

Enero 2026

Por qué importa profesionalizar los modelos de predicción

Evidencia a partir de la demanda diaria del transporte urbano

Motivación: ¿Para qué predecir?

Sea en un trabajo académico, una consultoría profesional o una planificación operativa, la necesidad de predecir comportamientos futuros es constante. Ahora bien, predecir no implica acertar, sino minimizar el margen de error en la estimación: conocer el pronóstico del tiempo sirve para decidir si llevar el paraguas o dejarlo en casa; conocer el volumen de ventas del próximo año sirve para calcular la cantidad de insumos y horas hombre necesarias para satisfacer la demanda. En todos los casos, lo que interesa es disminuir la incertidumbre en la estimación. La elección de la metodología adecuada resulta crucial para separar señal de ruido.

Datos y enfoque de análisis

Para ilustrar este punto, se presenta un ejercicio aplicado a la predicción de demanda diaria del transporte público de colectivos en la ciudad de Santa Fe.

El objetivo es comparar el desempeño de nuevos métodos de predicción —basadas principalmente en técnicas de *machine learning*— respecto de enfoques convencionales como promedios históricos o extrapolaciones lineales. No se trata de desestimar estos últimos —que en muchos casos sirven como referencia inicial— sino de mostrar las ventajas concretas de utilizar herramientas más sofisticadas.

Contar con predicciones precisas de la demanda diaria permite optimizar la asignación de recursos, mejorando la eficiencia operativa y la satisfacción de los usuarios. Este informe muestra, de manera práctica, cómo las técnicas modernas reducen sustancialmente los errores de estimación respecto de enfoques básicos.

La demostración se organiza en tres pasos: **primero** se describe la dinámica de la demanda de transporte en la ciudad de Santa Fe; **luego** se contrastan las predicciones obtenidas mediante métodos convencionales basados en promedios históricos; y, **por último**, se presentan los resultados de enfoques alternativos de predicción. El trabajo cierra con una sección de **reflexiones finales**.

El transporte urbano por colectivos en Santa Fe

La ciudad de Santa Fe cuenta con un sistema de transporte urbano por colectivos compuesto por 18 líneas. Desde la implementación del sistema SUBE en 2019, la adopción ha

sido progresiva, con un marcado crecimiento desde 2020 en adelante, influido por la recuperación post-pandemia.

La demanda semanal promedio por línea se encuentra entre 30.000 y 40.000 boletos¹, con excepciones destacadas como la línea 5 (70.000) y la línea 18 (50.000). Las líneas 20, 21, 22 y Ronda B presentan los menores niveles de demanda, con una actividad irregular durante todo el período analizado.

Debido a su nivel de demanda, el análisis se realizará teniendo en cuenta dicha línea, siendo replicable el mismo trabajo para cualquier otra línea del sistema².

La Línea 5 como caso de estudio

La línea 5 conecta el extremo noroeste con el extremo sur de la ciudad de Santa Fe. Si bien la distancia directa entre ambos puntos es de aproximadamente 11 km, el recorrido efectivo —tanto en la ida como en la vuelta— duplica esa extensión debido a las calles y desvíos que atraviesa. En su trayecto, la línea recorre diversos barrios de la ciudad, integrando zonas residenciales y sectores de alta circulación. La siguiente figura ilustra el recorrido completo.

Figura 1. Recorrido del colectivo línea 5 en la ciudad de Santa Fe.



Fuente: elaboración propia.

Antes de iniciar cualquier estudio de predicción —y en general, cualquier análisis más complejo— es fundamental revisar algunas estadísticas descriptivas de la variable objetivo. Estas medidas permiten dimensionar la magnitud de la demanda y reconocer patrones básicos que

¹Datos recolectados por red SUBE (Ministerio de Transporte).

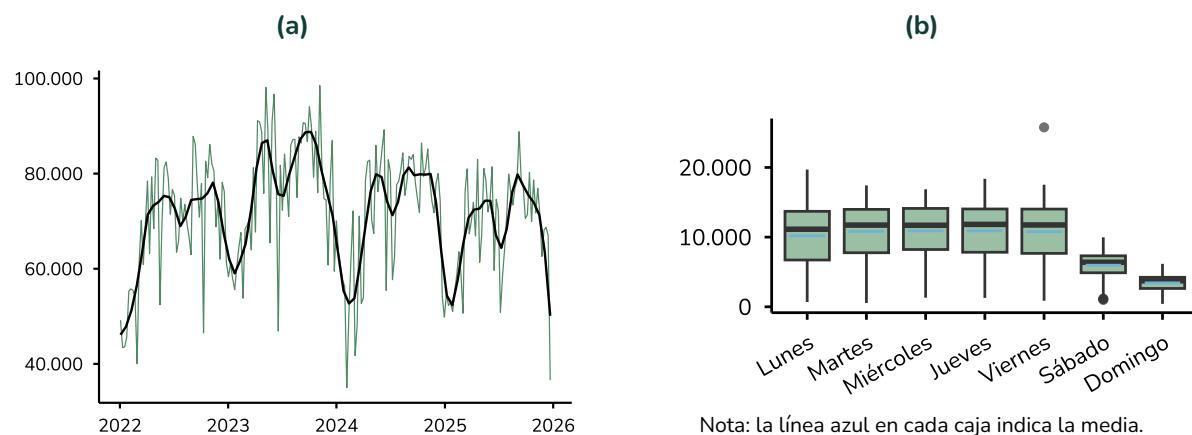
²De hecho, se ha realizado este trabajo alcanzando resultados similares en cuanto a la mejora en la predicción, con algunas heterogeneidades. Dicho trabajo se encuentra a disposición de consulta.

luego orientan el modelado.

Tal como se ha señalado, la Línea 5 se consolida como la de mayor demanda dentro del sistema de transporte urbano santafesino. Durante el período 2022–2025, el promedio anual de boletos diarios se mantuvo cercano a los 10.000, mientras que el promedio semanal se ubicó en torno a los 70.000 boletos.

La distribución por día de la semana muestra un comportamiento estable de martes a viernes, con medias diarias superiores a los 12.000 boletos y medianas cercanas a los 13.500. El lunes presenta una leve diferencia respecto de los demás días hábiles, con una media algo menor (11.739 boletos). En contraste, la demanda cae de manera marcada durante el fin de semana: los sábados promedian alrededor de 6.800 boletos diarios y los domingos apenas 4.000. Las figuras siguientes ilustran estos patrones.

Figura 2. Evolución de la cantidad de boletos semanales (A) y frecuencia de boletos diarios por día (B). Línea 5 de la ciudad de Santa Fe, período 2022-2025



En este momento entra en juego nuestro trabajo. Si se observa el año 2025, y nublando brevemente los ojos, se podrá comprobar que resulta una aproximación de los tres años previos. En efecto, la silueta semanal (véase Figura 2a) del año 2025 parece ubicarse por encima del año 2022, bastante por debajo del año 2023 y levemente por debajo del año 2024.

Frente a esta observación surge la siguiente reflexión: si uno estuviese parado en el día 31 de diciembre de 2024 ¿podría estimar la demanda de transporte simplemente como el promedio de los últimos tres años?; ¿qué tanto error implicaría esa aproximación? ¿existen métodos más adecuados que limitarse a observar promedios diarios, semanales o anuales? La siguiente sección busca responder estas cuestiones, explorando alternativas de modelado que capturen mejor la dinámica real de la demanda.

Un primer intento: predicciones basadas en promedios

Se comienza con métodos tradicionales que podrían servir para estimar la demanda de transporte. Para ello se consideraron seis alternativas tradicionales o simples para estimar bo-

letos diarios del año 2025, mostrando los resultados en términos semanales para facilitar la comparación y reducir la variabilidad diaria. En primer lugar, se podría pensar que los días 1 de enero de cada año se comportan de manera similar, al igual que los días 2 de enero, y así sucesivamente. Siguiendo esta lógica, el primer método asigna a cada fecha el promedio histórico de boletos emitidos en ese mismo día (**promedio día del año**). Los dos métodos siguientes tienen un razonamiento similar, pero en lugar de agrupar por día del año, se agrupa por día del mes (todos los 1 de cada mes se comportan iguales, y así sucesivamente, **promedio día del mes**) y por día de la semana (todos los lunes se comportan igual, y así sucesivamente, **promedio día de la semana**).

Posteriormente vienen métodos un poco más sofisticados, pero igualmente tradicionales. ¿Qué sucede si en vez de promediar por día del año, mes o semana, se promedia utilizando una ventana móvil? En este caso, se consideraron dos alternativas: un promedio móvil de 7 días y otro de 30 días. Esto implica que para predecir la demanda de un día determinado, se toma el promedio de los últimos 7 o 30 días previos. El primero (**promedio móvil de 7 días**) permite capturar mejor la dinámica reciente de la serie, mientras que el segundo (**promedio móvil de 30 días**) suaviza las fluctuaciones y refleja tendencias más estables en el tiempo.

Finalmente, se incluyó una sexta alternativa basada en el (**promedio por trimestre**). Este método supone que los patrones de demanda se repiten en ciclos de tres meses, lo que permite captar cierta estacionalidad más amplia sin necesidad de modelos complejos. En conjunto, estas seis estrategias conforman un conjunto de predicciones base que sirven como referencia inicial para evaluar el comportamiento de la demanda y comparar su desempeño frente a los valores observados.

Ahora bien, ¿cómo se suele medir la eficacia de un modelo al predecir? La respuesta está en el uso de métricas de error que permiten cuantificar la diferencia entre los valores observados y los valores estimados por cada método. Entre las más utilizadas se encuentran:

- RMSE (*Root Mean Squared Error*): mide la magnitud promedio de los errores al cuadrado. Penaliza fuertemente los errores grandes y es útil para evaluar la precisión general del modelo.
- MAE (*Mean Absolute Error*): calcula el promedio de los errores absolutos, ofreciendo una medida más intuitiva de la desviación promedio entre predicción y realidad.
- sMAPE (*Symmetric Mean Absolute Percentage Error*): expresa el error en términos relativos, comparando la diferencia entre observado y predicho respecto a su magnitud. Es especialmente útil para comparar modelos en series con diferentes escalas.

Estas métricas permiten establecer un ranking de desempeño entre los distintos métodos base considerados. Además de las métricas comunes (RMSE, MAE y sMAPE), se agregaron dos medidas relativas: **RMSE_rel** y **MAE_rel**. Ambas expresan el error en relación con la media diaria de boletos emitidos durante 2025, lo que permite interpretar los resultados en términos porcentuales y comparar la magnitud del error respecto del nivel promedio de la serie. En la

tabla siguiente se presentan los resultados obtenidos al aplicar cada uno de los métodos sobre la serie de boletos emitidos durante 2025.

Cuadro 1. Métodos convencionales de predicción. Métricas absolutas (RMSE, MAE, sMAPE) y relativas (RMSE_rel, MAE_rel) respecto de la media diaria.

| Modelo | RMSE | MAE | sMAPE | RMSE_rel | MAE_rel |
|------------------------|-------|-------|-------|----------|---------|
| Promedio día de semana | 2.928 | 2.034 | 0.228 | 0.300 | 0.209 |
| Promedio móvil 7d | 4.042 | 3.515 | 0.409 | 0.415 | 0.361 |
| Promedio móvil 30d | 4.169 | 3.696 | 0.425 | 0.428 | 0.379 |
| Promedio mes del año | 4.252 | 3.563 | 0.410 | 0.436 | 0.366 |
| Promedio trimestre | 4.317 | 3.617 | 0.414 | 0.443 | 0.371 |
| Promedio día del año | 5.571 | 4.567 | 0.495 | 0.572 | 0.469 |

Tal como se definieron las métricas, en este caso se observa que el método basado en el **promedio por día de semana** presenta el menor error en todas las medidas, tanto absolutas (RMSE, MAE, sMAPE) como relativas (RMSE_rel y MAE_rel), representando entre 1.000 y 1.500 boletos semanales de diferencia con respecto al resto de métodos (según RMSE), y entre 18 y 28 puntos porcentuales en términos de sMAPE.

Así, una predicción utilizando este método para una semana determinada del año 2025 erra en aproximadamente 2.928 boletos semanales en términos de RMSE. En términos relativos, esto equivale a cerca del 30 % del nivel promedio diario de la serie (cuando se consolida a nivel semanal). Dicho de otro modo, el error típico equivale a perder de vista aproximadamente tres de cada diez boletos respecto de lo que se espera en promedio. Asimismo, en términos de MAE, se erra en 2.034 boletos semanales, y la predicción se desvía en promedio un 22.8 % respecto del valor observado (métrica sMAPE).

A decir verdad, un sMAPE como el obtenido con dicho modelo se considera un resultado bastante bueno para un método tan simple, ya que en la práctica los modelos de predicción suelen mostrar errores porcentuales más elevados. En la literatura aplicada, se suele considerar que valores de sMAPE por debajo de 20–30 % indican un modelo razonablemente aceptable, mientras que cifras superiores al 50 % reflejan predicciones poco útiles.

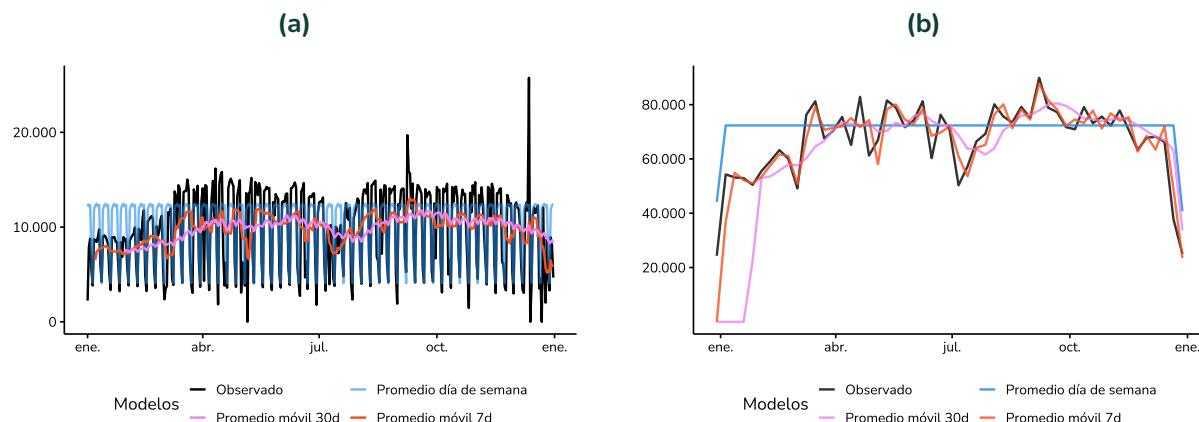
No obstante, este método adolece de una cuestión importante que probablemente ya se haya advertido. Recuérdese que este enfoque asigna el mismo valor a todos los lunes, lo mismo que al resto de los días de la semana. De esta manera, al sumar los boletos en formato semanal, todas las semanas terminan con la misma cantidad (y lo mismo ocurre si se agregan en cualquier otro bloque de 7 días). En otras palabras, este método no logra capturar la dinámica completa de la serie, ya que los picos asociados a eventos especiales, feriados o variaciones estacionales quedan fuera de su alcance.

Para ilustrar esta limitación, en la siguiente figura se muestran los tres mejores modelos

en términos de la métrica de referencia³, incluyendo al método aparentemente ganador junto con los otros dos enfoques más competitivos.

Estos son: **promedio por día de semana**, **promedio móvil de 7 días** y **promedio móvil de 30 días**. En el gráfico se observa la evolución semanal de la demanda observada junto con las predicciones generadas por estos tres enfoques.

Figura 3. Métodos convencionales. Boletos observados y predichos a nivel diario (a) y semanal (b). Línea 5 de la ciudad de Santa Fe, año 2025



Visto de esta manera, el método más eficiente para predecir valores diarios —el promedio por día de semana— pierde potencia cuando se lo analiza en términos de dinámica agregada. Al asignar valores fijos a cada día, genera semanas idénticas entre sí, lo que impide detectar variaciones reales entre semanas consecutivas. En cambio, los métodos móviles, si bien presentan errores más altos en promedio, permiten visualizar mejor los cambios graduales, las subidas y bajadas asociadas a eventos puntuales y las transiciones estacionales. Esta capacidad de adaptación se vuelve evidente en la figura, donde las curvas móviles acompañan con mayor sensibilidad la evolución semanal, aunque sin alcanzar plenamente los niveles observados.

Por su naturaleza, estos métodos tampoco logran captar la volatilidad diaria (véase Figura 3a), pero vistos desde una lógica semanal (véase Figura 3b) adquieren mayor relevancia, ya que siguen de cerca la dinámica general de la serie. Sin embargo, aparece otra limitación: no representan con eficacia los meses de inicio y cierre del período predicho, debido a la forma en que se construyen las ventanas móviles.

En suma, se observa una tensión entre precisión promedio y capacidad de seguimiento: el método más exacto no siempre es el más expresivo, y viceversa.

³Se utiliza RMSE como criterio principal de selección, dado que penaliza más fuertemente los errores grandes y es la medida más extendida en la evaluación de modelos de predicción.

Más allá de los promedios: nuevas técnicas de predicción

Los modelos avanzados que se presentan a continuación fueron entrenados de manera robusta, aunque no con un nivel de optimización exhaustivo. Se aplicó validación cruzada con ventanas rodantes, lo que supone un grado adicional de rigor frente a los enfoques tradicionales, pero sin llegar a un ajuste máximo de hiperparámetros ni a un despliegue computacional intensivo. En otras palabras, los resultados aquí mostrados son perfectibles: con mayor inversión de tiempo y recursos podrían obtenerse métricas aún mejores. Sin embargo, el propósito de este ejercicio es demostrar que adoptar métodos más avanzados ya representa una ganancia sustantiva frente a las aproximaciones simples. Un paso adicional, naturalmente, sería profundizar este enfoque mediante modelos más finos y optimizados.

Ahora sí, a los resultados. La comparación con enfoques más avanzados permite dimensionar la mejora alcanzada. Mientras que los métodos tradicionales ofrecían predicciones con un error de al menos 4.000 boletos semanales —con excepción del mejor de ellos—, los modelos aquí presentados erran por menos de 2.700 boletos semanales, aproximadamente. En efecto, todos son mejores que el mejor de ellos. Las nuevas técnicas de predicción logran reducir de manera sistemática los errores y capturar mejor la dinámica de la serie. En la Tabla 2 se resumen los resultados de cinco modelos entrenados y evaluados mediante validación cruzada.

Cuadro 2. Nueva técnicas de predicción. Métricas absolutas (RMSE, MAE, sMAPE) y relativas (RMSE_rel, MAE_rel) respecto de la media diaria.

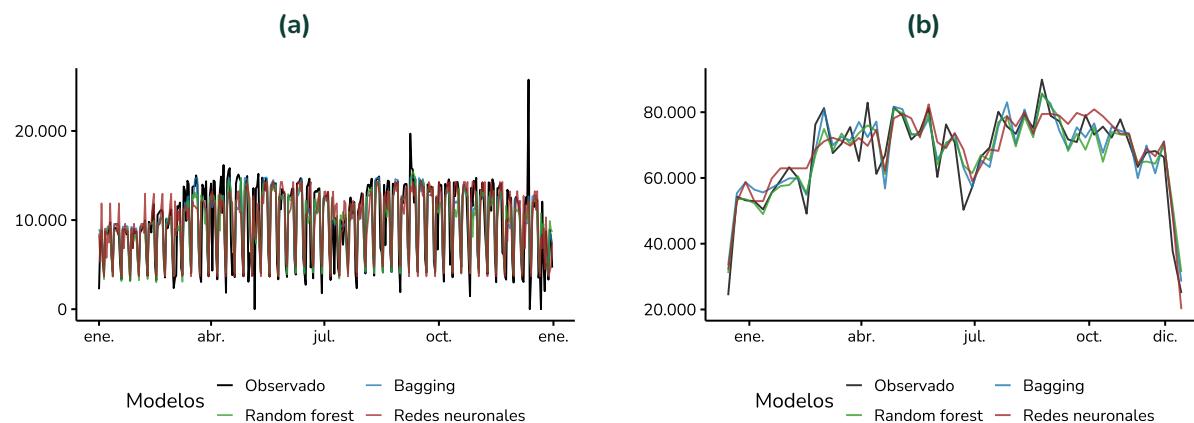
| Modelo | RMSE | MAE | sMAPE | RMSE_rel | MAE_rel |
|------------------|-------|-------|--------|----------|---------|
| Random forest | 2.106 | 1.597 | 18,3 % | 20,4 % | 15,5 % |
| Bagging | 2.113 | 1.612 | 18,7 % | 20,5 % | 15,6 % |
| XGBoost | 2.218 | 1.691 | 19,6 % | 21,5 % | 16,4 % |
| Regresión | 2.265 | 1.790 | 22,1 % | 21,9 % | 17,3 % |
| Redes neuronales | 2.653 | 2.109 | 24,0 % | 25,7 % | 20,4 % |

Aunque no interesa hablar en términos técnicos, vale la pena señalar que los modelos **Bagging** y **Random Forest** se diferencian en cómo combinan varios árboles de decisión, pero en esencia ambos buscan lo mismo: suavizar las predicciones y evitar que un único árbol se equivoque demasiado. Por eso aparecen tan cerca en los resultados, casi como un empate técnico. Más atrás, **XGBoost** y **Regresión** forman otro bloque intermedio, mientras que las **redes neuronales** quedan más lejos, mostrando que en este caso no lograron adaptarse tan bien a la dinámica de la demanda.

La Figura 4 muestra cómo se comportan los modelos seleccionados frente a la serie observada. En el panel (a), con datos diarios, se aprecia que Bagging y Random Forest siguen de manera bastante fiel la dinámica de la demanda, mientras que las redes neuronales presentan desvíos más notorios. En el panel (b), con agregación semanal, las diferencias se suavizan y se

observa con mayor claridad la tendencia general, destacando nuevamente la superioridad de los modelos basados en árboles frente al resto.

Figura 4. Nuevas técnicas de predicción. Boletos observados y predichos a nivel diario (a) y semanal (b) Línea 5 de la ciudad de Santa Fe, año 2025



Cierre: implicancias de mejorar la predicción

El análisis desarrollado en este trabajo permitió comparar distintos enfoques de predicción de la demanda de transporte a partir de un caso empírico concreto, contrastando modelos convencionales vs modelos avanzados basados en técnicas modernas de *machine learning*. Los resultados muestran que, aun cuando los métodos tradicionales pueden ofrecer aproximaciones razonables en determinados períodos, los modelos avanzados exhiben una mayor precisión y, especialmente, una dinámica de predicción más estable a lo largo del tiempo.

Más allá de las diferencias estadísticas entre modelos, las brechas de predicción observadas tienen implicancias operativas concretas. En contextos de gestión del transporte urbano, errores sistemáticos o predicciones inestables pueden traducirse en una asignación ineficiente de unidades, subutilización de recursos en semanas de baja demanda, o congestión y deterioro del servicio en picos no anticipados. En este marco, la mayor precisión y estabilidad de enfoques avanzados en materia predictiva no representa únicamente una mejora técnica, sino una herramienta clave para la planificación operativa y la toma de decisiones basada en evidencia.

En consecuencia, la adopción de estas nuevas técnicas de predicción se presenta como un paso necesario para fortalecer los procesos de planificación y evaluación en sistemas de transporte, contribuyendo a un uso más eficiente de los recursos disponibles y a una mejor calidad del servicio ofrecido a los usuarios.