

# **PEF 1: Comparación de Modelos para la Predicción de la Volatilidad**

## **Tarea 1.**

### **Miembros del equipo.**

1. Alberto De la Peña García
2. Edgar Alejandro Gonzalez Lira
3. Ana Paula Martinez Santos
4. José Pablo Tamez Ayala

### **Tema.**

Comparación de Modelos para la Predicción de la Volatilidad: Un Estudio Empírico Aplicado al Mercado Financiero.

### **Asesor.**

Dr. Martín Lozano Banda.

### **Objetivo General 1.**

Replicar y extender modelos de predicción de volatilidad propuestos en la literatura a través de la implementación de nuevas metodologías y variables explicativas con el fin de comparar e identificar el enfoque más robusto y preciso.

### **Objetivo Específico 1.1.**

Implementar y contrastar metodologías de Machine Learning adicionales.

### **Objetivo Específico 1.2.**

Ampliar el set de variables explicativas y contrastar su poder predictivo frente a los modelos anteriores.

### **Objetivo Específico 1.3.**

Escalar a una base de datos más amplia y reciente para evaluar el desempeño de los modelos bajo diferentes escenarios de mercado.

## **Objetivo General 2.**

Aplicar el mejor modelo de predicción de volatilidad en contextos financieros prácticos para evaluar su valor agregado frente a medidas tradicionales de volatilidad.

### **Objetivo Específico 2.1.**

Identificar y seleccionar diferentes contextos financieros en los que la predicción de volatilidad pueda aportar valor.

### **Objetivo Específico 2.2.**

Diseñar una metodología de implementación que permita contrastar el uso de la volatilidad predicha por los modelos frente a medidas tradicionales.

## **Tarea 2.**

### **A machine learning approach to volatility forecasting. Christensen et al. (2023)**

Este artículo evalúa de forma sistemática qué tan competitivos son diversos algoritmos de *machine learning* (ML) para pronosticar la varianza realizada (RV) frente a la familia de modelos HAR/HAR-X. Aporta una comparación entre métodos de regularización (ridge, lasso, elastic net), árboles (bagging, random forest, gradient boosting) y redes neuronales. También

establece evidencia de que el ML extrae información incremental a partir de covariables adicionales, así como un marco de interpretabilidad basado en *Accumulated Local Effects* (ALE) para entender importancia y relaciones no lineales entre predictores.

### Datos y variables.

Utilizan 29 acciones del índice Dow Jones Industrial Average (DJIA), con horizonte de un día adelante, en el periodo 2001-2017. La RV se construye con datos de alta frecuencia. Además de los rezagos diarios, semanales y mensuales de la RV (RVD, RVW, RVM), incorporan la volatilidad implícita (IV), un indicador de anuncio de resultados (EA), un índice de incertidumbre de política económica (EPU), VIX, la tasa T-bill a 1 mes (US1M), *momentum* a 1 semana (M1W), retorno del Hang Seng (HSI), volumen en dólares (DVOL), y el índice ADS de condiciones de negocio. Para organizar la información definen tres conjuntos de datos: M1 (sólo rezagos de RV), M2 (subconjunto “óptimo” vía BIC) y M3 (conjunto amplio de 12+ variables).

### Metodología.

Comparan HAR/HAR-X contra: ridge, lasso, elastic net; bagging, random forest, gradient boosting; y varias arquitecturas de redes neuronales “*off-the-shelf*” con ajuste mínimo de hiperparámetros. Evalúan *out-of-sample* con MSE relativo al HAR y emplean *Model Confidence Set* (MCS) al 90% para comparar robustamente. También estudian versiones en log-volatilidad (L-HAR, etc.).

### Resultados principales.

- Desempeño predictivo: Los métodos de ML superan consistentemente al HAR; entre ellos, random forest y redes neuronales suelen liderar.
- Información incremental: Frente a HAR-X, el ML extrae más señal de covariables adicionales y maneja mejor correlaciones altas, baja señal/ruido y no linealidades.
- Estados del mercado: ML mantiene ventajas tanto en episodios tranquilos como turbulentos; HAR-X tiende a deteriorarse en alta volatilidad. Con log-volatilidad, L-HAR-X mejora, pero L\_NN vuelve a dominar en regímenes medio-altos.

### Limitaciones y alcances.

El estudio se centra en acciones del DJIA y en horizontes de un día y persisten preguntas sobre la transferibilidad a otros activos/mercados y el *tuning* de hiperparámetros más agresivo.

Existen diferentes oportunidades de extensión para nuestro proyecto, entre ellas:

- Extender con nuevos métodos de *boosting* moderno (XGBoost/LightGBM), redes recurrentes/transformers para series, y regulación con búsqueda sistemática de hiperparámetros.
- Ampliar variables explicativas de análisis de sentimiento.
- Evaluar más horizontes de tiempo.

## **The impact of sentiment and attention measures on stock market volatility. Audrino et al. (2020)**

Este artículo estudia si medidas de sentimiento (tono en redes/noticias) y atención (volumen de mensajes/búsquedas) aportan poder predictivo adicional para la volatilidad realiazda.

### **Datos y variables.**

- Volatilidad realizada (índice Dow Jones y 18 acciones de EE.UU.) a partir de datos intradía de Oxford-Man Institute.
- Redes sociales: Twitter (muestra histórica 1% de todos los tuits, 2012-2016) y Stock-Twits, de donde se extraen sentimiento (promedios y dispersión diaria) y volumen de mensajes como proxy de atención. El sentimiento se suaviza con filtro de Kalman para reducir ruido.
- Noticias: RavenPack News Analytics (Event Sentiment Score, Composite Sentiment Score, News Impact Score), ponderados por relevancia del artículo para la empresa.
- Atención web: Google Trends (búsquedas) y Wikipedia page views, tanto específicos de empresa como de mercado (“stock market”, “financial markets”, etc.)
- Covariables económicas y financieras: factores Fama-French, ratios precio-fundamental, retornos de índices, VIX, tasas e indicadores de liquidez, entre otros.

### **Metodología.**

- Modelo base: HAR de Corsi para volatilidad realizada (captura persistencia mediante promedios diarios/semanales/mensuales).
- Extensión: Sentiment-HAR, que agrega sentimientos y medidas de atención junto con las variables económicas. Dado el alto número de predictores, emplean adaptive lasso para selección y estimación, con pruebas de significancia conjunta y post-selección conservadoras.

## Resultados principales.

Evidencia in-sample:

- Las variables de sentimiento/atención son conjuntamente significativas (F-tests) sobre la volatilidad futura en casi todos los activos (17/18 acciones y el índice).
- La atención domina al sentimiento como forma funcional: el volumen (mensajes/búsquedas) es el predictor más relevante; el VIX aparece sistemáticamente entre las variables seleccionadas.

Desempeño out-of-sample:

- Horizonte 1 día:

## Revisión Literaria. [ALT]

La predicción de la volatilidad financiera ha sido un campo de investigación en constante evolución desde la introducción de los modelos ARCH y GARCH (Engle, 1982; Bollerslev, 1986).

Aunque estos enfoques capturaron por primera vez la heterocedasticidad condicional en los retornos, pronto se reconocieron limitaciones importantes, como su incapacidad para reproducir memoria larga, efectos de salto y patrones no lineales.

En respuesta, Corsi (2009) propuso el modelo HAR, que combina componentes de volatilidad en distintas frecuencias (diaria, semanal y mensual) y se consolidó como un benchmark ampliamente utilizado en la literatura. Sin embargo, la creciente complejidad de los mercados y la disponibilidad de datos de alta frecuencia han impulsado el desarrollo de nuevas metodologías, particularmente aquellas basadas en aprendizaje automático (Machine Learning, ML).

En este sentido, Christensen et al. (2023) ofrecen un aporte fundamental al demostrar que algoritmos de ML —incluyendo regularización, métodos basados en árboles y redes neuronales— superan de manera sistemática al HAR en la predicción de volatilidad realizada.

Sus resultados, obtenidos con datos de alta frecuencia de acciones del Dow Jones Industrial Average, muestran reducciones de hasta un 12% en el error cuadrático medio (MSE) frente al benchmark tradicional, incluso en entornos enriquecidos con predictores adicionales como volatilidad implícita, volumen de negociación y factores macroeconómicos.

Otros estudios recientes refuerzan esta evidencia. Zhang et al. (2022), en *Journal of Financial Econometrics*, documentan que los modelos de deep learning aplicados al análisis de "intra-day commonality" —patrones compartidos de volatilidad a nivel intradía— ofrecen mejoras sustanciales respecto al HAR y a algoritmos de árboles.

Por su parte, Kumar et al. (2025), en *International Review of Economics & Finance*, proponen modelos híbridos que integran descomposición de la señal mediante VMD con redes ANN, LSTM y GRU, concluyendo que estas arquitecturas son más precisas y robustas, especialmente bajo condiciones de riesgo extremo. Estos aportes consolidan el consenso en torno a que el ML—en particular sus variantes profundas e híbridas— representa la frontera actual en predicción de volatilidad.

Desde una perspectiva de síntesis, Qiu et al. (2025) presentan un panorama comparativo de metodologías tradicionales y de ML, destacando cómo arquitecturas modernas, como Transformers, se perfilan como alternativas sólidas en este campo.

De manera complementaria, Mansilla-Lopez et al. (2025) desarrollan una revisión sistemática que clasifica las contribuciones del ML a lo largo de dos décadas, identificando vacíos persistentes en la integración de variables de sentimiento y en la aplicación de procedimientos estadísticos de comparación robusta como el Model Confidence Set (MCS). Si bien estos trabajos no aportan resultados empíricos directos, su utilidad radica en proveer un marco conceptual que orienta la selección metodológica y posiciona la investigación en el estado del arte.

En línea con el Objetivo General 2, la literatura también ha explorado aplicaciones prácticas de la volatilidad predicha en contextos financieros. En el ámbito de derivados, Poteshman (2000) demuestra que la incorporación de predicciones más precisas de volatilidad mejora la calibración de superficies de volatilidad implícita, reduciendo sesgos en la valoración de opciones. Más recientemente, van Lent (2023) examina cómo los modelos de ML pueden predecir características de la superficie de volatilidad implícita (nivel, pendiente, curvatura), ofreciendo un marco metodológico útil para contrastar la información de volatilidad realizada y su valor agregado frente a métricas tradicionales de opciones.

Por otra parte, la incorporación de información proveniente de noticias y análisis de sentimiento ha emergido como un área clave.

(Brauneis y Sahiner, 2024), en *Asia-Pacific Financial Markets*, muestran que en criptomonedas los modelos HAR y de ML enriquecidos con indicadores de sentimiento logran capturar dinámicas no lineales que los enfoques tradicionales ignoran. De forma más amplia, Audrino et al. (2020) y estudios recientes en *Journal of Applied Econometrics* evidencian que métricas de atención y análisis textual incrementan significativamente la capacidad predictiva de la volatilidad, lo que se traduce en aplicaciones concretas para la gestión de riesgos, la asignación de capital y el trading algorítmico.

En conjunto, la literatura converge en dos hallazgos centrales.

- 1) Primero, que los métodos de ML superan de manera consistente a los modelos tradicionales como HAR y GARCH en precisión y robustez predictiva, validando el Objetivo General 1 de esta tesis.
- 2) Segundo, que la predicción de volatilidad, especialmente al integrar información proveniente de volatilidad implícita, noticias o sentimiento, aporta un valor agregado significativo en aplicaciones financieras prácticas, justificando el Objetivo General 2.

No obstante, persisten vacíos relevantes: la falta de consenso sobre métricas de evaluación más allá del MSE, la necesidad de explorar de forma más sistemática la interacción entre volatilidad realizada e implícita en contextos de ML, y la escasez de replicaciones en bases de datos más recientes y amplias. Estos vacíos constituyen precisamente el espacio de contribución original que busca llenar la presente investigación.

## Bibliografía

- Audrino, F., Sigrist, F., & Ballinari, D. (2020). The impact of sentiment and attention measures on stock market volatility. *International Journal of Forecasting*, 36(2), 334–357.
- Christensen, K., Siggaard, M., & Veliyev, B. (2023). A machine learning approach to volatility forecasting. *Journal of Financial Econometrics*, 21(5), 1680–1727.
- Kumar, S., Rao, A., & Dhochak, M. (2025). Hybrid ML models for volatility prediction in financial risk management. *International Review of Economics & Finance*, 98, 103915.
- Mansilla-Lopez, J., Mauricio, D., & Narváez, A. (2025). Factors, forecasts, and simulations of volatility in the stock market using machine learning. *Journal of Risk and Financial Management*, 18(5), 227.
- Poteshman, A. M. (2000). Forecasting future volatility from option prices. *Available at SSRN 243151*.
- Qiu, Z., Kownatzki, C., Scalzo, F., & Cha, E. S. (2025). Historical perspectives in volatility forecasting methods with machine learning. *Risks*, 13(5), 98.
- Zhang, C., Zhang, Y., Cucuringu, M., & Qian, Z. (2022). Volatility forecasting with machine learning and intraday commonality. arXiv. *arXiv Preprint arXiv:2202.08962*.