

GRADO EN CIENCIA DE DATOS



VNIVERSITAT
ID VALÈNCIA

TRABAJO FIN DE GRADO

PREDICCIÓN DE LA RESPUESTA INDIVIDUAL A LA
PREGUNTA ENPS EN DISTINTAS EMPRESAS Y
ANÁLISIS DE LOS DISTINTOS MODELOS
APLICADOS

AUTOR:

JORGE TERRASA VIDAL

TUTOR:

VALERO LAPARRA PÉREZ-MUELAS



VNIVERSITAT
DE VALÈNCIA



Escola Tècnica Superior
d'Enginyeria **ETSE-UV**

TRABAJO FIN DE GRADO

PREDICCIÓN DE LA RESPUESTA INDIVIDUAL A LA PREGUNTA ENPS EN DISTINTAS EMPRESAS Y ANÁLISIS DE LOS DISTINTOS MODELOS APLICADOS

AUTOR: JORGE TERRASA VIDAL

TUTOR: VALERO LAPARRA PÉREZ-MUELAS

Declaración de autoría:

Yo, Jorge Terrasa Vidal, declaro la autoría del Trabajo Fin de Grado titulado “Predicción de la respuesta individual a la pregunta eNPS en distintas empresas y análisis de los distintos modelos aplicados” y que el citado trabajo no infringe las leyes en vigor sobre propiedad intelectual. El material no original que figura en este trabajo ha sido atribuido a sus legítimos autores.

Valencia, 18 de julio de 2025

Fdo: Jorge Terrasa Vidal

Resumen:

El Employee Net Promoter Score (eNPS) es una métrica sintética que captura, mediante una sola pregunta, el compromiso y la satisfacción de los empleados y cuyo valor se interpreta como barómetro del clima laboral. Este Trabajo Fin de Grado aborda la predicción individual (escala 1–10) de esa respuesta aprovechando datos reales proporcionados por la plataforma HappyForce, que recopila de forma continua el Happiness Index y distintos *scores* sobre la experiencia del empleado.

El estudio diseña un flujo completo de *People Analytics*: depuración y alineación temporal de registros, construcción de conjuntos de entrenamiento, validación y prueba sin fuga de datos, y generación de dos enfoques de modelado, observaciones independientes y secuencias de cuatro etapas, para contrastar el valor de la información histórica. Sobre el enfoque independiente se implementa una arquitectura MLP y, sobre el enfoque secuencial, una arquitectura LSTM y otra MLP, aplicando búsqueda automática de hiperparámetros a todos los modelos base para optimizar su rendimiento.

Los resultados confirman que incorporar dinámica temporal mejora todas las métricas de regresión; el LSTM optimizado se erige como el mejor modelo ($MSE = 1,41$; $MAE = 0,82$; $R^2 = 0,65$), superando a ambos enfoques del MLP con un coste computacional asumible. Este nivel de precisión permite anticipar variaciones significativas en la satisfacción y facilita intervenciones proactivas de RR.HH., situando la predicción del valor eNPS individual como una herramienta estratégica para impulsar el bienestar y la retención del talento.

En conjunto, el trabajo demuestra la viabilidad de integrar técnicas avanzadas de *Machine Learning* en entornos de *feedback* continuo, aporta un análisis comparativo de modelos secuenciales e independientes y sienta las bases para desplegar soluciones prescriptivas en plataformas de experiencia del empleado.

Abstract:

The Employee Net Promoter Score (eNPS) is a synthetic metric that captures employee engagement and satisfaction through a single question, and its value is interpreted as a barometer of the work environment. This Final Degree Project addresses the individual prediction (scale 1–10) of this response by leveraging real data provided by the HappyForce platform, which continuously collects the Happiness Index and various employee experience scores.

The study designs a complete People Analytics workflow: record cleaning and temporal alignment, training, validation and testing set construction without data leakage, and generation of two modeling approaches, independent observations and four-stage sequences, to contrast the value of historical information. An MLP architecture is implemented on the independent approach, and an LSTM and an MLP architecture are implemented on the sequential approach, applying automatic hyperparameter search to all base models to optimize their performance.

The results confirm that incorporating temporal dynamics improves all regression metrics; the optimized LSTM emerges as the best model ($\text{MSE} = 1.41$; $\text{MAE} = 0.82$; $R^2 = 0.65$), outperforming both MLP approaches at an affordable computational cost. This level of accuracy makes it possible to anticipate significant variations in satisfaction and facilitates proactive HR interventions, positioning individual eNPS prediction as a strategic tool for boosting well-being and talent retention.

Overall, the work demonstrates the feasibility of integrating advanced Machine Learning techniques into continuous feedback environments, provides a comparative analysis of sequential and standalone models, and lays the groundwork for deploying prescriptive solutions in employee experience platforms.

Agradecimientos:

En primer lugar, quiero agradecer a mi familia, que me ha apoyado durante todos estos años.

En segundo lugar, quiero agradecer a los amigos que he hecho durante la carrera, que nos hemos ayudado y apoyado en todo momento para poder llegar hasta aquí.

En tercer lugar, quiero agradecer a mi tutor, Valero, la paciencia que ha tenido conmigo, además de la ayuda y guía que me ha proporcionado a lo largo de este trabajo.

Por último, quiero agradecer a HappyForce la oportunidad de trabajar con ellos y las facilidades que me han dado para llevar a cabo este TFG.

Índice general

1. Introducción	15
1.1. Introducción	15
1.2. Motivación	16
1.3. Objetivos	17
2. Estado del arte, materiales y métodos	19
2.1. Estado del arte	19
2.1.1. Introducción al eNPS	19
2.1.2. Aplicaciones del eNPS en la literatura	19
2.1.3. Modelos predictivos aplicados al eNPS y métricas de clima laboral	20
2.1.4. Plataformas de <i>feedback</i> continuo y datos en tiempo real	20
2.1.5. Avances recientes en modelado del eNPS	20
2.1.6. Tendencias emergentes en <i>People Analytics</i> (2025)	21
2.2. Datos del estudio	21
2.2.1. ¿Qué datos se utilizan?	21
2.3. Metodología	24
2.3.1. Preparación de los datos	24
2.3.2. Unificación temporal y preparación contextual	24
2.3.3. Aspectos comunes del entrenamiento y evaluación	25
2.3.4. Modelado sin secuencias	25
2.3.5. Modelado con secuencias	26
3. Resultados y comparación de modelos	31
3.1. Resultados	31
3.2. Enfoque sin secuencias	31
3.2.1. MLP básico	32
3.2.2. MLP hypertuning	34
3.3. Enfoque LSTM con secuencias	35
3.3.1. LSTM básico	36

3.3.2. LSTM hypertuning	38
3.4. Enfoque MLP con secuencias	39
3.4.1. MLP básico	40
3.4.2. MLP hypertuning	42
3.5. Comparación global de modelos	44
4. Conclusiones y trabajo futuro	47
4.1. Conclusiones	47
4.2. Trabajo futuro	47
Bibliografía	49

Capítulo 1

Introducción

1.1. Introducción

En un entorno empresarial cada vez más competitivo y dinámico, la satisfacción y lealtad de los empleados se han convertido en factores cruciales para el éxito organizacional. La capacidad de atraer y retener talento no solo afecta la productividad y el rendimiento, sino que también influye en la cultura organizacional y la reputación de la empresa.

En este contexto, el Employee Net Promoter Score (eNPS) emerge como una herramienta valiosa para medir y gestionar el compromiso y la satisfacción de los empleados. Basado en el concepto del Net Promoter Score (NPS), utilizado ampliamente para evaluar la lealtad de los clientes, el eNPS permite a las organizaciones obtener una visión clara y cuantificable de la percepción de sus empleados.

El eNPS se calcula a partir de una pregunta simple pero poderosa: En una escala del 1 al 10, ¿qué probabilidades hay de que recomiendes tu empresa como un buen lugar para trabajar?

1	2	3	4	5	6	7	8	9	10
Detractores						Pasivos		Promotores	

Figura 1.1: Ejemplo de representación de las posibles respuestas a la pregunta sobre el eNPS[1]

Las respuestas se clasifican en promotores, pasivos y detractores, y el resultado proporciona una métrica que puede oscilar entre -100 y +100.

La fórmula del eNPS es la siguiente:

$$\text{eNPS} = \% \text{Promotores} - \% \text{Detractores} \quad (1.1)$$

Un eNPS positivo indica una mayor proporción de empleados satisfechos y comprometidos, mientras que un eNPS negativo sugiere la necesidad de intervenciones significativas para mejorar el ambiente laboral.

Dado el impacto del eNPS en la estrategia empresarial, surge la necesidad de desarrollar modelos predictivos que permitan anticipar el valor del eNPS en diferentes empresas.

La capacidad de predecir la respuesta eNPS (1–10) no solo facilita la identificación temprana de problemas potenciales, sino que también permite a las organizaciones implementar acciones proactivas para mejorar la satisfacción y el compromiso de los empleados.

La colaboración con una empresa para la realización de este TFG ha sido un aspecto clave debido a la necesidad de hacer uso de datos reales y actualizados. La empresa que ha proporcionado los datos utilizados a lo largo del proyecto es HappyForce[2], una empresa dedicada a proporcionar soluciones de recursos humanos enfocadas en la mejora del compromiso y la satisfacción de los empleados.

Este trabajo de fin de grado se centra en la predicción del valor que dará cada empleado a la pregunta eNPS en la escala 1–10 en distintas empresas y el análisis del modelo predictivo utilizado. Se explorarán diversas técnicas de aprendizaje automático y análisis de datos para construir un modelo robusto y preciso. Además, se examinarán los factores que más influyen en la predicción del valor de la pregunta eNPS, proporcionando una comprensión profunda de los elementos que determinan la satisfacción y lealtad de los empleados.

1.2. Motivación

En la actualidad, la gestión efectiva del talento humano se ha convertido en un componente esencial para el éxito de cualquier organización. La capacidad de atraer, retener y motivar a los empleados no solo impacta directamente en la productividad y eficiencia, sino que también juega un papel crucial en la construcción de una cultura organizacional sólida y en la mejora de la reputación de la empresa. En este contexto, la satisfacción y el compromiso de los empleados se han vuelto indicadores clave del bienestar organizacional y de la salud a largo plazo de la empresa.

El Employee Net Promoter Score (eNPS) es una métrica poderosa que proporciona una evaluación cuantitativa del nivel de satisfacción y lealtad de los empleados. Un eNPS positivo sugiere un alto grado de compromiso y satisfacción, lo cual es indicativo de un entorno de trabajo saludable y atractivo. Por el contrario, un eNPS negativo puede señalar problemas subyacentes que requieren atención urgente para evitar una alta rotación de personal y un descenso en la moral.

El uso de datos proporcionados por HappyForce ofrece una oportunidad única para explorar y analizar de manera profunda las dinámicas internas que afectan el eNPS. La plataforma HappyForce recoge *feedback* continuo y en tiempo real de los empleados, lo que permite una visión precisa y actualizada del clima laboral. Al aplicar técnicas avanzadas de aprendizaje automático y análisis de datos a estos valiosos conjuntos de datos, es posible desarrollar modelos predictivos que anticipen el voto eNPS, facilitando así una gestión proactiva de los recursos humanos.

La motivación detrás de este trabajo de fin de grado radica en la posibilidad de proporcionar a HappyForce una herramienta más sofisticada y precisa para entender y mejorar la satisfacción de los empleados. Un modelo predictivo robusto no solo permitirá identificar problemas potenciales antes de que se conviertan en crisis, sino que también ofrecerá conclusiones accionables para fomentar un ambiente laboral positivo y motivador. En última instancia, este enfoque contribuirá a crear entornos de trabajo más saludables y productivos, beneficiando tanto a los empleados como a las organizaciones.

1.3. Objetivos

El propósito general de este Trabajo Fin de Grado es predecir el valor que cada empleado otorgará, en la escala 1–10, a la pregunta eNPS y analizar los modelos desarrollados.

Objetivo general

- Desarrollar y validar un sistema de *Machine Learning* capaz de anticipar la respuesta individual (1–10) a la pregunta eNPS con la mayor precisión posible, utilizando datos reales de la plataforma HappyForce.

Objetivos específicos

1. Preparación de los datos

- Limpiar y transformar en una tabla única los registros de *scores*, Happiness Index (HI) y respuestas eNPS.
- Generar variables *dummies* de calendario (mes y trimestre) para el enfoque de observaciones independientes.
- Construir el enfoque de secuencias temporales de longitud 4, garantizando un intervalo máximo de 180 días entre pasos, creando la variable `delta_days` (diferencia de días entre cada respuesta de la variable objetivo).
- Normalizar todas las variables numéricas mediante escalado Min–Max.
- Dividir los datos por `employeeId` en conjuntos de entrenamiento (70 %), validación (15 %) y prueba (15 %) para evitar fuga de datos.

2. Construcción de modelos predictivos

- Implementar un modelo MLP para observaciones independientes.
- Implementar dos modelos sobre secuencias: MLP y LSTM.
- Aplicar *Keras Tuner* (búsqueda aleatoria, 50 combinaciones) para optimizar hiperparámetros en las tres arquitecturas base.

3. Evaluación y visualización de resultados

- Calcular MSE, MAE y R^2 para cada modelo.
- Generar curvas de aprendizaje y gráficos de dispersión (valor real vs. predicho) para cada modelo.

4. Comparación y selección de enfoques

- Comparar las métricas de los seis modelos (básico y optimizado, con y sin secuencias).
- Identificar el modelo con mejor equilibrio entre precisión y complejidad para un posible despliegue en HappyForce.

Capítulo 2

Estado del arte, materiales y métodos

El presente capítulo reúne el marco conceptual y los cimientos técnicos del trabajo. Se abre con un estado del arte que posiciona la investigación dentro de la literatura existente (Sección 2.1), continúa con la descripción de las fuentes de datos y los criterios de preprocesado inicial (Sección 2.2) y concluye con la metodología empleada para entrenar y evaluar los distintos modelos predictivos (Sección 2.3).

2.1. Estado del arte

2.1.1. Introducción al eNPS

El Employee Net Promoter Score (eNPS) es una adaptación del Net Promoter Score (NPS), una métrica ampliamente utilizada en marketing para medir la lealtad del cliente. En el ámbito de los recursos humanos, el eNPS se emplea para evaluar el compromiso y la satisfacción de los empleados, mediante una única pregunta: “¿Recomendarías tu empresa como un buen lugar para trabajar?” [3].

Las respuestas se clasifican en promotores (9–10), pasivos (7–8) y detractores (1–6), y el eNPS se calcula como la diferencia porcentual entre promotores y detractores. Esta simplicidad y facilidad de interpretación han favorecido su uso en entornos corporativos para monitorizar el clima laboral de forma continua y cuantificable.

2.1.2. Aplicaciones del eNPS en la literatura

Aunque el eNPS es una herramienta popular en el entorno empresarial, su tratamiento académico todavía es limitado. Sin embargo, estudios recientes han comenzado a explorar su validez como indicador de clima laboral y su relación con variables como la rotación, el desempeño o la cultura organizacional [4].

Autores como Frederick F. Reichheld y Rob Markey han defendido que el eNPS puede ser un predictor indirecto del éxito organizacional, ya que un alto compromiso interno tiende a traducirse en una mejor experiencia del cliente [5].

2.1.3. Modelos predictivos aplicados al eNPS y métricas de clima laboral

En el ámbito del *Machine Learning*, existe una creciente tendencia a aplicar modelos predictivos a métricas relacionadas con los recursos humanos, un campo conocido como *People Analytics*. Recientemente se ha abordado el uso de algoritmos como regresión lineal, árboles de decisión, redes neuronales y XGBoost para predecir la satisfacción laboral, la rotación de empleados o el ausentismo.

Aunque el uso específico del eNPS como variable dependiente aún no es habitual en la literatura científica, existen datos similares, como encuestas de clima, puntuaciones de desempeño o índices de felicidad, pueden ser modelados eficazmente con técnicas supervisadas.

Cabe destacar que, junto a los trabajos que modelan directamente el eNPS agregado, también existen estudios centrados en predecir la **respuesta individual** a la pregunta eNPS (1–10). Kumar y Sharma (2024), por ejemplo, demuestran que las redes neuronales profundas capturan con mayor precisión los factores que influyen en dicha valoración a nivel micro [6]. El presente TFG adopta este mismo enfoque individualizado, alineando el marco teórico con la metodología desarrollada.

2.1.4. Plataformas de *feedback* continuo y datos en tiempo real

Sistemas como HappyForce, Officevibe o Glint han ganado relevancia por su capacidad de recopilar datos frecuentes, anónimos y estandarizados de los empleados. Estos sistemas permiten generar *datasets* ricos y actualizados que pueden utilizarse para desarrollar modelos predictivos útiles en la toma de decisiones estratégicas de RR. HH. [7].

El presente trabajo se inserta dentro de esta línea emergente, proponiendo un modelo basado en redes neuronales que, a partir de datos como el Índice de Felicidad (HI) y los *scores* de preguntas internas, permita predecir con antelación el voto futuro al eNPS de un empleado, facilitando una gestión más proactiva del compromiso laboral.

2.1.5. Avances recientes en modelado del eNPS

La literatura más reciente ha comenzado a explorar el uso de arquitecturas *deep learning* para comprender y predecir el Net Promoter Score. Un ejemplo destacado es el trabajo de Xiaowei Shi and Qiang Wei (2024), que introduce el modelo RNSC (Review, Note and Semantic Consistency) para estimar el NPS a partir de texto libre y notas internas, demostrando que los enfoques jerárquicos superan a los métodos clásicos de regresión lineal y *gradient boosting* [8].

En el ámbito específico del eNPS, Kumar y Sharma (2024) aplican redes neuronales profundas para evaluar la satisfacción interna y concluyen que la combinación de *embeddings* semánticos y variables estructuradas mejora la precisión frente a modelos tradicionales [6]. Estos estudios refuerzan la conveniencia de emplear modelos no lineales cuando se dispone de históricos suficientemente largos, como ocurre con los registros de HappyForce.

2.1.6. Tendencias emergentes en *People Analytics* (2025)

Los informes de tendencias publicados en 2025 coinciden en señalar una evolución de *People Analytics* desde la mera elaboración de informes descriptivos descriptivo hacia modelos prescriptivos basados en IA, con impacto directo en la toma de decisiones estratégicas de RR. HH. [9, 10]. Entre las líneas emergentes destacan:

- **Medición de energía y bienestar.** Los sistemas actuales no sólo monitorizan horas trabajadas, sino también patrones de fatiga y picos de productividad, favoreciendo intervenciones personalizadas[11].
- **El bienestar como Indicador Clave de Desempeño (KPI) del negocio.** Herramientas de *pulse surveys* y microaprendizaje integradas en las suites de RR. HH. permiten vincular el eNPS con métricas de rotación y desempeño [12].
- **Ética y privacidad.** Kirsty y otros analistas subrayan la necesidad de marcos éticos que garanticen la trazabilidad y la equidad algorítmica, especialmente al modelar indicadores sensibles como el eNPS [7].

Estas tendencias apoyan la pertinencia de este TFG, al situar la predicción del eNPS individual dentro de un panorama en el que las organizaciones demandan modelos cada vez más proactivos, interpretables y respetuosos con la privacidad.

2.2. Datos del estudio

Para llevar a cabo este trabajo de fin de grado, se han utilizado datos proporcionados por la empresa HappyForce, como se ha comentado anteriormente. Estos datos, obtenidos a través de la aplicación de HappyForce, incluyen información detallada sobre la experiencia laboral de los empleados, su nivel de satisfacción y su disposición a recomendar la empresa como un buen lugar para trabajar, entre otros aspectos que desarrollaremos a continuación.

A lo largo de este apartado, se describirán las características de los datos utilizados, incluyendo las variables recopiladas, el proceso de recolección de datos y su relevancia para la predicción del valor a la pregunta sobre el Employee Net Promoter Score (eNPS).

Este análisis detallado del conjunto de datos constituye la base sobre la que se construyen, entrenan y validan los modelos predictivos desarrollados en este trabajo. Garantiza que todas las variables relevantes queden correctamente preprocesadas y alineadas, lo que resulta esencial para evaluar con rigor el rendimiento de los distintos enfoques propuestos.

2.2.1. ¿Qué datos se utilizan?

HappyForce recoge distintos datos diariamente de los empleados que se han registrado en la aplicación móvil que proporciona a las distintas empresas con las que opera. A partir de esta base se pueden diferenciar tres tipos de datos disponibles: información básica del empleado, el Índice de Felicidad (Happiness Index, HI) y los “scores”.

Información básica

Aunque inicialmente se recopilaban datos como fecha de nacimiento, género y fecha de contratación, estos no se han utilizado en el análisis debido a que no aportan un valor significativo al objetivo del proyecto, además de garantizar la privacidad de los empleados. Sin embargo, otros datos básicos, como el id de la empresa y el id del empleado, resultaron fundamentales para vincular registros específicos a cada organización de manera anónima y preservar la integridad del análisis.

El id de la empresa es un código alfanumérico único para cada compañía, lo que permite identificar los datos a nivel organizacional sin comprometer la privacidad de los empleados. Adicionalmente, el id del empleado, único dentro de cada empresa pero no globalmente, se utilizó únicamente como un identificador técnico para relacionar registros dentro del conjunto de datos.

companyId	employeeId	gender	hireDate	birthDate	churnDate	churnReason
5faee32997df24000430d586	lzk	FEMALE	2024-08-08	2024-01-15		
630f6cc2db165720d5303baf	n5r0	FEMALE	2024-07-05	1998-03-23		
5b3ddd08ce75e500040f8c76	2B3		2024-02-26	1985-06-21		
6447e9ee7d2ae02f98fcd87e	XQj	FEMALE	2024-02-22	2024-02-22		
6447e9ee7d2ae02f98fcd87e	xPj	FEMALE	2024-02-22	2024-02-22		
6447e9ee7d2ae02f98fcd87e	5QV	MALE	2024-02-22	2024-02-22		
64f6e6563638254e9b8ee192	g5r	MALE	2024-02-21	1994-08-14		
5f0db44c900e790004bcd2aa	aO1	MALE	2024-02-21	1994-05-19		
5f0db44c900e790004bcd2aa	Wra	MALE	2024-02-21	2000-11-30		
637c8822cf5fc14ed2527269	Zmd	MALE	2024-02-21	2001-01-06		
64f6e6563638254e9b8ee192	EKD	MALE	2024-02-20	1993-04-22		
64f6e6563638254e9b8ee192	Vd6	FEMALE	2024-02-19	1994-08-17		
63e23864cc40256f39c5bfde	NY6	FEMALE	2024-02-19	1965-06-03		
63e23864cc40256f39c5bfde	JRx	FEMALE	2024-02-19	2024-02-20		

Figura 2.1: Extracto del csv con los datos básicos

Happiness Index, HI

El *Happiness Index* (HI) es una métrica de autopercepción emocional que recoge diariamente el estado de ánimo de los empleados a través de una escala ordinal de cuatro niveles: 1 (triste), 2 (regular), 3 (contento) y 4 (muy feliz). Esta información se obtiene mediante la interacción voluntaria del trabajador con la aplicación móvil de HappyForce, lo que permite una recolección continua y longitudinal de datos emocionales a nivel individual.

Cada registro del HI incluye el identificador del empleado, el identificador de la empresa, la fecha y el valor declarado por el trabajador para ese día.

Desde un punto de vista metodológico, el HI representa una fuente rica de información subjetiva que complementa otras métricas más estructuradas como los *scores*. Su simplicidad facilita la recogida frecuente de datos sin generar fricción con los usuarios, y su formato ordinal permite explorar técnicas tanto estadísticas como basadas en aprendizaje automático para capturar correlaciones significativas con otras variables de interés.

En este proyecto, los registros diarios del HI se han promediado a nivel trimestral únicamente para generar una representación estacional compacta que sirva como variable de entrada en la comparación de modelos.

companyId	employeeId	date	vote
58b5f7cb917c350004dc0b09	o6	2017-05-22	3
58b5f7cb917c350004dc0b09	p5	2017-06-13	3
58b5f7cb917c350004dc0b09	MW	2017-06-20	1
58b5f7cb917c350004dc0b09	o6	2017-06-20	3
58b5f7cb917c350004dc0b09	n6	2017-06-20	3
58b5f7cb917c350004dc0b09	53	2017-06-20	3
58b5f7cb917c350004dc0b09	Gv	2017-06-20	4
58b5f7cb917c350004dc0b09	v5	2017-06-21	1
58b5f7cb917c350004dc0b09	MW	2017-06-21	1
58b5f7cb917c350004dc0b09	o6	2017-06-21	2
58b5f7cb917c350004dc0b09	Ee	2017-06-21	3
58b5f7cb917c350004dc0b09	Gv	2017-06-21	3

Figura 2.2: Extracto del csv con datos del HI

Scores

Los *scores* representan las respuestas cuantitativas que los empleados proporcionan a una serie de preguntas estructuradas que forman parte del ciclo de evaluación del eNPS. Estas preguntas están organizadas en torno a factores clave del entorno laboral, identificados en el conjunto de datos mediante el campo `factorName`. Algunos de los factores más representativos incluyen *Autonomy*, *Mastery*, entre otros, y agrupan preguntas relacionadas con dimensiones específicas como la libertad para tomar decisiones, el desarrollo profesional o la percepción de progreso.

Cada *score* se asocia a una pregunta concreta, identificada por su `questionId` y su enunciado (`questionTitle`), y se registra como una puntuación numérica sobre una escala fija. Estas puntuaciones permiten analizar de forma segmentada las percepciones de los empleados sobre distintos aspectos de su experiencia laboral, y son esenciales para comprender las causas subyacentes del compromiso o descontento.

Además, las respuestas incluyen metainformación como el identificador del empleado, el de la empresa, y la fecha, lo que permite contextualizar los datos sin comprometer la privacidad. Esta estructura facilita el uso de técnicas de análisis estadístico y aprendizaje automático para modelar cómo cada dimensión del entorno influye en la probabilidad de recomendar la empresa como un buen lugar para trabajar.

En este proyecto, los *scores* han sido tratados como variables predictoras clave del valor eNPS (1–10), tanto por su granularidad como por su conexión directa con factores organizacionales específicos. Se han explorado agregaciones temporales, normalizaciones y transformaciones estadísticas que permiten capturar su comportamiento de forma efectiva en el modelo predictivo.

companyId	employeeId	scoreId	factorId	questionId	date	vote
5b1ee1a38b1b030004b8ec59	aG	5dd6d642103b810004915e04	5dd6d642103b810004915dfb	5dd6d642103b810004915dfc	2020-02-03	10
5b1ee1a38b1b030004b8ec59	W62	5dd6d642103b810004915e04	5dd6d642103b810004915dfb	5dd6d642103b810004915dfc	2020-02-12	10
5b1ee1a38b1b030004b8ec59	a8G	5dd6d642103b810004915e04	5dd6d642103b810004915dfb	5dd6d642103b810004915dfc	2020-02-17	10
5b1ee1a38b1b030004b8ec59	Xj9	5dd6d642103b810004915e04	5dd6d642103b810004915dfb	5dd6d642103b810004915dfc	2020-02-26	8
5b1ee1a38b1b030004b8ec59	vvy	5dd6d642103b810004915e04	5dd6d642103b810004915dfb	5dd6d642103b810004915dfc	2020-02-29	10
5b1ee1a38b1b030004b8ec59	PPM	5dd6d642103b810004915e04	5dd6d642103b810004915dfb	5dd6d642103b810004915dfc	2020-03-02	8
5b1ee1a38b1b030004b8ec59	O2j	5dd6d642103b810004915e04	5dd6d642103b810004915dfb	5dd6d642103b810004915dfc	2020-03-02	8
5b1ee1a38b1b030004b8ec59	7D	5dd6d642103b810004915e04	5dd6d642103b810004915dfb	5dd6d642103b810004915dfc	2020-03-02	9
5b1ee1a38b1b030004b8ec59	d3q	5dd6d642103b810004915e04	5dd6d642103b810004915dfb	5dd6d642103b810004915dfc	2020-03-03	8
5b1ee1a38b1b030004b8ec59	JEL	5dd6d642103b810004915e04	5dd6d642103b810004915dfb	5dd6d642103b810004915dfc	2020-03-04	9
5b1ee1a38b1b030004b8ec59	r7e	5dd6d642103b810004915e04	5dd6d642103b810004915dfb	5dd6d642103b810004915dfc	2020-03-16	7

Figura 2.3: Extracto del csv con datos de los Scores

2.3. Metodología

2.3.1. Preparación de los datos

Para el desarrollo del modelo predictivo del voto del eNPS, se emplearon datos de empleados que incluyen puntuaciones de satisfacción (HI), respuestas a preguntas estructuradas (*scores*) y el valor del eNPS individual. En primer lugar, se realizó un proceso de limpieza y transformación: se eliminaron duplicados, valores nulos y los empleados con menos de 20 entradas en el registro de datos de la tabla *scores*, se combinaron identificadores por empleado y empresa, y se unificaron los conjuntos de datos mediante claves comunes. Posteriormente, se trabajaron dos enfoques diferenciados: observaciones independientes y secuencias temporales.

2.3.2. Unificación temporal y preparación contextual

Una parte fundamental del tratamiento de los datos consistió en vincular correctamente los distintos eventos temporales asociados a cada empleado. En particular, fue necesario alinear cronológicamente las entradas del *Happiness Index* (HI), los *scores* y las respuestas del eNPS.

Se definió un intervalo temporal de referencia de 90 días anterior a cada registro del eNPS, debido a que el eNPS es una métrica trimestral. A partir de este criterio, se seleccionaron únicamente los registros de HI y scores que cayeran dentro de ese intervalo, garantizando así que el modelo utilizara únicamente información disponible antes de la respuesta objetivo.

Para ello, se transformaron las fechas a formato *datetime*, se ordenaron cronológicamente los registros por empleado, y se calculó la diferencia temporal entre respuestas al eNPS sucesivas. Si el intervalo entre respuestas consecutivas del eNPS superaba los 90 días, se asignaba este valor (90) por defecto para crear el ciclo de cada valor individual del eNPS.

Este tratamiento permitió construir un conjunto de observaciones donde cada punto de entrenamiento se basaba únicamente en información previa al voto de eNPS, respetando la lógica causal del problema. Esta fase fue especialmente relevante para la posterior división de datos en observaciones individuales o secuencias temporales, que se detallan en los siguientes apartados.

2.3.3. Aspectos comunes del entrenamiento y evaluación

En todos los modelos desarrollados, se empleó una estrategia común de división de datos y evaluación con el objetivo de asegurar la comparabilidad entre enfoques.

Previo al entrenamiento de los modelos, todas las variables numéricas utilizadas fueron normalizadas mediante el método de escalado *Min-Max*. Este procedimiento transforma cada variable al rango $[0, 1]$, aplicando la fórmula:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2.1)$$

donde x es el valor original, y x_{min} y x_{max} corresponden a los valores mínimo y máximo de cada variable.

Los conjuntos de datos fueron particionados en tres subconjuntos: entrenamiento (70 %), validación (15 %) y prueba (15 %). Esta división se realizó de forma estratificada por `employeeId`, asegurando que cada empleado estuviera presente únicamente en uno de los conjuntos, y así evitar fuga de datos.

Para todos los modelos se utilizaron las métricas estándar de regresión: el error cuadrático medio (MSE), el error absoluto medio (MAE) y el coeficiente de determinación R^2 . Durante el entrenamiento, se monitorizó la pérdida en el conjunto de validación mediante el `callback ModelCheckpoint`, que almacenó la mejor versión del modelo en cada caso.

Además, para facilitar la interpretación de los resultados, se generaron visualizaciones de las curvas de aprendizaje (evolución de MSE y MAE por época) y gráficos de dispersión entre los valores reales y los valores predichos, tanto para modelos basados en observaciones independientes como en secuenciales.

2.3.4. Modelado sin secuencias

Para uno de los enfoques desarrollados, se trabajó con observaciones independientes por empleado, sin tener en cuenta la evolución temporal de las mismas.

De las fechas se extrajeron variables temporales como el mes y el trimestre. Estas variables se transformaron mediante codificación *one-hot*, utilizando la función `get_dummies`, para poder ser interpretadas por los modelos de aprendizaje automático.

Modelo MLP base

Se construyó un modelo base de red neuronal multicapa (*Multilayer Perceptron*, *MLP*) mediante la API de Keras. El modelo constaba de:

- Capa de entrada con dimensión igual al número de características.
- Primera capa oculta: 64 neuronas con activación *ReLU* + *Dropout* de 0.3.
- Segunda capa oculta: 32 neuronas con activación *ReLU* + *Dropout* de 0.2.
- Capa de salida: 1 neurona sin función de activación (regresión).

El modelo fue entrenado durante 50 épocas, siguiendo el procedimiento de evaluación, métricas y *callbacks* descritos anteriormente.

Modelo MLP con Hypertuning

Con el objetivo de mejorar el rendimiento del modelo base, se aplicó una búsqueda automatizada de hiperparámetros utilizando la librería *Keras Tuner*. Se definió una función `build_model` que permitía configurar arquitecturas personalizadas en función de distintos parámetros ajustables:

- Número de neuronas en la primera (entre 32 y 128, cada 32) y segunda capa oculta (entre 16 y 64, cada 16).
- Inclusión o no de la segunda capa oculta.
- Tasas de *Dropout* en cada capa incluida (entre 0.1 y 0.5, cada 0.1).

Se utilizó el método `RandomSearch` con un máximo de 50 combinaciones distintas. Cada modelo fue entrenado durante 10 épocas, seleccionando el que obtuvo menor `val_loss` para ser reentrenado completamente durante 50 épocas.

El modelo final obtenido con el hypertuning fue:

- Capa de entrada con dimensión igual al número de características.
- Primera capa oculta: 32 neuronas con activación *ReLU* + *Dropout* de 0.1.
- Segunda capa oculta: 64 neuronas con activación *ReLU* + *Dropout* de 0.3.
- Capa de salida: 1 neurona sin función de activación (regresión).

2.3.5. Modelado con secuencias

Además del enfoque con observaciones independientes, se exploró una segunda estrategia centrada en el modelado secuencial de las respuestas de los empleados. En este caso, se consideró la evolución temporal del comportamiento de cada individuo mediante el uso de ventanas deslizantes de longitud fija.

El conjunto de datos utilizado se basó en registros históricos de empleados con al menos cuatro observaciones válidas. Para cada uno, se generaron secuencias temporales de longitud 4 que incorporaban tanto las variables promediadas de satisfacción (HI) como las puntuaciones de los factores (*scores*). Además, se introdujo una nueva variable denominada `delta_days`, definida como el número de días transcurridos entre registros consecutivos del eNPS.

Para garantizar la coherencia temporal de estas secuencias, se estableció una restricción que descartó aquellas observaciones en las que los intervalos temporales entre registros consecutivos fueran superiores a 180 días. Esta restricción se justifica porque la variable objetivo (`vote_y`) representa el estado de un empleado aproximadamente cada 90 días; si el intervalo entre observaciones supera los 180 días (seis meses), es razonable suponer que el estado anterior ya no refleja con fiabilidad el contexto actual. Además, la inclusión de estas secuencias atípicas introduce ruido y reduce la coherencia temporal necesaria para que modelos secuenciales como LSTM puedan aprender patrones longitudinales de forma eficaz.

Para cumplir estrictamente esta condición, se implementó un procedimiento iterativo mediante un bucle que eliminó progresivamente aquellas observaciones que excedían el intervalo permitido, recalculando los valores de `delta_days` tras cada eliminación hasta asegurar que todas las secuencias restantes cumplieran con el criterio temporal establecido. Finalmente, los valores de `delta_days` resultantes fueron normalizados con el mismo método Min-Max descrito previamente, y añadidos como una característica adicional al conjunto de variables predictoras, antes de proceder a la generación final de las secuencias empleadas por el modelo.

Tanto el modelo LSTM como el modelo MLP descritos a continuación fueron entrenados sobre el mismo conjunto de datos secuenciales previamente detallado.

Modelo LSTM base

Se implementó un modelo basado en redes neuronales recurrentes del tipo *Long Short-Term Memory* (LSTM) utilizando la API de Keras. La arquitectura consistió en:

- Capa de entrada con forma $(4, n_features)$ correspondiente a la secuencia de 4 pasos.
- Primera capa oculta: *LSTM* con 32 unidades + *Dropout* de 0.3.
- Segunda capa oculta: 32 neuronas con activación *ReLU* + *Dropout* de 0.2.
- Capa de salida: 1 neurona sin función de activación (regresión).

Se entrenó durante 50 épocas utilizando el mismo esquema de evaluación y visualización que en los modelos anteriores.

Modelo LSTM con Hypertuning

Al igual que en el enfoque sin secuencias, se aplicó un proceso de búsqueda de hiperparámetros mediante *Keras Tuner*. Se definió una función `build_model` que permitía generar arquitecturas variables en función de los siguientes hiperparámetros:

- Número de capas *LSTM* (una o dos).
- Número de unidades por capa (entre 32 y 128, cada 32).
- Número de unidades en la capa densa intermedia (entre 16 y 64).
- Tasas de *Dropout* para cada capa *LSTM* y la capa *Densa* (entre 0.2 y 0.5, cada 0.1).

La búsqueda se realizó mediante el algoritmo `RandomSearch`, evaluando 50 combinaciones distintas durante 10 épocas cada una. El modelo con menor pérdida en validación fue seleccionado y reentrenado completamente durante 50 épocas.

El modelo final obtenido con el hypertuning fue:

- Capa de entrada con forma $(4, n_features)$ correspondiente a la secuencia de 4 pasos.

- Primera capa oculta: *LSTM* con 64 unidades + *Dropout* de 0.3.
- Segunda capa oculta: 32 neuronas con activación *ReLU* + *Dropout* de 0.2.
- Capa de salida: 1 neurona sin función de activación (regresión).

Modelo MLP base sobre datos secuenciales

Además del modelo LSTM, se implementó un modelo de red neuronal multicapa (MLP) entrenado sobre los mismos datos secuenciales, pero transformados en vectores unidimensionales. Cada secuencia de longitud fija ($4 \times n_features$) fue aplanada en un vector de dimensión $4 \cdot n_features$, lo que permitió procesarla mediante una arquitectura densa convencional.

El modelo MLP sobre secuencias aplanadas funciona como referencia intermedia: permite comprobar, primero, si incorporar información temporal, aunque sea sin preservar el orden interno, mejora el rendimiento frente al MLP sin secuencias y, segundo, qué valor adicional aporta el LSTM al mantener la dinámica completa de las secuencias.

La arquitectura empleada para el MLP fue la siguiente:

- Capa de entrada con dimensión igual al vector aplanado.
- Primera capa oculta: 128 neuronas con activación *ReLU* + *Dropout* de 0.4.
- Segunda capa oculta: 64 neuronas con activación *ReLU* + *Dropout* de 0.3.
- Capa de salida: 1 neurona sin función de activación (regresión).

Se mantuvo el mismo procedimiento de entrenamiento, con 50 épocas y los métodos de evaluación ya descritos.

Modelo MLP con Hypertuning

Con el objetivo de mejorar el rendimiento del modelo MLP aplicado sobre datos secuenciales aplanados, se realizó una búsqueda de hiperparámetros utilizando la librería *Keras Tuner*. Se definió una función `build_model` que permitía configurar arquitecturas personalizables en función de distintos parámetros:

- Número de neuronas en la primera (entre 32 y 128, cada 32) y segunda capa oculta (entre 16 y 64, cada 16).
- Inclusión o no de la segunda capa oculta.
- Tasas de *Dropout* en cada capa incluida (entre 0.1 y 0.5, cada 0.1).

La búsqueda se realizó utilizando el algoritmo `RandomSearch`, con un total de 50 combinaciones distintas. Cada modelo fue entrenado durante 10 épocas y evaluado en el conjunto de validación. La métrica objetivo fue la pérdida (`val_loss`).

Una vez identificada la mejor combinación de hiperparámetros, el modelo resultante fue:

- Capa de entrada con dimensión igual al vector aplanado.
- Primera capa oculta: 64 neuronas con activación *ReLU* + *Dropout* de 0.4.
- Capa de salida: 1 neurona sin función de activación (regresión).

La arquitectura final fue reentrenada durante 50 épocas sobre el conjunto de entrenamiento.

Comparación de enfoques

Ambos enfoques fueron evaluados bajo las mismas métricas para permitir una comparación directa. Mientras que el modelo sin secuencias trata cada observación como independiente y desprovista de contexto temporal, el modelo con secuencias capta relaciones longitudinales entre valoraciones sucesivas. Esta diferencia en el diseño se refleja en los resultados que se analizan en la sección siguiente.

Capítulo 3

Resultados y comparación de modelos

3.1. Resultados

En este capítulo se presentan y analizan los resultados obtenidos con los modelos desarrollados para predecir la respuesta individual (1–10) a la pregunta eNPS. El capítulo se estructura en tres bloques:

1. **Enfoque sin secuencias (3.2)**: modelos MLP entrenados sobre observaciones independientes, sin información temporal explícita.
2. **Enfoque LSTM con secuencias (3.3)**: modelo recurrente que explota la dinámica de las respuestas mediante ventanas de longitud 4.
3. **Enfoque MLP con secuencias (3.4)**: modelo denso aplicado a las mismas secuencias (aplanadas) para servir de línea base frente al LSTM.

Para cada bloque se muestran:

- Las curvas de aprendizaje (MSE y MAE) en entrenamiento y validación.
- Las métricas finales, MSE, MAE y R^2 , en escala original (1–10).
- El gráfico de dispersión entre valores reales y predichos.

Tras analizar individualmente cada enfoque, se cierra el capítulo con una comparación de los seis modelos (3.5), destacando el que ofrece los mejores resultados.

3.2. Enfoque sin secuencias

En este bloque se evalúan dos versiones de MLP entrenadas sobre observaciones independientes, sin incorporar información temporal explícita. La primera configuración (3.2.1) corresponde al diseño base, mientras que la segunda (3.2.2) aplica optimización automática de hiperparámetros para explorar mejoras de precisión sin cambiar la naturaleza del enfoque.

3.2.1. MLP básico

Curvas de entrenamiento

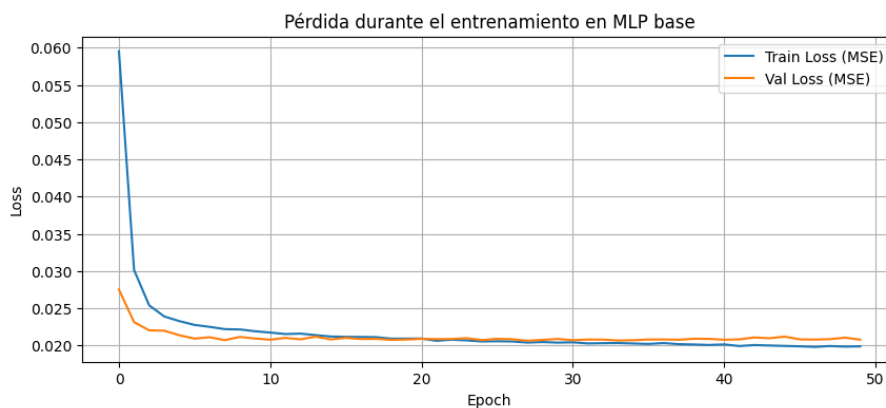


Figura 3.1: Evolución de MSE - MLP básico sin secuencias.

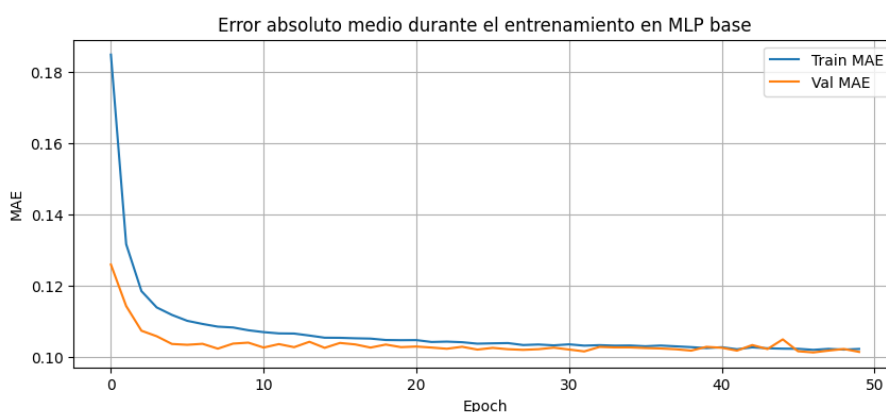


Figura 3.2: Evolución de MAE - MLP básico sin secuencias.

Las curvas de las Figuras 3.1 y 3.2 descienden con rapidez en las primeras épocas y luego se estabilizan, manteniéndose muy próximas en entrenamiento y validación; esta evolución coherente sugiere que el modelo aprende sin sobreajuste y que el proceso de optimización está bien controlado.

Métricas en el conjunto de prueba

MSE	MAE	R^2
1.6751	0.9192	0.6163

Tabla 3.1: Métricas finales en el MLP base

En cuanto a las métricas finales, el MSE de 1.6751 equivale a un RMSE (raíz del MSE) de 1.2943, de modo que la predicción se desvía, en promedio, $\pm 1,3$ puntos sobre la escala 1–10. El MAE de 0.9192 refuerza esa lectura: la diferencia media absoluta también queda

por debajo de un punto. El $R^2 = 0.6163$ significa que el modelo reduce en torno al 61.6 % el error que se cometería si asignásemos a todos los empleados la misma cifra, la media de las respuestas.

Dispersión del valor real vs. predicho

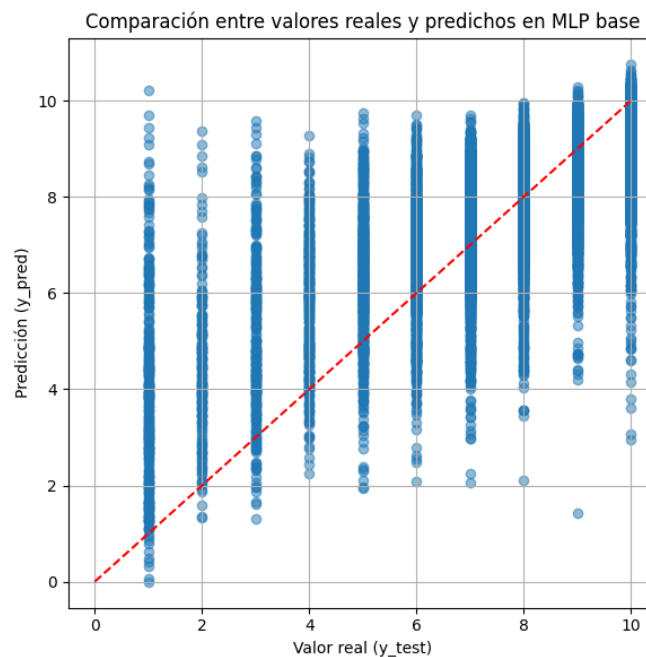


Figura 3.3: Valor real vs. predicho - MLP básico sin secuencias.

La dispersión de la Figura 3.3 confirma estos resultados: la mayoría de los puntos se alinean con la diagonal, aunque se observa ligera sobreestimación en los valores más bajos (1–4) y cierta saturación al acercarse a 10. En conjunto, este MLP básico ofrece una referencia sólida, pero su rendimiento queda limitado por la ausencia de información temporal.

3.2.2. MLP hypertuning

Curvas de entrenamiento

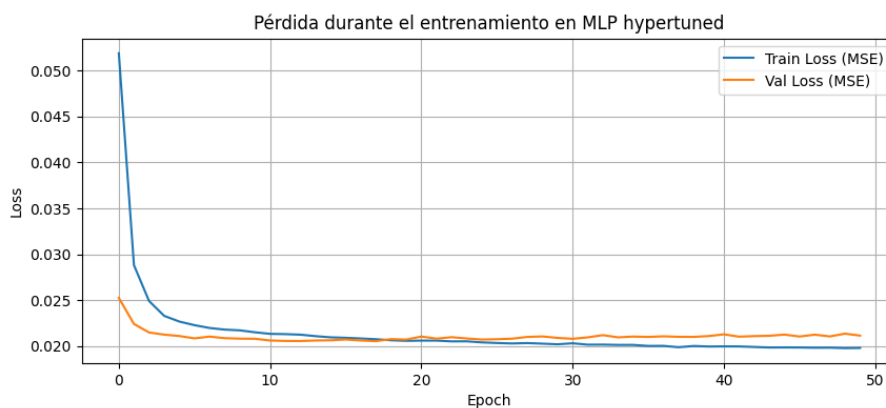


Figura 3.4: Evolución de MSE - MLP optimizado sin secuencias.

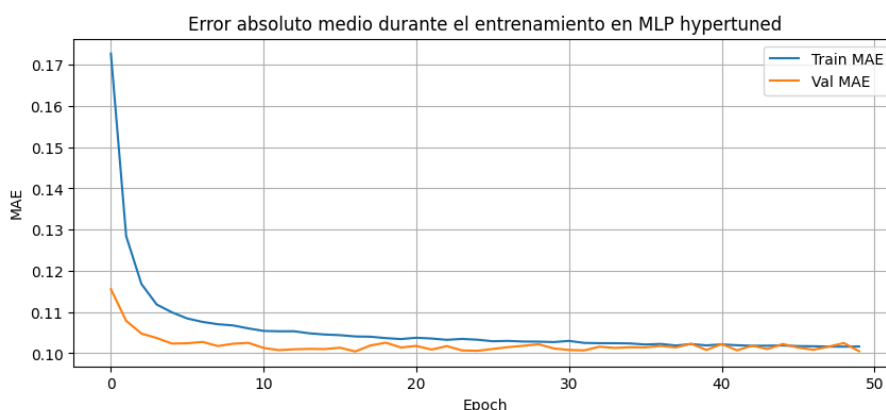


Figura 3.5: Evolución de MAE - MLP optimizado sin secuencias.

Las curvas de las Figuras 3.4 y 3.5 muestran un descenso rápido durante las primeras épocas y se estabilizan sin divergencia notable entre entrenamiento y validación; este comportamiento evidencia un proceso de optimización bien regulado y confirma que la búsqueda automática de hiperparámetros no condujo a sobreajuste.

Métricas en el conjunto de prueba

MSE	MAE	R^2
1.6676	0.9167	0.6181

Tabla 3.2: Métricas finales en el MLP optimizado

El MSE de 1.6676 supone un RMSE de 1.2914, ligeramente mejor que el modelo base; el MAE desciende ligeramente a 0.9167. El $R^2 = 0.6181$ equivale a una reducción del 61.8 % del error frente al predictor que devuelve la media, apenas 0.2 puntos porcentuales

por encima del 61.6 % logrado por el MLP básico. La lectura es la misma que en el modelo base, pero mejorando un poco las métricas.

Dispersión del valor real vs. predicho

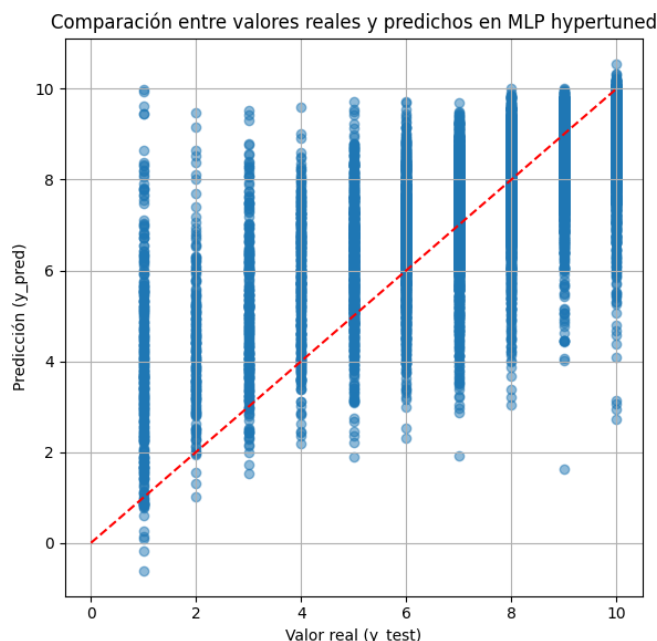


Figura 3.6: Valor real vs. predicho - MLP optimizado sin secuencias.

La dispersión de la Figura 3.6 muestra un patrón muy similar al del modelo base: la mayoría de los puntos se alinea con la diagonal, existiendo una ligera sobreestimación en el rango 1–4 y cierta saturación cerca de 10. La búsqueda de hiperparámetros ha logrado una mejora modesta sin incrementar la complejidad, lo que sugiere que, en este enfoque sin secuencias, el margen de ganancia es limitado.

Síntesis del enfoque sin secuencias

La comparación entre el MLP básico y su versión optimizada confirma que la búsqueda automática de hiperparámetros aporta una mejora, aunque modesta. El MAE se reduce de 0.9192 a 0.9167 y el MSE de 1.6751 a 1.6676, mientras que el R^2 pasa de 0.6163 a 0.6181. Si bien el incremento absoluto es inferior a un punto porcentual, estos resultados validan la utilidad del hypertuning para afinar la arquitectura sin incrementar la complejidad del modelo, dejando la configuración base como una referencia sólida y el modelo optimizado como su versión ligeramente mejorada.

3.3. Enfoque LSTM con secuencias

En este bloque se analizan dos versiones de LSTM que aprovechan explícitamente la información temporal mediante secuencias de longitud 4. La primera configuración (3.3.1) es el modelo LSTM base, mientras que la segunda (3.3.2) incorpora optimización automática de hiperparámetros manteniendo la misma filosofía de modelado secuencial.

3.3.1. LSTM básico

Curvas de entrenamiento

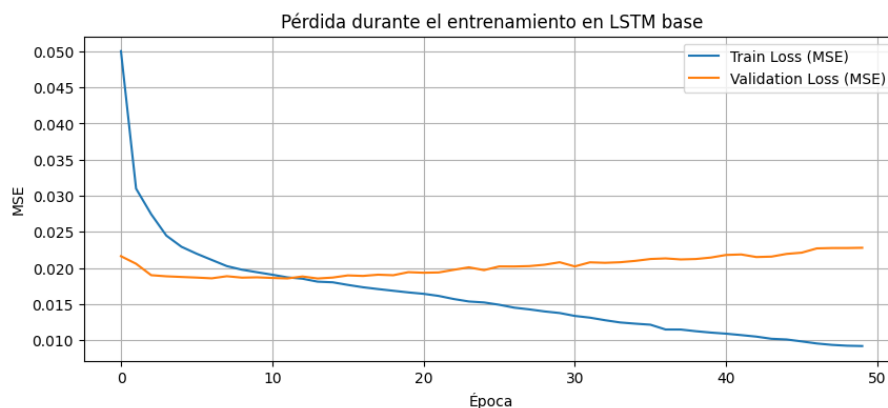


Figura 3.7: Evolución de MSE - LSTM básico con secuencias.

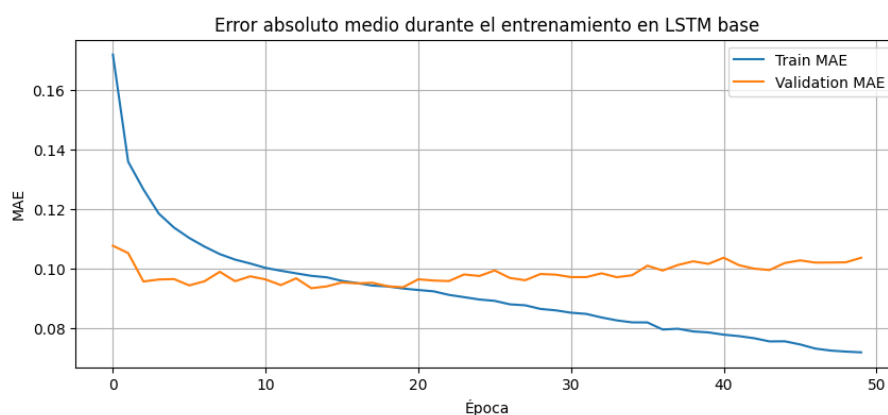


Figura 3.8: Evolución de MAE - LSTM básico con secuencias.

Las curvas de las Figuras 3.7 y 3.8 descienden rápidamente durante las primeras diez épocas y, a partir de la época 20, la línea de validación se aplanan mientras la de entrenamiento continúa cayendo; esta ligera separación sugiere un comienzo de sobreajuste, probablemente debido a la mayor capacidad del LSTM y a los cambios naturales que presentan los datos secuenciales.

Métricas en el conjunto de prueba

MSE	MAE	R^2
1.4121	0.8239	0.6533

Tabla 3.3: Métricas finales en el LSTM base

Por lo que se refiere a las métricas obtenidas, el MSE de 1.4121 corresponde a un RMSE de 1.188 puntos, de modo que la predicción se desvía, en promedio, $\pm 1,2$ puntos sobre la

escala 1–10. El MAE de 0.8239 refleja un error medio de unas ocho décimas, quedando por debajo de un punto. El $R^2=0.6533$ indica que el modelo reduce cerca del 65.3 % del error frente a predecir siempre la media, lo que evidencia la ventaja de incorporar información secuencial.

Dispersión del valor real vs. predicho

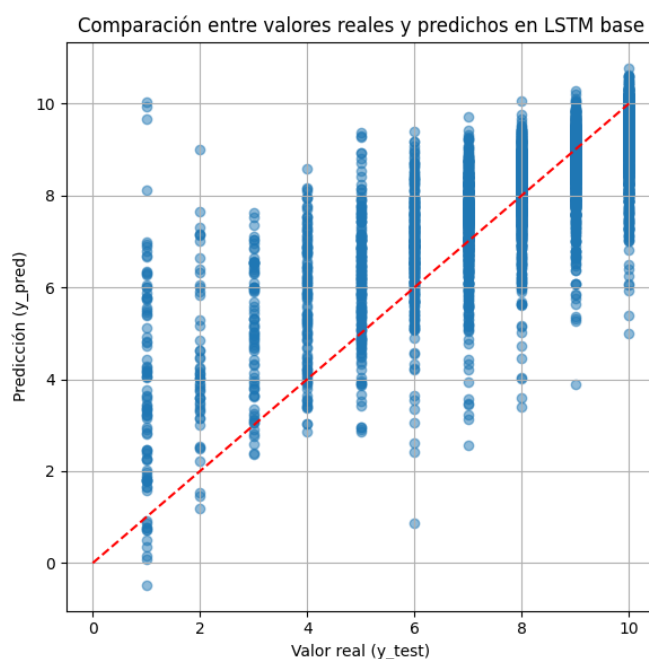


Figura 3.9: Valor real vs. predicho - LSTM básico con secuencias.

La dispersión de la Figura 3.9 muestra una alineación general con la diagonal, aunque persiste cierta sobreestimación en valores reales más bajos (1–4) y ligera compresión cerca del máximo de la escala. Con todo, el uso de secuencias temporales permite captar mejor la dinámica de las respuestas, logrando métricas más favorables que las obtenidas sin información temporal explícita.

3.3.2. LSTM hypertuning

Curvas de entrenamiento

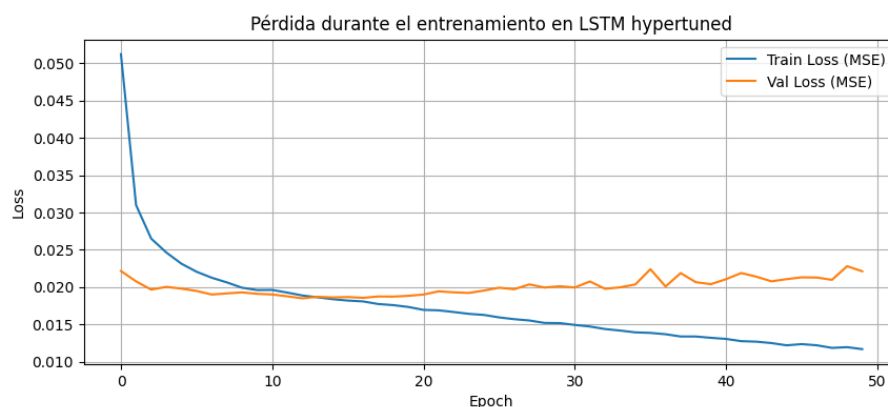


Figura 3.10: Evolución de MSE - LSTM optimizado con secuencias.

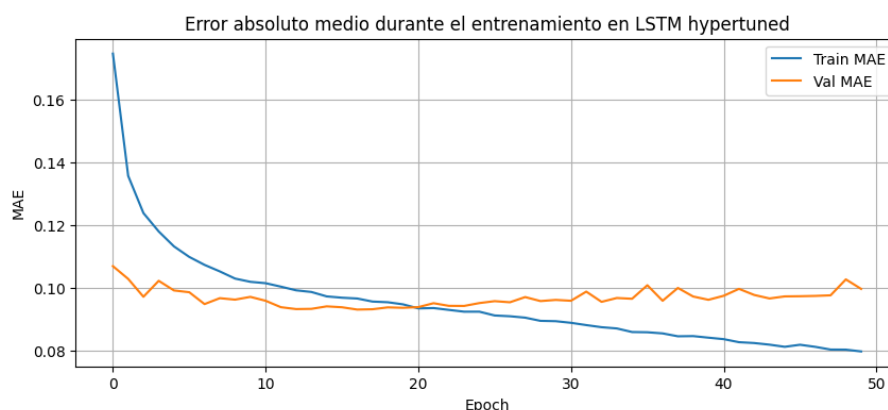


Figura 3.11: Evolución de MAE - LSTM optimizado con secuencias.

Las curvas de las Figuras 3.10 y 3.11 descienden con rapidez durante las diez primeras épocas. A partir de la época 15, la línea de validación se estabiliza, mientras que la de entrenamiento continúa bajando de forma paulatina; esta pequeña separación refleja un inicio de sobreajuste, que sigue la misma lógica que en el modelo base. La diferencia en la separación de las líneas con respecto al modelo base sugiere que la configuración elegida por el hypertuning mejora ligeramente las curvas de aprendizaje.

Métricas en el conjunto de prueba

MSE	MAE	R^2
1.4081	0.8180	0.6543

Tabla 3.4: Métricas finales en el LSTM optimizado

El MSE de 1.4081 implica un RMSE de 1.186 puntos, una mejora sutil sobre el modelo base; el MAE baja a 0.818, mejorando ligeramente la cifra del LSTM base. Con un R^2

$=0.6543$, el modelo reduce cerca del 65.4% el error en comparación con asignar a todos los casos el valor medio, logrando la mejor capacidad explicativa dentro de este enfoque. El análisis es muy similar al del modelo base, pero con unas métricas sutilmente mejores.

Dispersión del valor real vs. predicho

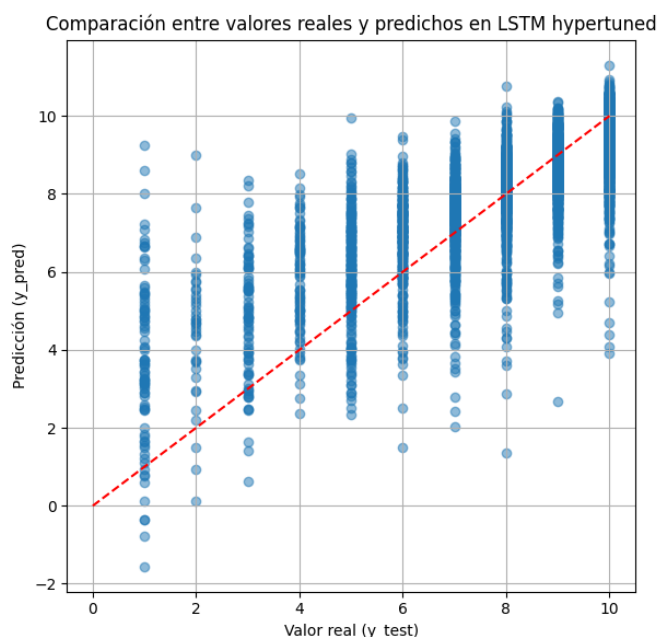


Figura 3.12: Valor real vs. predicho - LSTM optimizado con secuencias.

La dispersión de la Figura 3.12 se ajusta bien a la diagonal, con menor dispersión que en el modelo base. Aún se observa cierta sobreestimación en el rango 1–4 y ligera compresión al acercarse a 10, pero en menor grado. En conjunto, el LSTM optimizado aprovecha la búsqueda de hiperparámetros para obtener una mejora moderada en todas las métricas, al tiempo que mantiene controlado el sobreajuste.

3.4. Enfoque MLP con secuencias

En este bloque se estudian dos variantes de MLP entrenadas sobre las mismas ventanas temporales de longitud 4 que emplean los modelos recurrentes. Cada secuencia se aplanan en un único vector, de modo que la red densa recibe información histórica, aunque sin procesar el orden interno paso a paso. La subsección (3.4.1) presenta la configuración básica y la (3.4.2) su versión optimizada mediante búsqueda automática de hiperparámetros.

3.4.1. MLP básico

Curvas de entrenamiento

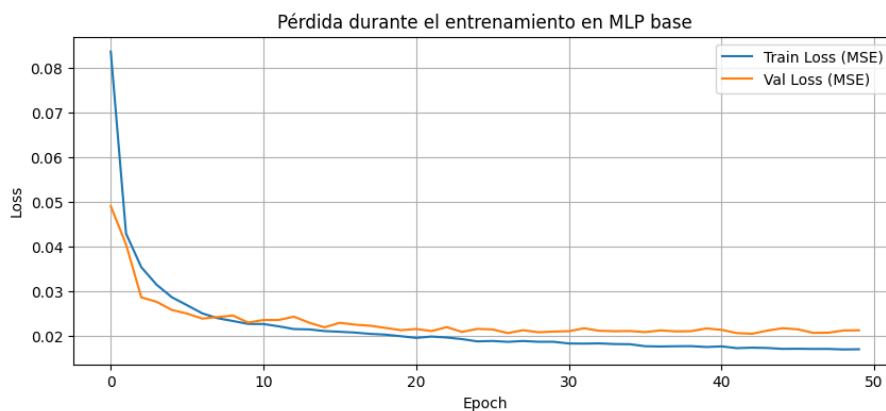


Figura 3.13: Evolución de MSE - MLP básico con secuencias.

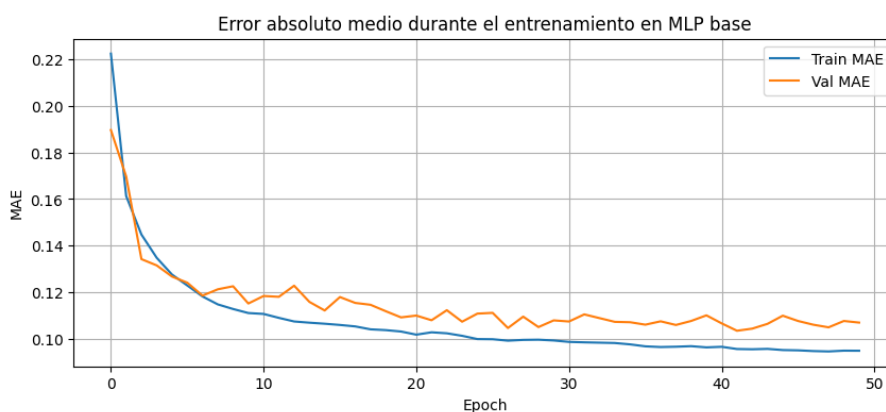


Figura 3.14: Evolución de MAE - MLP básico con secuencias.

Las curvas de las Figuras 3.13 y 3.14 descienden con rapidez en las primeras diez épocas y, a partir de la época 15, la línea de validación se estabiliza ligeramente por encima de la de entrenamiento; la escasa distancia entre las curvas de entrenamiento y validación indica que el modelo converge sin mostrar sobreajuste significativo y mantiene una capacidad de generalización aceptable.

Métricas en el conjunto de prueba

MSE	MAE	R^2
1.5298	0.9024	0.6244

Tabla 3.5: Métricas finales en el MLP base sin secuencias

Por lo que respecta al resultado de las métricas, el MSE de 1.5298 equivale a un RMSE de 1.237 puntos en la escala 1–10, de modo que la predicción se desvía, en promedio, $\pm 1,25$

puntos; el MAE de 0.902 confirma un error medio por debajo de un punto sobre la escala 1–10. El $R^2 = 0.6244$ muestra que el modelo reduce un 62.4% aproximadamente el error frente a predecir la media, evidenciando la utilidad de incorporar información secuencial aunque esté aplanada.

Dispersión del valor real vs. predicho

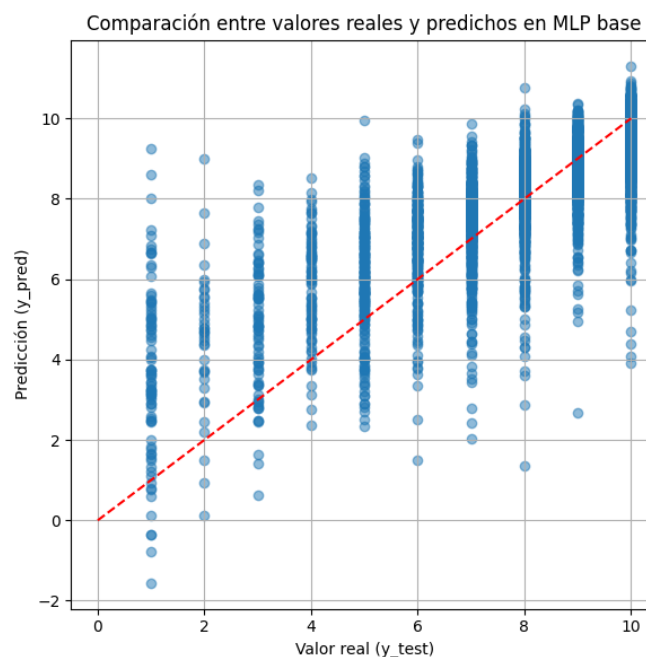


Figura 3.15: Valor real vs. predicho - MLP básico con secuencias.

La dispersión de la Figura 3.15 se alinea razonablemente con la diagonal, revelando predicciones consistentes en la zona central de la escala. Persiste ligera sobreestimación en los valores reales más bajos (1–4) y cierta compresión cerca del máximo, pero el ajuste global es más estrecho que en el modelo sin secuencias, lo que respalda la aportación de las ventanas temporales aplanadas.

3.4.2. MLP hypertuning

Curvas de entrenamiento

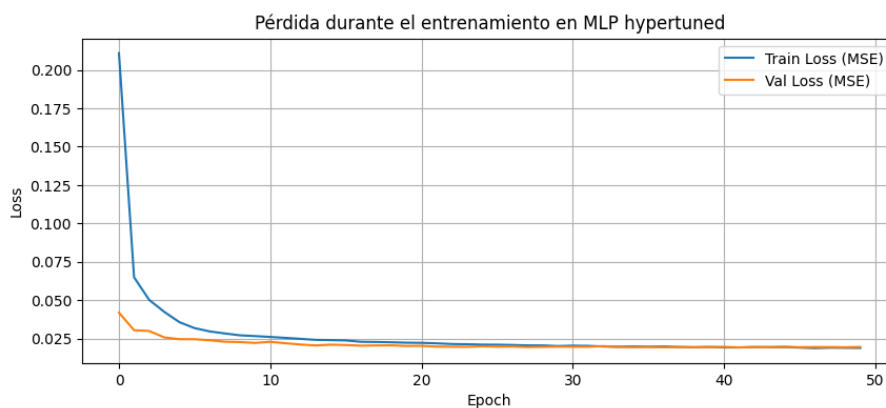


Figura 3.16: Evolución de MSE - MLP optimizado con secuencias.

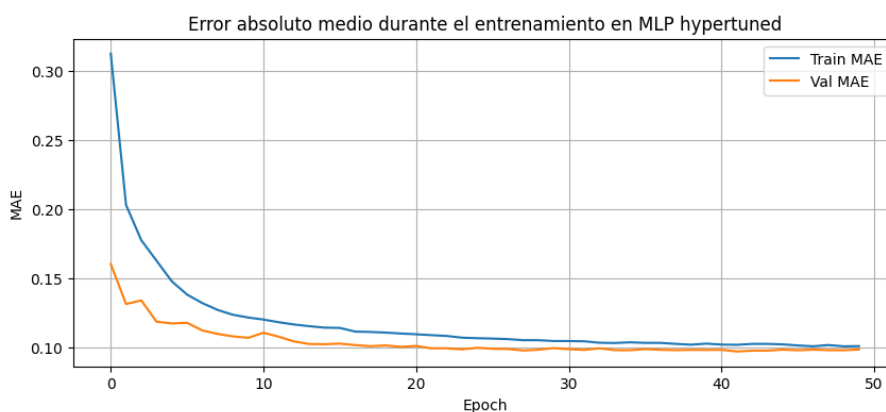


Figura 3.17: Evolución de MAE - MLP optimizado con secuencias.

Las curvas de las Figuras 3.16 y 3.17 descienden de forma sostenida durante las primeras quince épocas y se estabilizan con un margen muy reducido entre entrenamiento y validación, lo que indica que el ajuste de hiperparámetros ha logrado un compromiso equilibrado sin sobreajuste apreciable.

Métricas en el conjunto de prueba

MSE	MAE	R^2
1.4819	0.8486	0.6362

Tabla 3.6: Métricas finales en el MLP optimizado sin secuencias

El MSE desciende a 1.4819 ($RMSE = 1.217$ puntos) y el MAE a 0.8486, lo que supone una mejora respecto al MLP secuencial básico. El $R^2 = 0.6362$ confirma una reducción del error del 63.6 % frente al predictor de la media y refleja un incremento de capacidad

explicativa de casi dos puntos porcentuales dentro de este enfoque. El comportamiento es muy parecido al de la versión base, aunque exhibe una ligera mejora en las métricas.

Dispersión del valor real vs. predicho

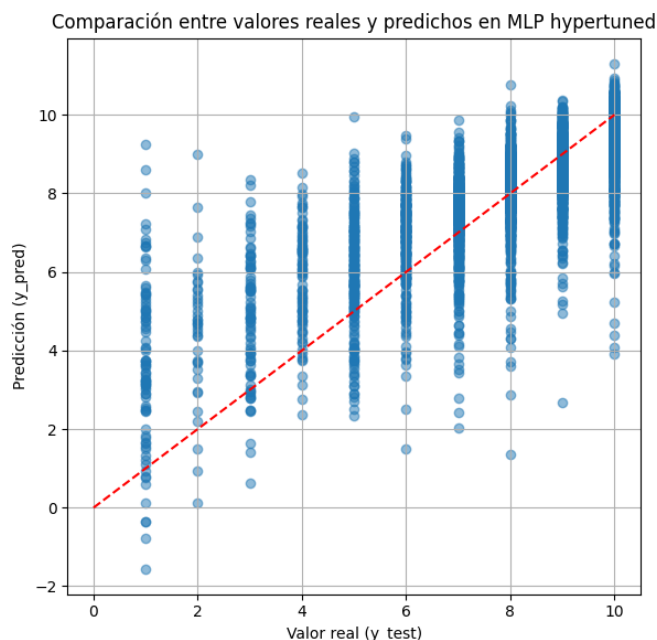


Figura 3.18: Valor real vs. predicho - MLP optimizado con secuencias.

La dispersión de la Figura 3.18 mantiene la alineación con la diagonal y muestra una menor dispersión en el rango central; la leve sobreestimación para valores reales 1–4 y la compresión cerca de 10 persisten, pero se atenúan. En conjunto, el hypertuning aporta una mejora tangible en las tres métricas sin aumentar la complejidad del modelo, consolidando este MLP optimizado como la mejor variante densa dentro del enfoque secuencial.

Síntesis del enfoque con secuencias

Los cuatro modelos basados en ventanas temporales superan a los que trabajan con observaciones independientes, lo que confirma la utilidad de la información histórica. Dentro de las redes densas, la versión optimizada reduce el MSE de 1.5298 a 1.4819 y el MAE de 0.9024 a 0.8486, mientras que el R^2 sube de 0.6244 a 0.6362.

En las redes recurrentes, el ajuste de hiperparámetros mejora el MSE de 1.4121 a 1.4081 y el MAE de 0.8239 a 0.8180, con un incremento de R^2 de 0.6533 a 0.6543; las diferencias son muy pequeñas, pero mejoran, lo que indica que la arquitectura LSTM base ya estaba cerca de su mejor configuración.

Comparando las mejores variantes de cada arquitectura, el LSTM optimizado alcanza las métricas más favorables (MSE = 1.4081, MAE = 0.8180, R^2 = 0.6543), mientras que el MLP optimizado con secuencias se queda en MSE = 1.4819, MAE = 0.8486 y R^2 = 0.6362. Por ello, el LSTM optimizado se consolida como el modelo más preciso dentro del enfoque secuencial.

3.5. Comparación global de modelos

Tras las síntesis parciales de cada enfoque, esta sección contrasta los seis modelos evaluados en un único marco comparativo. Primero se presenta una tabla que reúne los valores exactos de MSE, MAE y R^2 ; a continuación se incluyen tres gráficos de barras, uno por métrica, que permiten apreciar de forma visual las diferencias de rendimiento. El apartado concluye con una síntesis que identifica el modelo con la mejor combinación de precisión y coste computacional, recomendándolo como candidato para su implementación piloto en la plataforma HappyForce.

Tabla Comparativa

Modelo	MSE	MAE	R^2
MLP básico sin secuencias	1.6751	0.9192	0.6163
MLP hypertuning sin secuencias	1.6676	0.9167	0.6181
MLP básico con secuencias	1.5298	0.9024	0.6244
MLP hypertuning con secuencias	1.4819	0.8486	0.6362
LSTM básico con secuencias	1.4121	0.8239	0.6533
LSTM hypertuning con secuencias	1.4081	0.8180	0.6543

Tabla 3.7: Métricas finales de cada modelo

La Tabla 3.7 recoge los valores exactos que se representan en los gráficos siguientes. Cabe recordar que, para el MSE y el MAE, se busca el valor más bajo y, para el R^2 , el valor más alto, de modo que se obtenga una mejor precisión.

Gráfico de barras MSE

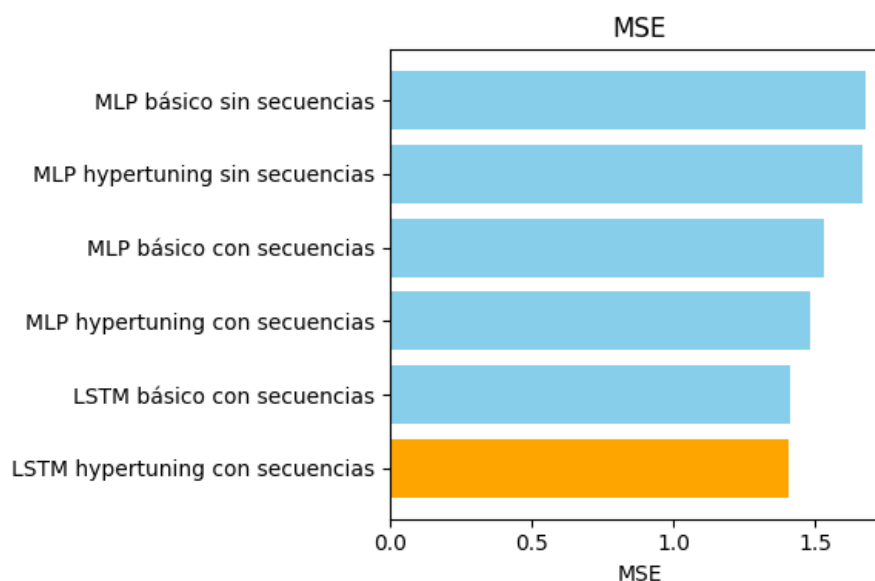


Figura 3.19: Comparativa de MSE entre los seis modelos.

Como podemos observar en la Figura 3.19, el valor más bajo corresponde al LSTM optimizado (1.4081), seguido muy de cerca por el LSTM básico (1.4121). Los modelos MLP con secuencias, optimizado y base, registran MSE de 1.482 y 1.530, mientras que los MLP sin secuencias, ambos, superan 1.66.

Gráfico de barras MAE

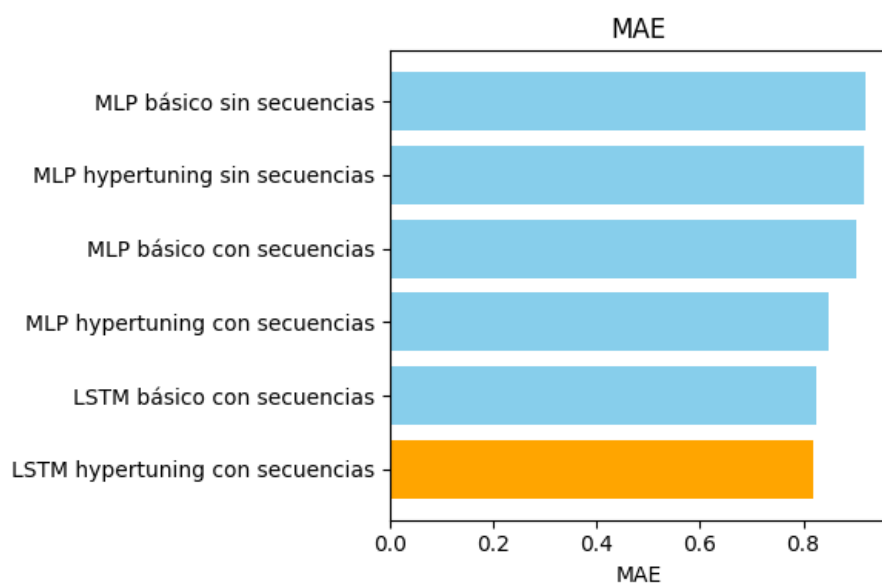


Figura 3.20: Comparativa de MAE entre los seis modelos.

Respecto a la Figura 3.20, el LSTM optimizado ofrece el MAE mínimo (0.818), con el LSTM básico muy próximo (0.824). El mejor MLP secuencial, el optimizado, alcanza 0.849, mientras que la variante base se sitúa en 0.902, casi al nivel de los MLP sin secuencias, que superan ligeramente 0.91.

Gráfico de barras R^2

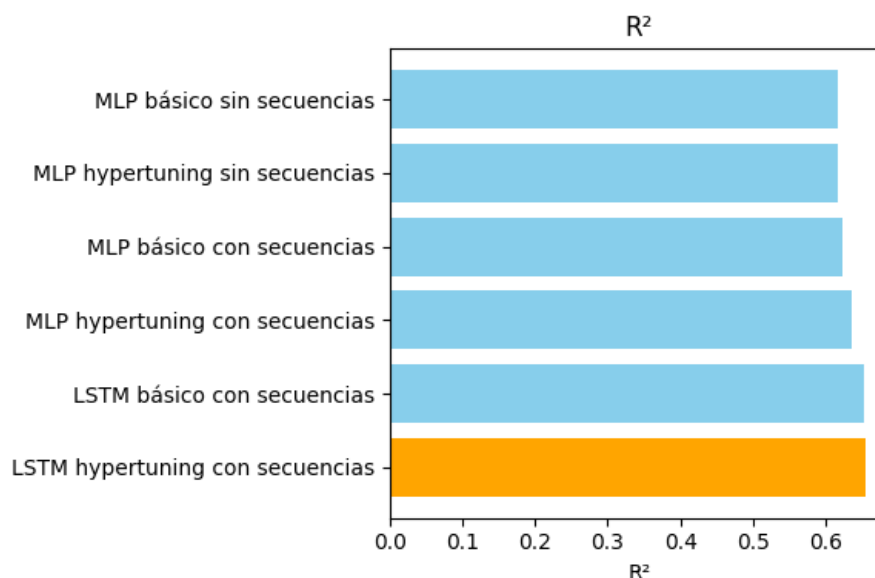


Figura 3.21: Comparativa de R^2 entre los seis modelos.

En cuanto a la Figura 3.21, el LSTM optimizado lidera con 0.6543, seguido del LSTM básico (0.6533). Los MLP con secuencias, optimizado y base respectivamente, alcanzan 0.636 y 0.624, mientras que los MLP independientes se establecen ligeramente por debajo de 0.62.

Selección del modelo final

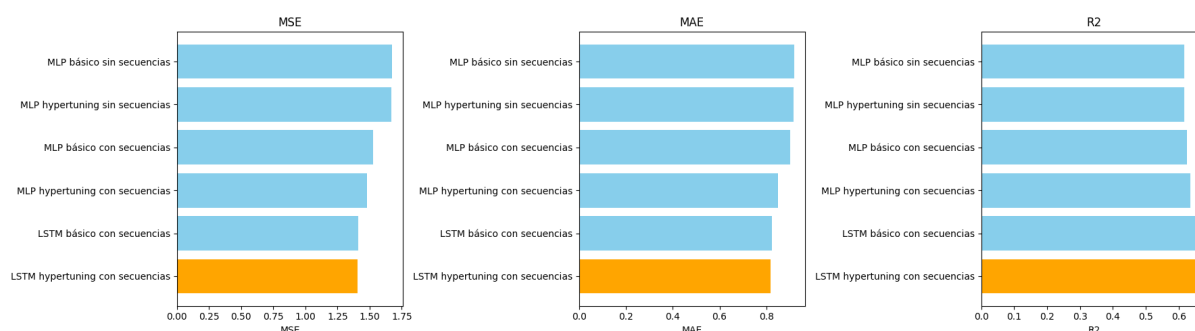


Figura 3.22: Comparativa general entre los seis modelos.

En las tres métricas principales, el LSTM optimizado con secuencias ocupa la primera posición, con márgenes consistentes sobre las demás alternativas. El entrenamiento requiere algo más de tiempo que los modelos MLP, pero la mejora absoluta de 0.03 en MAE, 0.08 en MSE y 0.02 en R^2 frente al mejor MLP, el secuencial, justifica el esfuerzo computacional adicional. Se selecciona, por tanto, el LSTM optimizado como la opción más precisa y robusta para un despliegue piloto en HappyForce.

Capítulo 4

Conclusiones y trabajo futuro

4.1. Conclusiones

A lo largo de este trabajo se han analizado los datos proporcionados por HappyForce mediante un enfoque secuencial y un enfoque con observaciones independientes, con el fin de obtener el modelo que prediga, con la mayor precisión posible, la respuesta individual (1–10) a la pregunta eNPS.

A continuación se detallan las principales conclusiones extraídas durante el trabajo:

1. **Valor de la información temporal.** Los cuatro modelos que integran ventanas de longitud 4 superan al enfoque sin secuencias en las tres métricas (MSE, MAE, R^2), lo que confirma que la dinámica histórica aporta un valor adicional para predecir la valoración eNPS individual.
2. **Rendimiento del modelo ganador.** El LSTM optimizado alcanzó las mejores cifras globales (MSE = 1.4081, MAE = 0.8180, $R^2 = 0.6543$), superando al mejor MLP secuencial en 0.03 MAE, 0.08 MSE y 0.02 R^2 . Esta mejora compensa el leve incremento de tiempo de entrenamiento observado en la práctica.
3. **Eficacia del *hypertuning*.** La búsqueda automática de hiperparámetros aportó mejoras notables en los dos modelos MLP y refinó ligeramente el LSTM, validando su utilidad sin incrementar la complejidad conceptual de las arquitecturas.
4. **Calidad de las predicciones.** El error medio absoluto del modelo ganador ($\pm 0,82$ puntos en la escala 1–10) se sitúa dentro de un rango aceptable para la práctica de RR.HH., permitiendo detectar variaciones tempranas en la satisfacción con un grado de precisión útil para la toma de decisiones.

4.2. Trabajo futuro

Como línea de trabajo futuro se propone integrar el modelo seleccionado en la plataforma HappyForce como una herramienta adicional de apoyo a RR.HH.; su despliegue incluiría un ciclo de reentrenamiento automático, mensual o trimestral, alimentado con los datos que la aplicación recoge a diario, de modo que el sistema mantenga su precisión a medida que cambian las dinámicas de los empleados y se incorpore nueva información sin intervención manual.

Bibliografía

- [1] HappyForce. enps: Qué es y cómo funciona, 2023.
- [2] HappyForce. Happyforce – mejora el compromiso de tus empleados, 2025.
- [3] Frederick F. Reichheld. The one number you need to grow. *Harvard Business Review*, 81(12):46–54, 2003.
- [4] Mark A. Huselid. The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of Management Journal*, 38(3):635–672, 1995.
- [5] Frederick F. Reichheld and Rob Markey. *The Ultimate Question 2.0*. Harvard Business Review Press, 2011.
- [6] Arjun Kumar and Priya Sharma. Employee net promoter score: Deep neural approach for organizational development. *International Journal of Organizational Analytics*, 2024.
- [7] Kirsty Graham. Upgrading the employee experience in a hybrid world. *People Management*, 2022.
- [8] Xiaowei Shi and Qiang Wei. Rnsc: A hierarchical deep learning model for net promoter scoring. *Knowledge-Based Systems*, 2024.
- [9] AIHR. People analytics: An essential guide for 2025, 2025.
- [10] EmployeeConnect. Enhanced use of people analytics: Transforming hr decision-making in 2025, 2025.
- [11] The Guardian. Time or energy? how managing people is changing in the workplace, 2025.
- [12] The Guardian. Work-life balance: why quality data and personal development tools are crucial for employee wellbeing, 2025.