



Pronóstico de generación de energía reactiva en una planta de productos lácteos

Estudiantes:

Javier Andrés Causil Martínez

Andrés Borja Goez

Monografía presentada para optar al título de Especialista en Analítica y Ciencia de Datos

Asesor

Gabriel Dario Uribe Guerra

Universidad de Antioquia

Facultad de Ingeniería

Especialización en Analítica y Ciencia de Datos

Medellín, Antioquia, Colombia

2024

Índice general

1. Descripción del problema	1
1.1. Problema de negocio	1
1.2. Aproximación desde la anítica	2
1.3. Origen de los datos	2
1.4. Métricas de desempeño	2
1.5. Modelos	3
1.5.1. ARIMA	3
1.5.2. RNN	3
1.5.3. LSTM	3
2. Objetivos	4
2.1. Objetivos	4
2.2. Objetivos específicos	4
3. Datos	5
3.1. Datos originales	5
Bibliography	6

Índice de figuras

3.1. Energía reactiva	6
---------------------------------	---

©MONOGRAFÍA

MONOGRAFIA

1 Descripción del problema

Una empresa de productos lácteos requiere disminuir los costos asociados al consumo de energía en una planta de producción. Se aplicarán modelos para pronosticar la generación de energía en una planta de producción. Se aplicarán modelos para pronosticar la generación de energía reactiva que afecta el consumo pero no agrega valor a la empresa. Se cuenta con datos historial de consumo de energía eléctrica desde el año 2020 y con datos capturados por sensores instalados en las máquinas de la planta desde el año 2023. Las métricas de desempeño son las asociadas a la precisión de los modelos y el valor de la facturación mensual.

1.1. Problema de negocio

Una empresa de productos lácteos requiere disminuir los costos asociados al consumo de energía en una planta de producción.

Según **CELSIA** (2022), la energía reactiva es un tipo de energía eléctrica absorbida o inyectada a la red por algunos equipos que para su funcionamiento necesitan un campo magnético, tales como motores, transformadores, ascensores, sistemas de bombeo de agua, motores de aireación de piscinas, iluminación eficiente, entre otros. La unidad de medida de este tipo de energía es $kVarh$. En complemento, de acuerdo con (EPM, 2023), la energía reactiva se puede entender como una energía que ocupa espacio de las redes eléctricas, pero no es útil a la hora de hacer trabajo. Como esta energía que ocupa espacio de las redes eléctricas, pero no es útil a la hora de hacer trabajo. Como la energía reactiva satura las redes, es necesario para las empresas reducirla a su mínima expresión para evitar problemas en la calidad de la energía, sobrecargas e ineficiencias que redundaría en mayores costos para prestar el servicio.

De acuerdo con lo indicado por el personal técnico de la Planta, en los últimos años, el nivel de generación de energía reactiva en la planta ha ido en aumento, y aunque la información de energía global se conoce, no es posible identificar la totalidad de las fuentes que la están generando, lo cual tiene un impacto significativo sobre los costos de producción. La energía reactiva puede ser generada por múltiples fuentes (motores, equipos eléctricos, instalaciones, etc.) y puede variar según el estado de la maquinaria, la tecnología de las máquinas, el mantenimiento y otras variables no identificadas, por lo cual no es posible discriminar el aporte de cada fuente al sobre costo establecido por el prestador de servicios de energía, lo cual dificulta la toma de decisiones para optimizar su uso.

La Comisión Reguladora de Energía y Gas CREG, mediante Resoluciones 15 de 2018 (CREG, 2018), 199 de 2019 (CREG, 2020) ha emitido los lineamientos con los cuales se determina la variable M que establece el nivel en el que se encuentra una empresa, de acuerdo

con la cantidad energía reactiva generada en un determinado periodo de tiempo. Esta variable es la que utilizan los prestadores de servicios de energía para establecer las penalizaciones aplicadas por sobrepasar los límites establecidos en la normativa. En el año 2024 la Empresa se encuentra en el nivel 6 de 12 niveles posibles y ha venido en aumento, donde el nivel 12 es el de mayor sanción. Lo anterior se ve reflejado en un aumento en los pagos realizados a EPM asociados a la penalización por exceso de generación de energía reactiva. La situación se vuelve en un tema crítico para la gerencia de la Planta, al ver que los costos de producción aumentan por la generación de energía reactiva, y no se tiene establecido un plan para controlarla.

1.2. Aproximación desde la anítica

Se aplicarán modelos estadísticos y de machine learning para pronosticar la generación de energía reactiva que afecta el consumo energético, pero no agrega valor a la empresa. Se espera que, con los resultados del modelo, el equipo técnico pueda tomar decisiones para controlar la energía reactiva generada que lleve a la reducción del valor de la facturación mensual.

1.3. Origen de los datos

Para este proyecto se cuenta con información de la energía activa y reactiva por hora para el periodo comprendido entre febrero de 2020 y septiembre de 2024. Estos datos son proporcionados directamente por la empresa prestadora de servicios de energía eléctrica y son un instrumento para el control del pago de la facturación mensual.

Adicionalmente, se tienen mediciones de los sensores de los equipos, los cuales toman los datos por segundo de valores asociados al estado de los equipos para el seguimiento en la producción desde el año 2023, los cuales son almacenados en la nube a través del servicio de AWS.

1.4. Métricas de desempeño

Para evaluar nuestros modelos utilizaremos métricas de desempeño que nos permitan medir la precisión de los modelos. Se considerarán las siguientes métricas de desempeño del modelo utilizando la librería Scikit-learn [3]:

1. Error absoluto medio (MAE): Nos arroja la diferencia promedio absoluta entre los valores pronosticados y los valores reales. Es decir que tan distante están los valores pronosticados de los valores reales.

$$MAE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^{n-1} |y_i - \hat{y}_i|, \quad (1.1)$$

2. Error cuadrático medio (RMSE): Similar al MAE, pero considera los errores al cuadra-

do, penalizando más los errores grandes. Un RMSE bajo indica mayor precisión.

$$RMSE(y, \hat{y}) = \sqrt{\frac{1}{n} \sum_{i=1}^{n-1} (y_i - \hat{y}_i)^2}, \quad (1.2)$$

3. Error porcentual absoluto medio (MAPE): Expresa el error en términos porcentuales, siendo útil para comparar series de tiempo con diferentes escalas.

$$MAPE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^{n-1} \left| \frac{y_i - \hat{y}_i}{y_i} \right|. \quad (1.3)$$

1.5. Modelos

1.5.1. ARIMA

El modelo ARIMA (Autoregressive Integrated Moving Average) es un modelo estadístico que se utiliza para analizar y predecir series temporales. Se utilizará ARIMA para pronosticar la generación de energía reactiva en la planta de producción.

1.5.2. RNN

1.5.3. LSTM

2 Objetivos

2.1. Objetivos

Desarrollar un modelo utilizando técnicas estadísticas y de aprendizaje automático para pronosticar la generación de energía reactiva en una planta de productos lácteos, con datos del consumo energético en el periodo 2020 a 2024, que permita disminuir los costos asociados al consumo de energía en la planta.

2.2. Objetivos específicos

- Realizar los pasos contemplados en la metodología CRISP-DM para la creación de los modelos de analítica.
- Evaluar la precisión de los modelos aplicados considerando errores inferiores al 30 %.
- Establecer, mediante el uso del modelo, los días del mes que hacen que la generación de energía reactiva sobrepase los 12 días permitidos.

3 Datos

3.1. Datos originales

El conjunto de datos encrudo son archivos en formato xlsx, desde el 2020 hasta el 2024, donde se tiene el reporte por hora de:

- Energía Activa Consumo (kWh).
- Energía Activa Generación (kWh).
- Energía Reactiva Inductiva (kVarh).
- Energía Reactiva Capacitiva (kVarh).
- Fecha.

Se presentan unos datos atípicos enero de 2021 a junio 2021, explicados por paros, apagones en la empresa. Debido a que estás anomalías son provocadas por sucesos ajenos a un error en la codificación del dato, no se procede a eliminarlos, ya que son datos reales y pueden afectar el modelo.

Al conjunto de datos se hace una limpieza y agrupación en un solo archivo permitiendonos tener un conjunto de datos con una ventana de tiempo de febrero de 2020 a septiembre de 2024. Con la siguiente estructura de datos: En la tabla podemos [?] En la siguiente gráfica

Cuadro 3.1: Muestra de datos de energía

Fecha	Activa Consumo	Activa Generación	Reactiva Capacitiva	Reactiva Inductiva
2020-02-16	4948.0	0.0	0.0	1382.0
2020-02-17	14400.0	0.0	0.0	7072.0
2020-02-18	14896.0	0.0	0.0	7640.0
2020-02-19	14976.0	0.0	0.0	7800.0
2020-02-20	16104.0	0.0	0.0	8376.0

tenemos los datos originales gráficos

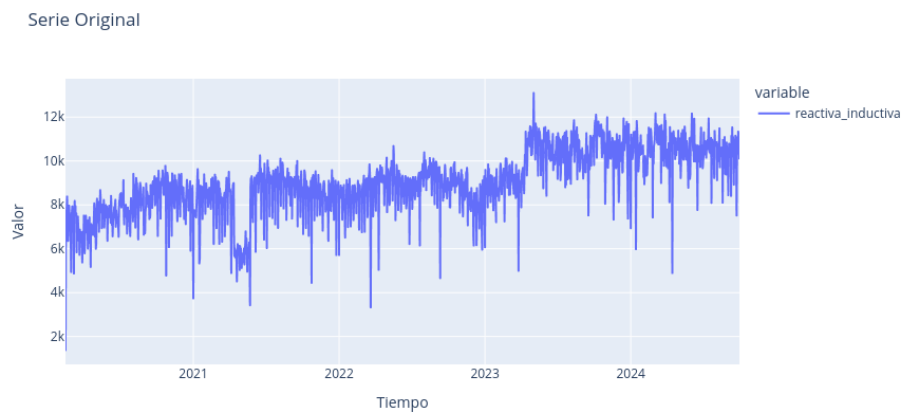
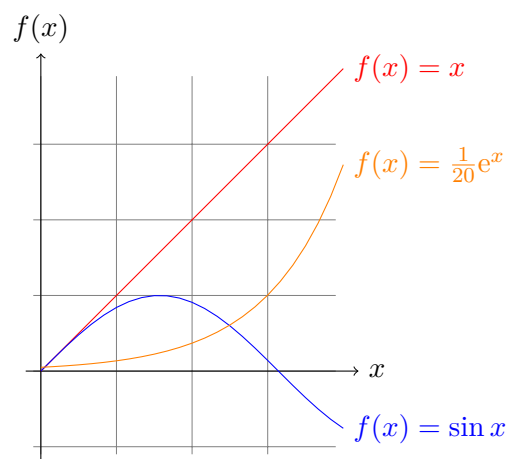
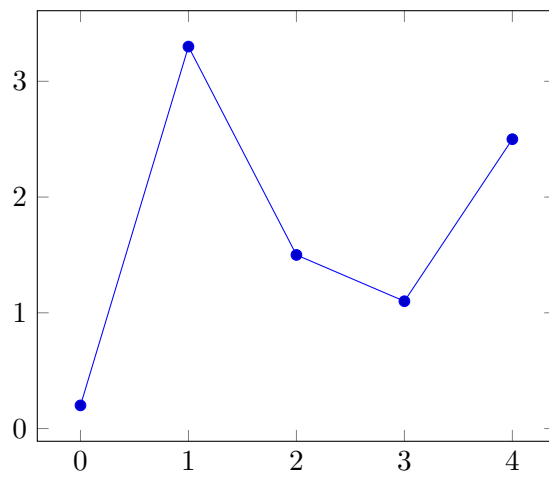


Figura 3.1: Energía reactiva



Bibliografía

- [1] D. Peña Sánchez de Riviera, *Análisis de series temporales*. 2^a ed. Madrid, España: Alianza Editorial, 2010.
- [2] Celsia, “*Conoce todo lo que debes saber sobre energía reactiva: ¿por qué se refleja en tu factura?*”, *Celsia*. [En línea]. Disponible: <https://www.celsia.com/en/blog-celsia/conoce-todo-lo-que-debes-saber-sobre-energia-reactiva-por-que-se-refleja-en-tu-factura-la-energia-reactiva>. Fecha de acceso: 29-10-2024.
- [3] “Metrics and scoring”, scikit-learn: Machine Learning in Python, 1.5 documentation. [Online]. Available: https://scikit-learn.org/1.5/modules/model_evaluation.html[Accessed: 13-Nov-2024].