

MACHINE LEARNING APLICADO A LA GESTIÓN DE CORROSIÓN Y PREDICCIÓN DE LA VELOCIDAD DE CORROSIÓN

Alan Cornejo, TGN, alan.cornejo@tgn.com.ar

Martín Melucci, TGN, martin.melucci@tgn.com.ar

Marianela Gomez, TGN, marianela.gomez@tgn.com.ar

Sinopsis

En los últimos años, se han producido avances significativos en el campo de la ciencia de datos y Machine Learning. Estos avances han venido acompañados de nuevas aplicaciones en diversas industrias como lo es la industria del Oil & Gas. En lo que respecta a la gestión de integridad de ductos estos desarrollos son de vital importancia, ya que ofrecen una serie de ventajas clave:

- **Análisis de Grandes Volúmenes de Datos en Tiempo Reducido:** La capacidad para analizar grandes cantidades de información en un período de tiempo más breve es fundamental. Esto permite tomar decisiones más rápidas y eficientes.
- **Detección de Patrones y Comportamientos Anómalos**
- **Identificación de Parámetros Clave en la Generación de Amenazas:** El aprendizaje automático puede ayudar a identificar los factores más influyentes en la generación de amenaza y focalizar las medidas de mitigación de manera precisa.
- **Capacidad de Predicción y Alerta Temprana:** Los modelos de aprendizaje automático pueden utilizarse para realizar predicciones y estimar valores futuros, como ser velocidad de crecimiento de corrosión.

El objetivo del presente trabajo es presentar un proceso de implementación de Machine Learning en el estudio de la amenaza de corrosión externa. A través de diversos modelos se estudian las zonas activas de corrosión y se predicen velocidades de crecimiento en función de diversas fuentes de datos. Además, se compartirá un caso de aplicación con el propósito de compartir los hallazgos encontrados, las conclusiones adquiridas y los pasos a seguir.

Introducción

El sistema de transporte de gas de TGN, con más de 11.000 km de extensión, moviliza gas a 17 provincias del centro y norte del país. El paso del tiempo y las condiciones del entorno pueden afectar la integridad de sus ductos, siendo la corrosión externa una de las principales amenazas. Hoy en día, el uso de herramientas de inspección interna inteligentes (ILI) permite detectar y dimensionar con precisión las anomalías volumétricas, características típicas de la corrosión. Si bien estas herramientas nos brindan información valiosa, la comparación de resultados entre diferentes inspecciones solo nos ofrece una aproximación de la velocidad de corrosión. Esto se debe a que dichas comparaciones se basan en dos momentos específicos, lo que limita su precisión y exige un análisis minucioso. Para una evaluación más completa en el periodo entre corridas y con el objetivo de evitar un comportamiento reactivo al momento de realizar la próxima inspección interna, es necesario considerar otros factores que pueden influir en la velocidad de corrosión externa, como los cambios en los potenciales de la protección catódica, el estado del revestimiento, las variables ambientales (precipitaciones y temperatura), características del suelo

y características de la cañería (espesor, grado del material, diámetro, tipo de soldadura longitudinal, fabricante y revestimiento externo).

En este contexto, el uso de la ciencia de datos y en específico, el machine learning emerge como una herramienta poderosa y prometedora para complementar los resultados de las inspecciones. La capacidad de procesar grandes volúmenes de datos y extraer patrones significativos permite mejorar la precisión en la estimación de la velocidad de corrosión, lo que a su vez facilita la toma de decisiones informadas en materia de mantenimiento predictivo y reparación de los ductos.

La aplicación del machine learning se divide en varias etapas: recolección, limpieza, procesamiento y el análisis exploratorio de los datos, la selección y entrenamiento de modelos, la aplicación de los modelos aprendidos y la validación de las predicciones.

A continuación, se compartirán los pasos transitados en el desarrollo de modelos predictivos de corrosión para su aplicación en nuestro sistema de gasoductos.

Definición del objetivo

En este desarrollo se busca predecir la velocidad de corrosión, específicamente la máxima velocidad por caño (12 metros de longitud promedio) teniendo en cuenta los factores que promueven esta amenaza.

Desarrollo

Adquisición de datos

A lo largo del proyecto se integraron datos provenientes de distintas fuentes (externa e internas) y formatos (GIS y hojas de cálculo). La siguiente figura presenta los datos utilizados durante el flujo de trabajo de predicción de velocidad de corrosión.

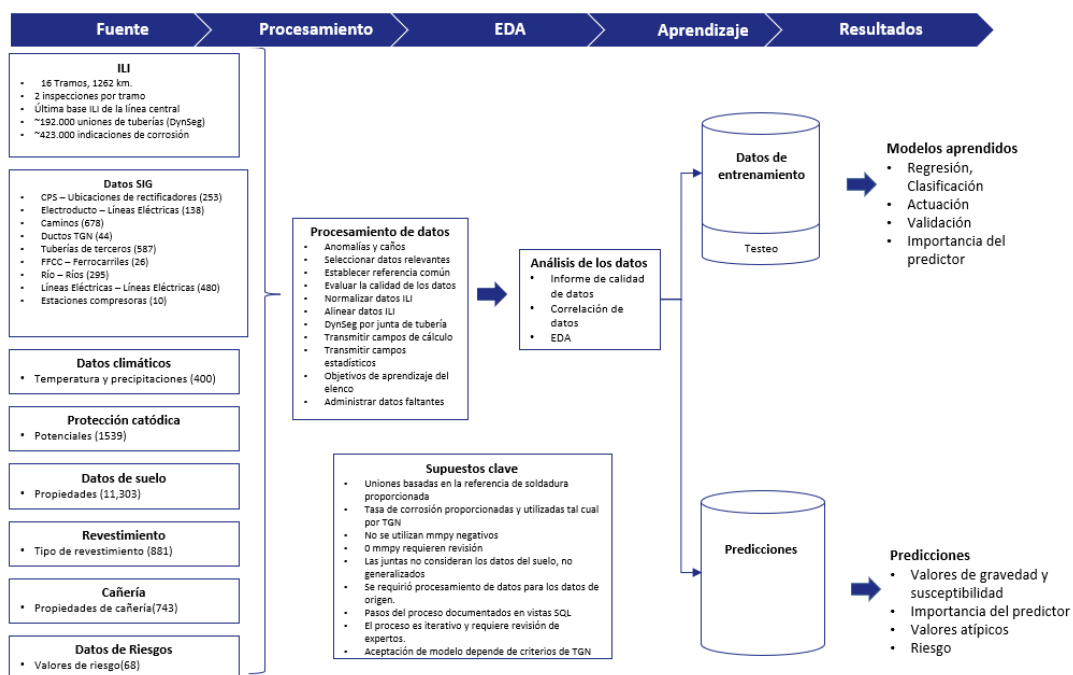


Figura 1. Flujo de trabajo del proyecto.

El primer paso, es recolectar los datos que sirven para estudiar la amenaza. En el contexto de cañerías enterradas, la corrosión externa se define como la degradación del material del caño debido a reacciones químicas con el suelo que lo rodea. Por lo tanto, al estudiar la corrosión externa interesan las características propias del ducto, del suelo y de cualquier medida que se utilice para monitorear su avance (resultados de inspecciones) y/o para contrarrestar su efecto (protección catódica y estado revestimiento). En este caso de estudio, se recolectaron los datos mencionados para 1916 km de gasoducto (16 tramos).

La extensión de analizada en el modelo de la corrosión mediante machine learning representa aproximadamente 17% de sistema de transporte de TGN. Con el objetivo de asegurar su representatividad del resto del sistema se seleccionaron gasoductos de distinto año de construcción, diámetro y ubicación geográfica.

En la siguiente figura se presenta la ubicación geográfica y características constructivas de los tramos de gasoductos utilizados para la realización de este proyecto. Se observa una amplia distribución geográfica, fechas de construcción entre 1960 y 2010; y diámetros de 16, 24 y 30 pulgadas.

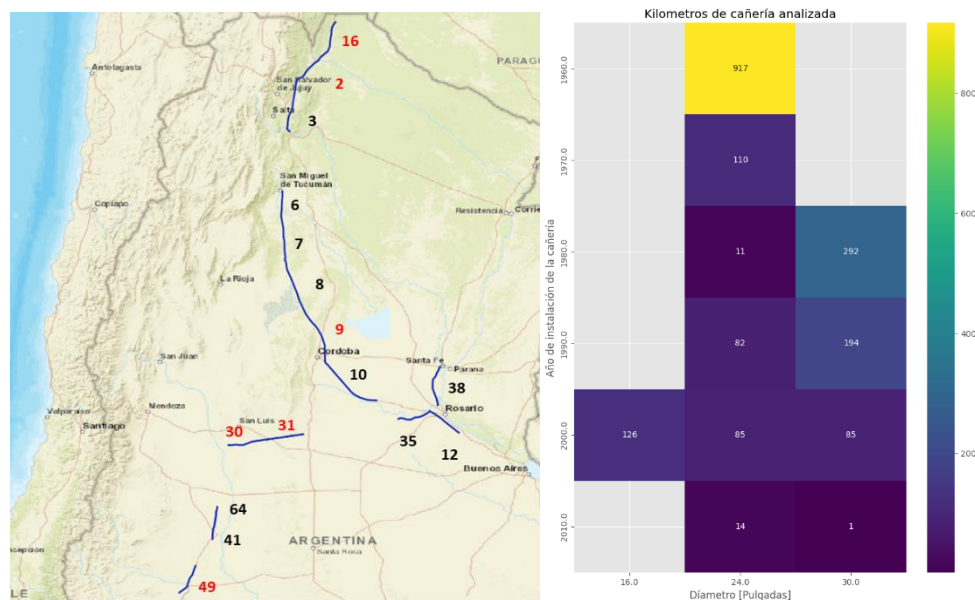


Figura 2. Ubicación geográfica y características constructivas de los gasoductos seleccionados.

Las etapas de procesamiento y análisis de datos son las etapas que conllevan mayor cantidad de tiempo dado que la calidad de los datos de entrada es la principal componente de la efectividad de las predicciones. En el proyecto se utilizaron las siguientes fuentes de datos:

- Datos ILI de reportes de proveedores de servicios de inspección interna.
- Datos de ductos, ferrocarriles y líneas eléctricas en base de datos nacionales.
- Datos de clima georreferenciados de institutos nacionales.
- Mediciones de potenciales, nivel de riesgo, cruces de ríos y cruces de caminos en bases de datos generadas por TGN.

Procesamiento

Dado que la totalidad de estos datos provienen de distintos orígenes (GIS, SAP, reportes externos), se requirió la integración de estos utilizando una única referencia. En la etapa de procesamiento se seleccionaron los datos relevantes, se evaluó la calidad de los datos eliminando datos erróneos, se gestionaron datos nulos y se normalizaron en caso de ser necesario (datos ILI). Durante esta etapa se encontró con la dificultad de obtener datos sobre las condiciones climáticas con suficiente resolución para obtener un perfil apropiado a la segmentación planteada.

Otro de los inconvenientes encontrados fue la falta de homogeneidad en la información por lo que se definieron estrategias para completar datos faltantes. Asimismo, se encontró variación en los formatos de las inspecciones de los reportes ILI que generó etapas de procesado adicionales.

Como se puede observar en la Tabla 1 existen datos faltantes o nulos los cuales deben ser tratados y en algunos casos se realizan presunciones para operar de forma correcta con los modelos.

Tabla 1. Ejemplo de análisis de calidad de datos para 6 de las variables utilizadas.

Variable	Métricas	Cantidades	Gráfico	Faltantes
Revestimiento del caño	1. Sin Dato 2. Cinta Plástica 3. Epoxi Liquido (y 6 otros)	3 (0,0%) 18.399 (9,6%) 1.085 (0,6%) 172.412 (89,8%)		1 (0,0%)
Susceptibilidad a la corrosión	1. Susceptibilidad baja 2. Susceptibilidad alta	189.486 (98,7%) 2.414 (1,3%)		0 (0,0%)
Máxima profundidad en última corrida [%]	Media: 5,52 Desviación estándar: 8,57	63 valores distintos		0 (0,0%)
Máxima precipitación [mm/año]	Media: 137,737 Desviación estándar: 48,626	19 valores distintos		44 (0,0%)
Máxima temperatura [°C]	Media: 31,128 Desviación estándar: 1,385	16 valores distintos		44 (0,0%)
Diferencia entre potenciales OFF-NATURAL [mV]	Media: 461,75 Desviación estándar: 219,161	733 valores distintos		30.597 (15,9%)

Se tomaron las siguientes estrategias a la hora del procesamiento de los datos:

- Se utilizaron los soldaduras o caños como mínima unidad de longitud del modelo.

- Las tasas de corrosión proporcionadas y utilizadas por TGN: se estimaron a partir de análisis de crecimiento de defectos de corrosión entre corridas ILI.
- No se utilizan velocidades de corrosión negativas: En los casos en los que se reportaron menores profundidades de un defecto en la última corrida, se considera la velocidad de corrosión nula.
- Los datos de suelo se alinearon con la traza y se aplicaron a la soldadura o caño más cercano. La alineación de los datos se realizó mediante la unión de consultas SQL. Este proceso es iterativo y requirió revisión de expertos: luego de cada alineación, se revisaron los datos aplicados a cada soldadura para asegurar que la correcta referenciación entre datos de distintas fuentes.

Como finalización de esta etapa, se obtuvieron 41 variables por cada junta de cañería en los gasoductos estudiados.

Análisis exploratorio de datos

Esta etapa del proceso consta en un estudio inicial de los datos para detectar posibles fallas en la adquisición de datos, tratamiento e identificación de datos faltantes, identificación de casos atípicos (outliers) y comprender cuales son las principales variables que correlacionan con la velocidad de corrosión. La Figura 3 presenta la distribución de velocidad de corrosión donde se observa que las mayores velocidades se ubican en revestimientos tipo epoxi líquido y mantas termocontraíbles, seguidos por esmalte asfáltico. En el caso de las mantas es posible que el despegue propio de la instalación y el apantallamiento característico de este revestimiento genere problemas de corrosión los cuales son de difícil detección; en el caso del revestimiento epoxi este es característico de zonas en las cuales se observan grandes cantidades de defectos y fueron sujetas a extensas obras de refrorrido, esto nos indica que la velocidad de corrosión calculada es previa a estas obras y no es representativa de la condición actual de la cañería. Es por estas razones que se debe complementar el análisis con la utilización del revestimiento original del ducto.

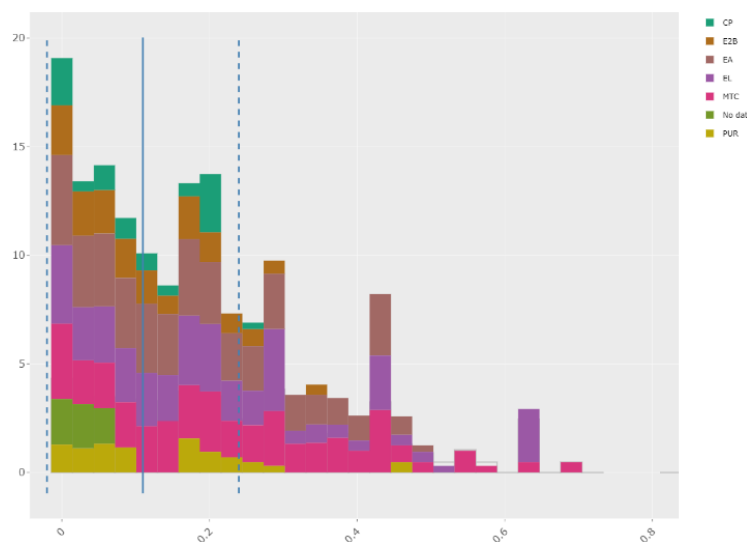


Figura 3. Histograma de velocidades de corrosión y tipos de revestimiento. Línea azul: valor medio; Líneas punteadas: valor medio más y menos una desviación estándar; CP: Cinta plástica; EA: esmalte asfáltico; EL: Epoxi líquido; PEB: Polietileno bicapa; PET: Polietileno tricapa; PUR: Poliuretano; MTC: manta termocontraíble.

La velocidad de corrosión promedio se ubicó cercana al 0,1 mm por año. Por otro lado, los caños que no cuentan con información presentan las menores velocidades en los gasoductos estudiados. Estos valores al ser nulos deben ser tratados posteriormente al análisis exploratorio de datos.

En cuanto a la cantidad de defectos por tipo de revestimiento se observa en la Figura 4 que la mayoría se concentra en los revestimientos tipo epoxi líquido, esmalte asfáltico, polietileno tricapa. Por esta razón, es posible que se obtenga menor desempeño en la predicción de velocidades de corrosión en los restantes revestimientos presentes en el sistema.

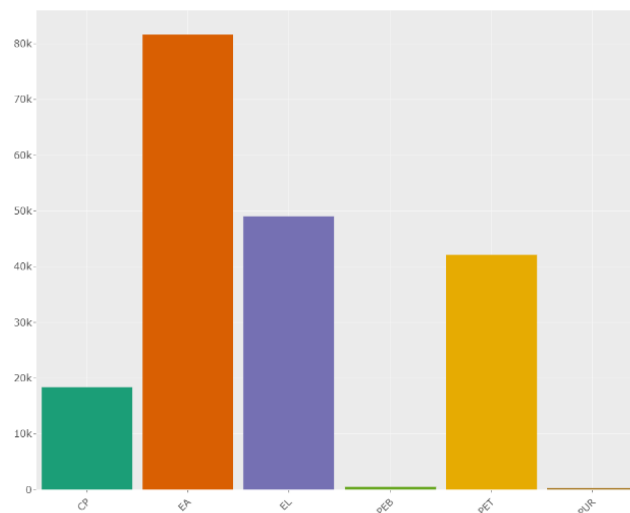


Figura 4. Cantidad de defectos de corrosión por tipo de revestimiento. CP: Cinta plástica; EA: Asfáltico; EL: Epoxi líquido; PEB: Polietileno bicapa; PET: Polietileno tricapa; PUR: Poliuretano; MTC: manta termocontraíble.

En el EDA, el análisis de correlación de las variables con la velocidad de corrosión juega un papel crucial para:

- Identificar relaciones: La correlación permite cuantificar la incidencia y dirección de la asociación entre variables. Un alto coeficiente de correlación positivo indica una fuerte relación directa, mientras que uno negativo indica una relación inversa. Aunque dos variables estén correlacionadas no indica causalidad.
- Detectar outliers: Los valores atípicos pueden distorsionar la correlación. Al analizar la correlación, es importante identificar y considerar posibles outliers para obtener una imagen precisa de la relación entre variables.
- Seleccionar variables: Al construir modelos, la correlación puede ayudar a seleccionar las variables más relevantes y evitar la multicolinealidad, un problema que surge cuando las variables están altamente correlacionadas entre sí.

Para analizar esta correlación una forma útil es realizar una representación gráfica que cuantifica la proporcionalidad entre las variables en el rango de 1 a -1 (correlación de Pearson). En la Figura 5 se presenta el estudio de correlación entre todas las variables consideradas inicialmente para el modelo, donde se observa correlaciones de valor 1 en el caso de algunas pares de variables. Esto indica que estas variables son combinaciones lineales de otras por lo que una de estas fue retirada o combinadas con otras para reducir la dimensionalidad.



Correlación de variables

Variable	Correlación
DIFERENCIA DE ANCH	0.00
CANTIDAD DE ANOMALIAS	0.02
CANTIDAD DE ANCH PREVIAS	0.08
CT	-0.02
FECHA DE REVESTIMIENTO	0.40
CONSECUENCIA	-0.15
PRESION DE DISEÑO	-0.55
DIAMETRO	0.10
DISTANCIA A PC	-0.20
DISTANCIA A LÍNEA ELÉCTRICA	0.20
DISTANCIA A ELECTRODUTO	-0.05
DISTANCIA A EQUIPO RECTIFICADOR	-0.20
DISTANCIA A CRUCE DE RÍO	0.32
DISTANCIA A CAMINO	0.32
DISTANCIA A FCC	0.15
Distance_TONPL_IT	0.05
Distance_TPNL_IT	0.08
EC_LOSS_VOL_IT	0.05
EC_Vol_500m	0.15
GRADO	-0.05
RANGO DE INSPECCIÓN	-0.08
LONGITUD DE JUNTA	-0.20
Kuratos	0.08
MAX PROF	0.02
MAX PROF PREV	0.30
MAX RECTIFICACIONES	-0.15
MAX TEMPERATURAS	0.30
FIN DE MEDICIÓN	0.18
EQUIPOS DE CP	0.18
INICIO DE MEDICIÓN	0.18
MIN RECTIFICACIÓN	-0.30
MIN TDS	0.10
MPY_L_IT	-0.05
CAMBO DE POTENCIALES	0.15
WT	0.18
ODOMETRO	-0.05
AÑO DE INSTALACION	-0.05
POTENCIAL NATURAL	-0.35
RANGO DE PRECIPITACION	-0.08
RANGO DE TDS	0.12
Slow_Mean_Distance	-0.08
Slow_Min_Distance	-0.05
Slowness	-0.05
TIME	-0.08

Figura 6. Correlación entre variables seleccionadas con la velocidad de corrosión.

Al analizar los datos incorporados al proyecto se observa una alta cantidad de variables que pueden incidir el proceso de corrosión. Con la finalidad de reducir la dimensionalidad y optimizar recursos es una buena práctica realizar un análisis de componentes principales (PCA).

El PCA es una técnica que reduce la dimensionalidad de los datos al transformar las variables originales en nuevas variables (componentes principales), estas nuevas variables son una combinación de las originales. Algunos puntos clave del PCA son:

- Reducir la cantidad de variables mientras se conserva la mayor cantidad posible de información.
- Los componentes principales son combinaciones lineales de las variables originales.
- Los primeros componentes principales representan las relaciones más importantes entre las variables originales.

En la Figura 7 se presentan los 10 componentes principales que conservan la mayor varianza del conjunto de datos. Se decidió mantener el 80% de la varianza en la información como límite para definir la cantidad de componentes principales a utilizar en el modelo. Al aplicar este criterio se definieron en total 14 componentes principales que representan al 81% de la varianza, disminuyendo la dimensionalidad de la información de 41 a 14 variables. Esta simplificación ayuda a mejorar la eficiencia computacional de los modelos de machine learning, ya que se necesita menos tiempo para procesar un número menor de variables.

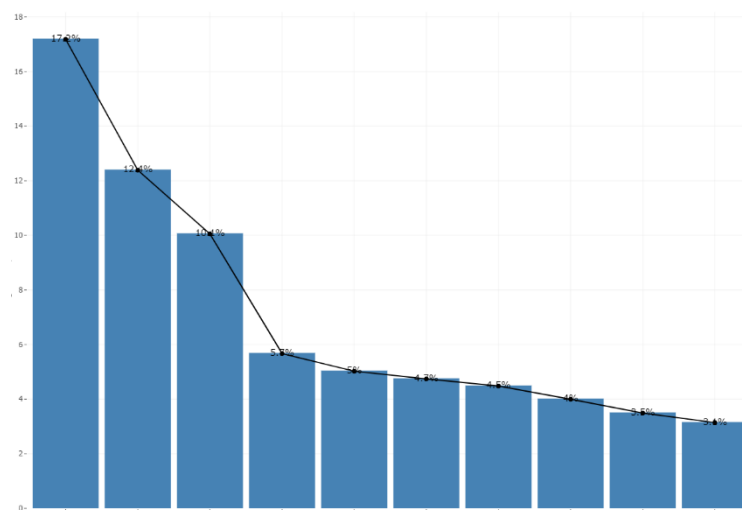


Figura 7. Análisis de los 10 componentes principales. Escala representa el porcentaje de varianza que conserva cada componente principal.

La Figura 8 presenta la composición de dos de los 14 componentes seleccionados para realizar la etapa de modelado. En el componente principal 1 se representan predominantemente datos climáticos y de revestimiento, a contraposición en el componente principal 2 predominan datos constructivos de los gasoductos como diámetro, sistema y tipo de soldadura longitudinal. De esta forma se puede mantener la varianza de los datos utilizados y disminuir la complejidad del modelo. Además, la implementación del PCA ayuda a mitigar el problema de la colinealidad, ya que los componentes principales son independientes entre sí. Sin embargo, el uso de PCA también tiene sus limitaciones al perderse la interpretación directa de las variables.

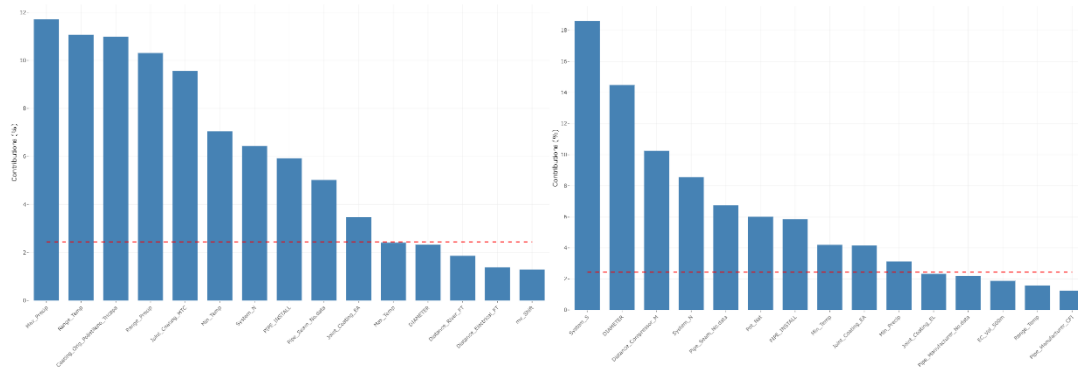


Figura 8. Composición de los componentes principales 1 y 2.

Aprendizaje

Habiendo realizado la limpieza y exploración de los datos se procedió a la etapa de selección del modelo a aplicar. Se puede abordar el problema de la predicción de la velocidad de corrosión desde dos puntos de vista distintos dentro de los algoritmos disponibles de machine learning: modelos de regresión y modelos de clasificación. En la Figura 9 se presenta el flujo de trabajo durante la etapa de modelado. En esta etapa se debe tener en cuenta que se requieren datos de entrenamiento y datos de testeo. En este caso en particular se dividieron los datos recolectados en 75% para que el modelo aprenda (o entrene) y 25% para testear los modelos y evaluar los resultados alcanzados.

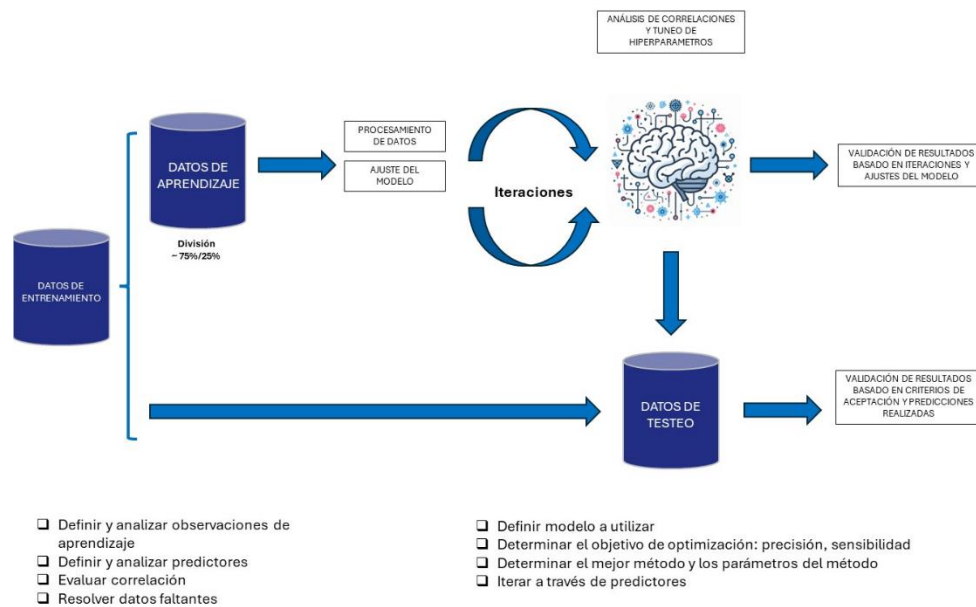


Figura 9. Flujo de trabajo en la etapa de modelado.

En el caso de los modelos de clasificación se planteó segmentar las cañerías en función de dos categorías basadas en su velocidad de corrosión (“corrosión activa” y “corrosión inactiva”). Para realizar esta clasificación se tuvo en cuenta las tolerancias de las distintas dimensiones reportadas por las herramientas ILI para determinar el umbral de clasificación (fijado en 0,05 mm por año). De esta forma los valores de velocidad de corrosión mayores al umbral fueron considerados como casos de corrosión activa y estas categorías fueron alimentadas al modelo como objetivo de las

predicciones. Este enfoque podría ser útil si el objetivo principal es identificar cañerías que tienen un alto riesgo de corrosión para tomar medidas preventivas o de monitoreo. En este caso se utilizó una cantidad reducida de datos del sistema debido a que se priorizaron los modelos de regresión y se empleó el algoritmo KNN para la clasificación.

Por otro lado, se planteó un modelo de regresión donde se buscó asignar a cada conjunto de valores de entrada una velocidad de corrosión correspondiente y permitir predecir un valor numérico continuo. Además, los modelos de regresión pueden ayudar a entender cómo cada predictor contribuye a la velocidad de corrosión, lo que puede ser útil para optimizar, controlar el proceso de corrosión y hacer más eficiente la campaña de verificación de defectos en campo.

En esta etapa se establecieron distintos valores para los hiperparámetros y se probaron 20 distintas combinaciones de estos. En la Figura 10 se presenta el error cuadrático medio (RMSE) y el coeficiente de determinación (R^2). Se utilizaron modelos XGBoost donde se varió la cantidad de “árboles”, su profundidad, el coeficiente de aprendizaje, cantidad de iteraciones y tamaño de muestra. Luego, se seleccionó el Modelo 16 por su buen ajuste a los datos de entrenamiento y que en simultaneo presenta un R^2 alto por lo que ayuda a evitar un sobreajuste manteniendo la mayor varianza posible. Este modelo se utilizó para continuar el análisis resultados sobre los datos de prueba o testeo.

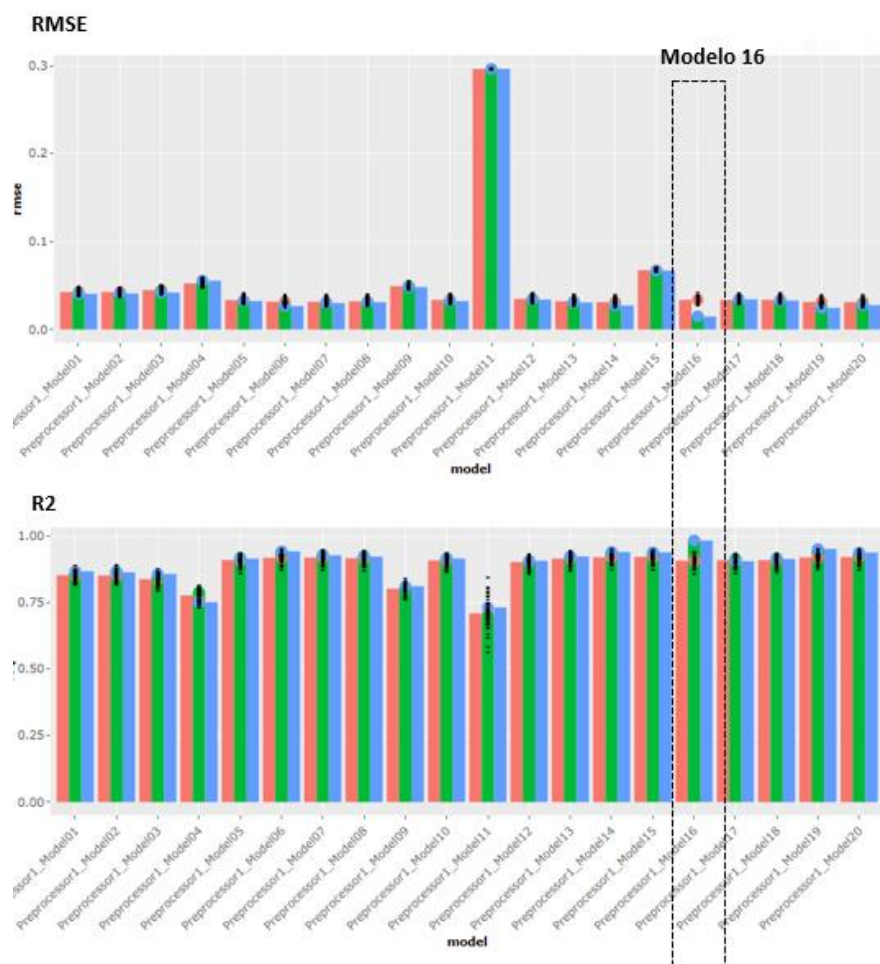


Figura 10. Error cuadrático medio y coeficiente de determinación en los modelos de regresión estudiados. Barra verde: desempeño en datos de aprendizaje; barra celeste: desempeño en datos de testeo o prueba; barra naranja: desempeño en subconjunto de los datos.

Análisis de resultados

Una forma de visualizar los resultados del modelo de clasificación se detalla en la Figura 11 en donde se nos presentan los resultados de las predicciones teniendo en cuenta las clasificaciones previamente definidas.

- Verdaderos positivos (TP): 1544 predicciones fueron clasificadas correctamente como positivas. Estas predijeron correctamente la existencia de corrosión activa.
- Falsos positivos (FP): 14 predicciones fueron clasificadas incorrectamente como positivas.
- Falsos negativos (FN): 57 predicciones fueron clasificadas incorrectamente como negativas.
- Verdaderos negativos (TN): 928 predicciones fueron clasificadas correctamente como negativas.

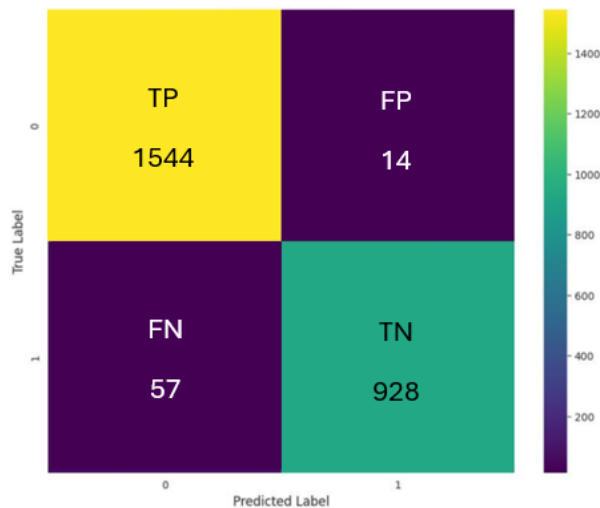


Figura 11. Matriz de confusión obtenida de los resultados del modelo clasificación.

Para finalizar el estudio del modelo de clasificación, el análisis se centra en las siguientes métricas:

- Precisión: Es la proporción de verdaderos positivos (TP) sobre la suma de verdaderos positivos y falsos positivos (FP). En otras palabras, mide la precisión del modelo al predecir la clase positiva. En este caso, la precisión para la clase de corrosión activa es del 96% y para la clase corrosión inactiva es del 99%.
- Recall: Es la proporción de verdaderos positivos (TP) sobre la suma de verdaderos positivos y falsos negativos (FN). Mide la capacidad del modelo para encontrar todos los casos positivos. En este caso, el recall para la clase de corrosión activa es del 99% y para la clase corrosión inactiva es del 94%.
- F1-score: Es una medida combinada de precisión y recall. Es útil cuando hay un desequilibrio entre las clases, ya que esta métrica es una combinación de las dos anteriores. El F1-score para la clase de corrosión activa es del 98% y para la clase de corrosión inactiva es del 96%.
- Accuracy (exactitud): Es la proporción de predicciones correctas sobre el total de predicciones y es del 97%.

Estas métricas brindan una idea del rendimiento general del modelo en términos de su capacidad para predecir correctamente las clases, se observa un rendimiento confiable en la predicción de puntos con corrosión activa. Cabe mencionar que este modelo fue aplicado a un subconjunto de los datos por lo que es posible obtener menor rendimiento al aplicar a todo el dataset.

En la Figura 12 se presenta los resultados para el modelo de regresión, donde se compara la velocidad de corrosión objetivo (calculada mediante métodos tradicionales) con la velocidad de corrosión predicha por el modelo. Como se puede ver existe una buena semejanza entre los valores. Por otro lado, al realizar un estudio de métricas se observa un RSME 0,03 que indica que las predicciones del modelo son, en promedio, muy cercanas a los valores reales. El R^2 de 0,91 indica un correcto ajuste explicando que describe el 91% de la variabilidad en los datos.

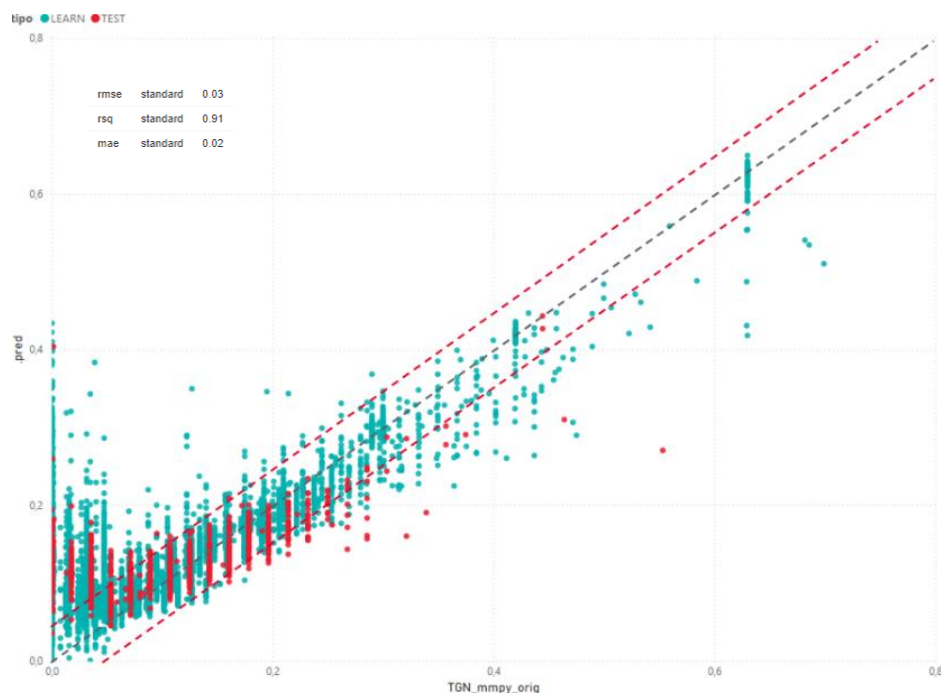


Figura 12. Velocidad de corrosión actual vs Velocidad de corrosión modelada. En rojo los datos utilizados en el aprendizaje y en verde los datos usados como prueba (test).

Al analizar las predicciones sobre los datos de testeo se observa que el 92% de estas cumplen con una tolerancia propuesta por TGN de 0,05 mm por año. Asimismo, se observa que la separación aleatoria del dataset generó un conjunto de datos de prueba y de aprendizaje que representan la totalidad de la información. El modelo de regresión presentado tiene un rendimiento consistente, con un error bajo y un buen ajuste.

En la Figura 13 se presenta el estudio de incidencia de las variables utilizadas en el modelo, esta es una herramienta para entender que condiciones están impulsando la velocidad de corrosión y planificar acciones de mitigación en caso de ser necesarias. Se observa que los principales predictores son la temperatura, la distancia a la planta compresora aguas arriba, precipitaciones y los potenciales de protección catódica.

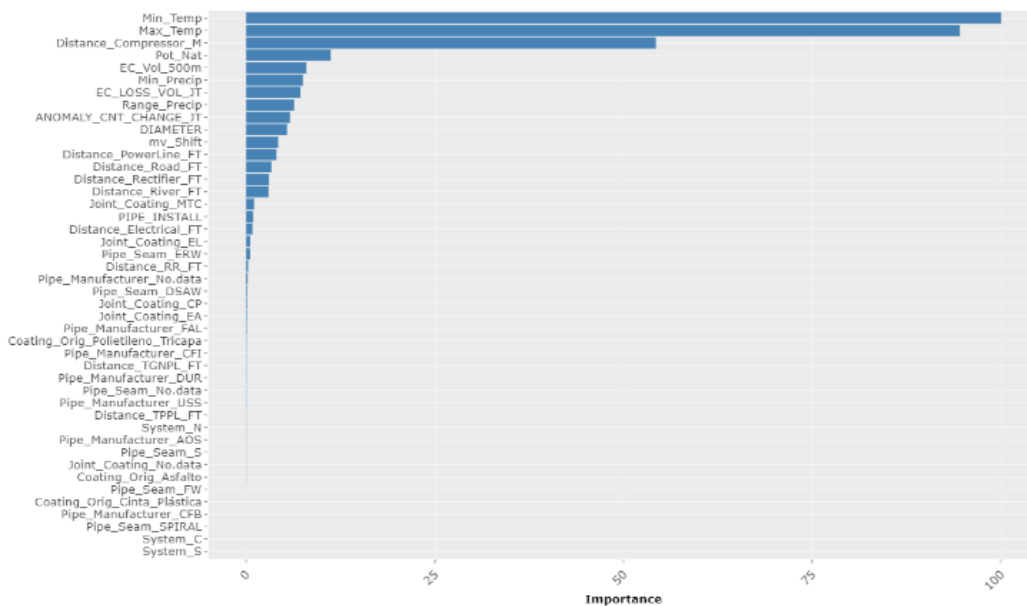


Figura 13. Peso de cada predictor en el modelo seleccionado.

Esta ponderación de predictores de la velocidad de corrosión son un promedio en todos los tramos analizados, su peso puede variar de tramo a tramo o dentro de secciones de un tramo. El peso de los predictores puede depender del tramo y modelo elegido.

Conclusiones

El proyecto de ML demostró ser una herramienta útil para la predicción de la velocidad de crecimiento de la corrosión. La identificación de las variables más relevantes permitió enfocar los esfuerzos de inspección y mantenimiento en los puntos críticos. La predicción precisa de la velocidad de corrosión permite tomar decisiones más informadas sobre la gestión de activos y la planificación de reparaciones.

Al comenzar un proyecto de machine learning es importante definir objetivos claros y realizar un estudio de los datos disponibles ya que si no se cuenta con información confiable el resultado del modelo no será representativo de la realidad.

Se observaron los siguientes beneficios de la aplicación de Machine learning en la gestión de integridad de cañería son:

- Rápido sistema de toma de decisiones
- Posibilidad de tener resultados más precisos
- Integración de grande caudales de información
- Automatización de procesos
- Transformación de datos a insights
- Capacidad de realizar acciones preventivas de manera programada y priorizada.
- Identificación de tendencias

Si bien se observa un buen desempeño del modelo, es crucial realizar un análisis más profundo para confirmar la ausencia de sesgo y la generalización del modelo a nuevos datos. Por este motivo se realizará como siguiente etapa 10 verificaciones en campo para validar los resultados del modelo y se incorporan nuevos tramos al estudio.

Bibliografía

NACE SP 502. Pipeline External Corrosion Direct Assessment Methodology

Applied Predictive Modeling. Max Kuhn y Kjell Johnson. 2013

The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World. Pedro Domingos. 2015

Analytics & Machine Learning for Pipeline Integrity & Risk. Michael Gloven

Breve Cv

Alan Cornejo

Ingeniero Mecánico, graduado de la Universidad Tecnológica Nacional (UTN FRBA), con más de 9 años de experiencia en la industria del petróleo y gas. Desde 2015, me he dedicado a tareas relacionadas con la gestión de inspecciones internas, la definición de verificaciones en campo, el análisis de aptitud para el servicio y la implementación de medidas de mitigación.

Martin Melucci

Ingeniero en Materiales, graduado en la Universidad Nacional de Mar del Plata. Se desempeñó como Consultor de Integridad en GIE desde el 2021 hasta el 2023. Desde ese año a la actualidad se desempeña como Analista de Integridad en el Departamento de Integridad de TGN.

Marianela Gomez

Ingeniero Química, graduada en la Universidad Nacional de Mar del Plata. Se desempeñó como Ingeniera de Integridad en Sintec desde el 2014 hasta al 2017. Desde ese año a la actualidad se desempeña como Analista de Integridad en el Departamento de Integridad de TGN.