

# Propuesta TFG: Análisis de Series Temporales y Técnicas de Extracción de Tópicos para la Gestión Dinámica de Carteras

---

## Abstracto

Este trabajo tiene como objetivo mejorar las estrategias de gestión de carteras mediante la integración de modelos predictivos de series temporales con técnicas avanzadas de análisis de texto. Al analizar datos estructurados (precios de acciones) y no estructurados (noticias financieras, redes sociales), se busca generar una estrategia de inversión más efectiva y dinámica. Se emplearán modelos como ARIMA, LSTM y Prophet para las predicciones financieras, y técnicas de análisis de texto como el análisis de sentimiento, la extracción de tópicos y la detección de eventos para identificar las dinámicas del mercado. La optimización de carteras será el enfoque central, integrando tanto las previsiones del mercado como los indicadores derivados de los datos no estructurados.

---

## 1. Introducción

### 1.1. Objetivo/s

El objetivo principal de este trabajo es desarrollar una estrategia de gestión de carteras que integre la predicción de series temporales con análisis de datos no estructurados, como noticias financieras y redes sociales. Se pretende mejorar la toma de decisiones en inversiones mediante el uso combinado de estas técnicas.

### 1.2. Justificación del tema / objeto del estudio

El crecimiento del volumen de datos no estructurados, como noticias y redes sociales, tiene un impacto directo en los mercados financieros. La combinación de estos datos con modelos de series temporales permite mejorar la precisión en la gestión de carteras, lo cual puede generar un valor significativo para los inversores al incorporar factores de riesgo y tendencia que los modelos tradicionales no capturan. Dada la evolución de las

tecnologías de análisis de datos, esta propuesta busca explorar un enfoque multidimensional para optimizar decisiones financieras.

### 1.3. Metodología general del TFG

El trabajo combinará metodologías cuantitativas para el análisis de series temporales y el procesamiento de lenguaje natural (NLP) para el análisis de texto. Se recopilarán datos de precios de acciones y noticias financieras entre 2016 y 2023, con los cuales se entrenarán modelos de predicción y se realizarán análisis de sentimientos y extracción de tópicos. Los resultados obtenidos se integrarán en un modelo de optimización de carteras, el cual será evaluado mediante backtesting y métricas de rendimiento, como el ratio de Sharpe.

### 1.4. Estructura

El TFG se estructurará de la siguiente manera:

Capítulo 1: Introducción, que aborda los objetivos, la justificación y la metodología general.

Capítulo 2: Revisión de la literatura, con un marco teórico sobre predicción de series temporales, análisis de texto, y gestión de carteras.

Capítulo 3: Análisis empírico, donde se detallan los datos, la metodología específica utilizada y los resultados obtenidos.

Capítulo 4: Conclusiones y futuras líneas de investigación.

---

## 2. Revisión de la literatura / marco teórico

En esta sección se explorará la literatura relacionada con tres áreas clave:

Modelos de series temporales en finanzas: Incluyendo técnicas tradicionales (ARIMA, GARCH) y enfoques de machine learning (LSTM, Prophet).

Análisis de texto en el contexto financiero: Revisión de técnicas como el análisis de sentimiento (VADER, BERT), la extracción de tópicos (LDA, NMF) y la detección de eventos.

Gestión de carteras: Se revisarán teorías como la Teoría Moderna de Portafolios y su evolución hacia estrategias de inversión basadas en datos.

---

### 3. Análisis empírico

#### 3.1. Datos

Datos estructurados: Se utilizarán precios de acciones y otros indicadores financieros obtenidos de fuentes como Yahoo Finance y Bloomberg, abarcando el periodo 2016-2023.

Datos no estructurados: Noticias financieras, comentarios en redes sociales (Twitter, Reddit) y reportes de resultados corporativos extraídos mediante técnicas de web scraping y APIs públicas.

#### 3.2. Metodología

Modelos de series temporales: Se emplearán ARIMA, GARCH y LSTM para la predicción de precios futuros.

Análisis de texto: Se utilizarán técnicas de NLP para analizar los datos no estructurados. Esto incluye:

Análisis de Sentimiento: Herramientas como VADER o BERT.

Extracción de Tópicos: Modelos como LDA para identificar temas recurrentes en los textos.

Detección de Eventos: Identificación de eventos importantes y su correlación con las fluctuaciones del mercado.

Optimización de carteras: Se aplicarán los resultados de las predicciones de series temporales y el análisis de texto a un modelo de gestión de carteras, utilizando métricas de riesgo-rendimiento como la relación riesgo/beneficio y el VaR.

### 3.3. Resultados

Presentación de los resultados de los modelos de predicción de series temporales.

Análisis de las tendencias derivadas del análisis de texto, correlacionadas con los movimientos del mercado.

Resultados de la gestión de carteras optimizada comparados con modelos tradicionales, incluyendo el backtesting y las métricas de rendimiento.

---

## 4. Conclusiones

Resumen de los hallazgos clave sobre cómo la integración de modelos de predicción y análisis de texto mejora la gestión de carteras.

Discusión de las limitaciones encontradas durante la investigación, como la calidad de los datos no estructurados o las dificultades en la interpretación de resultados de NLP.

Propuestas para futuras investigaciones, incluyendo la exploración de nuevas técnicas de análisis de datos o la aplicación de los modelos a otros tipos de activos financieros.

---

## Bibliografía

Liu, B. (2015). *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press.

Manning, C., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 7(1), 77–91.

Patterson, C. S. (2010). *The Risk of Active Portfolio Management*. The Research Foundation of CFA Institute.

Se incluirán las referencias utilizadas, abarcando estudios académicos sobre modelos de series temporales, análisis de texto en finanzas, y gestión de carteras, además de fuentes de datos financieros y herramientas tecnológicas empleadas en el desarrollo del trabajo.

Esta propuesta detalla un plan integral para abordar el uso de técnicas avanzadas de análisis de datos en la gestión de carteras, manteniendo la conexión con el análisis de series temporales y el procesamiento de datos no estructurados.