

Sistema de Detección de Anomalías

Manual de Usuario

1.	Introducción.....	2
1.1	Objetivo de la Aplicación	2
1.2	Audiencia Objetivo.....	2
2.	Preparación de los datos.....	2
3.	Inicio de Sesión	2
4.	Página de Inicio	3
5.	Perfil del Cliente.....	4
6.	Detección de Anomalías	5
7.	Ficha Técnica.....	8
8.	Desconexión y Cierre de Sesión	8

1. Introducción

Electro Dunas, empresa peruana especializada en distribución de energía eléctrica, opera en las provincias de Ica, Huancavelica y Ayacucho, abarcando 5,402 km² y atendiendo a 264,480. Desde su integración al Grupo Energía Bogotá en agosto de 2019, pertenece a un conglomerado líder en energía y gas natural en Colombia, Perú y Brasil. El GEB, comprometido con generar valor, bienestar comunitario y sostenibilidad ambiental, refleja la visión y enfoque de Electro Dunas. La empresa se desenvuelve en dos segmentos de mercado: el regulado y el de competencia. Este último se compone de clientes libres, tanto propios como terceros. A los clientes libres propios se les facturan los precios de generación según sus acuerdos contractuales, además de los cargos regulados por transmisión. Por otro lado, los clientes terceros son facturados de acuerdo con los cargos regulados por transmisión y/o distribución, en función de la utilización que hagan del sistema eléctrico de Electro Dunas.

Con un crecimiento significativo de clientes no regulados, la empresa se propone utilizar analítica de datos para identificar posibles anomalías en el comportamiento de sus clientes no regulados. El proyecto se enfoca en desarrollar un Producto Mínimo Viable (PMV) que visualice datos históricos, resuma comportamientos, identifique anomalías y proporcione alertas, con el objetivo de ser adoptado como una herramienta eficaz en los flujos operativos de Electro Dunas.

1.1 Objetivo de la Aplicación

El tablero de control tiene como objetivo el análisis y detección de anomalías en el comportamiento del consumo de energía de los clientes no regulados propios y terceros de Electro Dunas. Esta herramienta facilita la visualización del consumo histórico, la detección de anomalías y la predicción de comportamientos futuros, permitiendo una gestión más eficiente.

1.2 Audiencia Objetivo

Área operativa de Electro Dunas.

2. Preparación de los datos

La preparación de los datos se realizó mediante varios pasos cuidadosos para asegurar la integridad y la utilidad de la información. Lo primero, los archivos CSV proporcionados por la empresa fueron consolidados en un único conjunto de datos unificado. Este proceso incluyó la lectura de cada archivo individual, la adición de una columna para identificar el archivo original y la combinación de todos estos archivos en un solo DataFrame. Además, se integró información adicional sobre el sector económico de los clientes, permitiendo enriquecer el conjunto de datos con columnas relevantes como el sector económico, el año, el mes, el día de la semana y la hora de cada registro de consumo. La combinación de estos datos se realizó mediante una fusión basada en los identificadores de los clientes.

Para garantizar la calidad y completitud de los datos, se evaluó el porcentaje de completitud de cada serie temporal dentro del período de estudio especificado, de enero de 2021 a abril de 2023. Este análisis permitió identificar y eliminar los registros que excedían una fecha límite predefinida, asegurando que todas las series temporales estuvieran correctamente alineadas. Además, se realizaron muestras por bootstrap, extrayendo subconjuntos de datos repetidamente para evaluar la variabilidad y la robustez de los modelos de detección de anomalías. Estos pasos aseguraron que los datos utilizados en el análisis fueran precisos, completos y adecuados para identificar y clasificar anomalías en el consumo energético de los clientes no regulados de Electro Dunas.

3. Inicio de Sesión

Los archivos de datos y de la aplicación de encuentran en este repositorio publico

<https://github.com/dianavillalba/Anomalias.git>

Carpetas de repositorio

- informes: Prototipo Fachada, Tabla de requerimientos, Desarrollo y Prueba de los Modelos, Validación del prototipo con base en los requerimientos y Manual de Usuario.
- src: Jupyter Notebooks
- statics: imágenes, versión minimizada del archivo de CSS de Bootstrap, versión minimizada del archivo JavaScript de jQuery, archivo JavaScript personalizado y archivo CSS personalizado
- templates: Plantillas HTML de cada página

Ejecución

Puedes ejecutar el código desde sitios como (Google Colab, Jupiter Notebook o VSC) una vez escojas el ambiente donde vas a desplegar asegúrate de usar una versión de python superior a la 3.5 para verificar la versión puedes ejecutar el siguiente comando:

```
!python --version
```

una vez tengas la versión necesaria instala las siguientes dependencias como sigue

```
!pip install -r requirements.txt
```

A partir de acá ya puedes ejecutar la aplicación C:\Projects\Anomalías>Python app.py

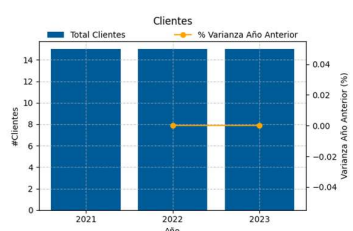
Abre tu navegador y copia el siguiente enlace: <http://127.0.0.1:5000>

4. Página de Inicio

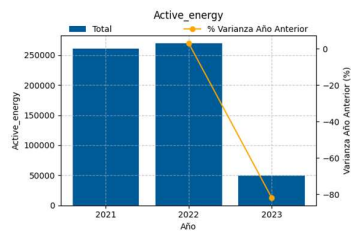
La Página de inicio presenta una interfaz intuitiva para facilitar la navegación y proporcionar acceso rápido a las funcionalidades principales de la aplicación.

Sección: Clientes y Consumo de Energía

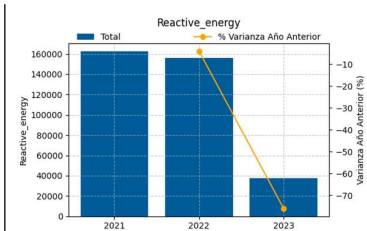
Esta sección del tablero de control proporciona un análisis visual del número de clientes y el consumo de energía, tanto activa como reactiva, a lo largo de los años. Las barras azules representan el total por año y la línea naranja indica el porcentaje de variación respecto al año anterior.



Número total de clientes no regulados por año.



Consumo total de energía activa (kWh) por año



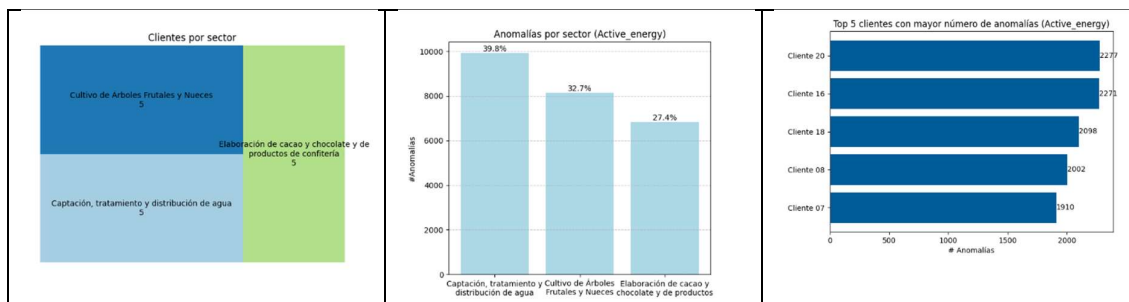
Consumo total de energía reactiva (kVarh) por año

Sección: Clientes y anomalías por Sector

Esta sección cuenta con herramientas interactivas y visualizaciones de las anomalías en el consumo de energía de los clientes por sector y clientes. Se divide en cuatro componentes principales. El Panel de Búsqueda utiliza los campos de fecha para filtrar los datos mostrados en los gráficos, permitiendo un análisis temporal específico. El componente de Clientes por Sector proporciona una visión general de la distribución de clientes en diferentes sectores económicos. El análisis de Anomalías por Sector ayuda a identificar qué sectores económicos presentan más anomalías en el consumo de energía activa, permitiendo focalizar esfuerzos en áreas específicas. Finalmente, el gráfico de los Top 5 Clientes

con Mayor Número de Anomalías facilita la identificación de clientes individuales que requieren atención especial debido a un alto número de anomalías en su consumo de energía.

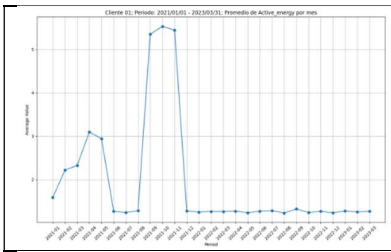
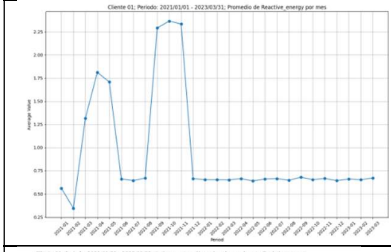
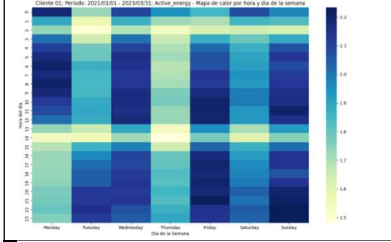
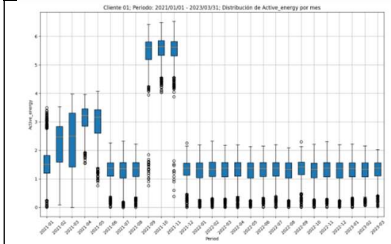
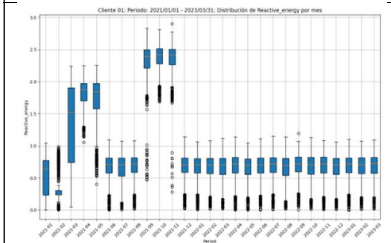
<p>Buscar</p> <p>Fecha Inicial</p> <input type="text" value="2021/01/01"/> <p>Fecha Final</p> <input type="text" value="2023/03/31"/> <p>Submit</p>	<p>Entre la fecha inicial y final Haga clic en el botón de "Submit"</p>
---	---



5. Perfil del Cliente

Esta página del tablero de control permite un análisis detallado del consumo de energía de los clientes individuales. Está dividida en 8 secciones, cada una con una funcionalidad específica para ayudar en la visualización y comprensión de los datos:

<p>Buscar</p> <p>Cliente</p> <input type="text" value="Cliente 01"/> <p>Fecha Inicial</p> <input type="text" value="2021/01/01"/> <p>Fecha Final</p> <input type="text" value="2023/03/31"/> <p>Submit</p>	<p>Panel de Búsqueda: Este panel permite filtrar los datos por cliente y por un rango de fechas especificado. Introduzca el ID del cliente, la fecha inicial y la fecha final, y haga clic en el botón "Submit" para actualizar las visualizaciones.</p>
<p>Promedio de Active_energy por Día de la Semana y Mes:</p>	<p>Promedio de Active_energy por Día de la Semana y Mes: Este gráfico muestra el promedio de energía activa consumida por día de la semana y mes, desglosado por días de la semana. Ayuda a identificar patrones de consumo específicos en función de los días.</p>

	<p>Promedio de Active_energy por Mes: Este gráfico muestra el promedio mensual del consumo de energía activa para el cliente seleccionado. Permite visualizar las tendencias y variaciones a lo largo del tiempo.</p>										
	<p>Promedio de Reactive_energy por Mes: Similar al gráfico anterior, pero enfocado en la energía reactiva. Ayuda a comprender el comportamiento del consumo de energía reactiva del cliente.</p>										
	<p>Mapa de Calor por Hora y Día de la Semana (Active_energy): Este mapa de calor ilustra el consumo de energía activa por horas y días de la semana. Es útil para detectar patrones horarios de uso de energía.</p>										
	<p>Distribución de Active_energy por Mes: Un gráfico de caja que muestra la distribución mensual del consumo de energía activa. Permite identificar la variabilidad y detectar posibles anomalías en los datos.</p>										
	<p>Distribución de Reactive_energy por Mes: Similar al gráfico de caja anterior, pero enfocado en la energía reactiva. Facilita la identificación de la variabilidad mensual en el consumo de energía reactiva.</p>										
<table border="1"> <tr> <td>Sector</td> <td>Elaboración de cacao y chocolate y de productos de confitería</td> </tr> <tr> <td>Total Energía Activa</td> <td>37953.8706014499</td> </tr> <tr> <td>Total Energía Reactiva</td> <td>18283.44586627198</td> </tr> <tr> <td>Información Disponible Desde</td> <td>2021-01-01 00:00:00</td> </tr> <tr> <td>Información Disponible Hasta</td> <td>2023-03-23 23:00:00</td> </tr> </table>	Sector	Elaboración de cacao y chocolate y de productos de confitería	Total Energía Activa	37953.8706014499	Total Energía Reactiva	18283.44586627198	Información Disponible Desde	2021-01-01 00:00:00	Información Disponible Hasta	2023-03-23 23:00:00	<p>Tabla Resumen: Muestra un resumen de la información del cliente, incluyendo el sector económico al que pertenece, el total de energía activa consumida por el cliente durante el período analizado, y el total de energía reactiva consumida por el sector en el mismo período. Además, se proporciona la fecha y hora de inicio del período para el cual se tienen datos disponibles y la fecha y hora final del período para el cual se tienen datos disponibles.</p>
Sector	Elaboración de cacao y chocolate y de productos de confitería										
Total Energía Activa	37953.8706014499										
Total Energía Reactiva	18283.44586627198										
Información Disponible Desde	2021-01-01 00:00:00										
Información Disponible Hasta	2023-03-23 23:00:00										

6. Detección de Anomalías

Esta página del tablero está diseñada para visualizar y analizar anomalías en la energía activa a través de gráficos de series temporales. Se proporcionan herramientas para buscar y configurar parámetros,

así como secciones que muestran detalles sobre las anomalías detectadas y proyectadas. La pagina consta de 5 secciones:

<p>Buscar</p> <p>Cliente</p> <input type="text" value="Cliente 01"/> <p>Fecha Inicial</p> <input type="text" value="2021/01/01"/> <p>Fecha Final</p> <input type="text" value="2023/03/24"/> <p>Nivel Confianza Inicial (%)</p> <input type="text" value="50%"/> <p>Nivel Confianza Final (%)</p> <input type="text" value="50%"/> <p># Días para Pronóstico</p> <input type="text" value="1"/> <p>Submit</p>	<p>El Panel de Búsqueda permite filtrar los datos por cliente y por un rango de fechas especificado. Introduzca el ID del cliente, la fecha inicial, la fecha final, el nivel de confianza inicial y final para las proyecciones, el número de días a pronosticar entre 1 y 30 días, y haga clic en el botón "Submit" para actualizar las visualizaciones.</p>
	<p>La gráfica muestra la serie temporal de los datos históricos y proyecciones de la energía activa, así como las anomalías para cada uno de estos grupos.</p> <p>A continuación, se describen los elementos principales de la gráfica:</p> <ul style="list-style-type: none"> • La línea azul representa los datos históricos del consumo histórico de energía activa. • La línea verde muestra los datos pronosticados dependiendo del número de días seleccionados. • Los puntos rojos indican la presencia de anomalías en los datos. • La línea naranja representa los niveles de confianza dentro de los cuales se espera que los datos se mantengan.
	<p>La gráfica muestra la serie temporal de los datos históricos y proyecciones de la energía reactiva, así como las anomalías para cada uno de estos grupos.</p> <p>A continuación, se describen los elementos principales de la gráfica:</p> <ul style="list-style-type: none"> • La línea azul representa los datos históricos del consumo histórico de energía activa.

	<ul style="list-style-type: none"> La línea verde muestra los datos pronosticados dependiendo del número de días seleccionados. Los puntos rojos indican la presencia de anomalías en los datos. La línea naranja representa los niveles de confianza dentro de los cuales se espera que los datos se mantengan. 												
<p>Anomalías Históricas</p> <p>Fecha y Hora: 2023-03-23 20:00:00 Energía Activa: 1.5251029063661428</p> <p>Fecha y Hora: 2023-03-23 12:00:00 Energía Activa: 1.2563627461970763</p> <p>Fecha y Hora: 2023-03-23 05:00:00 Energía Activa: 1.6458880778175544</p> <p>Fecha y Hora: 2023-03-23 01:00:00 Energía Activa: 0.9467032893039586</p> <p>Fecha y Hora: 2023-03-22 04:00:00 Energía Activa: 1.5529855457958763</p> <p>Fecha y Hora: 2023-03-21 23:00:00 Energía Activa: 1.2189456202855695</p> <p>Fecha y Hora: 2023-03-21 15:00:00 Energía Activa: 1.7755027885468593</p> <p>Fecha y Hora: 2023-03-21 13:00:00 Energía Activa: 1.155578536561639</p> <p>Fecha y Hora: 2023-03-20 16:00:00 Energía Activa: 1.4723375216390395</p> <p>Fecha y Hora: 2023-03-20 13:00:00 Energía Activa: 1.078018105899405</p>	<p>Tabla de Anomalías Históricas: esta sección lista las anomalías históricas detectadas. Cada fila incluye la fecha, la hora y el valor de energía en el momento en que se detectó una anomalía.</p>												
<p>Anomalías Proyectadas</p> <p>No se encontraron anomalías en el rango de fechas.</p>	<p>Tabla de Anomalías Proyectadas: Esta sección lista las proyecciones que son detectadas como anomalías. Cada fila incluye la fecha y la hora futura, así como el valor de energía de la proyección identificada como anomalía</p>												
<table> <tr> <td>Sector</td><td>Elaboración de cacao y chocolate y de productos de confitería</td></tr> <tr> <td>Total Energía Activa</td><td>37955.40191260177</td></tr> <tr> <td>Total Energía Reactiva</td><td>18284.11086939308</td></tr> <tr> <td>Posición Ranking Anomalías</td><td>8</td></tr> <tr> <td>Cantidad de anomalías detectadas actuales</td><td>1569</td></tr> <tr> <td>Cantidad de anomalías detectadas proyectadas</td><td>0</td></tr> </table>	Sector	Elaboración de cacao y chocolate y de productos de confitería	Total Energía Activa	37955.40191260177	Total Energía Reactiva	18284.11086939308	Posición Ranking Anomalías	8	Cantidad de anomalías detectadas actuales	1569	Cantidad de anomalías detectadas proyectadas	0	<p>Tabla Resumen: Muestra un resumen de la información del cliente, incluyendo el sector económico al que pertenece, el total de energía activa consumida por el cliente durante el período analizado, y el total de energía reactiva consumida por el sector en el mismo período. Además, se proporciona la posición en el ranking de anomalías, la cantidad de anomalías detectadas en los datos históricos y la cantidad de anomalías detectadas en el pronóstico</p>
Sector	Elaboración de cacao y chocolate y de productos de confitería												
Total Energía Activa	37955.40191260177												
Total Energía Reactiva	18284.11086939308												
Posición Ranking Anomalías	8												
Cantidad de anomalías detectadas actuales	1569												
Cantidad de anomalías detectadas proyectadas	0												

Ejemplo: Visualizar las anomalías históricas y los pronósticos de 15 días. En el panel de búsqueda, seleccione el Cliente 01, la fecha inicial 2023-03-01, la fecha final 2023-03-23, nivel de confianza inicial y final para las proyecciones del 30%, y el número de días a pronosticar 15.



7. Ficha Técnica

Esta sección contiene información general del tablero de control

Objetivo:

El tablero de control tiene como objetivo el análisis y detección de anomalías en el comportamiento del consumo de energía de los clientes no regulados propios y terceros de ElectroDunas. Esta herramienta facilita la visualización del consumo histórico, la detección de anomalías y la predicción de comportamientos futuros, permitiendo una gestión más eficiente y proactiva del suministro eléctrico.

Datos utilizados:

- Periodo disponible: Desde el 1 de enero de 2021 hasta el 1 de abril de 2023.
- Variables: Energía activa (kWh), energía reactiva (kVarh), Sector económico e ID cliente.
- Cantidad de clientes: 15 clientes analizados con alta completitud de datos ($\geq 99\%$).

Metodología y Modelos:

- Detección de anomalías: implementación de modelos no supervisados y series de tiempo tales como DBSCAN y KernelCPD.
- Predicción consumo: SVR (Support Vector Regression) y LSTM (Long Short-Term Memory).
- Clasificación anomalías futuras: Random Forest y Gradient Boosting.

Interactividad y Funcionalidades:

- Filtros Dinámicos: Permiten explorar y segmentar los datos según diferentes criterios tales como el sector económico, intervalo de tiempo (fecha inicio y fecha fin), ID cliente, nivel de confianza del pronóstico y periodo de pronóstico.
- Alertas de Anomalías: Notificaciones visuales cuando se detectan patrones de consumo atípicos tanto a nivel histórico como patrones futuros.
- Cuadro resumen con la información de las anomalías detectadas según el periodo seleccionado con la información del registro exacto y su desviación.

Requerimientos Técnicos:

- Backend: Flask, Python.
- Frontend: Dash y Bootstrap.
- Modelos Analíticos: Scikit-learn, TensorFlow.

Área encargada:

Operaciones

Elaboración e implementación:

- Diana Villalba (diana.villalba@uniandes.edu.co)
- Santiago Pachon (s.pachonr@uniandes.edu.co)
- Andrés Felipe Valencia (af.valenciad@uniandes.edu.co)

8. Desconexión y Cierre de Sesión

Si ya deseas terminar de usar el tablero, solo basta con que des en el icono de inicio y luego se puede cerrar la ventana del explorador donde estas utilizando el tablero. Con ello, se cierra el tablero y se realiza la desconexión de manera segura.