

# Detección temprana de obstrucción en pipelines de larga distancia mediante el uso de algoritmos de machine learning

Marcelo Carmona Saavedra - Paula Alvarez Varas

## Resumen

Las obstrucciones en pipelines de transporte de sólidos (concentrados) generan pérdidas operacionales, riesgos de seguridad y costos elevados de mantenimiento. Proponemos un sistema de detección temprana basado en modelos supervisados (Regresión Logística, SVM, Random Forest, y XGBoost) entrenados con telemetría de presión, y caudal. El enfoque integra caracterización de transientes hidráulicos, para identificar y localizar eventos de incipiente sedimentación, reducción de sección y plugging. La propuesta se sustenta en avances recientes en la monitorización de pipelines de larga distancia y detección de obstrucción mediante dinámicas de presión. Los pipelines de gran extensión presentan desafíos complejos: variaciones de presión por factores hidráulicos, pérdidas de carga naturales y ruido en las mediciones. Una obstrucción puede generar patrones de presión sutiles y no lineales, difíciles de distinguir de fluctuaciones normales. La detección temprana es crítica para evitar riesgos ambientales, pérdidas económicas y daños estructurales. Este proyecto fue abordado como una tarea de clasificación binaria, donde la variable objetivo es si existe obstrucción o no en algún punto del trazado del perfil del pipeline. Se contó con una base de datos proveniente de 8 sensores de presión distribuidos a lo largo de una tubería de 143 km, junto con una base de referencia del comportamiento esperado de presiones obtenida mediante un modelo hidráulico específico del pipeline. Con estos datos se aplicó el método de características para flujos transitorios de modo de extraer parámetros clave como propagación de onda, celeridad y fricción. De esta forma, se genera un dataset sintético simulando distintos escenarios de obstrucción en diferentes puntos del trazado, con el fin de entrenar estos modelos seleccionados. La métrica principal de evaluación fue F1 score de modo de obtener un equilibrio entre detección correcta y falsas alarmas. Finalmente los resultados indicaron que Random Forest optimizado presentó mejor desempeño global.

El uso de técnicas de machine learning representa una estrategia necesaria para enfrentar los desafíos de monitoreo predictivo en pipelines, aportando valor técnico como económico a la gestión de sistemas hidráulicos complejos.

**Palabras clave:** pipeline, slurry, obstrucción, aprendizaje automático, transientes hidráulicos, SVM, Random Forest, XGBoost.

## 1. Introducción

Las obstrucciones en tuberías de transporte de sólidos son causadas principalmente por sedimentación, variaciones reológicas y condiciones transientes de flujo. Estos eventos no solo representan riesgos de seguridad y altos costos de mantenimiento, sino que su impacto económico es significativo por la detención de operaciones. La detección es compleja debido a que la hidráulica de la pulpa no es lineal y los patrones de presión sutiles son difíciles de diferenciar de las fluctuaciones normales de operación. El objetivo de esta investigación es desarrollar un modelo capaz de identificar estados de obstrucción de manera temprana integrando datos de presión y caudal. Tales bloqueos provocan pérdida de presión, aumento del requerimiento de bombeo y detenciones no programadas [1], [2].

El estado del arte muestra avances en la **detección y localización** de bloqueos por **métodos basados en pulso de presión**, con precisiones en ubicación por debajo del 5 %, condicionadas por la geometría del bloqueo [3]. Asimismo, los enfoques de *machine learning* (ML) supervisado—incluyendo Regresión

Logística, SVM, Random Forest y XGBoost—han logrado altas precisiones en la detección de anomalías (fugas, bloqueos) usando señales de **presión transiente** [4], [5].

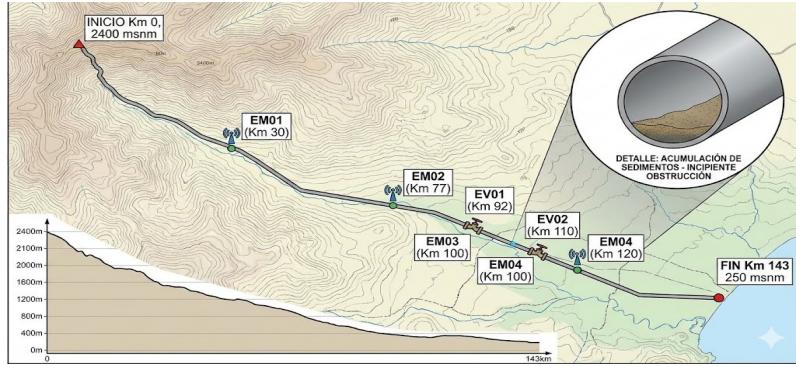


Figura 1: Perfil Longitudinal característicos de pipeline. Elaboración propia.



Figura 2: Rotura pipelines.

## 2. Metodología

El estudio se basó en un flujo de procesamiento de datos que incluyó:

- Adquisición de Datos: Se utilizó un dataset operacional real con 95,000 registros de presión y caudal, complementado con un dataset de 3,200 datos obtenidos de modelo hidráulico de pipeline en particular.
- Modelado Hidráulico: Se aplicó el "Método de las Características" para simular la propagación de ondas de presión, considerando variables como celeridad y fricción.
- Entrenamiento de Modelos: Se evaluaron cuatro clasificadores supervisados: Regresión Logística, Máquinas de Soporte Vectorial (SVM), Random Forest y XGBoost.
- Optimización: Se realizó un ajuste de hiperparámetros mediante Random Search para maximizar la métrica F1-score



Figura 3: No trivialidad de detección de obstrucción.

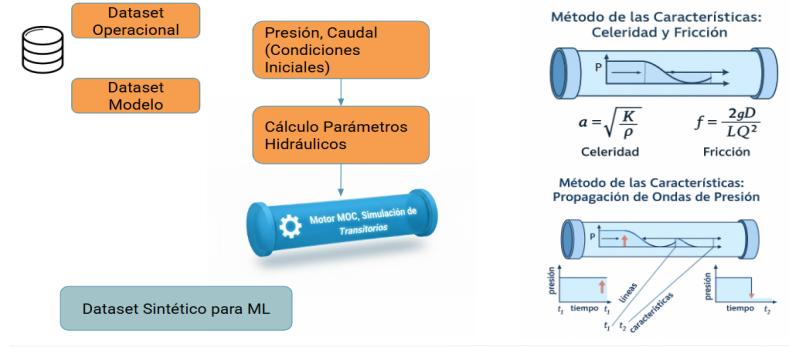


Figura 4: Flujo creacion dataset sintético.

## 2.1. Arquitectura de sensorización

**Sensores hidráulicos:** presión (PT), caudal (FT), temperatura, con muestreo de 1 a 5 s. [3].

## 2.2. Datos y etiquetado

Se integrará histórico operacional (variaciones de presión/caudal) *normal/obstrucción* se definirán por criterios hidráulicos (pérdida de carga incremental, respuesta transiente) [2].

## 2.3. Modelos supervisados

**Clasificación:** Regresión Logística, SVM, Random Forest y XGBoost sobre *features* estadísticos[4], [6].

## 2.4. Validación

Validación por métricas: *precision/recall/F1 tasa de falsas alarmas y error de localización (m)*.

## 3. Resultados

En la etapa inicial (baseline), el modelo Random Forest demostró el desempeño más sólido con una precisión del 92.08

- Random Forest: Incrementó de 0.915 a 0.926.
- SVM: Tuvo la mejora más drástica, pasando de un F1-score de 0.414 a 0.840.
- XGBoost: Alcanzó un score ajustado de 0.783.

La matriz de confusión para Random Forest confirmó una alta capacidad de clasificación, identificando correctamente 1.528 eventos de obstrucción frente a solo 13 falsos positivos.

En la literatura, se espera un  $F1 \geq 0,90$  para clasificar *normal/obstrucción* en segmentos críticos, consistente con rendimientos reportados para ML en anomalías hidráulicas  $\geq 95\%$ ) [4], [5], [7].

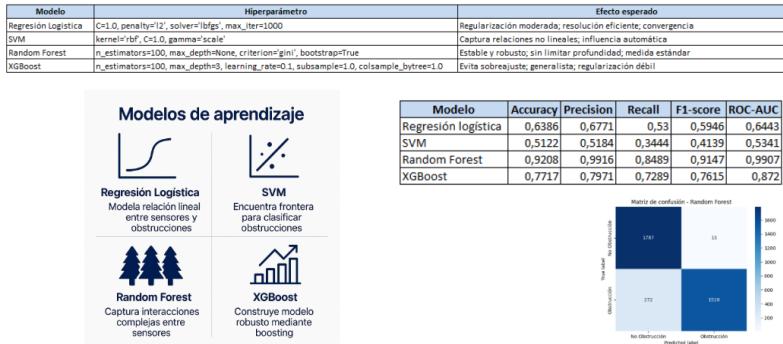


Figura 5: Entrenamiento Modelos baseline.

Modelo	Hiperámetros seleccionados	Score F1
Random Forest	n_estimators: 500, 'min_samples_split': 2, 'min_samples_leaf': 1, 'max_features': 'log2', 'max_depth': None	0,9264
XGBoost	subsample: 0.8, n_estimators: 300, 'max_depth': 5, 'learning_rate': 0.2, 'gamma': 0.1, 'colsample_bytree': 0.8	0,7828
Regresión Logística	cif_penalty: 'l1', 'cif_l1_ratio': 0.25, 'cif_C': np.float64(0.5994842503189409)	0,596
SVM	cif_kernel: 'rbf', 'cif_gamma': 'auto', 'cif_C': 100	0,84

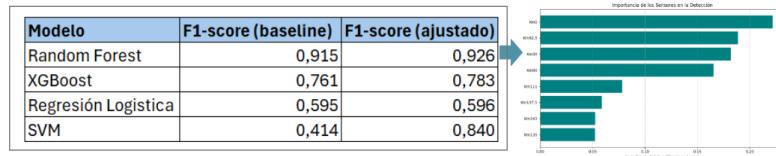


Figura 6: Ajuste de Hiperparámetros y Resultados.

## 4. Discusión y Conclusiones

Los resultados indican que los modelos basados en ensambles, específicamente Random Forest, son altamente efectivos para capturar las interacciones complejas y no lineales de los sensores hidráulicos en pipelines. La importancia de los sensores varía según su ubicación en el perfil del trazado, siendo críticos para la localización precisa del evento. Se concluye que la integración de telemetría con machine learning permite una detección robusta que podría mitigar las pérdidas económicas millonarias asociadas a detenciones prolongadas por obstrucciones.

### 4.1. Aporte

La fusión de transientes hidráulicos con ML supervisado suple limitaciones de enfoques umbrales/inspección intrusiva, ofreciendo **detección temprana** en *slurry pipelines*. La literatura respalda la **viabilidad, sensibilidad y escalabilidad** del enfoque [3], [6], [8].

### 4.2. Limitaciones

La variabilidad reológica (densidad, granulometría, reología no newtoniana) afecta firmas de presión; se requiere **calibración por campaña** y seguimiento de *drift*. La geometría del bloqueo influye en

intensidades de reflexión y en el error de localización; se mitiga con características tiempo-frecuencia y aprendizaje secuencial [3].

## 5. Conclusiones

La propuesta es **factible** y alineada con prácticas modernas de **monitorización de pipelines** de larga distancia. Se estima una mejora sustancial en **detección temprana**, reducción de tiempos de respuesta y soporte a decisiones operacionales (ajuste de velocidad, pigging), contribuyendo a seguridad y continuidad operativa [8], [9].

## Referencias

- [1] J. P. Chapman, «Long-Distance Slurry Pipelines,» en *Handbook of Pipeline Engineering*, Springer, 2023. dirección: [https://link.springer.com/rwe/10.1007/978-3-031-05735-9\\_37-1](https://link.springer.com/rwe/10.1007/978-3-031-05735-9_37-1).
- [2] S. Razvarz et al., «Blockage Detection in Pipeline,» en *Flow Modelling and Control in Pipeline Systems*, Springer, 2021. dirección: [https://link.springer.com/content/pdf/10.1007/978-3-030-59246-2\\_7.pdf](https://link.springer.com/content/pdf/10.1007/978-3-030-59246-2_7.pdf).
- [3] K. Yan, D. Xu, Q. Wang et al., «Experimental Investigation of Gas Transmission Pipeline Blockage Detection Based on Dynamic Pressure Method,» *Energies*, vol. 16, n.º 15, 2023. DOI: [10.3390/en16155620](https://doi.org/10.3390/en16155620). dirección: <https://www.mdpi.com/1996-1073/16/15/5620>.
- [4] S. S. Aljameel, D. M. Alomari et al., «An Anomaly Detection Model for Oil and Gas Pipelines Using Machine Learning,» *Computation*, vol. 10, n.º 8, 2022. DOI: [10.3390/computation10080138](https://doi.org/10.3390/computation10080138). dirección: <https://www.mdpi.com/2079-3197/10/8/138>.
- [5] N. Ullah, Z. Ahmed y J.-M. Kim, «Pipeline Leakage Detection Using Acoustic Emission and Machine Learning Algorithms,» *Sensors*, vol. 23, n.º 6, 2023. DOI: [10.3390/s23063226](https://doi.org/10.3390/s23063226). dirección: <https://www.mdpi.com/1424-8220/23/6/3226>.
- [6] B. Xiao, S. Miao et al., «Detecting the Backfill Pipeline Blockage and Leakage through an LSTM-Based Deep Learning Model,» *International Journal of Minerals, Metallurgy and Materials*, vol. 30, n.º 8, 2023. DOI: [10.1007/s12613-022-2560-y](https://doi.org/10.1007/s12613-022-2560-y). dirección: <https://link.springer.com/content/pdf/10.1007/s12613-022-2560-y.pdf>.
- [7] F. Saleem, Z. Ahmad y J.-M. Kim, «Real-Time Pipeline Leak Detection: A Hybrid Deep Learning Approach Using Acoustic Emission Signals,» *Applied Sciences*, vol. 15, n.º 1, 2025. DOI: [10.3390/app15010185](https://doi.org/10.3390/app15010185). dirección: <https://www.mdpi.com/2076-3417/15/1/185>.
- [8] Y. Wu, L. Gao, J. Chai, Z. Li et al., «Overview of Health-Monitoring Technology for Long-Distance Transportation Pipeline and Progress in DAS Technology Application,» *Sensors*, vol. 24, n.º 2, 2024. DOI: [10.3390/s24020413](https://doi.org/10.3390/s24020413). dirección: <https://www.mdpi.com/1424-8220/24/2/413>.
- [9] P. bibinitperiod Cooke, *From Flow to Future: Managing Long-Distance Slurry Pipelines*, 2025. dirección: <https://www.patersoncooke.com/2025/09/08/from-flow-to-future-managing-long-distance-pipelines/>.