

Manual de Usuario

Dashboard Mediciones de Gas – Contugas

Proyecto Aplicado en Analítica de Datos

MIAD

Universidad de los Andes

Elaborado por:

Diana Bayona, Juan Diego Pérez, Jorge Rodríguez y Santiago Gutiérrez

Tabla de Contenido

Introducción.....	3
1. Acceso al Dash	3
2. Análisis Histórico	4
3. Detección de Anomalías	5
4. Modelamiento.....	6
5. Recomendaciones de Uso	8
6. Consideraciones Finales.....	8
7. Contacto y Soporte.....	9
Especificaciones Técnicas	10
Anexos.....	11

Introducción

Este Dashboard es una herramienta analítica diseñada para brindar a los usuarios una vista clara, accesible y proactiva sobre el comportamiento operativo de la red de distribución de gas. Este sistema permite tomar decisiones informadas, identificar comportamientos anómalos de forma temprana y anticipar desviaciones relevantes en variables clave como presión, temperatura y volumen de consumo.

Su diseño visual intuitivo y su arquitectura modular hacen que pueda ser utilizado tanto por analistas técnicos como por personal operativo sin experiencia en ciencia de datos. El objetivo principal es transformar los datos crudos generados por sensores en información accionable y comprensible. A través de tres módulos principales (Análisis Histórico, Detección de Anomalías y Modelamiento), el sistema combina lo mejor de la visualización interactiva con algoritmos de inteligencia artificial para que los usuarios puedan:

- Auditar comportamientos pasados.
- Detectar situaciones atípicas que requieren intervención.
- Simular escenarios y validar comportamientos futuros con nuevos datos.

Este manual de usuario está enfocado en explicar paso a paso cómo aprovechar cada sección del dashboard, qué se puede interpretar en cada pantalla, y cómo aplicar los hallazgos en contextos reales de monitoreo y operación, como algunas recomendaciones y consideraciones finales.

1. Acceso al Dash

El Dashboard ha sido desplegado en la nube para facilitar su acceso desde cualquier dispositivo con conexión a internet. Sin la necesidad de ejecutar la aplicación localmente ni configurar entornos de Python en el computador del usuario.

<https://contugas-dashboard-production.up.railway.app/>

Este enlace apunta a una instancia activa alojada en Railway, lo que garantiza seguridad, disponibilidad y facilidad de acceso sin depender de infraestructura local. El usuario puede explorar el dashboard sin instalar ningún software adicional, manteniendo una experiencia fluida, visual e interactiva. Desde allí, se tendrá acceso completo a todas las funcionalidades del dashboard, incluyendo visualización de datos históricos, detección automática de anomalías y predicción basada en modelos entrenados.

Requisitos del usuario final

- Conexión a internet estable
- Navegador moderno actualizado (se recomienda Google Chrome o Firefox)
- Archivo Excel (.xlsx) válido en caso de querer realizar predicciones personalizadas

La interfaz se adapta automáticamente al tamaño de la pantalla, lo que facilita su uso tanto en computadores de escritorio como en dispositivos portátiles. Una vez dentro de la aplicación, el usuario encontrará tres módulos principales accesibles desde la parte superior.

Cada sección está pensada para ofrecer valor operativo inmediato, sin requerir experiencia técnica previa. Estas pestañas son:

- **Análisis Histórico:** Para examinar el comportamiento pasado de variables clave como presión, temperatura y volumen.
- **Detección de Anomalías:** Para identificar desviaciones significativas a través de modelos predictivos y no supervisados.
- **Modelamiento:** Para cargar nuevos archivos de datos y recibir predicciones basadas en modelos entrenados previamente.

La estructura del sistema ha sido diseñada para mantener una experiencia fluida, incluso en computadores con recursos limitados. Gracias a su arquitectura basada en Dash y FastAPI, la solución se ha desplegado exitosamente en la nube usando Railway, lo cual permite escalar a múltiples usuarios sin fricción y sin requerimientos locales.

2. Análisis Histórico

La pestaña de **Análisis Histórico** está diseñada para ofrecer una visualización detallada y comprensible del comportamiento pasado de las variables operativas más importantes: presión, temperatura y volumen. Su principal objetivo es proporcionar una base sólida para el análisis de tendencias, la validación de hipótesis y la identificación de patrones estacionales o inusuales a lo largo del tiempo.

Esta sección es especialmente útil para auditores, ingenieros de procesos y analistas de datos que deseen comprender cómo ha evolucionado el comportamiento de un cliente o grupo de clientes en un intervalo específico.

Componentes funcionales:

- **Selector de rango de fechas:** Permite delimitar el periodo de análisis. Ideal para enfocar la observación en eventos críticos o cambios operativos.
- **Selector de cliente:** Filtra los datos por el código del cliente. Cada cliente posee una serie temporal independiente.
- **Tipo de agregación:** Ofrece la posibilidad de visualizar los datos de forma agregada (por ejemplo, diaria) para una lectura más amigable y menos ruidosa.

Visualizaciones:

- **Histogramas individuales:** Por cada variable, se despliega un histograma que muestra la distribución de los datos en el rango seleccionado. Además de la forma general, incluye: Media y mediana; Cuartiles (Q1 y Q3); Valores máximo y mínimo; y desviación estándar

Estos elementos ayudan a interpretar la variabilidad y simetría del comportamiento de cada variable.

- **Serie de tiempo:** Para cada variable se despliega una gráfica lineal en el tiempo que incluye: Línea de tendencia o valor promedio y bandas de control (límites ± 2 desviaciones estándar)

Estas gráficas permiten identificar días o períodos específicos donde se presentaron desviaciones importantes, cortes, picos extremos o comportamientos cíclicos.

Aplicaciones prácticas:

- Validación de la calidad de los datos históricos
- Identificación de patrones anuales, mensuales o semanales
- Soporte para toma de decisiones sobre mantenimiento o revisiones operativas
- Establecimiento de umbrales base para comparaciones futuras

Gracias a su diseño visual claro y al uso de estadística descriptiva automatizada, esta pestaña permite una exploración profunda sin requerir conocimientos avanzados de análisis de datos.

3. Detección de Anomalías

La pestaña de **Detección de Anomalías** es el núcleo analítico del sistema, donde se combinan dos enfoques complementarios para identificar comportamientos atípicos en las series de tiempo de presión, temperatura y volumen:

- Modelos LSTM (Long Short-Term Memory):** Algoritmos de aprendizaje profundo que aprenden los patrones temporales esperados a partir de los datos históricos y predicen lo que debería ocurrir. Cualquier diferencia significativa entre la predicción y la realidad se considera una posible anomalía.
- Isolation Forest:** Algoritmo no supervisado que identifica outliers en los datos históricos sin necesidad de etiquetas. Su ventaja es que puede detectar comportamientos estructuralmente diferentes, incluso si no son extremos numéricamente.

Esta combinación permite identificar tanto desviaciones puntuales inesperadas como patrones anómalos más sutiles pero persistentes.

Filtros disponibles:

- **Cliente:** Selecciona el cliente cuyos datos se desean analizar.
- **Variable:** Presión, Temperatura o Volumen.
- **Rango de fechas:** Limita la ventana temporal para concentrarse en el periodo de interés.

Gráfica principal:

Una visualización clara y dinámica donde se observa:

- **Línea azul:** Valores reales registrados.
- **Línea naranja:** Predicciones generadas por el modelo LSTM.
- **Línea punteada naranja:** Predicciones futuras para 24 pasos adelante.
- **Puntos rojos:** Anomalías detectadas por el Isolation FOrrest

Esta vista permite comparar en detalle la fidelidad del modelo y dónde se rompen los patrones esperados.

Métricas y resultados:

- **RMSE (Root Mean Squared Error):** Cuantifica la magnitud promedio de los errores.
- **MAE (Mean Absolute Error):** Evalúa la precisión del modelo con menor sensibilidad a valores extremos.
- **Anomalías por modelo:** Cantidad de puntos detectados como anómalos por el Isolation Forest
- **Anomalías históricas:** Resultados de Isolation Forest aplicados a datos anteriores.
- **Coincidencias:** Casos donde ambos modelos detectan la misma anomalía.
- **% Coincidencia:** Mide la solidez del hallazgo cruzado.

Estas métricas ayudan a entender si el sistema está siendo demasiado sensible o poco reactivo, y facilitan calibraciones futuras.

Score Isolation Forest:

Se muestra un histograma que permite analizar la distribución del score de anomalía. Cuanto más extremo sea el score, mayor la probabilidad de tratarse de un valor atípico. Incluye también: Media, Q1, Q3 del score y visualización del umbral estimado para marcar outliers.

Tabla de anomalías:

En la parte inferior, una tabla sintetiza las últimas 10 anomalías detectadas, indicando:

- Fecha
- Variable afectada
- Cliente
- Score o magnitud de la desviación

Aplicaciones prácticas:

- Identificación de fallas incipientes en sensores o infraestructura
- Priorización de inspecciones técnicas basadas en evidencia
- Soporte para decisiones preventivas en la operación
- Análisis retrospectivo de incidentes

Gracias a este módulo, el equipo operativo puede dejar de depender exclusivamente de la supervisión manual y adoptar un enfoque proactivo, inteligente y basado en datos para la gestión de eventos atípicos.

4. Modelamiento

La pestaña de **Modelamiento** está diseñada para permitir la carga dinámica de nuevos datos operativos y obtener predicciones personalizadas mediante modelos LSTM previamente entrenados. Esta funcionalidad convierte al dashboard en una herramienta viva que no solo explora el pasado, sino que también anticipa el futuro.

Es especialmente útil para situaciones donde se desea validar el comportamiento de clientes con patrones recientes inusuales, simular escenarios de consumo ante cambios operativos o ambientales y evaluar la efectividad de decisiones de mantenimiento o intervención.

Subida de datos:

El usuario puede cargar un archivo Excel (.xlsx) que contenga registros recientes de mediciones. Para garantizar una predicción precisa y compatible con el modelo, el archivo debe cumplir con las siguientes condiciones:

- **Columnas requeridas:** Fecha, Presion, Temperatura, Volumen, Client_ID
- **Formato de fecha:** YYYY-MM-DD HH:MM:SS
- **Cantidad de datos:** Mínimo 24 registros por cliente para que el modelo pueda generar una predicción de los siguientes 24 pasos.
- **Cientes únicos:** Se recomienda que el archivo contenga los datos de un solo cliente por carga para asegurar precisión en el modelo seleccionado.

Una vez cargado, los datos se validan y se envían automáticamente al backend a través de la API construida en FastAPI. Esta API busca el modelo correspondiente al cliente en lstm_models.pkl, realiza las predicciones y devuelve los resultados al dashboard.

Probar conexión API:

El botón de prueba de conexión permite verificar que la API esté activa y disponible para recibir solicitudes. Esto es clave antes de cargar archivos reales.

Visualización de predicciones:

Una vez procesado el archivo, se despliega una gráfica donde el usuario puede observar las últimas mediciones reales cargadas y la predicción generada por el modelo LSTM para las siguientes 24 horas. Esta visualización ayuda a comparar la tendencia reciente con la expectativa futura, brindando información accionable para ajustar decisiones operativas, energéticas o comerciales.

Información explicativa:

El panel también incluye una descripción de cómo funciona el modelo LSTM, sus fortalezas y recomendaciones de uso, con el fin de que los usuarios no técnicos comprendan lo que ocurre "detrás del telón".

Aplicaciones prácticas:

- Anticipación de picos de consumo o caídas abruptas.
- Verificación de comportamiento tras cambios de infraestructura.
- Simulación de escenarios futuros para planes de contingencia.
- Inclusión en reportes periódicos de monitoreo.

Gracias a esta funcionalidad, el sistema deja de ser solo un visor y se transforma en una herramienta predictiva capaz de brindar ventaja competitiva y mejorar la toma de decisiones.

5. Recomendaciones de Uso

Para asegurar el máximo aprovechamiento del Dashboard es importante tener en cuenta una serie de buenas prácticas que garantizan tanto la precisión de los resultados como la fluidez en la experiencia de usuario.

- **Uso inmediato de la plataforma en la nube:** No es necesario ejecutar servidores locales ni instalar dependencias técnicas. La API y el motor de predicción ya están activos y conectados en la nube, por lo que el usuario puede cargar archivos y visualizar predicciones en tiempo real sin configuraciones adicionales.
- **Validar los archivos antes de cargarlos:** Los archivos Excel deben estar correctamente formateados. Verifica que: No haya columnas vacías o con nombres distintos; El formato de fecha esté unificado; Existan al menos 24 registros consecutivos por cliente; y evitar errores en el formato que pueden generar resultados erróneos o impedir el procesamiento del archivo.
- **Usar las visualizaciones como herramienta de decisión:** El dashboard no es solo una plataforma de monitoreo pasivo. Sus gráficas, histogramas y tablas están diseñadas para facilitar el análisis crítico. Te recomendamos: Comparar bandas de control para detectar variabilidad anormal; Analizar las coincidencias entre modelos para priorizar alertas; Verificar tendencias predichas antes de tomar acciones operativas.
- **Aprovechar el enfoque dual de detección de anomalías:** El uso conjunto de LSTM (modelo secuencial) y Isolation Forest (modelo estructural) brinda una detección más robusta. En contextos reales: Las anomalías detectadas por ambos modelos pueden considerarse más urgentes; Las que aparecen en solo uno pueden requerir validación adicional.
- **Incorporar el dashboard en rutinas de monitoreo:** Este sistema puede integrarse en ciclos diarios o semanales de revisión operativa. Gracias a su enfoque visual e interactivo, puede ser proyectado en salas de control, presentado en informes ejecutivos o integrado en workflows internos.

6. Consideraciones Finales

Esta sección resume las principales **ventajas, limitaciones y advertencias** relacionadas con el uso del dashboard, con el fin de brindar un entendimiento claro y realista sobre sus capacidades y condiciones de operación.

Ventajas del sistema:

- **Acceso inmediato desde la nube:** No requiere instalación, gracias a su despliegue en Railway.
- **Interfaz intuitiva:** Diseñada para usuarios no técnicos con navegación clara y explicaciones visuales.
- **Capacidad predictiva:** Utiliza modelos LSTM entrenados por cliente para anticipar comportamientos futuros.
- **Detección híbrida de anomalías:** Combina modelos supervisados (LSTM) y no supervisados (Isolation Forest).

- **Escalabilidad:** Puede ampliarse a múltiples usuarios o clientes sin modificar la arquitectura base.

Limitaciones y advertencias:

- **Formato de datos estricto:** Los archivos cargados deben contener columnas específicas y al menos 24 registros por cliente.
- **Predicción cliente a cliente:** Cada modelo está entrenado por cliente; si se carga un cliente no entrenado, no se generará predicción.
- **No evalúa causalidad:** El sistema detecta anomalías, pero no explica su causa. Se recomienda validación operativa.
- **Carga de archivos uno a uno:** Actualmente, el sistema acepta un solo archivo por cliente en cada carga.
- **Dependencia de la calidad del dato:** Anomalías mal interpretadas pueden ser causadas por errores de sensores o formatos.

Recomendaciones clave:

- Validar el archivo antes de subirlo.
- Consultar las coincidencias entre modelos como criterio de confianza.
- Revisar periódicamente los resultados para identificar patrones persistentes.
- Considerar el dashboard como un complemento analítico a la operación, no un reemplazo de la supervisión técnica.
- Se recomienda que un equipo humano valide inicialmente que las anomalías señaladas por el sistema reflejen situaciones reales. Esta retroalimentación es fundamental para ajustar los modelos y aumentar su precisión.
- Ante hallazgos relevantes, establecer un canal de comunicación con el equipo técnico para mejorar continuamente el desempeño del prototipo.

Esta sección sirve como guía para maximizar el beneficio del sistema y prevenir malentendidos en su operación diaria.

7. Contacto y Soporte

Para garantizar el uso eficiente del dashboard y resolver cualquier duda técnica o funcional, se ha dispuesto un canal de contacto directo con el equipo de desarrollo y soporte.

Soporte general y asistencia funcional:

- **Correos electrónicos:** d.bayonap@uniandes.edu.co, jd.perez11@uniandes.edu.co, ja.rodriguezp123@uniandes.edu.co, s.gutierrez11234@uniandes.edu.co
- **Horario de atención:** Lunes a viernes, 8:00 a.m. – 5:00 p.m. (GMT-5)
- **Tiempo de respuesta estimado:** 24 a 48 horas hábiles

Solicitudes de mejora, errores o evolución del sistema:

- Si se encuentra algún comportamiento inesperado, desea sugerir una funcionalidad nueva o reportar una mejora, se puede enviar una solicitud con el asunto “Dashboard Contugas – Mejora” a los mismos correos señalados en el soporte general.

Actualizaciones del sistema:

- Las mejoras y nuevas funcionalidades se documentan y comunican mediante actualizaciones de versión.
- Se recomienda consultar el dashboard de manera periódica para acceder a nuevas capacidades y correcciones.

Especificaciones Técnicas

Esta sección resume la arquitectura técnica y operativa del sistema para equipos de soporte, analistas técnicos o desarrolladores que deseen comprender los componentes que hacen posible el funcionamiento del dashboard.

Arquitectura general:

- **Frontend:** Dash + Plotly desplegado como app web interactiva.
- **Backend:** FastAPI para predicción en tiempo real de nuevos datos.
- **Motor de predicción:** Modelos LSTM serializados por cliente.
- **Algoritmo complementario:** Isolation Forest para detección no supervisada.
- **Hosting:** Railway para despliegue continuo (producción).

Estructura del sistema:

- App.py -> Aplicación principal Dash
- api.py -> Servidor FastAPI para predicciones
- Layout1.py -> Layout Análisis Histórico
- layout2.py -> Layout Detección de Anomalías
- layout3.py -> Layout Modelamiento
- lstm_models.pkl -> Modelos LSTM por cliente
- Save Model.ipynb -> Notebook de entrenamiento
- Datos_Contugas_Compacted.csv -> Dataset base
- requirements.txt -> Dependencias del proyecto

Requisitos de entorno (para despliegue local opcional):

- Python 3.8 o superior
- Dependencias del archivo requirements.txt

Flujo funcional:

1. El usuario interactúa con la aplicación vía navegador.
2. La capa frontend renderiza las gráficas y envía datos a la API.
3. La API consulta el modelo correspondiente, predice y retorna los resultados.
4. Se visualizan los resultados, incluyendo predicciones y anomalías.

Consideraciones sobre Railway:

- Railway permite desplegar tanto el backend como el frontend como un servicio conjunto.

- Las rutas están configuradas para resolver correctamente llamadas internas desde Dash hacia FastAPI.
- Permite actualizaciones rápidas y despliegue continuo desde repositorio Git.

Esta arquitectura modular y escalable garantiza facilidad de mantenimiento, extensibilidad y adaptabilidad para nuevos proyectos o integraciones futuras.

Anexos

- Prototipo Fachada Inicial: [Prototipo Fachada Contugas.docx](#)
- Reporte de Selección y Parametrización de Modelos: [Reporte de Selección y Parametrización de Modelos para Contugas ¶.docx](#)
- Notebook Reporte de Selección y Parametrización de Modelos: https://github.com/JuanD13Perez/contugas-dashboard/blob/main/model_Notebook/VA_Reporte%20de%20Seleccio%CC%81n%20y%20Parametrizacio%CC%81n%20de%20Modelos%20para%20Contugas.ipynb
- Tabla de Requerimientos: [Tabla de Requerimientos.docx](#)
- Validación Tabla de Requerimientos: [Validacion_Tabla_Requerimientos_Contugas.xlsx](#)
- Dashboard - Railway: <https://contugas-dashboard-production.up.railway.app/>
- Api Modelo - Railway: <https://contugas-dashboard-production-6acf.up.railway.app/docs>
- Estructura del Proyecto-GitHub: <https://github.com/JuanD13Perez/contugas-dashboard>