

Universidad Nacional de Asunción

Facultad de Ingeniería



Propuesta de Trabajo Final de Grado

“Diseño de un modelo de Deep Learning basado en la arquitectura Temporal Fusion Transformer para la predicción horaria a corto plazo de la demanda eléctrica nacional”

Por:

- Silvio José Aguilar Velazco
- Santiago Dionisio Vargas García

Carrera: Ingeniería Mecatrónica.

Tutores:

- Ing. Carlos Sauer
- Dr. Diego Stalder
- Ing. Félix Morales

Noviembre 2021

Introducción

La Administración Nacional de Energía (ANDE) tiene como una de sus prioridades principales la de satisfacer de forma adecuada las necesidades de energía eléctrica del país, teniendo en cuenta la previsión de la demanda de consumo eléctrico a nivel nacional, factor primordial para poder tomar las mejores y más eficientes decisiones en ámbitos técnicos, administrativos y operativos [1]. Podemos destacar dos principales problemas que se pueden presentar si no se realiza de forma correcta este proceso: por un lado, la adquisición excesiva e innecesaria de energía, que puede desencadenar en pérdidas económicas para la entidad, o por otro lado, no contar con la cantidad necesaria para satisfacer la demanda mínima, repercutiendo directamente en los gastos del consumidor ya que se deberá recurrir a generación de contingencia, e incluso en un caso peor, generar apagones generales a niveles relativamente grandes. Cabe destacar además que toda la energía eléctrica que no se consume a nivel nacional es cedida o vendida a países vecinos, teniendo en cuenta que estos acuerdos no son necesariamente bien aprovechados en beneficio de nuestro país.

La planificación que realiza la ANDE, en distintas escalas de tiempo (diario, mensual y anual), y para distintos escenarios cuenta con cierta incertidumbre, considerando distintos factores que afectan directa e indirectamente a la demanda que se quiere predecir.

Existen modelos de previsión de carga basados en la estadística, considerados tradicionales, como la regresión lineal, así como modelos semi-analíticos compuestos a partir del perfil de consumo de clientes, e incluso, modelos basados en redes neuronales profundas [5]. Estos modelos fueron utilizados previamente para la predicción de demanda eléctrica, siendo considerados impulsores del estudio en el área. En el mismo se establece que no existe un consenso que determine el uso particular de un modelo o método por encima de otro, generalmente suele ser resultado de las métricas de desempeño para seleccionar el mejor modelo, por lo cual la práctica común es determinar qué modelo se ajusta mejor para una determinada situación.

Como complemento, el estudio y aplicación de redes neuronales se ha convertido en una atracción para los investigadores del área, ya que han sido utilizadas para predicción de tareas que incluyen el pronóstico de carga a corto plazo (STLF: Short Term Load Forecasting) en microrredes [10].

En 2017 se introduce la arquitectura Transformer por parte de un grupo de investigadores de Google [6], creando una revolución en el área de procesamiento natural del lenguaje, implementando mecanismos de atención que son utilizados para determinar la importancia o peso de una palabra durante el proceso de traducción a otro idioma. La misma tiene una variante que se puede aplicar a la predicción de series temporales, denominada arquitectura Temporal Fusion Transformer(TFT) [7]. Esta última será tomada como línea de base para la realización de este Trabajo Final de Grado.

Objetivos

General

- Diseñar un modelo inteligente TFT (Temporal Fusion Transformer) para la predicción horaria a corto plazo de la demanda eléctrica nacional.

Específicos

- Actualizar datos históricos de suministro de energía proveída por la ANDE.
- Validar datos de variables meteorológicas a nivel país e imputar datos faltantes.
- Implementar un modelo de predicción de demanda de energía basado en la arquitectura TFT.
- Realizar un ajuste fino de los hiper-parámetros.
- Comparar diferentes modelos basados en la Deep Learning para forecasting (previsión), incluyendo trabajos finales de grado de compañeros de FIUNA relacionados al área.

Alcance y Justificación

Las redes del Sistema Interconectado Nacional abarcan un conjunto de líneas y subestaciones, incluyendo sus equipos asociados, que transportan la energía desde las plantas de generación a las subestaciones de transformación y finalmente al consumidor final. La planificación de su operación debe asegurar un suministro de energía suficiente y de bajo costo, teniendo en cuenta diversos aspectos y factores, como la cantidad de generadores inhabilitados o equipos que no se encuentran en funcionamiento, además de factores externos y no controlables como la inestabilidad climática o los días especiales durante el año (fines de semana y feriados). Es posible afirmar en base a esto, que la demanda de potencia a nivel global en el país condiciona directamente la contratación de potencia a las centrales eléctricas, que abarcaría la mayor parte del costo operativo de la ANDE.

Teniendo en cuenta el campo y el ámbito en el que se profundizará, existen trabajos realizados con anterioridad que incursionaron en el diseño de modelos para predicción de demanda energética, basándose en la LSTM (long-short term memory), una arquitectura evolucionada de las Redes Neuronales Recurrentes, además de otros que lo aplicaron en el entorno MatLab® Neural Network Toolbox, considerando factores meteorológicos y de calendario [9,10,11].

Si bien, en los modelos de predicción en el entorno de MatLab se arrojaron de forma satisfactoria tendencias de la curva en una buena parte de los días del año, la precisión de los mismos se veía afectada en picos altos y bajos en ciertos periodos, por motivos que pueden estar relacionados al clima o bien, información ineficaz a la hora de ser agrupada [10,11] . Además utiliza la temperatura como atributo para la predicción, sin tener en cuenta otros factores meteorológicos o fiestas festivas u otros factores que podrían influir en la demanda eléctrica.

Aprovechando la optimización realizada en el enfoque del modelo LSTM (Long Short-Term Memory)[9], el método a utilizar en este Trabajo Final de Grado será el de pronóstico de múltiples horizontes, mediante series temporales y Machine Learning, cuya función es predecir múltiples pasos en el futuro permitiendo al usuario acceso a información durante un periodo de tiempo a diferencia de las predicciones de un punto por paso. De forma más específica, el modelo con el que se trabajará será una red TFT [7]. Su diseño se realizará utilizando softwares de carácter open-source, en un entorno de Python con sus respectivas librerías específicas para Machine Learning, teniendo en cuenta también otras arquitecturas como soporte, además de utilizar como un constante respaldo otros métodos estadísticos junto con los datos históricos existentes del SIN.

El uso de la arquitectura TFT permitirá acceder a un nivel de explicación e interpretabilidad superior al modelo presentado un año antes en cuanto al comportamiento dinámico de las series temporales, mediante el uso de mapas de calor en base a los mecanismos de atención [12].

Se espera que este Trabajo Final de Grado pueda servir como herramienta eficiente de proyección, logrando así un equilibrio entre oferta y demanda en las distintas operaciones referentes al despacho de carga de la ANDE, ya que desde ese sector se podrá hacer uso de las curvas obtenidas por el modelo de la demanda predicha y tomar en base a las mismas las mejores decisiones operativas. Es importante mencionar que, así como se partió desde la concepción de la idea de este trabajo, el mismo puede servir como base para próximas investigaciones que buscan optimizar las mencionadas predicciones en distintos plazos.

Metodología y Plan de Trabajo

El modelo trabaja sobre observaciones pasadas así como sobre covariables estáticas de cada grupo de datos, los cuales se clasifican en “conocidos”, pudiendo ser un día de la semana o un día festivo, y “sin conocer”, como por ejemplo la temperatura, la humedad, la velocidad del viento, entre otros, con el fin de predecir valores. Además de la obtención de valores se quiere determinar los cuantiles de dichos valores, ya que la predicción es no determinista.

La arquitectura TFT está conformada por dos regiones: una de codificación y otra de decodificación (Ver Fig. 1). Entre los principales componentes de la arquitectura TFT se encuentran los mecanismos de atención, los cuales son productos escalares dentro de un espacio vectorial embebido donde se trata de obtener el efecto de una variable sobre otra y así determinar un peso que represente la importancia de dicha variable. Este conocimiento es aprendido por el modelo durante el entrenamiento, el cual al principio es por medio de un entrenamiento forzado donde el modelo puede conocer a priori los resultados. Como las entradas de estos mecanismos son vectores en un espacio embebido, se utilizan codificadores teniendo en cuenta variables estáticas.

La arquitectura de TFT ya fue testeada con el conjunto de datos de consumo de energía de Estados Unidos [11]. El mismo contiene el consumo de energía de 370 usuarios, del cual se utiliza el registro de una semana (168 horas) para predecir las siguientes 24 horas.

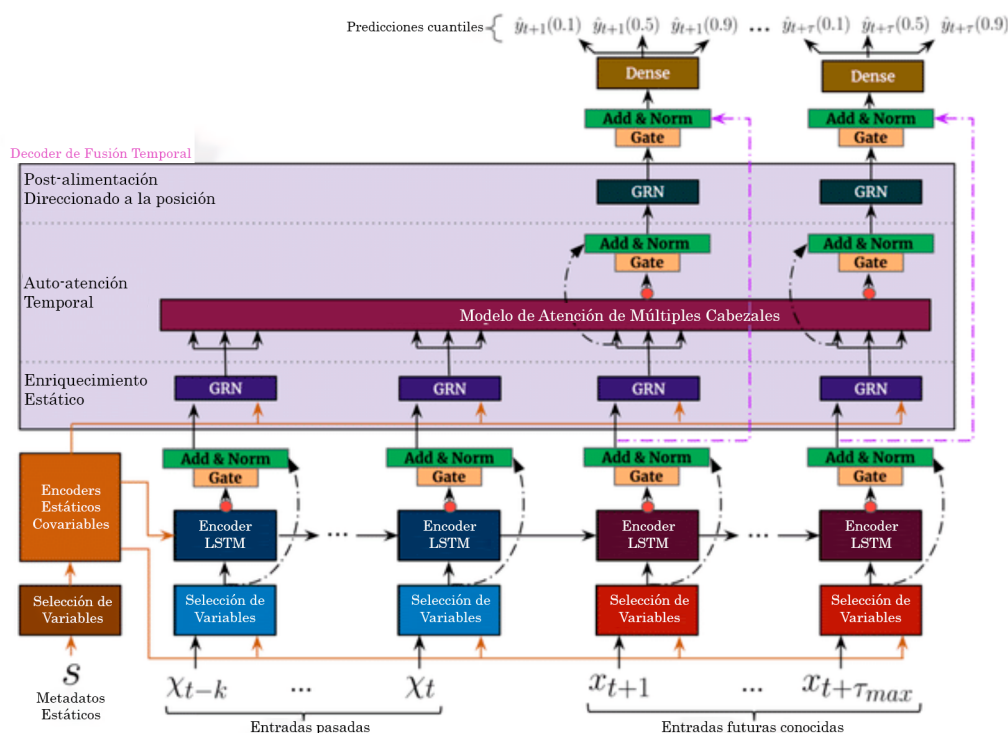


Figura 1: Arquitectura Temporal Fusion Transformer [7]

Adoptaremos la metodología por paquetes de trabajo (WP) y la metodología de trabajo ágil, enfocada a los objetivos duros de los WP, dividiendo los objetivos en otros objetivos a trabajarlos como sprints, a ser definidos al empezar cada WP. La duración trabajo se estima en 6 meses, cada etapa será dividida por una estimación de

[WP1]: Revisión bibliográfica (Duración estimada: 4 semanas)

[WP2]: Diseño de la red neuronal artificial(Duración estimada: 9 semanas)

[WP3]: Proceso de entrenamiento y validación (Duración estimada: 4 semanas)

[WP4]: Verificación y comparación de los resultados por medio de métodos estadísticos, y otros modelos neuronal (Duración estimada: 8 semanas)

[WP5]: Redacción del libro de tesis (Duración estimada: 25 semanas) tiempo.

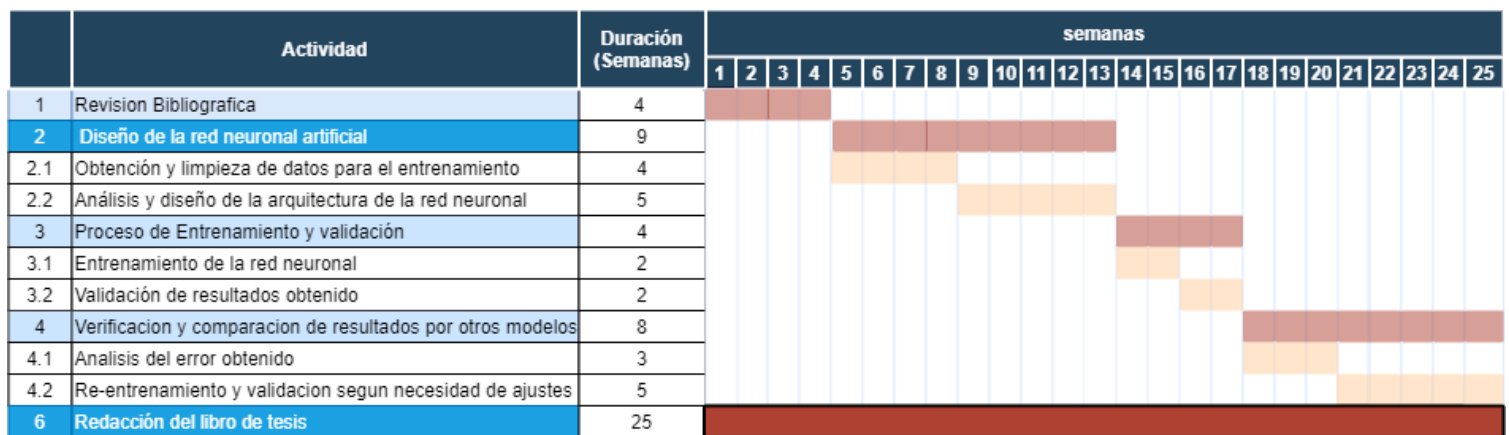


Figura 2: Diagrama de Gantt

Bibliografía

1. ANDE. (2016). *Plan maestro de generación*.
2. Azadeh, A. y Faiz, Z. (2011). *An integrated artificial neural network-computer simulation for optimization of complex tandem queue systems*.
3. Hou, Z. y Lian, Z. (2009). *An Application of Support Vector Machines in Cooling Load Prediction*.
4. Hasni, A. y Belkacem, D. (2011). *Applying Time Series Analysis Model to Temperature Data in Greenhouses*
5. Kuster, C., Rezgui, Y., & Mourshed, M. (2017). Electrical load forecasting models: a critical systematic review. *Sustainable cities and society*.
6. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., & kaiser, L. (2017, 12 6). Attention Is All You Need.
7. Lima, B., Arik, S. O., Loeffb, N., & Pfister, T. (2020). Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting.
8. Manchado, M. C. (n.d.). *Predicción de la demanda eléctrica española. Implementación de redes neuronales recurrentes en Python*.
9. Zhifeng, L. & Lianglun, Ch. (2020). *Electricity consumption prediction based on LSTM with attention mechanism*
10. Parra, G., & Verdún, J. (2015). *Metodología para estimación de demanda eléctrica a corto plazo de una subestación utilizando redes neuronales artificiales* [Biblioteca FP-UNA].
11. Barboza, O. A. (2014). *Automatización de previsión de demanda horaria de potencia eléctrica en el Sistema Interconectado Nacional Hourly Load Forecast Automation in the National Interconnected System* [Revista Científica de la UCSA].
12. Verdún Bello, M. E., & Morales Mareco, F. J. (2020). *Diseño de un Modelo Inteligente para la Predicción Horaria a Corto Plazo de la Demanda Eléctrica Nacional Basado*

en Aprendizaje Profundo y Propuesta de Implementación en el Despacho de Carga de la ANDE.

13. Bouktif, S., Fiaz, A., Ouni, A., & Serhani, M. (2018). *Optimal deep learning lstm model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches.*