entrenamiento y prueba con relación 70/30, pero con 140 registros para entrenamiento y 60 registros para prueba. En la Figura 3 se ilustra el esquema de particiones utilizado en el presente trabajo.

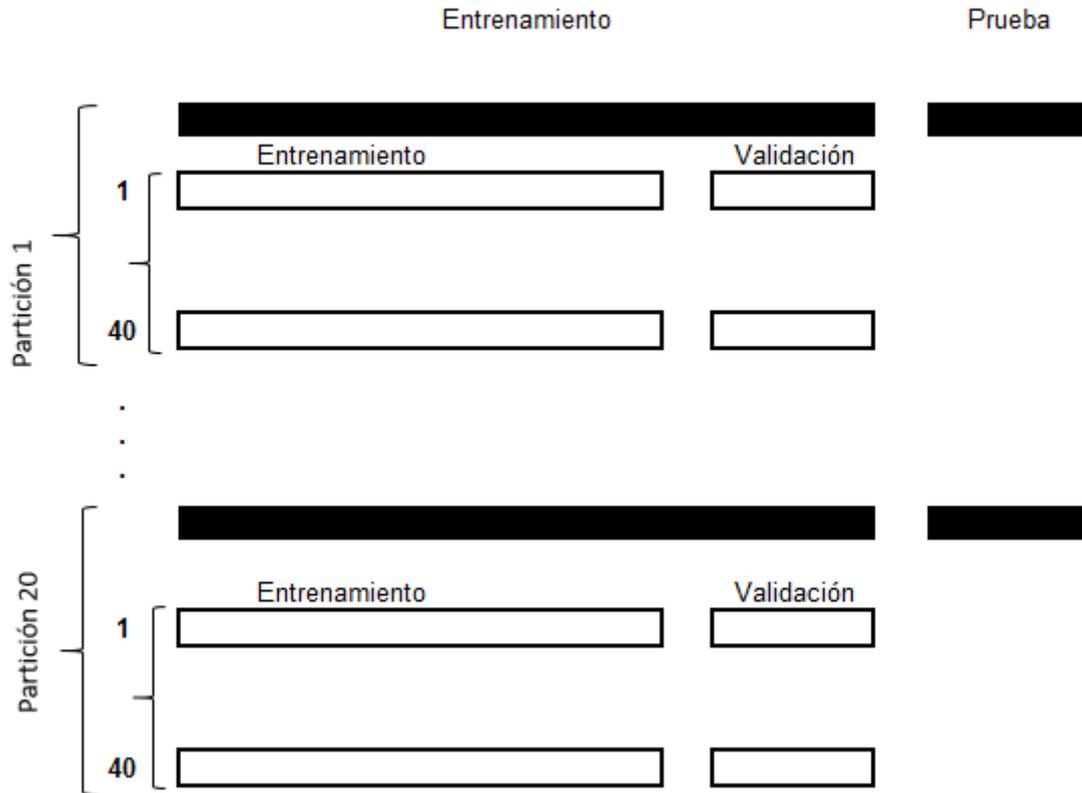


Figura 2. Esquema de particiones para la fase de entrenamiento y prueba.

Para realizar el ajuste de los hiperparámetros, tanto del MLP como del ELM, se utilizaron las 40 particiones de remuestreo. Luego, con los hiperparámetros determinados, para el ajuste de los pesos se utilizó la partición de entrenamiento madre sin remuestreo. Para el modelo MLP se exploró la cantidad de neuronas (1, 3, 5, 10 y 11), la velocidad de aprendizaje (0,15, 0,30 y 0,60), el momento (0,15, 0,30 y 0,60), y el número de épocas (200 y 500), por cada partición utilizada para el entrenamiento de la red. Para el modelo ELM se exploró la cantidad de neuronas en 100, 500, 1000, 1500 y 2000.

Para medir la el desempeño de cada prueba se utilizó el cálculo de la correlación:

$$R^2 = (\sigma_{y_i y_i'}) / (\sigma_{y_i} \sigma_{y_i'}),$$

y el error cuadrático medio:

$$RMSE = \sqrt{(\sum (y_i - y_i')^2/N)}, \text{ para } i=1 \dots N$$

donde y_i son las salidas deseadas, y_i' son las estimaciones, $\sigma_{y_i y_i'}$ es la covarianza de las salidas deseadas y las estimadas, σ_{y_i} es la desviación típica de la salida deseada, $\sigma_{y_i'}$ es la desviación típica de la salida estimada y N es el número de instancias o casos de prueba.

3.2. Resultados

Los criterios para evaluar el desempeño miden diferentes aspectos de la solución encontrada: la correlación muestra el grado asociación entre variables y el sentido (positiva o negativa), mientras que el error mide el grado de acierto en las estimaciones. Se esperaría que a mayor correlación se obtenga menor error.

En la Tabla 1 se muestra por columna el desempeño de cada método en la fase de prueba, medido por la correlación y el error cuadrático medio. En cada fila se observan las 20 particiones de prueba para ambos métodos. En la anteúltima fila se muestra la media de los valores por columna y en la última fila el desvío por cada columna.

Considerando el valor promedio de ambas medidas podemos decir que el modelo MLP fue mejor que ELM para predecir el DMI, aunque tuvo un mayor tiempo de aprendizaje, del orden de 60 veces más al ELM.

Tabla 1. Valores de correlación y error cuadrático medio de cada red neuronal artificial en la fase de prueba

Partición	R ²		RMSE	
	MLP	ELM	MLP	ELM
1	0,87	0,67	9,52	9,98
2	0,93	0,76	6,32	12,19
3	0,91	0,74	4,22	9,06
4	0,89	0,66	6,84	10,96
5	0,83	0,62	9,58	13,03
6	0,90	0,64	6,44	10,55
7	0,90	0,87	7,03	5,62
8	0,74	0,65	12,25	9,05
9	0,81	0,65	8,65	11,77
10	0,78	0,92	5,68	3,65
11	0,95	0,56	7,37	11,85
12	0,83	0,79	7,39	8,33
13	0,86	0,83	6,75	5,43
14	0,95	0,67	7,85	8,79
15	0,88	0,52	5,65	13,04
16	0,87	0,28	8,09	11,12
17	0,75	0,68	6,78	9,67
18	0,95	0,42	6,40	15,35
19	0,67	0,63	6,96	6,93
20	0,98	0,47	6,05	16,13
μ	0,86	0,65	1,74	10,13
σ	0,08	0,15	7,29	3,20

Para el MLP los hiperparámetros que mejor desempeño tuvieron en la fase de entrenamiento fueron: 5 neuronas en la capa oculta, 0,15 para la velocidad de aprendizaje, término de momento en 0,15 y 200 épocas de entrenamiento. Para el ELM el mejor desempeño se obtuvo con 2000 neuronas. A continuación se analiza con mayor detalle graficando los mismos.

En la Figura 4 se muestran los valores de DMI observados (correctos) y los estimados por el MLP (a la izquierda) y el ELM (a la derecha), para cada una de las 13 muestras del conjunto de prueba. Podemos ver que ambos modelos logran una buena aproximación de los valores DMI.

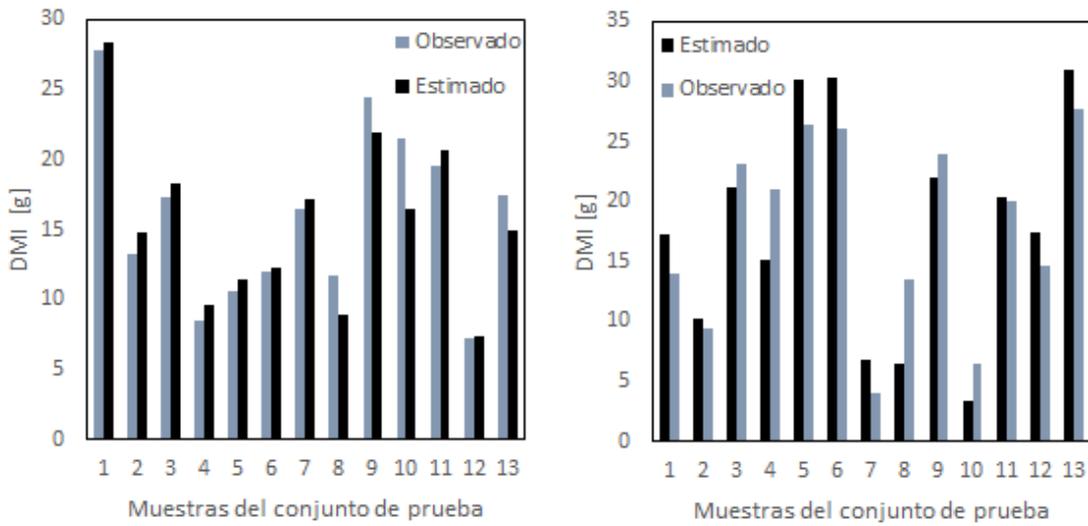


Figura 4. A la izquierda DMI estimado y observado con MLP para la partición 3, con $R^2 = 0.91$ y $RMSE = 4.22$. A la derecha se muestran los resultados ELM para la partición 10 con $R^2 = 0.92$ y $RMSE = 3,65$.

La Figura 5 compara los valores de DMI correctos y estimados por el MLP (a la izquierda) y el ELM (a la derecha) para los valores más altos de correlación según la Tabla 1. La nube de puntos se observa cercana a la recta indicando una correlación positiva, lo que nos muestra la bondad de estimación obtenida con ambos modelos. En el caso particular del ELM, coincidió que el valor más alto de correlación tiene asociado el menor valor de error, situación que no siempre sucede.

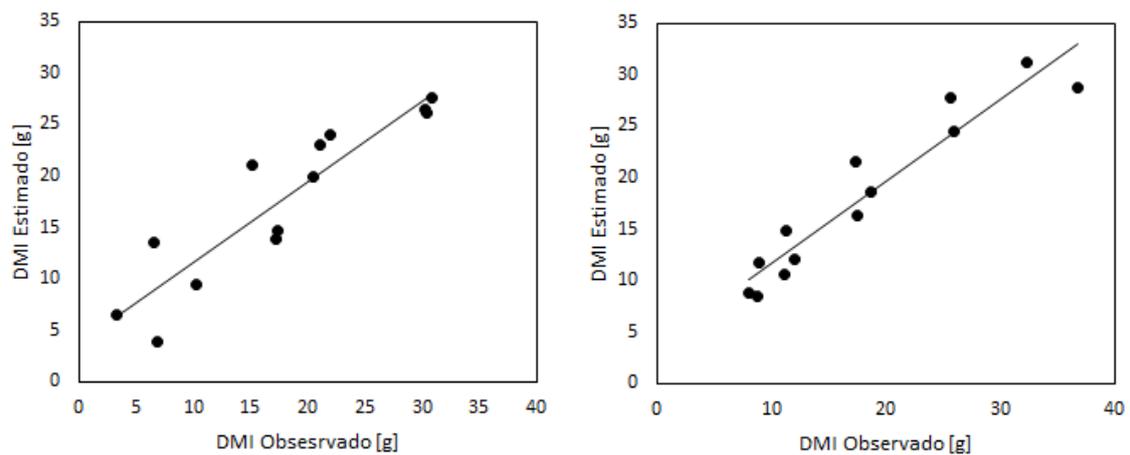


Figura 5. Relación entre el DMI estimado y correcto para el MLP a la izquierda, con $R^2 = 0.96$ y $RMSE = 3,76$, y para el ELM a la derecha, con $R^2 = 1$ y $RMSE = 1,33$ con los datos originales.

Para una mejor comparación con el modelo lineal del trabajo de Galli et al. (2011), se presentan en la Figura 6 ambos modelos de red neuronal, con los parámetros finalmente seleccionados como los de mejor desempeño, graficando los 44 registros originales. Como se puede observar, se obtuvieron resultados muy buenos a partir del

grupo de datos originales. Con MLP se obtuvo $R^2=0.96$ y un $RMSE=3,76$ y con ELM un $R^2=1$ y un $RMSE=1,33$.

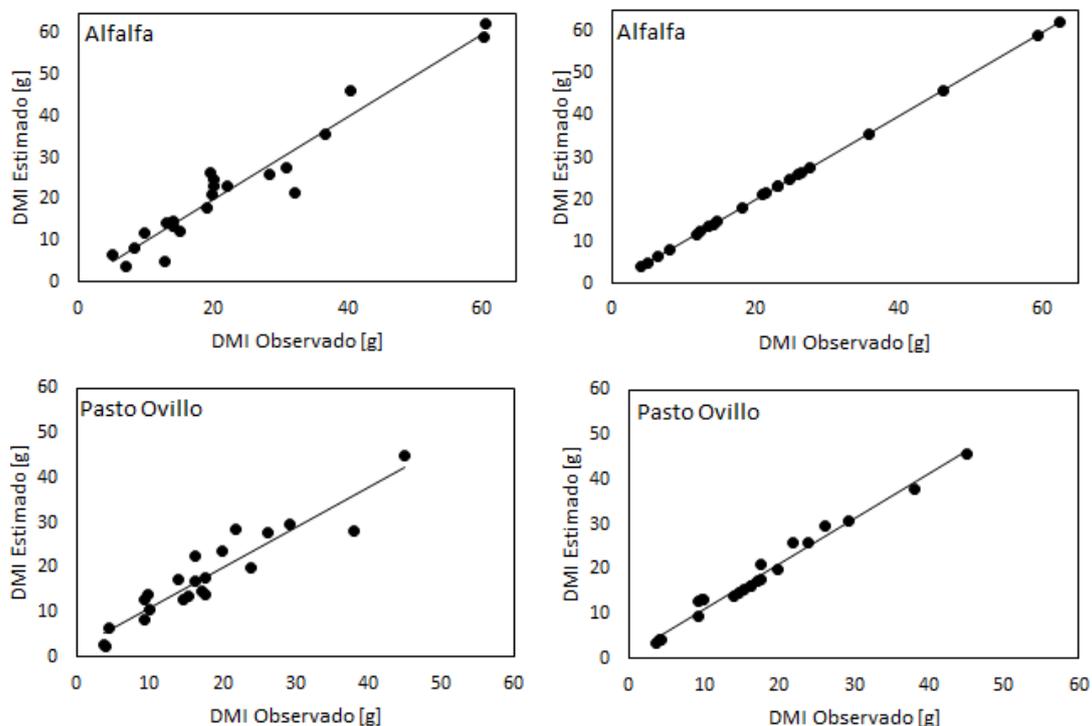


Figura 6. Relación entre DMI observado y estimado, a la izquierda para el MLP y a la derecha para el ELM

4. Conclusiones

En este trabajo se han presentado dos tipos de redes neuronales artificiales que abordan el problema de la automatización de la estimación del consumo de rumiantes. Los resultados que se presentan indican que es posible estimar el consumo de materia seca en el pastoreo de ovejas mediante regresión no lineal multivariada utilizando redes neuronales artificiales, y además apoyan la investigación de trabajos previos sobre la riqueza de la información contenida en el sonido de las masticaciones de las ovejas al ingerir forraje. Se proyectan trabajos futuros sobre la problemática en otros rumiantes, como es el caso de vacas, teniendo en cuenta que actualmente se están realizando experimentos para obtener más registros y con mayor duración. Para realizar una mejor comparación con el modelo lineal se requiere un estudio más profundo, dado que el método de validación empleado en ambos trabajos no es exactamente igual.

5. Referencias

Galli, J.R., Laforcada, H., Milone, D., Cangiano, C.A. Evaluación de un programa de computación para la clasificación automática de sonidos masticatorios. *Rev. Arg. Prod. Anim.* 25 (Sup. 1), 190–191. 2005.

Galli, J.R., Cangiano, C.A., Demment, M., Laca, E.A. Acoustic monitoring of chewing and intake of fresh and dry forages in steers. *Anim. FeedSci. Tech.* 128, 14–30. 2006a.

Galli, J.R., Cangiano, C.A., Pece, M.A., Larripa, M.J., Laca, E.A. Uso del sonido en el análisis de la tasa de consumo de bovinos. *Rev. Arg. Prod.Anim.* 26 (Sup.1), 165–167. 2006b.

Galli, J. R., Cangiano C. A., Milone D. H., Laca E. A.: Acousting monitoring of short-term ingestive behavior and intake in grazing sheep. *Livestock Science* Volume 140, Issues 1–3, Pages 32–412. 2011.

Good, Phillip I. *Resampling Methods. A Practical Guide to Data Analysis.* 3rd Edition 2005.

Huang, G-B., Zhu Q-Y, C.-K. Siew, “Extreme learning machine: A new learning scheme of feedforward neural networks,” in *Proceedings of ZOO4 International Joint Conference on NeumlNetworkz (IJCNN’2004)*, (Budapest, Hungary), July 2004.

Huang, G., Zhu, Q., Siew, C.: *Extreme learning machine: Theory and applications.* *Neurocomputing* (70), 489 – 501 2006.

Huang G., Bin, Senior Member, IEEE, Hongming Zhou, Xiaojian Ding, and Rui Zhang "Extreme Learning Machine for regression and Multiclass classification", *IEEE Transactions on Systems Man, and Cybernetics-Part B: CYBERNETICS*, VOL. 42, NO. 2, APRIL 2012

Haykin, S. *Neural networks: A comprehensive foundation.* Second Edition. Prentice-Hall, Inc. 1999.

Milone, H.L. Rufiner, Galli J.R., Laca E.A., Cangiano C.A. “Computational Method for Segmentation and Classification of Ingestive Sounds in Sheep,” *Computers and Electronics in Agriculture, Elsevier*, no. 65, pp. 228-237, 2009.

Paliwal, M., Kumar, U.: *Neural networks and statistical techniques: A review of applications.* *Expert Systems with Applications* 36(1), 2–17 2009

Rao, C.R., Mitra, S.K. *Generalized Inverse of Matrices and its Applications*, Wiley, New York, 1971.

Serre, D. *Matrices: Theory and Applications*, Springer, New York, 2002.

Vellido, A., Lisboa, P., Vaughan, J.: *Neural networks in business: a survey of applications (1992–1998).* *Expert Systems with Applications* 17(1), 51–70 1999.

Vergara, G., Cózar, J., Romero-González, C., Gámez, J.A., Soria-Olivas, and E.: Comparing elm against mlp for electrical power prediction in buildings. In: Ferrández Vicente, J.M., Álvarez Sánchez, J.R., de la Paz López, F., Toledo-Moreo, F.J., Adeli, H.

(eds.) *Bioinspired Computation in Artificial Systems*, Lecture Notes in Computer Science, vol. 9108, pp. 409–418. Springer International Publishing 2015

WallisDeVries, M.F., Laca, E.A., Demment, M.W. From feeding station to patch: scaling up food intake measurements in grazing cattle. *Appl. Anim. Behav. Sci.* 60, 301–315. 1998

Ives Chauvin & David E. Rumelhart. *Backpropagation: Theory, Architectures, and Applications* 1986