

Metodología de modelación del desempeño de pavimentos urbanos utilizando inteligencia artificial

Urban pavement performance modeling methodology using artificial intelligence

INFORMACIÓN DEL ARTÍCULO

Historial del
artículo:

Recibido
14-11-2023
Aceptado
03-12-2024
Publicado
30-04-2025

Palabras Clave:
Machine Learning
Deep Learning
Desempeño
Modelos
Pavimentos
urbanos.

Article history:

Received
14-11-2023
Accepted
03-12-2024
Available
30-04-2025

Keywords:
Machine Learning
Deep Learning
Performance
Models
Urban pavements

Salvador Pérez Jara¹, Aleli Osorio Lirio², Héctor Allende Cid³

Universidad Técnica Federico Santa María, Santiago, Chile, salvador.perez@sansano.usm.cl

² Universidad Técnica Federico Santa María, Santiago, Chile, aleli.osorio@usm.cl

³ Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile, hector.allende@pucv.cl

Resumen

Los modelos de desempeño de pavimentos son una herramienta sumamente útil para las agencias encargadas de la gestión vial, permitiendo estimar el comportamiento y designar la medida correctiva óptima para mantenimiento de las vías. Investigaciones recientes han presentado resultados favorables al utilizar técnicas de Machine y Deep Learning en la modelación del comportamiento de pavimentos. En esta investigación se plantea una metodología de desarrollo de modelos de desempeño de primera fase para pavimentos urbanos, gestionados a nivel de red, para la predicción de la condición a corto plazo utilizando técnicas de Machine y Deep Learning en su confección. Específicamente se utiliza la regresión de bosques aleatorios (RFR), regresión de vectores de soporte (SVR), regresor de refuerzo de gradiente (GBR), redes neuronales artificiales (ANN) y redes neuronales recurrentes (ANN). Con estos algoritmos se confeccionan modelos para predecir el Índice de Condición de Pavimentos Urbanos de Chile (ICPU) utilizando una base de datos sintética de elaboración propia para las simulaciones. En cuanto a los resultados obtenidos, las iteraciones ofrecen resultados favorables para la predicción a corto plazo, obteniendo bajos errores promedio de predicción (0.4% para el algoritmo GBR) y resultados de RMSE próximos a cero (0.063 para el algoritmo GBR), lo cual resulta de interés para cada una de las alternativas de algoritmos mencionadas anteriormente, pues las sitúan como herramientas recomendables para realizar predicciones sobre el desempeño de los pavimentos urbanos.

Abstract

Pavement performance models are an extremely useful tool for road management agencies, allowing them to estimate pavement performance and designate the optimal corrective measure for road maintenance. Recent research has presented favorable results when using Machine and Deep Learning techniques in pavement performance modeling. This research proposes a methodology for developing first-phase performance models for urban pavements, managed at the network level, for short-term condition prediction using Machine and Deep Learning techniques in their preparation. Specifically, random forest regression (RFR), support vector regression (SVR), gradient boosting regression (GBR), artificial neural networks (ANN) and recurrent neural networks (ANN) are used. With these algorithms, models are built to predict the Chilean Urban Pavement Condition Index (ICPU) using a synthetic database of our own elaboration for the simulations. Regarding the results obtained, the iterations offer favorable results for short-term prediction, obtaining low average prediction errors (0.4% for the GBR algorithm) and RMSE results close to zero (0.063 for the GBR algorithm), which is of interest for each of the algorithm alternatives mentioned above, since they place them as recommendable tools to make predictions on the performance of urban pavements.

* Corresponding author at: Salvador Pérez Jara, Universidad Técnica Federico Santa María, Santiago, Chile
E-mail address; salvador.perez@sansano.usm.cl

RIOCI

journal homepage: <https://rioc.ufro.cl/index.php/rioc/index>

Vol. 13, no. 01, pp. 1–9, Abril 2025

1. Introducción.

Las redes de pavimento juegan un rol fundamental dentro de la sociedad, pues más allá de cumplir su función primordial, permitir el traslado de personas y carga valiosa vía terrestre, son un elemento sumamente necesario para el desarrollo integrado de una nación (Solminihac et al., 2018). Por este motivo, resulta de gran relevancia el preservar sus características funcionales y estructurales en el tiempo, pues, de esta forma, se pueden sostener adecuadamente las comunicaciones, actividades económicas, sociales y todas aquellas que permiten el continuo desarrollo de un país.

Por tales motivos, las agencias de gestión de infraestructura vial deben velar por instaurar políticas de mantenimiento que permitan preservar las redes de pavimento en un nivel de servicio adecuado y, para ello, requieren de herramientas que permitan inferir cual es el mejor momento para intervenir una determinada vía buscando la optimización tanto del rendimiento material (durabilidad del pavimento) como del recurso económico destinado a estas actividades.

Para tomar estas determinaciones, las Agencias de gestión se valen de múltiples herramientas que ayudan a tomar la decisión de cuándo y cómo intervenir, entre ellas los modelos de desempeño, los cuales son expresiones que predicen la evolución del estado del pavimento en el tiempo, en base al conocimiento de sus condiciones al momento de puesta en servicio y en cualquier momento de su vida útil (Solminihac et al., 2018).

Respecto al desarrollo de modelos de desempeño, esta es una actividad que sufre ajustes conforme aparecen nuevos enfoques de modelación. Es así como en primera instancia se intentó explicar el comportamiento de los pavimentos mediante un enfoque empírico con modelos regresivos (lineales multivariados) y deterministas (Watanatada et al., 1987) que asocian características similares como estructura, tránsito o clima. Otros autores decidieron utilizar un enfoque probabilístico para la confección de modelos, apareciendo como opción viable los modelos con formulaciones bayesianas y también con "Cadenas de Markov" (Solminihac et al., 2014; Osorio, 2015). Últimamente, la tendencia es la de incursionar con técnicas de Inteligencia Artificial (IA), esencialmente mediante la interacción de algoritmos de Machine y Deep Learning tales como "Vector Support Regression (VSR)", "Random Forest Regression (RFR)", "Artificial Neuronal Networks (ANN)" y "Recurrent Neuronal Networks (RNN)" (Mazari y Rodríguez, 2016; Marcelino et al., 2019; Hossain, et al., 2019; Wang, et al.,

2020; Zhou, et al., 2021; Montecinos, 2022).

La utilización de estas técnicas ha demostrado la obtención de modelos con resultados de predicción dignos de atención, puesto que la versatilidad de los algoritmos, a diferencia de los modelos determinísticos y probabilísticos, permite la modelación no solo de múltiples escenarios, sino que permite también incorporar atributos explicativos que no necesariamente obedecen a comportamientos numéricos (variables categóricas como zonas climáticas o jerarquía de vías) o, prescindir, por ejemplo, de matrices de transición de probabilidad cuya obtención se dificulta si es que no se poseen datos representativos de la red vial o no se cuenta con un panel de expertos que ayude en su confección. Además, dado el alto potencial de procesamiento de datos se pueden generar predicciones con un notable nivel de precisión a corto y mediano plazo.

En esta investigación, se plantea una metodología para la utilización de diferentes algoritmos de Machine y Deep Learning en la construcción de modelos predictivos del desempeño de pavimentos, empleando para ello una base de datos sintética que simula el comportamiento en el tiempo de pavimentos urbanos de alta jerarquía de la Región Metropolitana de Santiago de Chile representado en la variación del Índice de Condición de Pavimentos Urbanos (ICPU) (Osorio, 2015). Las pruebas realizadas consisten en predecir el ICPU del año "t" utilizando información de diversos atributos de los años previos "t-1", "t-2" y "t-3", tales como Transito Medio Diario Anual (TMDA), ICPU y características estructurales como espesor de carpeta y Módulo Resiliente de la subrasante, en lo que equivale a la utilización de un enfoque dinámico de modelación según se plantea en investigaciones previas (Marcelino, et al., 2019; Yang, et al., 2021). Dichos análisis permitieron dilucidar que los algoritmos RFR, VSR, ANN, RNN y GBR poseen buena capacidad predictiva y de aprendizaje para la resolución de este tipo de problemas.

Por último, cabe destacar que este trabajo se enmarca en el desarrollo del Proyecto FONDECYT 11201150 "Evaluación y Modelación del desempeño de Pavimentos Urbanos usando Machine Learning", el cual persigue los objetivos de automatizar la evaluación de pavimentos mediante el empleo de algoritmos de Deep Learning enfocándolos a la detección y cuantificación de deterioros usando imágenes de bajo costo y, el correspondiente a lo desarrollado en este escrito, que busca evaluar si la confección de modelos de desempeño utilizando algoritmos de Inteligencia Artificial entrega resultados de mayor precisión que los modelos existentes para el análisis de pavimentos urbanos a nivel de red.

El éxito que pueda obtenerse en las dos aristas de investigación

comentadas puede generar un gran impacto en la gestión de pavimentos urbanos, puesto que permitiría incorporar aspectos valiosos de la gestión de pavimentos a agencias que no necesariamente cuentan con los recursos para adquirir tecnología de punta que les permita monitorear y evaluar la condición de la red de forma continua, además de entregar una herramienta de predicción que ayude a tomar mejores decisiones de priorización y mantenimiento de la red de pavimentos urbanos.

2. Metodología

Tras la revisión de múltiples trabajos previos, se ideó una metodología tanto para la confección de una base de datos adecuada como para la implementación de diferentes algoritmos que sean aplicables a la resolución de problemas de regresión.

2.1 Identificación de Algoritmos de IA:

Proviene netamente de realizar un proceso de revisión de investigaciones previas, pues la utilización de herramientas de Inteligencia Artificial (IA) se ha potenciado con el correr de los años. En el caso de la confección de modelos de desempeño, se encuentran estudios desde comienzos de los años 2000 en donde ya se aprecian las primeras investigaciones con uso de ANN (Lin, et al., 2003). Tomando en consideración lo anterior, los algoritmos seleccionados para esta investigación son RFR, GBR, VSR, ANN y RNN, los cuales se caracterizan por ser algoritmos especializados en problemas de predicción mediante regresión. En este aspecto, se destaca la presencia del algoritmo GBR o regresión de refuerzo de gradiente, el cual no se ha visto utilizado con regularidad en relación a los otros algoritmos mencionados.

2.2 Selección de variables o Atributos Explicativos:

El deterioro de los pavimentos no puede ser explicado por una sola variable, pues son muchos los factores que inciden en la progresión de este, algunos con mayor preponderancia que otros, pero que, combinados, pueden repercutir en el deterioro rápido de un pavimento. Por este motivo es que los distintos factores que inciden en el deterioro como agrietamientos, depresiones, baches o la rugosidad suelen acoplarse en un índice combinado que refleje su estado de deterioro actual (Osorio, 2015). Ahora bien, el deterioro de los pavimentos se produce tanto por la solicitación de tránsito para la cual presta servicio como por las inclemencias climáticas que puedan afectarlo y el cómo responda a las solicitudes tiene que ver con aspectos relevantes de diseño tales como características del suelo de subrasante, espesores de las distintas capas constituyentes del paquete estructural y la calidad de los materiales utilizados en su confección.

Con lo anterior en contexto, tienden a utilizarse como variables explicativas para la progresión de métricas de desempeño aspectos como el tránsito (caracterizado como TMDA o Ejes Equivalentes), parámetros estructurales como el Número estructural del Asfalto u espesor de la carpeta, la edad del pavimento, aspectos climáticos como precipitación, temperatura, heladas y nieve y, por último, series de tiempo de alguna métrica de condición, tránsito o clima. Esto se ve refrendado en una investigación de revisión del estado del arte, en la cual se reafirma el uso de estos atributos como los principales utilizados para realizar predicciones sobre el desempeño de los pavimentos (Yang, et al., 2021).

En el caso puntual de este trabajo, se descartan variables climáticas por estar situado en una zona geográfica puntual como la Región Metropolitana de Santiago, en donde el contexto supone un clima mediterráneo (Osorio, 2015). Por otra parte, se considera además como atributo de entrada las jerarquías de las vías, las cuales, dependiendo de su función, se pueden clasificar como “Expresas”, “Troncales” o “Colectoras”. Junto con lo anterior, se utiliza también series de tiempo de flujos vehiculares expresados en Tránsito Medio Diario Anual (TMDA) y de ICU previos, ambos con rezagos “t-1”, “t-2” y “t-3”, módulo resiliente y espesor de carpeta asfáltica.

Lo mencionado en el párrafo anterior no quiere decir que no se consideren los efectos de precipitaciones o inundaciones, que son las principales inclemencias temporales que afectan a los pavimentos urbanos en la región metropolitana, puesto que la forma de modelar el desempeño en esta metodología se basa en la construcción de diversas familias de pavimentos. En ese sentido, juega un rol fundamental el saber, por ejemplo, si las avenidas analizadas cuentan con sistemas de colectores de aguas lluvias y si estos se encuentran o no en buen estado. De esta forma, si la avenida cuenta con sistema de colectores o solo sistema de cunetas formarán dos familias de pavimentos diferentes, de la misma manera que ocurre con la materialidad, en donde se pueden definir múltiples familias dependiendo del material de construcción del pavimento.

2.3 Confección de Base de Datos:

Ya establecidos los atributos explicativos de la variable dependiente o de salida, se debe configurar la base de datos que se utilizará para realizar las pruebas con los diferentes algoritmos planteados. En ese sentido, podemos subdividir las variables en atributos estáticos, es decir, que se mantienen constantes durante la vida útil del pavimento tales como la clasificación funcional, el espesor de carpeta asfáltica y el módulo resiliente, y dinámicos, que varían de un lag temporal a otro, como el TMDA o las series de ICU anteriores al horizonte de predicción. En la Tabla 1 se exponen mayores antecedentes de las variables explicativas.

Tabla 1. Muestra estadísticas descriptivas de la base de datos utilizadas.

Variable	Tipo	Rezago	Valor Min	Valor Max	Promedio
M_R [MPa]	Numérica	-	31	162	95
Espesor		-			
Asfalto [mm]	Numérica	-	40	200	-
Clasificación Funcional	Categórica	-	-	-	-
TMDA [veh/día]	Numérica	3	432	12852	6362
ICPU	Numérica	3	1	10	-

En cuanto a utilizar variables dinámicas, esto es por expresa recomendación de Marcelino, et al., 2019, puesto que el deterioro de los pavimentos es un proceso incremental que guarda estrecha relación con el historial previo. Por tales motivos, utilizar secuencias temporales ayuda a tener predicciones del desempeño más precisas.

Con respecto a la confección de una base de datos sintética, esto tiene su fundamento en que la implementación de esta metodología busca instaurarse en la gestión de pavimentos urbanos a nivel de red en Chile, en donde los últimos trabajos que se han realizado en cuanto a predecir el desempeño de los pavimentos datan del año 2015 con la creación del Sistema de Gestión de Pavimentos Urbanos (SGPUC) en conjunto con su métrica combinada de condición denominada Índice de Condición de Pavimentos Urbanos (ICPU) (Osorio, et al., 2014). Esto trae consigo la desventaja de que no existe una base de datos lo suficientemente robusta como para alimentar adecuadamente un algoritmo de IA. Por este motivo, es que por medio de modelos de desempeño creados para esta métrica con métodos Markovianos (Osorio, et al., 2018), se decidió elaborar un set de datos de 835 muestras a raíz de obtener un análisis regresivo (ecuación 1) al modelo de desempeño de asfalto para vías de alta jerarquía presentado en la Figura 1. De esta forma se conformaron sets de series de tiempo que incluyen el $ICPU_0$ (variable a predecir) y sus correspondientes rezagos anuales $ICPU_{t-3}$, $ICPU_{t-2}$ y $ICPU_{t-1}$.

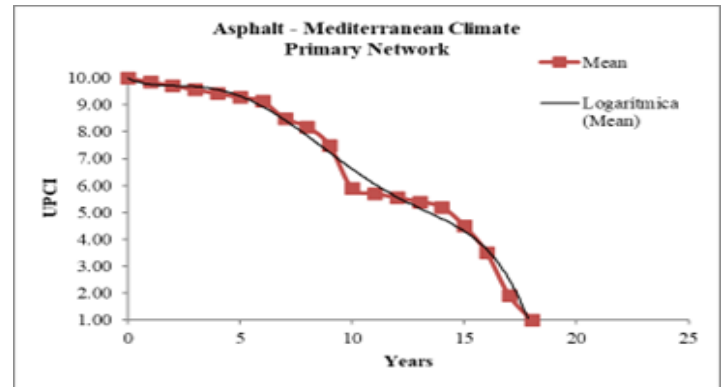


Figura 1. Modelo de desempeño para pavimentos asfálticos de alta jerarquía en clima mediterráneo. (Fuente: Osorio, 2015).

$$ICPU = 10 - 8 \cdot 10^{-5} \cdot t^5 + 0.0032 \cdot t^4 - 0.0438 \cdot t^3 + 0.1996 \cdot t^2 - 0.3923 \cdot t \quad (1)$$

En cuanto a los datos de TMDA, se estudió la variación anual en la matrícula de vehículos motorizados examinando la data del Instituto Nacional de Estadísticas de Chile (INE) estableciendo un crecimiento anual promedio de 4%. De esta forma, se construyeron unidades muestrales con TMDA oscilante entre 200 y 14000 [veh/día] estableciendo rangos definidos según la clasificación funcional de cada vía y, al igual que con los datos de ICPU, se construyeron series de tiempo de variación de tránsito que acompañe los rezagos temporales “t-3”, “t-2” y “t-1”. Cabe destacar que el valor de TMDA es un único valor, el cual engloba las distintas clases de vehículos (livianos y pesados).

En cuanto a los rangos de variación de TMDA utilizados para cada clasificación jerárquica de vías, estos se explicitan en la Tabla 2.

Tabla 2. Muestra rangos de variación de TMDA según Clasificación Funcional de las vías.

Clasificación Funcional	Rango Inferior	Rango Superior
Colectora	200	2,000
Troncal	2,000	9,000
Expresa	9,000	14,000

En lo que concierne a los datos estáticos de espesor de carpeta asfáltica y módulo resiliente de la subrasante, estos se hacen variar aleatoriamente entre 40 y 200 [mm] y 30 y 165 [MPa] respectivamente, valores que resultan típicos en el diseño y manufactura de pavimentos urbanos en Chile.

2.4 Procesamiento de la base de datos:

Esta etapa de la metodología consiste en trabajar el set de datos de forma tal que sea procesable para los algoritmos de Machine y Deep Learning que se utilizan. En ese sentido, se trabajó el set de datos en el lenguaje de programación Python junto con la librería Scikit Learn.

La primera parte del procesamiento consistió en trabajar los atributos categóricos de forma tal que se encuentren en un lenguaje reproducible para los algoritmos. Por dicha razón, a los atributos de Clasificación Funcional se les aplicó “One Hot Encoder”, de esta forma, se crean variables numéricas ficticias que toman el valor 1 cuando la entrada es coincidente y 0 cuando no (Montecinos, 2022).

Por otra parte, las variables numéricas no son coincidentes en orden de magnitud, lo cual puede perjudicar el aprendizaje. Por dicha razón, se deben normalizar los datos para facilitar el aprendizaje de los algoritmos y así evitar sesgo ante la variación de valores máximos y mínimos (Fausset, 1994; Montecinos, 2022). Para cumplir con lo anterior, se normaliza de acuerdo con la regla de máximos y mínimos que se muestra en la ecuación 2.

$$x_{norm} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (2)$$

Por último, se subdivide el set de datos en conjunto de entrenamiento y validación. Para ello, se utiliza la herramienta “Train test Split” de Scikit Learn dividiendo los conjuntos en proporción 80/20 respectivamente.

2.5 Configuración de Algoritmos:

Con respecto a la configuración de los algoritmos, esta se basa en definir arquitecturas y seleccionar hiperparámetros, las cuales también se trabajan por medio de la librería Scikit Learn. Esto con miras a obtener una configuración que sea adecuada para la resolución del problema en cuestión. En ese sentido, se detallará como primer punto la configuración de los algoritmos de ML utilizados.

En el caso del algoritmo RFR es importante definir el “Número de Estimadores”, el cual hace referencia a la cantidad de árboles de decisión que incorporará el modelo, la “Semilla de Aleatoriedad” y la “Mínima Cantidad de Muestras para una División”, el cual hace referencia a como se subdivide el conjunto de datos y cuanta información recibe cada árbol de decisión. Para el caso del algoritmo GBR, se configuran los mismos hiperparámetros, a excepción de la “Mínima Cantidad de Muestras para una División” (Scikit Learn, 2011).

En lo referente al algoritmo VSR, la configuración difiere, pues en su caso se debe definir el kernel del modelo, el cual permite indicar si se está trabajando datos lineales o no lineales, y también los parámetros “C” y “Epsilon”. En el caso de “C”, esta es una constante mayor que cero que determina el equilibrio entre la regularidad de la función y la cuantía hasta la cual se tolera desviaciones mayores a los vectores de soporte, en tanto que “Epsilon” controla el error cometido por la función de regresión al aproximar a las bandas de soporte. El resumen de hiperparámetros se observa en la Tabla 3. (Scikit Learn, 2011).

Tabla 3. Muestra Hiperparámetros utilizados en la configuración de algoritmos de ML

Hiperparámetro	RFR	GBR	VSR
N estimators	100	100	-
Random State	2016	2016	-
Min Samples leaf	8	-	-
Kernel	-	-	“rbf”
C	-	-	3.0
Epsilon	-	-	0.15

En el caso de los algoritmos de DL, estos tratan de redes neuronales, por lo que se debe conformar su arquitectura y calibrar hiperparámetros. En el caso del modelo con ANN, se configura utilizando Scikit Learn y su estructura se compone de una capa de entrada, 2 capas densas de 7 neuronas cada una y la capa de salida que entrega el resultado de la regresión, lo que resume una estructura clásica de red “Feed Forward”. En el caso del modelo RNN, también se configura utilizando la librería Scikit Learn y su estructura consiste en la capa de entrada, una capa LSTM de 12 neuronas y la capa de salida que entrega el resultado de la regresión. Con respecto al resto de hiperparámetros que utilizan los modelos neuronales confeccionados, estos pueden observarse en la Tabla 4. (Scikit Learn, 2011).

Tabla 4. Muestra Hiperparámetros utilizados en la configuración de algoritmos de DL.

Características Hiperparámetros	e	Modelo ANN	Modelo RNN
Función de Activación		Tanh, Relu	Relu
Learning Rate		0.01	0.01
Función de Pérdida		MSE	MSE
Tamaño de Lote		128	128
Épocas de Entrenamiento		200	500

Habiendo configurado los algoritmos, procede la etapa de entrenamiento con el conjunto de datos confeccionado para dicha tarea y, posteriormente, la etapa de prueba, en la cual se estudia el rendimiento de los modelos evaluando métricas de desempeño como la raíz del error cuadrático medio (RMSE), el error promedio o la “Accuracy” con el conjunto de datos elaborado específicamente para las pruebas de validación. De esta forma se puede estudiar la capacidad predictiva a corto plazo de los distintos modelos configurados y analizar el potencial de cada herramienta.

2.6 Construcción de Modelo de Desempeño

Posterior a la configuración, entrenamiento y pruebas de los algoritmos tanto de ML como de DL, resta la labor de utilizar los modelos para realizar predicciones con el fin de poder evaluar su rendimiento ya no solo a corto, sino que también a mediano y largo plazo. Para ello, se le debe entregar un valor semilla al algoritmo con aquellos atributos que el modelo necesita de entrada. En este caso puntual, dichos atributos corresponden a las series de tiempo de ICPU previos y de TMDA junto con el espesor de carpeta asfáltica y módulo resiliente de la subrasante.

La construcción del modelo o curva de desempeño se basa en una recursión en donde el modelo predice lo que ocurre al año siguiente en función de la predicción del año anterior. De esta forma, la predicción correspondiente al año “t” pasa a ser el dato “t-1” para la predicción del año siguiente y así sucesivamente. Por lo demás, para la variación del tránsito, tal como se explicó anteriormente, este se hace crecer en un 4% conforme avanzan los ciclos de predicción que se deseen.

Teniendo en consideración lo anterior, se elabora un modelo de desempeño el cual se contrasta con una curva tipo generada con datos reales de calles y avenidas situadas en la comuna de Maipú, Santiago, Chile. Los cuales fueron obtenidos realizando auscultación manual en terreno y agrupados de tal forma que simularan el deterioro de una misma calle bajo la lógica del

Tabla 5. Muestra las métricas de error obtenidas como resultados de las predicciones.

Métrica	RFR	GBR	VSR	ANN	RNN
RMSE	0.078	0.063	0.362	0.227	0.266
Error Prom %	0.60	0.40	13.13	5 .17	7.07
Accuracy	0.999	0.999	0.947	-	-

“método de ventanas” (Videla, 1992) el cual consiste en definir sectores homogéneos en cuanto a sus variables más representativas (estructura, tránsito, geometría y clima). En ese sentido, los datos adquiridos correspondieron a calles y avenidas pertenecientes a la red vial primaria, con un ancho de pista de 3.5 [m], mismas condiciones de drenaje (cunetas) y tránsito homogéneo.

Por último, mencionar que de esta prueba solo se tienen resultados para el algoritmo RFR, estando en desarrollo los resultados para el resto de los algoritmos entrenados al momento de la elaboración de este manuscrito.

3. Análisis de resultados

3.1 Pruebas para la predicción a corto plazo:

Con respecto a los resultados de las predicciones, estos se observan en la Tabla 5 y Figura 2, en donde se analizan las métricas de raíz del error cuadrático medio RMSE, Error Promedio y Accuracy de cada uno de los modelos en el caso de los algoritmos de ML.

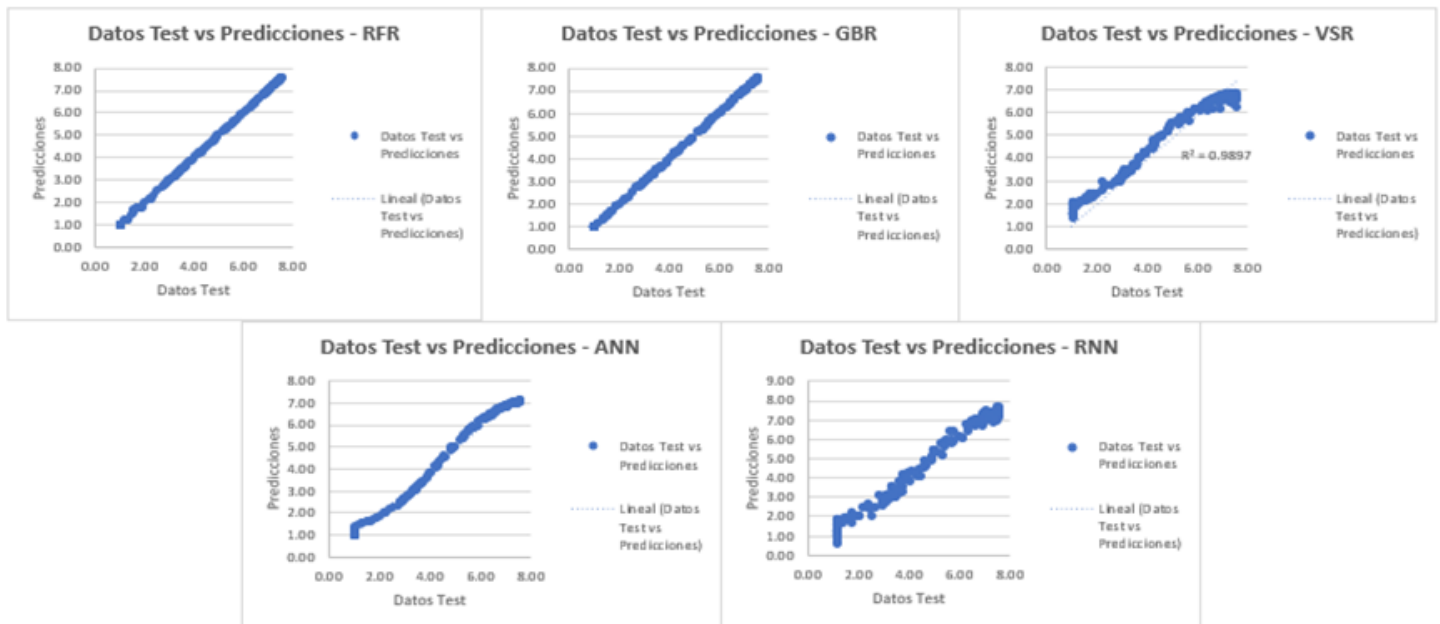


Figura 2. Muestra resultados gráficos de la comparativa entre las predicciones y los datos de testing.

De los resultados obtenidos, es evidente que los algoritmos con mejor desempeño en la tarea son los algoritmos de ML, específicamente RFR y GBR. Con respecto al desempeño del primero mencionado, esto resulta coincidente con los resultados que se obtuvieron en otras investigaciones como las realizadas por Marcelino, et al., 2019 y Montecinos, 2022, en las cuales se da cuenta que el rendimiento del algoritmo RFR está por sobre otras alternativas como ANN o VSR.

En cuanto al resultado del algoritmo VSR, este en ningún caso es malo, sin embargo, se observa que sus contenedores se ajustan mejor al problema planteado, esto queda demostrado al percatarse que de los 3 algoritmos de ML es el que tiene mayor nivel de error.

Con respecto al algoritmo GBR, este es el que presenta mejor resultado, muy similar al algoritmo RFR, pero ligeramente superior. Sobre la naturaleza del mismo, GBR también trabaja con árboles de decisión. Sin embargo, a diferencia de RFR, la generación de árboles de decisión se da en forma de secuencia de forma tal que cada árbol corrige los errores de su predecesor, en tanto que RFR sencillamente promedia el resultado de cada árbol como el resultado de la predicción.

Con respecto a los modelos de DL, trabajados con ANN y RNN, el análisis es similar al realizado para VSR. No se obtienen malos resultados, sin embargo, tanto RFR como GBR se ajustaron de mejor forma al contexto del problema. De todas formas, las predicciones obtenidas se ajustan y poseen un error promedio menor al 10%.

3.2 Resultados preliminares de predicción a mediano y largo plazo

De acuerdo con lo expuesto en el punto 2.6, se elabora una

curva de desempeño utilizando el modelo RFR configurado y se contrasta con una “curva tipo” elaborada con datos reales de unidades muestrales de calles y avenidas ubicadas en la comuna de Maipú, Santiago, Chile, las cuales fueron agrupadas siguiendo el “método de ventanas” expuesto por Videla, 1992 y a las cuales se les estimó su edad utilizando la expresión (1) expuesta en el punto 2.3. Cabe mencionar que se realizaron predicciones simulando un deterioro progresivo de 16 ciclos, es decir, simulando un período de análisis de 16 años.

El resultado obtenido se expone en la figura 3.

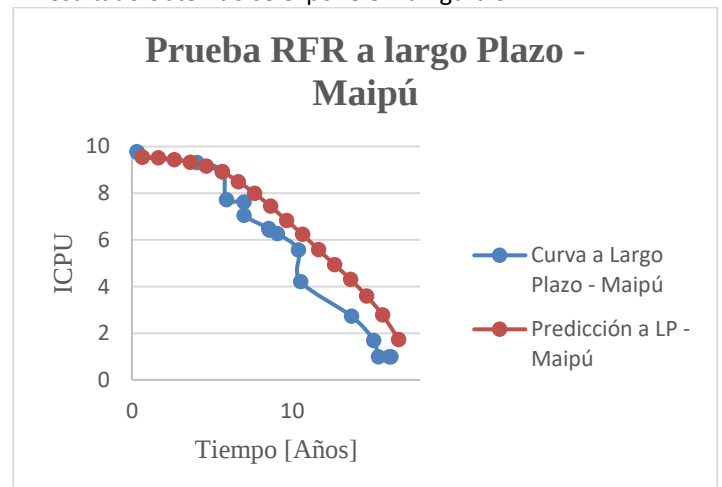


Figura 3. Muestra contraste entre la curva de desempeño generada por algoritmo RFR y la curva tipo generada en base a los datos reales obtenidos de las auscultaciones realizadas en Maipú.

De la figura se desprende que el modelo RFR posee una muy buena capacidad predictiva para horizontes bajos a medios de predicción, pues hasta un horizonte de 5 años se observa que

ambas curvas siguen un comportamiento similar. Sin embargo, a medida que el horizonte de predicción se hace mayor, las curvas comienzan a alejarse, observando que la curva generada por el modelo RFR sobrestima la condición del pavimento.

4. Conclusiones

En este estudio se utilizó una base de datos sintética para poner a prueba una metodología de desarrollo de modelos de desempeño para la gestión de pavimentos urbanos a nivel de red en la Región Metropolitana de Santiago de Chile utilizando algoritmos de Inteligencia Artificial. Puntualizando, se probó la metodología con 5 tipos de algoritmos distintos, 3 de ML y 2 de DL, obteniéndose buenos resultados de predicción a corto plazo con cada uno de ellos, pero con resultados destacables para los algoritmos GBR y RFR según los resultados expuestos en la Tabla 5.

Respecto a la metodología de modelación, se refrenda el uso de atributos explicativos como las solicitudes de tránsito, historial de métrica de condición previo y la contribución de atributos como parámetros de diseño y características del suelo de fundación. Por lo demás, se reafirma también que el utilizar un enfoque dinámico contribuye a obtener buenos resultados, puesto que se nutre a los modelos con información histórica que guía de mejor forma a los modelos hacia la solución. Por otra parte, se destaca la versatilidad de este tipo de algoritmos al poder incluir atributos categóricos. Esta característica no hace nada más que ampliar los horizontes de modelación, puesto que se pueden modelar múltiples escenarios en un mismo algoritmo, por ejemplo, podrían incluirse sin ningún problema aspectos como la materialidad del pavimento o discretizar por zonas climáticas, pudiendo incorporar múltiples familias de pavimentos en un solo modelo.

En cuanto a la construcción de curvas de desempeño, por el momento solo se posee el desarrollo de un caso muy acotado correspondiente a la predicción del modelo RFR para un caso puntual de estudio correspondiente a calles y avenidas de alta jerarquía, cuyos datos fueron obtenidos mediante campañas de auscultación manual en la comuna de Maipú, Santiago, Chile. Aún así, puede mencionarse que el resultado muestra una alta capacidad predictiva del modelo hasta un horizonte de 5 años utilizando sus propios datos de predicción. Sin embargo, dicho rendimiento decae cuando el horizonte de predicción es superior. Aun así, la curva desarrollada por el modelo muestra un comportamiento lógico, siendo decreciente y obteniendo datos razonables en lo referente a la escala de desarrollo del ICPU.

Profundizando en lo expuesto en el párrafo anterior, cabe

destacar que la baja en el rendimiento del modelo puede explicarse por múltiples factores. El primero de ellos tiene que ver con la base de datos de entrenamiento, la cual, además de ser sintética, proviene de un análisis regresivo a un modelo probabilístico elaborado por Osorio, 2015, el cual, de por sí ya posee incertidumbre y que también sufre una merma en su rendimiento para predicciones a largo plazo. Cabe destacar que esto ocurre independiente de la metodología utilizada para la construcción de los modelos, pues tanto para modelos regresivos, probabilísticos y basados en herramientas de Inteligencia Artificial se observa un desempeño menor en la predicción de horizontes amplios. Además de esto, se encuentra presente también la aleatoriedad propia que implica la variación de los diferentes atributos explicativos que conforman la base de datos, por lo que la distorsión en los resultados obtenidos puede explicarse con estos antecedentes.

Sobre el procesamiento de información y tiempo de entrenamiento de los modelos, este fue realizado con equipo de hardware estándar, procesador Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz 1.80 GHz junto con 8 GB de memoria RAM sin tarjeta gráfica, tardando no más de 15 minutos como tiempo máximo de entrenamiento, el cual fue empleado para la red neuronal recurrente (RNN). Para el resto de los algoritmos el tiempo de entrenamiento no sobrepasa el minuto. Referente al tiempo que tardan en dar una predicción, este no pasa de unos cuantos segundos.

Por último, se debe mencionar que queda abierta la puerta para poner a prueba el rendimiento del resto de los algoritmos en horizontes de predicción más amplios, tal como se realizó con el modelo RFR y, en definitiva, obtener conclusiones más amplias con la construcción de modelos de desempeño de distintas familias de pavimentos y contrastarlos también con modelos previos de naturaleza diferente ya confeccionados (probabilísticos o determinísticos). También, queda pendiente realizar un análisis de sensibilidad a los modelos confeccionados a miras de tener conocimiento de la real importancia de los atributos explicativos utilizados y como influyen en las predicciones realizadas por los modelos, aristas de investigación que se abordarán en el futuro cercano, pues competen una tarea fundamental en el desarrollo del Proyecto FONDECYT 11201150 "Evaluación y Modelación del desempeño de Pavimentos Urbanos usando Machine Learning".

5. Agradecimientos

Los autores agradecen el apoyo financiero de ANID CHILE, Proyecto FONDECYT 11201150 y al Programa de Incentivos a la Investigación Científica (PIIC) de la Dirección de Postgrado de la UTFSM.

6. Referencias

- Fausset, L. (1994). What is a Neural Net?. Fundamentals of neural networks: architectures, algorithms, and applications D. Fowley, B. Baker, and A. Dworkin, eds., Prentice Hall, Englewood Cliffs, NJ, 3-4.
- Hossain, M. I., Gopiseti, L. S. P., & Miah, M. S. (2019). International roughness index prediction of flexible pavements using neural networks. Journal of transportation engineering, part B: pavements, 145(1), 04018058.
- Lin, J. D., Yau, J. T., & Hsiao, L. H. (2003, January). Correlation analysis between international roughness index (IRI) and pavement distress by neural network. In 82nd Annual Meeting of the Transportation Research Board (Vol. 12, No. 16, pp. 1-21).
- Marcelino, P., & Antunes, L. E., & Castilho Gomes, M.(2019). Machine learning approach for pavement performance prediction. International Journal of Pavement Engineering, 1-14.
- Mazari, M., & Rodriguez, D. D. (2016). Prediction of pavement roughness using a hybrid gene expression programming-neural network technique. Journal of Traffic and Transportation Engineering (English Edition), 3(5), 448-455.
- Montecinos, A. (2021). METODOLOGÍA Y APLICACION DE ALGORITMOS DE MACHINE LEARNING EN EL DESARROLLO DE MODELOS DE DESEMPEÑO DE PAVIMENTOS AEROPORTUARIOS EN CHILE.