INSTITUCIÓN UNIVERSITARIA TECNOLÓGICA DE SANTANDER

HERRAMIENTAS DE INTELIGENCIA ARTIFICIAL Y CIENCIA DE DATOS

ANÁLISIS EXPLORATORIO Y MODELADO PREDICTIVO PARA LA GESTIÓN DEL RIESGO DE RETRASOS EN NUTRIFRESH S.A.S.

Informe técnico final · Grupo 16

Catalina Herrera Gómez

Karolain Yadixa Huerfano Reyes

Sara Lucía Lozano Rueda

Docente: José Luis González Serrano

Girón, Santander · 8 de noviembre de 2026

Resumen ejecutivo

NutriFresh S.A.S. es una empresa de distribución de alimentos perecederos que, según el caso académico, afronta aproximadamente un 20 % de entregas tardías y un incremento del 15 % en costos logísticos. Este informe documenta el diagnóstico exploratorio y un piloto de clasificación del riesgo de retraso a partir de datos académicos; no constituye una validación de la operación real de la empresa.

Se utilizó como fuente central el conjunto de pedidos registrado entre el 1 de octubre y el 14 de diciembre de 2023. De 228 registros iniciales se eliminaron ocho duplicados exactos, quedando 220 pedidos: 168 puntuales, 49 retrasados y tres devueltos. Para el clasificador se excluyeron las devoluciones y se trabajó con 217 registros. Las correlaciones de distancia con tiempo de recorrido ($r = 0{,}973$), costo ($r = 0{,}965$) y combustible ($r = 0{,}951$) fueron las asociaciones descriptivas principales.

Se desarrollaron un árbol de decisión y una regresión logística, además de una línea base mayoritaria. La evaluación cronológica reservó 43 pedidos de diciembre para prueba; de ellos, 11 tenían retrasos confirmados. El árbol detectó tres retrasos (recall = 27,3 %), generó quince falsas alertas y obtuvo ROC-AUC = 0,402. La regresión logística tampoco mostró desempeño aceptable. Por tanto, no se recomienda despliegue operativo. El valor del trabajo radica en la trazabilidad, la identificación de límites de calidad y la propuesta de fortalecer la captura de información previa al despacho.

Contenido

1. Introducción

2. Contexto del problema

3. Objetivos del análisis

4. Metodología de trabajo (CRISP-DM)

5. Descripción de fuentes y transformaciones

6. Análisis exploratorio de datos

7. Modelado predictivo y evaluación

8. Interpretación de resultados

9. Recomendaciones gerenciales

10. Limitaciones del análisis

11. Conclusiones

12. Referencias bibliográficas (APA 7)

13. Anexos y evidencias

1. Introducción

La ciencia de datos permite examinar información operativa, identificar patrones y evaluar modelos que apoyen decisiones en contextos complejos. En logística de alimentos perecederos, el cumplimiento de plazos, la conservación de la cadena de frío y el control de costos requieren relacionar pedidos, rutas y condiciones externas.

Este informe consolida la secuencia de actividades del caso académico NutriFresh S.A.S. y describe un análisis exploratorio respaldado por cálculos reproducibles, seguido de un ejercicio de clasificación supervisada. El propósito no es demostrar que la inteligencia artificial mejora automáticamente la operación, sino establecer qué resultados se sostienen con la información suministrada, cuáles decisiones metodológicas fueron necesarias y qué límites impiden recomendar un modelo en producción.

El trabajo se organiza con la referencia CRISP-DM (Chapman et al., 2000), la exploración gráfica y estadística planteada por Heckert y Filliben (2003) y la evaluación rigurosa de clasificadores en presencia de clases desbalanceadas (Saito & Rehmsmeier, 2015).

2. Contexto del problema

2.1. Caracterización de la empresa

De acuerdo con el caso suministrado, NutriFresh S.A.S. distribuye frutas, verduras, lácteos y carnes frías en Atlántico, Bolívar y Magdalena. Atiende entre 500 y 700 pedidos diarios mediante 25 camiones refrigerados. Estos datos describen el escenario empresarial propuesto; no son estadísticas inferidas de los registros analizados (Caso empresarial: Optimización logística en NutriFresh S.A.S., s. f.).

2.2. Problemas de negocio

- Retrasos: el caso estima que cerca del 20 % de los pedidos se entrega fuera del tiempo pactado.
- Costos: el caso describe un aumento del 15 % en gastos de transporte, asociado a recorridos y combustible.
- Demanda: se presentan ruptura de inventarios en algunas zonas y sobreabastecimiento en otras.
- Fragmentación: pedidos, rutas GPS y factores externos no se encuentran integrados sistemáticamente.

2.3. Pregunta orientadora

¿En qué medida la información de los pedidos disponible antes del despacho permite distinguir pedidos con retraso confirmado de aquellos entregados a tiempo, y cómo pueden utilizarse los hallazgos descriptivos para priorizar mejoras de captura y gestión de datos?

El proyecto se concentra en retrasos. El estudio de costos es descriptivo; no se implementó un optimizador de rutas ni un pronóstico de demanda, y por ello no se atribuyen esos resultados a la solución.

3. Objetivos del análisis

3.1. Objetivo general

Analizar las características logísticas registradas para NutriFresh S.A.S. y evaluar un primer modelo supervisado de clasificación de retrasos, con el fin de documentar resultados reproducibles, identificar limitaciones y formular recomendaciones sustentadas para la gestión operativa.

3.2. Objetivos específicos

1. Examinar la procedencia, estructura, integridad y completitud de los registros académicos.
2. Caracterizar distribuciones, valores extremos, tasas de incidencias y relaciones entre variables relevantes.
3. Definir y justificar la variable objetivo, el tratamiento de devoluciones y las restricciones temporales de las predictoras.
4. Entrenar y contrastar un árbol de decisión y una regresión logística con una evaluación cronológica y métricas adecuadas.
5. Interpretar críticamente los resultados y proponer acciones de mejora proporcionales a la evidencia disponible.

4. Metodología de trabajo: CRISP-DM

Se adoptó el esquema CRISP-DM, cuyo ciclo de vida agrupa comprensión del negocio, comprensión de datos, preparación, modelado, evaluación y despliegue. Las fases son iterativas y se ajustan a los resultados de las anteriores (Chapman et al., 2000; IBM, s. f.). En este ejercicio la fase de despliegue se limita a la comunicación académica, sin integración de sistemas empresariales.

Fase	Aplicación comprobable en el proyecto
Comprensión del negocio	Selección del retraso como problema principal; distinción entre riesgo operativo y costos.
Comprensión de datos	Revisión de los archivos suministrados; selección y auditoría del conjunto de pedidos de 2023.
Preparación	Eliminación de ocho duplicados, estandarización de ciudades y peso; definición del objetivo.
Modelado	Entrenamiento de árbol de decisión, regresión logística y predictor mayoritario.
Evaluación	Validación interna por bloques temporales y prueba cronológica independiente de 43 pedidos.
Despliegue/comunicación	Informe, código, gráficos, recomendaciones y futuras actividades académicas; sin despliegue operativo.

4.1. Revisión de decisiones de las unidades anteriores

La primera unidad definió el retraso como objetivo binario y justificó un árbol de decisión. Sin embargo, su primer modelo práctico cambió a regresión de minutos sobre 300 casos simulados. La segunda unidad inició un análisis descriptivo sobre 500 registros también simulados, cuestionado por el docente. La entrega visual de octubre utilizó el conjunto académico de 2023 y proporcionó las primeras cifras que pudieron contrastarse con esa fuente. El modelo consolidado de este informe recuperó el problema de clasificación binaria.

Estas diferencias se registran como parte de la evolución metodológica y no como resultados intercambiables.

4.2. Herramientas y trazabilidad

Los cálculos se desarrollaron en Python, con pandas y NumPy para tratamiento de datos, Matplotlib para gráficos y scikit-learn para clasificación. Se emplearon pipelines de imputación, codificación y estandarización ajustados solamente sobre datos de entrenamiento, conforme a las prácticas recomendadas para evitar fuga de información (scikit-learn developers, s. f.-a). Los scripts, el notebook y las métricas se incluyen en anexos.

5. Descripción de datos y transformaciones

5.1. Delimitación de la fuente

El conjunto principal utilizado es «nutrifresh_pedidos_dataset (2).csv», proporcionado durante la Unidad 2, con información académica fechada entre el 1 de octubre y el 14 de diciembre de 2023 (nutrifresh_pedidos_dataset.csv, s. f.). En la misma unidad también se suministró otro archivo de nombre semejante, pero con estructura y periodo diferentes (2026). No se combinaron ambos conjuntos porque no contienen la misma definición de variables ni registros equivalentes.

El caso original presenta cuatro categorías de información: pedidos, rutas, factores externos y costos operativos. Los cinco CSV breves facilitados durante la Unidad 1 ejemplifican su estructura; no se integraron forzosamente al archivo de 2023 al no disponer de una clave de vinculación pedido-ruta verificable.

Fuente/granularidad	Campos destacados	Aplicación en el proyecto
Pedidos (pedido individual)	ID, destino, producto, peso, fecha/hora, estado	Unidad principal para análisis y clasificación
Rutas (ruta o trayecto)	Distancia, tiempo, combustible, costo	Análisis descriptivo; sin unión forzada por falta de clave
Factores externos (ciudad y fecha)	Clima, precipitación, temperatura, tránsito	Análisis de asociación; sin garantía de predicción previa
Costos operativos (mes)	Mantenimiento, combustible, horas extra	Contexto de costos; no imputados a pedidos individuales

5.2. Perfil y calidad de los registros seleccionados

Control	Resultado	Tratamiento/alcance
Registros originales	228	Lectura del CSV
Duplicados exactos	8	Excluidos para EDA y clasificación
Pedidos después de depurar	220	168 a tiempo, 49 retrasados, 3 devueltos
Peso faltante	6 / 220 (2,7 %)	Análisis con n válido; en modelado imputación interna
Distancia faltante	6 / 220 (2,7 %)	Excluida del modelo previo al despacho
Tránsito faltante	6 / 220 (2,7 %)	EDA con 214 observaciones válidas
Peso con sufijo «kg»	4	Conversión a numérico
Tiempo entrega < recorrido	65	Requiere aclarar definición; no calcular retraso por diferencia

Los valores faltantes del EDA se conservaron y cada análisis utilizó sus casos completos. En el entrenamiento, la imputación del peso se aprendió exclusivamente con el conjunto de entrenamiento para no transferir información de prueba. Los atípicos se conservaron, pues el criterio estadístico por sí solo no demuestra errores de registro (Heckert & Filliben, 2003).

5.3. Variable objetivo y transformaciones

Para la clasificación se utilizó «retraso confirmado»: 1 cuando estado_entrega = Retrasado y 0 cuando estado_entrega = Entregado a tiempo. Tres devoluciones se analizaron descriptivamente como incidencias y se excluyeron del conjunto binario de entrenamiento. Así se evita asumir que «Devuelto» constituye automáticamente un retraso, sin fecha pactada que lo demuestre.

La muestra de clasificación quedó en 217 pedidos (49 retrasados y 168 a tiempo). Para el análisis gerencial se mantuvo un indicador diferente de incidencia, definido como retrasos más devoluciones: 52/220 = 23,6 %. Estas tasas no deben intercambiarse ni confundirse con el 20 % mencionado en el caso base.

Los indicadores costo por kilómetro y combustible por kilómetro se calcularon solo cuando los kilómetros eran válidos y mayores que cero. Para la clasificación se eligieron ciudad de destino, tipo de producto, peso, hora y día de semana. Las variables categóricas se codificaron con One-Hot Encoding y el peso se imputó con mediana; el preprocesamiento y el modelo se integraron en un pipeline.

5.4. Restricción de disponibilidad temporal

Se excluyeron como predictoras el estado final, tiempo de entrega, tiempo realmente recorrido, combustible consumido y costo definitivo, porque pueden conocerse después de la operación. Los kilómetros efectivamente recorridos y las condiciones observadas de clima o tránsito tampoco se incluyeron, por falta de evidencia de disponibilidad antes del despacho. Distancias planificadas y pronósticos previos podrían incorporarse únicamente si se registraran con fecha y hora adecuadas. Esta cautela evita una fuga de información que haría artificialmente optimista la evaluación.

6. Análisis exploratorio de datos

El EDA combinó distribución, medidas de tendencia central, detección de atípicos mediante rango intercuartílico, correlaciones de Pearson y Spearman y análisis segmentado. La exploración no equivale a causalidad ni demuestra capacidad predictiva futura (Heckert & Filliben, 2003).

6.1. Resultados descriptivos

Variable	n	Media	Mediana	Atípicos IQR
Tiempo de recorrido (min)	220	145,98	145,50	1
Distancia (km)	214	140,72	142,55	0
Combustible (gal)	220	26,57	26,30	0
Peso (kg)	214	302,05	276,80	14
Costo trayecto (COP)	220	$661.759	$682.000	0

Se encontró un recorrido de 286 minutos fuera del límite superior IQR y catorce pesos atípicos. Los valores extremos no deben asociarse sin comprobación a derrumbes, siniestros, sobrecarga o errores de registro, puesto que esos motivos no están documentados en la base.

6.2. Estado de entrega e incidencias

Estado	Pedidos	Porcentaje sobre 220
Entregado a tiempo	168	76,4 %
Retrasado	49	22,3 %
Devuelto	3	1,4 %
Total	220	100,0 %

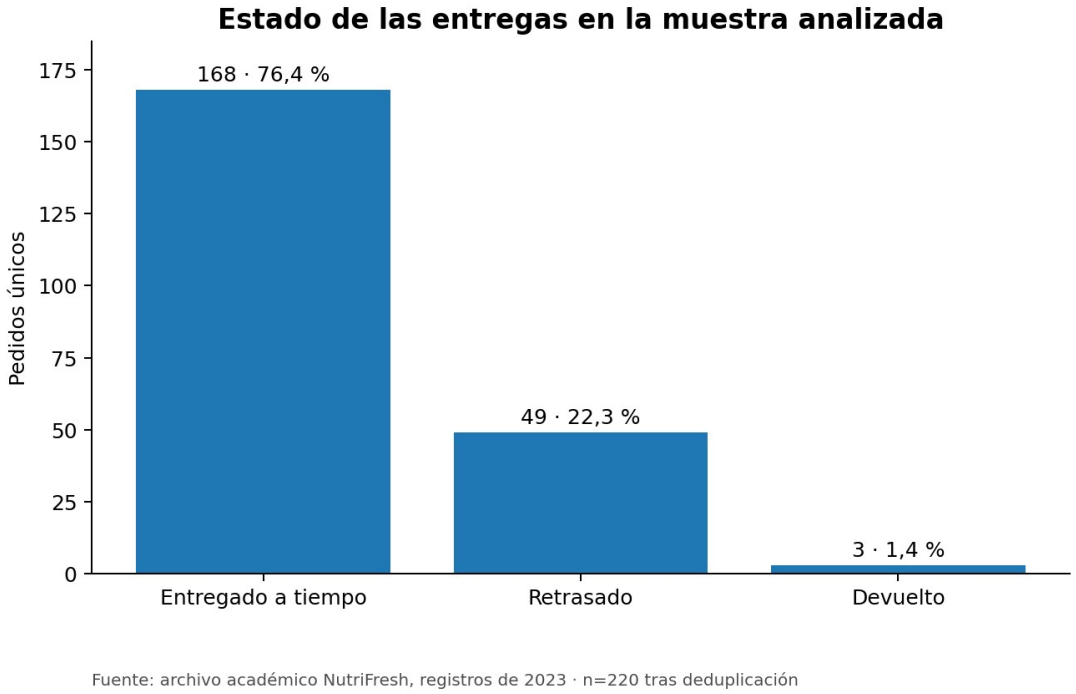


Figura 1. Estados de entrega del archivo académico de 2023 (n = 220). Elaboración propia.

6.3. Relaciones numéricas

Variables relacionadas	Pearson (r)	Lectura estadística
Distancia - tiempo recorrido	0,973	Positiva muy fuerte
Distancia - costo	0,965	Positiva muy fuerte
Distancia - combustible	0,951	Positiva muy fuerte
Tránsito - tiempo entrega	0,284	Positiva débil
Precipitación - tiempo entrega	0,272	Positiva débil
Temperatura - precipitación	-0,673	Negativa moderada/fuerte

La distancia presenta asociaciones fuertes con duración, consumo y costo; el costo absoluto más alto no prueba menor eficiencia, por lo que se recomiendan comparaciones por distancia, carga y tipo de vehículo. Por otro lado, las correlaciones de tránsito y precipitación con los tiempos registrados son débiles. Estos valores describen asociaciones en una muestra académica, no efectos causales.

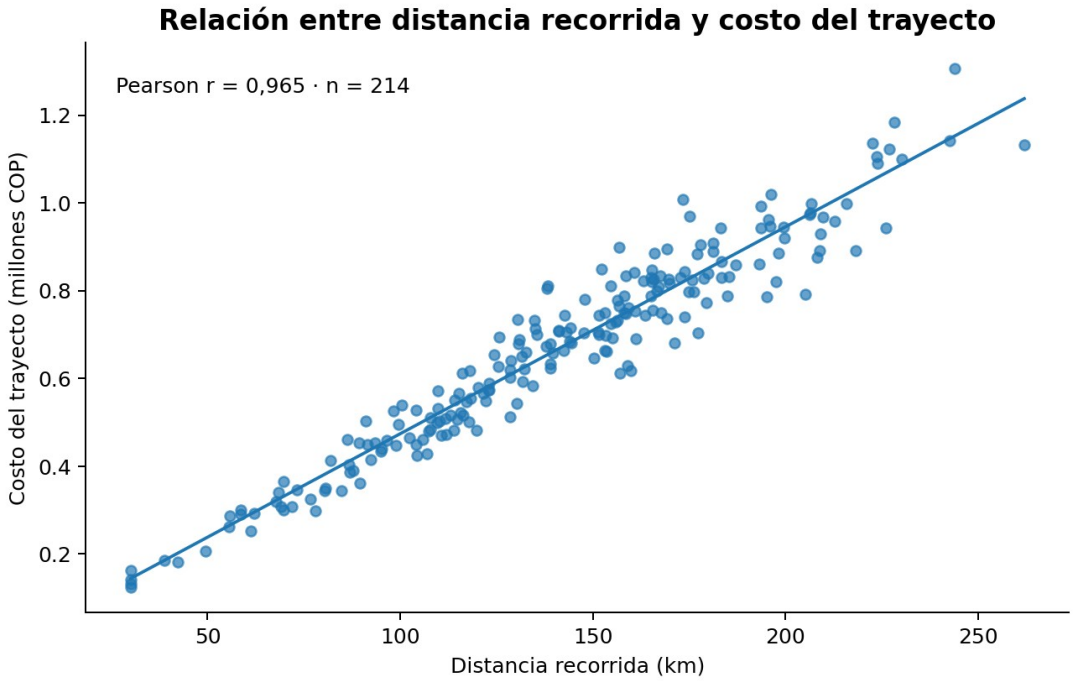
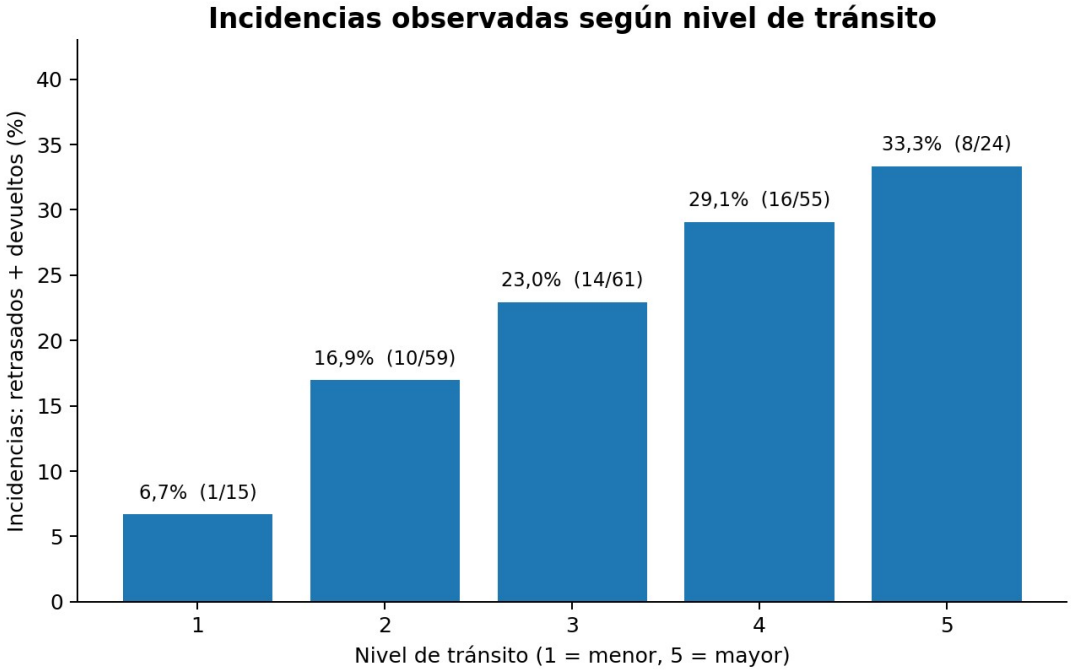


Figura 2. Distancia y costo de trayecto (n = 214; r = 0,965). Elaboración propia.

6.4. Comparaciones por tránsito, clima y destino

Nivel de tránsito	Incidencias / n válido	Tasa
1	1/15	6,7 %
2	10/59	16,9 %
3	14/61	23,0 %
4	16/55	29,1 %
5	8/24	33,3 %

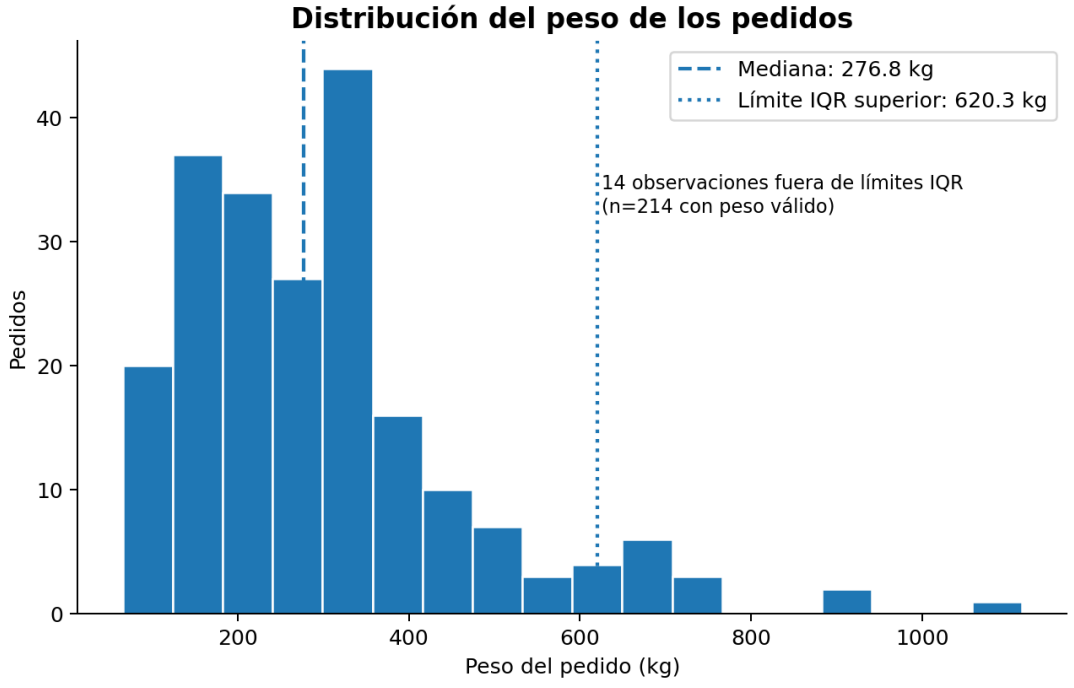


Fuente: archivo académico NutriFresh, registros de 2023 · n=214 con tránsito disponible · asociación, no causalidad

Figura 3. Incidencias según tránsito (n = 214). Elaboración propia.

En lluvia intensa se observaron 7 incidencias entre 21 pedidos (33,3 %) y en condiciones soleadas 20 entre 101 (19,8 %). Sincelejo registró 6/18 (33,3 %), Cartagena 14/50 (28,0 %) y Soledad 7/38 (18,4 %). Las tasas deben interpretarse con el tamaño de cada grupo y no demuestran que el clima, la ciudad o la congestión causen el incumplimiento.

6.5. Peso e indicadores relativos



Fuente: archivo académico NutriFresh, registros de 2023 · n=220 tras deduplicación

Figura 4. Distribución y atípicos IQR del peso de pedidos (n = 214). Elaboración propia.

El peso presenta media de 302,05 kg y mediana de 276,80 kg; catorce observaciones se encuentran fuera de los límites IQR. En 214 pedidos con kilómetros válidos se obtuvo una mediana global aproximada de $4.711 COP/km y 0,190 gal/km. Estos indicadores son exploratorios y deben ajustarse por factores operativos antes de utilizarlos para evaluar ahorro o eficiencia.

7. Modelado predictivo y métricas

7.1. Conjunto de clasificación y diseño

La etapa de clasificación utilizó 217 pedidos (49 positivos, 168 negativos). El preprocesamiento incluyó hora, día de semana, ciudad, producto y peso; evitó variables posteriores al despacho. El conjunto de entrenamiento contiene 174 registros desde el 1 de octubre al 29 de noviembre de 2023 (38 retrasos); la prueba cronológica reservó 43 pedidos del 1 al 14 de diciembre (11 retrasos). Se aplicaron tres ventanas cronológicas internas dentro del entrenamiento sin utilizar el conjunto final de prueba para seleccionar parámetros.

7.2. Técnicas y configuraciones

Se compararon: (a) clasificador mayoritario que siempre predice puntualidad; (b) árbol de decisión con criterio Gini, max_depth = 3, min_samples_leaf = 10, class_weight = balanced y random_state = 42; (c) regresión logística con regularización L2, C = 1,0, class_weight = balanced y max_iter = 2000. La salida de clase se generó con umbral 0,5. Los límites de profundidad y tamaño de hoja constituyen pre-poda para reducir el sobreajuste (Breiman et al., 1984; scikit-learn developers, s. f.-b).

En los tres modelos la codificación, estandarización e imputación se aplicaron mediante pipeline ajustado a entrenamiento. Se evitó usar el conjunto de prueba para entrenar o recalibrar el umbral (scikit-learn developers, s. f.-a).

7.3. Métricas y validación

Debido al desbalance, se examinaron accuracy, precision, recall, F1, ROC-AUC, Average Precision y matriz de confusión (scikit-learn developers, s. f.-c). Recall mide la fracción de retrasos identificados; precision, la proporción de alertas verdaderas; F1 resume ambas. Las curvas precision-recall son informativas cuando la clase de interés es minoritaria (Saito & Rehmsmeier, 2015).

Validación interna (promedio tres ventanas)	Árbol	Regresión logística
F1	0,306	0,259
ROC-AUC	0,581	0,507

Prueba (n = 43)	Línea base	Árbol	Logística
Accuracy	74,4 %	46,5 %	51,2 %
Precision	0,0 %	16,7 %	18,8 %
Recall	0,0 %	27,3 %	27,3 %
F1	0,000	0,207	0,222
ROC-AUC	0,500	0,402	0,375
Average Precision	0,256	0,241	0,215

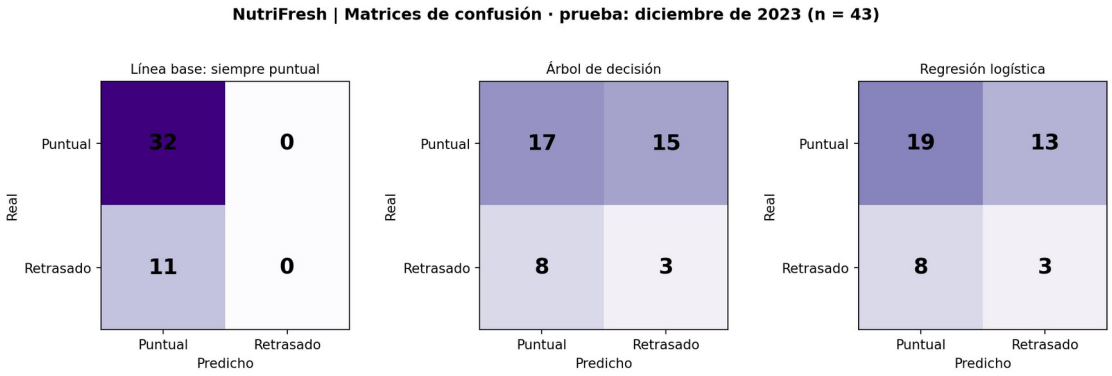


Figura 5. Matrices de confusión sobre prueba cronológica (n = 43). Elaboración propia.

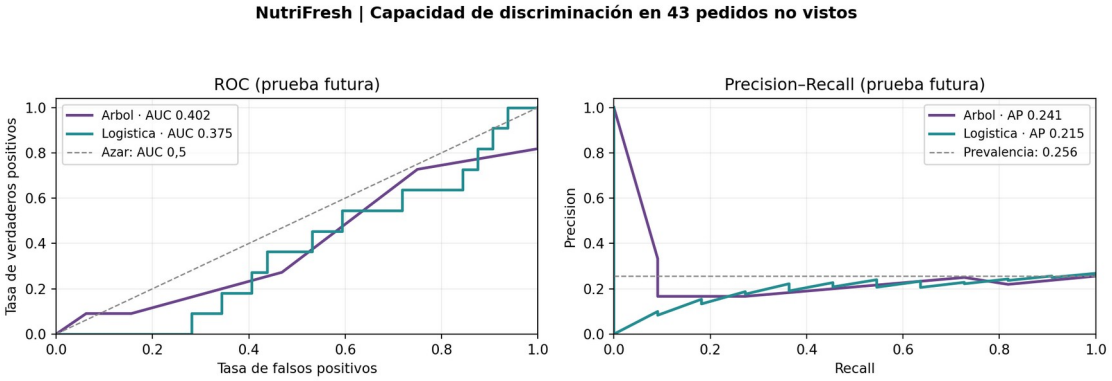


Figura 6. Comparación ROC y precisión-recall en el periodo de prueba. Elaboración propia.

8. Interpretación de resultados

8.1. ¿Qué identifica el árbol?

En diciembre había 32 pedidos puntuales y 11 retrasados. El árbol clasificó correctamente 17 puntuales y tres retrasados, pero emitió 15 falsas alertas y dejó ocho retrasos sin identificar (TN = 17; FP = 15; FN = 8; TP = 3). En consecuencia, su recall de 27,3 % y precision de 16,7 % son insuficientes para automatizar acciones preventivas. El ROC-AUC = 0,402 es inferior al valor 0,5 de referencia, lo que muestra deficiente discriminación en esta prueba.

8.2. Por qué accuracy no es suficiente

La línea base obtuvo 74,4 % de accuracy, pero predijo todos los pedidos como puntuales: no detectó ninguno de los once retrasos reales. Esta comparación muestra por qué las métricas enfocadas en la clase de interés son indispensables y por qué un porcentaje alto de acierto global no garantiza utilidad operativa.

8.3. Conexión con el EDA y el modelo simulado anterior

En el EDA, distancia y costo presentaron correlaciones fuertes, pero eso no implica que se pueda anticipar el atraso utilizando solo características previas del pedido. Además, varias variables asociadas con el tiempo registrado no cuentan con versiones de pronóstico disponibles antes del despacho. El ejercicio de regresión de la Unidad 1 presentó MAE aproximado de 3,15 min y R^2 cercano a 0,93, pero esos resultados provinieron de 300 registros simulados mediante una fórmula lineal y no son comparables con el clasificador evaluado sobre el CSV de 2023.

La diferencia entre desempeño interno y prueba posterior puede estar asociada con variabilidad de la muestra y escasez de señales previas al despacho. Estas explicaciones son hipótesis, no causas confirmadas. El modelo es reproducible como experiencia académica, pero no apto para producción.

9. Recomendaciones gerenciales

Prioridad 1. Gobernanza y captura de datos

Unificar identificadores de pedido y ruta; registrar hora pactada, salida, llegada, motivo de retraso y devolución, capacidades del vehículo y definiciones de tiempos.

Indicador sugerido: Porcentaje de pedidos con campos críticos completos y válidos.

Prioridad 2. Información disponible antes del despacho

Registrar distancia planificada y predicciones meteorológicas/de tráfico con marcas de tiempo para garantizar que las variables existían en el momento de emitir la alerta.

Indicador sugerido: Cobertura de variables previas al despacho por pedido.

Prioridad 3. Tablero segmentado de cumplimiento

Separar retrasos confirmados de devoluciones; presentar numeradores y denominadores por destino, producto, semana, clima y tránsito.

Indicador sugerido: Tasa de retraso confirmado, tasa de incidencias y recuento por segmento.

Prioridad 4. Control relativo de costos

Comparar COP/km y gal/km en recorridos homogéneos de distancia, carga y flota; auditar excepciones antes de atribuir ineficiencia.

Indicador sugerido: Mediana y dispersión de COP/km y gal/km en segmentos comparables.

Prioridad 5. Revalidación predictiva

Recopilar nuevos registros de mayor cobertura, repetir validación cronológica y probar pilotos controlados únicamente después de demostrar desempeño aceptable y estimar costos de errores.

Indicador sugerido: Recall, precision, AP, ROC-AUC y estabilidad mensual fuera de muestra.

Ninguna de las recomendaciones implica un ahorro ya obtenido. Deben considerarse acciones propuestas, sujetas a validación de factibilidad y seguimiento operativo.

10. Limitaciones del análisis

Carácter académico de los datos

El conjunto suministrado no permite afirmar que refleje toda la operación real; además existen dos versiones de pedidos con periodos y esquemas diferentes y no deben mezclarse.

Cobertura temporal y representatividad

Los 220 pedidos únicos abarcan octubre-diciembre de 2023, una muestra pequeña frente a los 500-700 pedidos diarios descritos en el contexto empresarial.

Clase minoritaria y tamaño de prueba

Solo 49 positivos en 217 registros de clasificación; la prueba contiene once retrasos, por lo que las estimaciones son inestables ante pequeñas variaciones.

Definiciones de tiempos y variables

Hay 65 casos con tiempo_entrega_min < tiempo_recorrido_min; faltan promesa pactada, GPS y clave pedido-ruta, que permitirían calcular retrasos en minutos o integrar las fuentes sin ambigüedad.

Señales previas al despacho

No se acreditó disponibilidad temporal de distancia prevista, pronósticos externos o costos finales; incluir mediciones posteriores generaría fuga de información.

Generalización y causalidad

La clasificación tuvo bajo recall y ROC-AUC; las asociaciones del EDA no prueban causas ni permiten estimar ahorros, mejoras en satisfacción o reducción efectiva de atrasos.

11. Conclusiones

La primera conclusión es metodológica: el proceso permitió delimitar de manera explícita una pregunta de clasificación binaria —retraso confirmado frente a entrega puntual— y distinguirla de otras incidencias, especialmente las devoluciones. La decisión de excluir tres devoluciones del entrenamiento evitó introducir una etiqueta ambigua; asimismo, la exclusión de variables que solo se conocen después de la entrega protegió la validez del escenario predictivo anterior al despacho.

La segunda conclusión se refiere al valor y al límite del EDA. En los 220 pedidos únicos examinados, el 22,3 % aparece retrasado y el 1,4 % fue devuelto. La distancia presentó asociaciones fuertes con costo (r = 0,965), tiempo de recorrido (r = 0,973) y combustible (r = 0,951). Estas relaciones ofrecen criterios para segmentar el seguimiento de costos y recorridos. En contraste, las tasas mayores observadas bajo tránsito alto o lluvia intensa deben entenderse como señales exploratorias que requieren validación ajustada; no se demostró causalidad.

La tercera conclusión es predictiva: el modelo de árbol de decisión y la regresión logística no lograron identificar los retrasos con precisión suficiente. En los 43 pedidos de prueba, ambos detectaron apenas tres de once retrasos (recall de 27,3 %). El árbol produjo quince alertas falsas y registró un ROC-AUC de 0,402. Estos resultados impiden recomendar el uso del piloto para automatizar decisiones de despacho. El valor del modelo radica en haber documentado de manera reproducible qué puede y qué no puede anticiparse con las variables actuales.

Finalmente, la trayectoria del proyecto permitió reconocer los riesgos de utilizar datos simulados como si representaran operaciones observadas, comparar coeficientes en distintas escalas o confundir ejecución de código con evidencia empresarial. La principal mejora futura consiste en registrar de forma consistente la promesa de entrega, la ruta planificada, las condiciones previstas y el resultado real; luego recopilar una muestra más amplia y evaluar de nuevo el desempeño sobre periodos independientes. Esta secuencia aporta un

aprendizaje académico verificable y una hoja de ruta razonable para madurar la analítica logística.

12. Referencias bibliográficas (APA 7)

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Wadsworth & Brooks/Cole.

Caso empresarial: Optimización logística en NutriFresh S.A.S. (s. f.). [Documento académico no publicado, suministrado en la asignatura; autoría no especificada].

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0: Step-by-step data mining guide*. CRISP-DM Consortium. https://www.kde.cs.uni-kassel.de/lehre/ws2012-13/kdd/files/CRISPWP-0800.pdf

Heckert, N. A., & Filliben, J. J. (2003). *NIST/SEMATECH e-handbook of statistical methods: Chapter 1, Exploratory Data Analysis*. National Institute of Standards and Technology. https://www.nist.gov/publications/nistsematech-e-handbook-statistical-methods-chapter-1-exploratory-data-analysis

IBM. (s. f.). *CRISP-DM help overview*. IBM Documentation. https://www.ibm.com/docs/en/spss-modeler/18.6.0?topic=dm-crisp-help-overview

nutrifresh_pedidos_dataset.csv (s. f.). [Conjunto de datos académico no publicado, registros de 2023; autoría no especificada].

Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE, 10*(3), e0118432. https://doi.org/10.1371/journal.pone.0118432

scikit-learn developers. (s. f.-a). *Common pitfalls and recommended practices*. scikit-learn Documentation. https://scikit-learn.org/stable/common_pitfalls.html

scikit-learn developers. (s. f.-b). *DecisionTreeClassifier*. scikit-learn Documentation. https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html

scikit-learn developers. (s. f.-c). *Metrics and scoring: Quantifying the quality of predictions*. scikit-learn Documentation. https://scikit-learn.org/stable/modules/model_evaluation.html

13. Anexos y evidencias

Los anexos permiten rastrear los resultados numéricos a los datos y al código. Las rutas de la tabla corresponden al paquete ZIP adjunto, que contiene los cuadernos y copias de los CSV con los nombres que cada notebook requiere. Los enlaces de Colab remiten a trabajos históricos del grupo; el clasificador consolidado es un archivo .ipynb independiente que no se presenta como publicado en línea.

Anexo	Archivo / ubicación	Contenido
A. Fuente académica	Datos/pedidos_2023.csv	Datos de 2023 empleados para EDA y modelado; no combinar con el CSV de 2026.
B. EDA verificado	EDA/analisis.ipynb	Notebook de visualizaciones de Unidad 2.
C. Modelo actualizado	Modelo/clasificador.ipynb	Pipeline, modelos, partición temporal y métricas.
D. Métricas inspeccionables	Modelo/metricas.json	Matrices de confusión y resultados de validación y prueba.
E. Figuras 1-4	Figuras/01 a 04 (.png)	EDA: estados, costo-distancia, tránsito y peso.
F. Figuras 5-6	Figuras/05 a 06 (.png)	Confusión; curvas ROC y precisión-recall.
G. Registros previos	Enlaces históricos	Historial de decisiones, errores y observaciones docentes.

Notebook histórico (Unidad 1, modelo base): Abrir cuaderno en Google Colab

Notebook histórico (Unidad 2, primer EDA): Abrir cuaderno en Google Colab

Notebook histórico (Unidad 2, visualizaciones): Abrir cuaderno en Google Colab

Control de reproducción: archivo principal SHA-256 = 76b8bf8dc9e538bc88cbf02ca22dd19aefc192b0cdf91d9ae8bf3c0783c771f2. La salida del modelo fue verificada con el archivo JSON adjunto; no se emplearon los resultados del modelo simulado de la Unidad 1 como evidencia de capacidad predictiva actual.

Material pendiente: la presentación, el video de sustentación y la bitácora digital serán incorporados por el grupo cuando se elaboren. No se adjunta un dashboard interactivo independiente; las visualizaciones comprobadas se incluyen en este informe y sus anexos. Estos elementos no se presentan como evidencias ya realizadas.