



UNIVERSIDAD INTERNACIONAL SEK

FACULTAD DE CIENCIAS AMBIENTALES

Trabajo de Fin de Carrera Titulado:

Artículo Científico:

“Predicción de consumo eléctrico mensual por cantones en el Ecuador por medio de redes neuronales: Caso de Estudio”

Realizado por:

JORGE ISMAEL GUACHIMBOZA DÁVALOS

Director del proyecto:

Ing. Edilberto Antonio Llanes Cedeño, PhD.

Como requisito para la obtención del título de:
**MAGISTER EN ECOEFICIENCIA INDUSTRIAL CON MENCIÓN EN EFICIENCIA
ENERGÉTICA**

Quito, 01 de octubre de 2020

DECLARACIÓN JURAMENTADA

Yo, JORGE ISMAEL GUACHIMBOZA DÁVALOS con cédula de identidad 1804005872, declaro bajo juramento que el trabajo aquí desarrollado es de mi autoría, que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración, cedo mis derechos de propiedad intelectual correspondientes a este trabajo, a la UNIVERSIDAD INTERNACIONAL SEK, según lo establecido por la Ley de Propiedad Intelectual, por su reglamento y por la normativa institucional vigente.



FIRMA
JORGE ISMAEL GUACHIMBOZA DÁVALOS
C.C: 1804005872

DECLARATORIA

El presente trabajo de investigación titulado:

“Predicción de consumo eléctrico mensual por cantones en el Ecuador por medio de redes neuronales: Caso de Estudio”

Realizado por:

JORGE ISMAEL GUACHIMBOZA DÁVALOS

Como Requisito para la Obtención del Título de:

MAGISTER EN ECOEFICIENCIA INDUSTRIAL

CON MENCIÓN EN EFICIENCIA ENERGÉTICA

Ha sido dirigido por el profesor

Edilberto Antonio Llanes Cedeño

Quien considera que constituye un trabajo original de su autor



FIRMA

LOS PROFESORES INFORMANTES

Los Profesores Informantes:

Rodolfo Jefferson Rubio Aguiar.

Diana Belén Peralta Zurita.

Después de revisar el trabajo presentado,
lo han calificado como apto para su defensa oral ante el
tribunal examinador



FIRMA

Rodolfo Jefferson Rubio Aguiar



FIRMA

Diana Belén Peralta Zurita

DEDICATORIA

A mi madre por apoyarme y
empujarme a ser mejor cada día.

A mis amores Gigi e Ismaelito por
ser mi fuerza y motivación.

A toda mi familia por cada palabra
de aliento, en especial a quienes ya
no están en este mundo, los extraño,
los amo.

AGRADECIMIENTO

A mis padres, a mi esposa y amigos.

A mi tutor Edilberto Llanes y a
todos mis profesores por todas sus
enseñanzas, apoyo y colaboración en
el transcurso de la carrera.

RESUMEN.

El presente trabajo muestra una regresión multivariable mediante aprendizaje profundo relacionado al consumo energético en la provincia de Pichincha-Ecuador en relación de sus cantones. Esto se realiza para establecer hábitos de consumo en relación a la población existente. Para ello se implementó una etapa de preprocesamiento de datos y posteriormente se implementó una red neuronal artificial de una capa oculta y 15 neuronas. Como resultados, se obtiene un modelo matemático con un error absoluto del 2.2%.

Palabras clave: Consumo de Potencia, Regresión Multivariable, Red Neuronal.

ABSTRACT.

The present work shows a multivariate regression through deep learning related to energy consumption in the province of Pichincha Ecuador in relation to its cantons. This is done to establish consumption habits in relation to the existing population. For this, a data preprocessing stage was implemented and later an artificial neural network of a hidden layer and 15 neurons. As a result, a mathematical model with an absolute error of 2.2% is obtained.

Keywords: Power consumption, Multivariate regression, Neural network.

Artículo Científico:

“Predicción de consumo eléctrico mensual por cantones en el Ecuador por medio de redes neuronales: Caso de Estudio”

Jorge I. Guachimboza-Davalos¹, Edilberto A. LLanes-Cedeño¹, Rodolfo J. Rubio-Aguar¹, Diana B. Peralta-Zurita¹, and Oscar F. Núñez-Barrionuevo¹

Universidad Internacional SEK, Quito-Ecuador
jiguachimboza.mee@uisek.edu.ec

RESUMEN.

El presente trabajo muestra una regresión multivariable mediante aprendizaje profundo relacionado al consumo energético en la provincia de Pichincha-Ecuador en relación de sus cantones. Esto se realiza para establecer hábitos de consumo en relación a la población existente. Para ello se implementó una etapa de preprocesamiento de datos y posteriormente se implementó una red neuronal artificial de una capa oculta y 15 neuronas. Como resultados, se obtiene un modelo matemático con un error absoluto del 2.2%.

Palabras clave: Consumo de Potencia, Regresión Multivariable, Red Neuronal.

1 Introducción

En la actualidad, una de las mayores preocupaciones gubernamentales es el calentamiento global. Ya que esto ocasiona un cambio abrupto en las condiciones ambientales del planeta, alterando las condiciones de vida de las personas, animales y plantas [1]. Este problema, es principalmente ocasionado por la desmedida emisión de gases contaminantes (dióxido de carbono CO₂, monóxido de carbono CO, hidrocarburos, óxidos de nitrógeno, azufre, entre

otros) de empresas que utilizan combustibles fósiles para el desarrollo de un producto o servicio [2,3]. Relacionado a esto, se encuentra la producción de emisiones debido a la demanda eléctrica de la población. La misma, que es un elemento clave para el desarrollo socioeconómico de cada país. Dicho esto, el 41% de las emisiones de CO₂ del mundo se atribuye a la generación de electricidad. Las mismas que aportan en su mayoría (alrededor del 99.5%) de CO₂.

En América Latina, existe un promedio de 0.65 toneladas de CO₂ por cada MWh generado[4,5]. En el Ecuador, este resultado es de 397.5g de CO₂ por KWh, donde el mayor valor corresponde al año 2010, ya que la generación de energía de fuentes no renovables fue del (52.2%) [6]. Actualmente, con el cambio de matriz productiva, el 60.85% de producción eléctrica es a base de energía renovable, donde el 58.53% está relacionado a fuentes Hidroeléctricas [7]. Esta producción está orientada en un 30.93% al consumo residencial y un 26.01% al sector Industrial [6]. Con ello, esta forma de uso tiene un factor importante para la proyección de consumo eléctrico y la planificación de nuevas fuentes de generación de electricidad [8]. Sin embargo, el uso de este recurso en hogares e industrias muestran una alta variación que depende del estilo de vida de los usuarios y el tipo de bien o servicio que se genera [15]. Es por ello, que es necesario implementar modelos de predicción que permitan buscar patrones de consumo de las personas en relación a una variedad de atributos (características del consumidor). Sin embargo, estos estudios han sido enfocados al consumo en hogares independientes gracias al crecimiento del internet de las cosas. Por lo tanto, no existen métricas adecuadas de tendencias de consumo organizadas por tipos de usuario domiciliario (residencial y rural) o de diferentes provincias de un país o región.

Además, el principal problema para obtener un pronóstico confiable del consumo de energía eléctrica consiste en elegir un modelo de pronóstico que utilice los datos iniciales en forma precisa. El patrón de demanda es muy complejo debido a la desregulación de los mercados energéticos. Por lo tanto, encontrar un modelo de pronóstico apropiado para una red eléctrica específica no es una tarea fácil. Aunque se han desarrollado muchos métodos de pronóstico, ninguno puede generalizarse para todos los patrones de demanda.

Para ello, los algoritmos de aprendizaje de máquina permiten realizar cálculos matemáticos computacionales para la extracción de conocimiento a partir de los datos pasados. Este criterio es muy utilizado, ya que busca emular cierto funcionamiento del cerebro humano para la toma de decisiones a base de la experiencia. Sin embargo, el análisis de datos tiene sus limitaciones debido a la deficiente estructura de almacenamiento de información por parte de los entes gubernamentales. Esto generalmente es ocasionado por el uso de variables categóricas (que no contienen números). Es por ello que es necesario codificar dichas variables a una forma numérica. Sin embargo, esto ocasiona aumentar el tamaño de la base de datos ya que se requiere de los “unos y ceros” para procesar la información necesaria lo cual puede confundir al modelo de aprendizaje, ya que podría entregar información inexistente al modelo. No obstante, con una adecuada metodología, se pueden realizar subprocesos de limpieza de datos, normalización, selección de características, entre otros que lleve a la obtención de un modelo adecuado. Como resultado, el modelo matemático puede aprender de verdadera información y desestimar las variables que sean de

menor aporte [10]. Con ello, se pueden obtener parámetros de consumo y pronosticar hacia el futuro la demanda eléctrica necesaria para satisfacer el incremento de este servicio.

Existen enfoques clásicos de aprendizaje de máquina que pueden representar el estudio de un fenómeno pero que dependen del tipo de datos a utilizar (categóricos y numéricos). Este modelo de aprendizaje necesita de la extracción de características y un algoritmo supervisado para generar un modelo. Por otro lado, el aprendizaje profundo se basa en uso de redes neuronales, ya que permite realizar las etapas anteriormente comentadas en un modelo computacional. Este nuevo enfoque se basa en redes neuronales [11]. Debido a que este criterio permite organizar por sí mismo lo aprendido (autoorganización), por ende, tienen la capacidad de ser tolerante a fallos, ya que mientras más veces entre en funcionamiento (itere) puede solventar su error (actualización de pesos). Es por ello, que no puede sufrir caídas repentinas del sistema [12].

Algunos trabajos como [13,14] presentan sistemas de medición de consumo eléctrico en domicilios para almacenarlos y analizarlos en la nube. Por otro lado, no presentan una solución de sector específico y se basa en tendencias de consumo por persona. En este caso, la aplicación está enfocada al usuario final y no a un análisis a nivel de estado. Es por ello, que el presente trabajo, realiza un análisis de consumo de diferentes por cantones de la provincia de Pichincha en el Ecuador con el objetivo de determinar parámetros y tendencias de consumo para presentar propuestas de demanda energética hacia el futuro. Para ello, se ha recopilado la información proveniente de la Agencia de Regulación y Control de Electricidad para implementar un

aprendizaje profundo mediante el uso de redes neuronales. Para ello, se realiza una etapa de preprocesamiento de datos, que implica un proceso de selección de características y normalización de la base de datos. Posteriormente, se realiza un análisis de las funciones de activación en las diferentes capas de una red neuronal para determinar el modelo adecuado que pueda brindar la mejor predicción de valores deseados de consumo eléctrico. Para esto, finalmente se realizan diferentes métricas para el cálculo del error de pronóstico a fin de validar el modelo. Como resultados relevantes, el modelo tuvo un error de pronóstico del 2.2%.

El resto del documento se encuentra estructurado de la siguiente manera: La sección 2, muestra el esquema metodológico de análisis de datos. La sección 2 presenta también el análisis de la red neuronal para brindar el mejor pronóstico de consumo eléctrico. La sección 3 se muestran los resultados obtenidos con los diferentes análisis de error. Finalmente, en la sección final se presentan las conclusiones y trabajos futuros.

2 Materiales y Métodos

La presente sección muestra, por un lado, la adquisición de datos (sección 2.1), preprocesamiento de la información usando la selección de características (sección 2.2) y la normalización de la base de datos (sección 2.3). Además, presenta el esquema de análisis de datos (sección 2.4).

2.1 Adquisición de datos

La recopilación de datos es sobre la base de la información presentada por el Ministerio de Electricidad y Energía Renovable en sus rendiciones de cuentas mensuales durante el año 2019. En estos reportes se explica a detalle datos referente a la empresa que distribuye la electricidad a las provincias, cantones y parroquias. En este caso específico se toma la información del consumo de los años 2019 y 2018. Todos los datos son almacenados en la matriz $Y \in \mathbb{R}^{m \times n}$ donde m representa el número de muestras y n los atributos de la base de datos. En este caso, $m = 2550$ y $n = 11$. Los atributos se refieren a: Año de consumo, mes, empresa que factura, provincia, cantón, parroquia, tipo de equipamiento con voltaje a 220V, número de clientes, energía facturada, incremento de consumo, consumo residencial, energía de subsidio generado, facturación del servicio y valor del subsidio.

2.2 Selección de características

La elección de un conjunto de entrenamiento para los algoritmos de predicción es un aspecto crucial, ya que el tiempo de respuesta debe ser el adecuado [15]. Normalmente, para una tarea de análisis supervisado de datos, una gran cantidad de atributos pueden ser candidatos para fines de caracterización. Sin embargo, muchos de estos atributos pueden ser irrelevantes o redundantes y, en consecuencia, el algoritmo puede sufrir un sobreajuste y algunas características importantes pueden verse afectadas. Por lo tanto, el rendimiento o aceptación del modelo podría no alcanzar lo esperado [16]. De hecho, en muchas aplicaciones que involucran grandes conjuntos de datos, los pronósticos no funcionan correctamente hasta que las características no deseadas se eliminan obligatoriamente [15].

En este sentido, la tarea de selección de características permite reducir el número de atributos lo cual representa un conjunto de datos que mejor se ajusta a un modelo o tarea de aprendizaje automático. Existe una amplia gama de criterios y algoritmos que se pueden utilizar para esta tarea. Sin embargo, dependiendo de la naturaleza de los datos y el objetivo del análisis de datos, algunos criterios de selección de características pueden resultar más adecuados. En este trabajo se toman en consideración los criterios más importantes. Por un lado, se toma en cuenta la correlación de variables en un modelo estadístico bajo los criterios de p_{valor} y R^2 con el uso del criterio “backward elimination”. Esta combinación indica que los cambios en los predictores están relacionados con los cambios en la variable de respuesta y que el modelo explica en gran parte la variabilidad de la respuesta. Con ello, se puede obtener un estimado de cómo el modelo representa al fenómeno estudiado [16].

2.3 Normalización de la base de datos

La normalización de índices significa ajustar los valores medidos en diferentes escalas respecto a una en común. Este proceso se realiza previo al proceso para obtener modelos matemáticos. Con esto, se evita que existan variables que, acorde su propia naturaleza del dato, aporte en gran cantidad al modelo al tener valores muy altos en relación al resto. Además, evita que el uso de las variables categóricas se vea mermado debido a su baja escala (0 a 1). Es por ello que existen diferentes métodos y formas de normalización.

2.4 Esquema de análisis de datos

El presente trabajo está orientado a encontrar el mejor modelo que se ajuste al estudio presentado donde pueda pronosticar el consumo eléctrico de la población ecuatoriana estableciendo las variables más representativas. En la Fig. 1 se muestra el esquema propuesto.

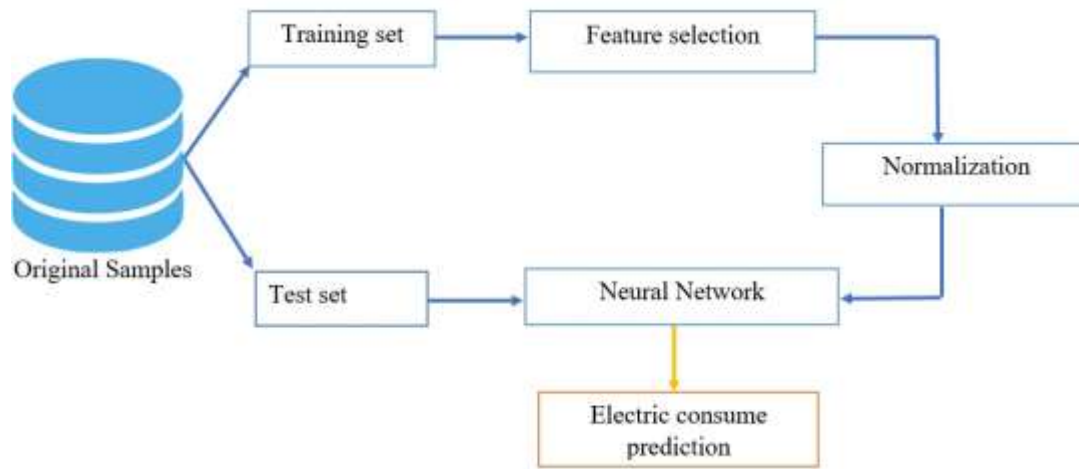


Fig. 1. Esquema de Análisis de datos

2.5 Entrenamiento de la red neuronal

La información obtenida por la red neuronal no puede transmitir todo lo aprendido, necesita una función de activación representando un potencial de activación potencial (una forma binaria de la probabilidad de suceso). En este sentido, la función de activación permite la actualización de pesos (aprendizaje de la neurona). En la ecuación (1) se muestra el funcionamiento de una red neuronal para predicción. Así también, su representación se observa en la Fig. 2.

$$Y = f\left(\sum_{i=0}^{i=m} \sum_{j=0}^{j=n} x_i w_{ij}\right) \quad (1)$$

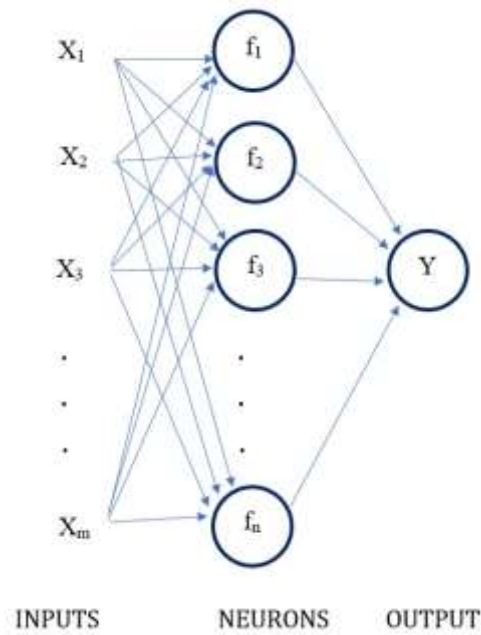


Fig. 2. Esquema de la Red Neuronal diseñada

3 Resultados

Los resultados del presente trabajo se encuentran divididos en: Selección de características y los resultados de la red neuronal.

3.1 Selección de características

La codificación de las variables categóricas se realiza en relación al cantón ya que por parroquia el modelo tendría una codificación muy alta y aportaría más información (33 parroquias rurales y 32 parroquias urbanas) que las variables numéricas. En consecuencia, la codificación de catones se muestra en la tabla 1.

Tabla 1. Codificación categórica de los cantones de Pichincha

Cantones	cod1	cod2	cod3	cod4	cod5	cod6	cod7
Cayambe	1	0	0	0	0	0	0
Quito	0	1	0	0	0	0	0
Pedro Vicente Maldonado	0	0	1	0	0	0	0
Puerto Quito	0	0	0	1	0	0	0
Rumiñahui	0	0	0	0	0	1	0
San Miguel de los Bancos	0	0	0	0	0	0	1

Posteriormente, se realiza un análisis de covarianza entre variables para conocer su índice de significancia y su aporte al modelo (R cuadrado y p valor). Como se puede apreciar, en la Fig. 3 las variables relacionadas al tipo de equipamiento y el subsidio generado no presentan información relevante.

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.968e+01  2.446e+01  2.031  0.0426 *
i..Canton2  -2.136e+01  4.352e+01 -0.491  0.6238
i..Canton3   8.331e+00  3.431e+01  0.243  0.8082
i..Canton4  -6.127e+01  4.065e+01 -1.507  0.1322
i..Canton5   5.440e+01  6.688e+01  0.813  0.4162
i..Canton6   1.893e+02  8.257e+01  2.293  0.0221 *
i..Canton7  -1.960e+01  9.091e+01 -0.216  0.8294
Equipamiento -1.672e-01  3.208e-01 -0.521  0.6023
Numero.Clientes -1.391e+00  8.446e-02 -16.471 <2e-16 ***
Energia.Facturada.k.wh  1.071e-01  1.294e-03  82.750 <2e-16 ***
Consumo.Incremental.kwh. -1.232e-02  8.861e-04 -13.902 <2e-16 ***
Consumo.Residencial.kwh. -7.286e-03  8.133e-04 -8.958 <2e-16 ***
Energia.Subsidio.PEC.kwh.  1.575e+02  1.341e+02  1.175  0.2404
ValorSubsidio.PEC.USD.  -1.750e+03  1.490e+03 -1.175  0.2404
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 269.3 on 818 degrees of freedom
Multiple R-squared:  0.9997,    Adjusted R-squared:  0.9997
F-statistic: 2.186e+05 on 13 and 818 DF,  p-value: < 2.2e-16

```

Fig. 3. Análisis de covarianza con una matriz de atributos Y .

En este sentido, se procedió a eliminar las variables menos relevantes con el objetivo de mejorar el modelo. En la Fig. 4 se muestra el resultado del análisis de covarianza donde indica que el ajuste del modelo no ha variado (R cuadrado) al eliminar las variables anteriormente comentadas.

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.251e+01  1.372e+01   2.369  0.0181 *
i..Canton2    -1.582e+01  4.325e+01  -0.366  0.7147
i..Canton3     1.156e+01  3.415e+01   0.339  0.7350
i..Canton4    -5.909e+01  4.064e+01  -1.454  0.1463
i..Canton5     5.637e+01  6.677e+01   0.844  0.3987
i..Canton6     1.905e+02  8.253e+01   2.308  0.0213 *
i..Canton7    -1.220e+01  9.088e+01  -0.134  0.8932
Numero.Clientes -1.344e+00  7.744e-02 -17.362 <2e-16 ***
Energia.Facturada.k.wh  1.084e-01  1.048e-03 103.433 <2e-16 ***
Consumo.Incremental.kwh. -1.197e-02  8.430e-04 -14.196 <2e-16 ***
Consumo.Residencial.kwh. -7.690e-03  7.775e-04  -9.890 <2e-16 ***
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 269.6 on 821 degrees of freedom
Multiple R-squared:  0.9997,    Adjusted R-squared:  0.9997
F-statistic: 2.837e+05 on 10 and 821 DF,  p-value: < 2.2e-16

```

Fig. 4. Análisis de covarianza con una matriz de atributos X .

3.2 Resultados de la Red neuronal

Una vez seleccionadas las variables para implementarlas en la red neuronal. Se establecen un mínimo de 10 neuronas para entrenar el modelo y probar su funcionamiento. Como resultado, se establecen 15 neuronas, las que en conjunto representan el menor error de dispersión de pronóstico (MSE) y el menor error de desviación media absoluta (MAD). Este valor es monetario. En la Tabla 2 se muestra dicho análisis (Véase Fig.3) .

Tabla 2. Análisis del Error en función del número de neuronas.

Número de Neuronas	MSE	MAD
10	4.5%	125
12	3.8%	75
15	2.2%	45
17	2.2%	45
20	2.2%	45

Con ello, se puede establecer una predicción del consumo eléctrico en relación al pago del retorno. En la Fig. 5 se muestra el pronóstico generado por la red neuronal.

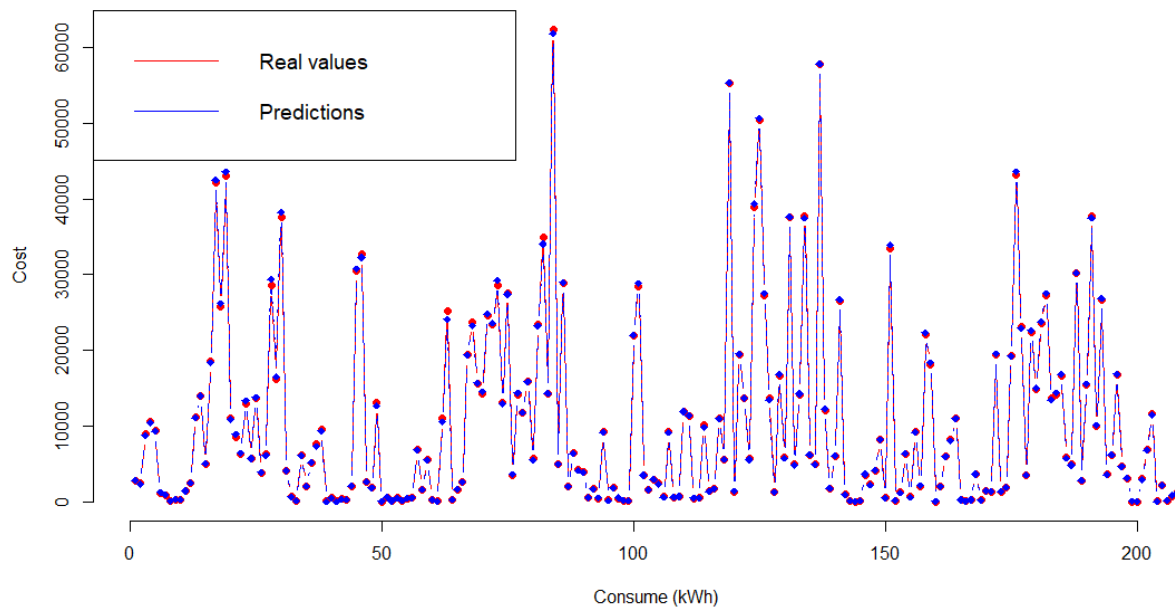


Fig. 5. Análisis de las Predicciones del costo obtenidas de la Red Neuronal

Finalmente, se realiza un resumen de todas las parroquias para tener un resumen final del consumo por cantón en relación al año 2019. Como se puede apreciar en la Fig. 6, el pronóstico es muy acertado al verdadero consumo.

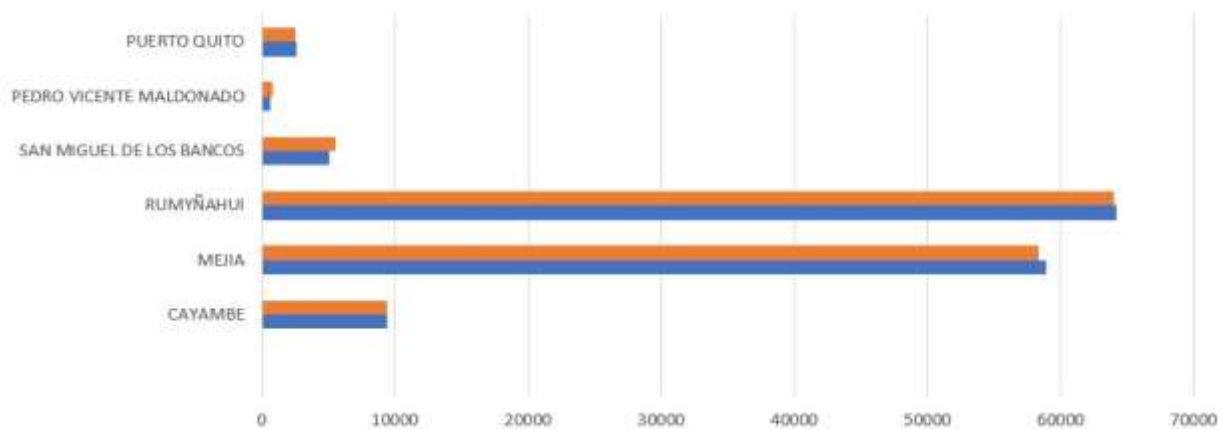


Fig. 6. Comparación de Resultados del Modelo de Regresión obtenido

4 Conclusiones y trabajos futuros

El presente trabajo pudo comprobar el funcionamiento del aprendizaje a partir de las redes neuronales. Con ello, la predicción de la variable objetivo fue muy cercano al valor real. En este sentido, se puede concluir que, se puede usar el modelo obtenido para definir estrategias hacia el futuro con el objetivo de mejorar el servicio eléctrico y establecer políticas claras para un mejor uso del recurso energético de país.

Como trabajos futuros se plantea realizar un análisis más profundo de la información en relación al consumo energético del país relacionado por parroquias

Referencias Bibliográficas:

1. Ye, X., Wu, X., Guo, Y.: Real-time quality prediction of casting billet based on random forest algorithm. In: Proceedings of the 2018 IEEE International Conference on Progress in Informatics and Computing, PIC 2018, pp. 140–143. Institute of Electrical and Electronics Engineers Inc., May 2019.
2. Liu, Y., Wang, W., Ghadimi, N.: Electricity load forecasting by an improved forecast engine for building level consumers. *Energy* 139, 18–30 (2017).
3. UEM en Funcionamiento - Ministerio de Educación. <https://educacion.gob.ec/ uem-en-funcionamiento/>
4. Zhang, X.M., Grolinger, K., Capretz, M.A., Seewald, L.: Forecasting residential energy consumption: single household perspective. In: Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018, pp. 110–117. Institute of Electrical and Electronics Engineers Inc., January 2019
5. Diao, L., Sun, Y., Chen, Z., Chen, J.: Modeling energy consumption in residential buildings: a bottom-up analysis based on occupant behavior pattern clustering and stochastic simulation. *Energy Build.* 147, 47–66 (2017)
6. Ministerio de Electricidad y Energía Renovable - Ente rector del Sector Eléctrico Ecuatoriano. <http://historico.energia.gob.ec/>
7. Parra Narváez, R.: Factor de emisión de CO₂ debido a la generación de electricidad en el Ecuador durante el periodo 2001–2014. In: *Avances en Ciencias e Ingeniería*, vol. 7, no. 2, December 2015
8. Baquero, M., Quesada, F.: Eficiencia energética en el sector residencial de la Ciudad de Cuenca, Ecuador. *Maskana* 7(2), 147–165 (2016)
9. Zhao, H.X., Magoules, F.: A review on the prediction of building energy consumption. *Renew. Sustain. Energy Rev.* 16(6), 3586–3592 (2012)
10. Toapanta-Lema, A., Gallegos, W., Rubio-Aguilar, J., Llanes-Cedeño, E., Carrascal García J., García-López, L., Rosero-Montalvo, P.D.: Regression models comparison for efficiency in electricity consumption in ecuadorian schools: a case of study. In: Botto-Tobar, M., Zambrano Vizuite, M., Torres-Carrión, P., Montes León, S., Pizarro Vásquez, G., Durakovic, B. (eds.) *Applied Technologies*, pp. 363–371. Springer, Cham (2020)
11. Wang, X., Lai, C., Yu, D., Xu, Y., He, Y.: Application of energy performance contracting in distribution transformer based on neural network prediction. In: 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), pp. 42–46, March 2019
12. Petkovic, I., Balaban, N.: Detecting defaulters for consumed electric energy with neural network clustering. In: 2008 6th International Symposium on Intelligent Systems and Informatics, pp. 1–4, September 2008
13. Rezaei, R., Sharghi, A., Motalebi, G.: A framework for analysis affecting behavioral factors of residential buildings' occupant in energy consumption. *J. Sustain. Archit. Urban Des.* 5(2), 39–58 (2018). http://jsaud.sru.ac.ir/article_895.html%0A, http://jsaud.sru.ac.ir/pdf/895_9f6e3441f7399851f2185e37696ed98e.html
14. Johnson, B.J., Starke, M.R., Abdelaziz, O.A., Jackson, R.K., Tolbert, L.M.: A method for modeling household occupant behavior to simulate residential energy consumption. In: IEEE PES Innovative Smart Grid Technologies Conference, ISGT 2014 (2014)
15. Zhai, Y., Song, W., Liu, X., Liu, L., Zhao, X.: A chi-square statistics based feature selection method in text classification. In: 2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS), pp. 160–163, November 2018

16. Peker, M., Arslan, A., Sen, B., Celebi, F.V., But, A.: A novel hybrid method for determining the depth of anesthesia level: combining relief feature selection and random forest algorithm. In: 2015 International Symposium on Innovations in Intelligent Systems and Applications (INISTA), pp. 1–8, September 2015