



# Trading de Earnings con Machine Learning y Opciones Financieras

Presentado por: Mario Velandia Rodriguez

Director de tesis: Giovanni Andrés Hernández Salazar

Tesis presentada como requisito para optar al título de: Magíster en Gerencia de Inversión

Universidad de la Sabana, Bogotá, Colombia. 2024.

## 1. RESUMEN

El Post-Earnings-Announcement-Drift (PEAD), según Fama (1998, p.286) es considerado el “*granddaddy of underreaction events*”, lo que refleja su reconocimiento del PEAD como una anomalía dentro de la hipótesis de mercados eficientes. A pesar de ser uno de los temas más investigados en la historia de la contabilidad (Lev, 1989, p. 153), aún no se ha encontrado una explicación que tenga una aceptación generalizada (Fink, 2021). Esto ha dado paso al desarrollo de modelos basados en inteligencia artificial (IA) los cuales han creado portafolios de PEAD que generan un retorno en exceso de hasta un 73% en un año (Meursault et al., 2023). Este documento se centró en crear un modelo de aprendizaje automático con XGBoost que obtuvo una tasa de precisión del 69% en la predicción de movimientos sobreestimados por los contratos de opciones cercanos a los *Earnings Announcements* de empresas del índice Russell 1000.

**Palabras claves:** Post Earnings Announcements Drift, PEAD, Inteligencia artificial aplicada a inversiones, Extreme Gradient Boosting XGBoost, Trading de earnings con opciones financieras

## 2. INTRODUCCIÓN

El Post-Earnings-Announcement-Drift (PEAD) contradice la hipótesis de mercados eficientes formulada por Fama (1991) en su sentido semifuerte. El PEAD describe la relación entre los *Earnings Announcements* inesperados con los retornos en exceso durante los meses subsiguientes (Ball & Brown, 1968; Bernard & Thomas, 1989) y muestra que esta información sorpresa no se asimila inmediatamente por el mercado. Portafolios compuestos por empresas con las mayores sorpresas en sus *Earnings Announcements* podían generar un retorno anualizado de entre 18% y 25% adicional al retorno del mercado (Bernard & Thomas, 1989; Foster et al., 1984). El PEAD sucede en todos los mercados financieros, tanto desarrollados como menos desarrollados manteniéndose vigente a la fecha (Griffin et al., 2010; Meursault et al., 2023)

Se han evaluado múltiples metodologías para explicar y pronosticar el PEAD. Alternativas basadas en variaciones en el cálculo de betas del *Capital-Asset-Pricing-Model* (CAPM), variaciones al modelo *Arbitrage-Pricing-Theory* (APT), factores en el comportamiento humano o mejoras en el pronóstico de utilidades no explica los retornos anormales del PEAD (Bernard & Thomas, 1989; Brav & Heaton, 2002; Meursault et al., 2023). La evidencia empírica muestra que el número de *Earnings Surprises* se ha mantenido a pesar de los avances en los modelos predictivos y el acceso a información actual (Shon & Zhou, 2011). Estas metodologías se basan en metodologías clásicas de estadística como los modelos de autocorrelación y modelos de series de tiempo.

La falta de efectividad de las metodologías clásicas para modelar y estimar los *Earnings Surprises* han dado paso a los modelos basados en inteligencia artificial (IA). Portafolios de PEAD que utilizan IA generan más retorno en exceso (73% en 1 año) que los portafolios de PEAD que solo usan

modelos de autorregresión y modelos de *time-series* (Meursault et al., 2023). Adicionalmente, se encontró que el modelo de aprendizaje automático, *Extreme Gradient Boosting* (XGBoost) tiene mejores resultados en predecir la dirección del *Post Earnings Announcements Drift* que modelos de *Deep Neural Networks* (DNN) y de *Support Vector Machine* (SVM) (Ye & Schuller, 2021).

Por definición, el PEAD indica que si la sorpresa fue positiva la dirección del *Drift* para esa acción sería al alza y si la sorpresa fue negativa la dirección del *Drift* sería a la baja. Sin embargo, esto se cumple cerca de un 60% de las veces (Shon & Zhou, 2011), por lo que este documento no busca pronosticar la dirección del PEAD sino que el objetivo de este documento es desarrollar un modelo de *Extreme Gradient Boosting* (XGBoost) que pronostique si el movimiento esperado por el mercado será mayor o será menor que el movimiento que realizará la acción después de su *Earnings Announcement*. Monetizando el efecto del PEAD en el mercado de opciones con estrategias de rango para generar Alfa (retorno en exceso) consistentemente en el tiempo, siempre y cuando exista la anomalía del PEAD.

El trading con opciones se justifica porque no solo ha venido creciendo en el tiempo alrededor de los *Earnings Announcements* (EA) (Roll et al., 2009) sino también porque en los EA la volatilidad se incrementa (Landsman & Maydew, 2002), potencializando la posibilidad de obtener ganancias con la venta de opciones por la caída en la volatilidad (vega negativo), el paso del tiempo (theta positivo) y sin necesidad de asumir necesariamente un riesgo direccional (delta neutral).

### 3. METODOLOGÍA

#### 3.1. Datos y Descripción de la Información

Se utilizaron tres proveedores de datos: Financial Modeling Prep (FMP) para el histórico de earnings y estados financieros, ORATS para el histórico de contratos de opciones (precios y griegas) y Yahoo Finance para el histórico de precios. Se analizó las 1009 empresas incluidas en el índice RUSSELL 1000 al 1 Noviembre del 2024, 108,549 *Earnings Announcements (EA)* abarcando la historia completa de cada empresa y 15,938 contratos de opciones de 2017-01-26 a 2024-09-11 para 680 de ellas. Depurando por *EA* que tuvieran *EPS* y *REVENUE* estimados por analistas y su histórico completo de fundamentales, se obtuvieron 73,874 registros de EA. Estos datan desde 1993-01-25 a 2024-11-26.

Los datos recolectados se clasificaron en cuatro categorías:

Categoría	Datos
Earning Announcement	<ul style="list-style-type: none"> <li>- Fecha de publicación</li> <li>- Tipo de anuncio: AMC o BMO</li> <li>- EPS y Revenue estimados</li> <li>- EPS y Revenue publicados</li> </ul>
Fundamentales de la Empresa	<ul style="list-style-type: none"> <li>- Operating Income*</li> <li>- Net Income*</li> <li>- Free Cash Flow*</li> <li>- Total Liabilities*</li> <li>- Gross Profit*</li> <li>- Return On Assets*</li> <li>- Price to Earnings Ratio*</li> <li>- Price to Sales Ratios*</li> <li>- Price to Book Ratio*</li> </ul>
Contrato de Opciones Mensual posterior a fecha del EA.	<ul style="list-style-type: none"> <li>- Fecha de Vencimiento</li> <li>- Precio Ask y Bid de Calls y Puts</li> <li>- Precio Spot</li> </ul>
Movimiento del Subyacente	<ul style="list-style-type: none"> <li>- Movimiento Esperado Straddle ATM en T0**</li> <li>- Movimiento Realizado entre T0 y T1**</li> </ul>
* Todos los datos se tomaron según el valor trimestral reportado en el anuncio	
** Fechas T0 y T1 definidas en el numeral 3.2.1 de este documento.	

### 3.2. Construcción de las Variables Fundamentales

A continuación, se describen las variables construidas para el modelo empleado. Calculadas a partir de los datos recolectados y descritos en el punto anterior.

#### 3.2.1. Fechas T0, T1 y T-1 según tipo de Earnings Announcement

Las fechas descritas a continuación se utilizaron en la construcción de las variables del modelo siendo su definición dependiente del tipo de *Earnings Announcement*.

Fecha	AMC (After Market Closes)	BMO (Before Market Opens)
T-1	Día de publicación del anuncio	Día anterior al anuncio.
T0	Día posterior al anuncio	Día de publicación del anuncio
T1	Último día de negociación del contrato con vencimiento posterior al Earning Announcement. Mayor 2 DTE.	

#### 3.2.2. EPS Surprise Normalizado por Precio de la Acción

Para el cálculo del *EPS Surprise* se realizó la resta aritmética entre el EPS reportado menos el EPS esperado por los analistas (Zhou & Shon, 2012). Debido a que la magnitud de este dato es específico para cada acción y por ende no tiene sentido compararlos entre sí, se procedió a normalizar el dato dividiéndolo por el precio de su respectiva acción. De esta forma el dato resultante es la proporción entre sorpresa en los EPS obtenida respecto al precio de cierre que cotizaba en la fecha T-1.

#### 3.2.3. Revenue Surprise Normalizado por Revenue Estimado

Para el cálculo del *Revenue Surprise* se realizó la resta entre el *Revenue* reportado menos el *Revenue* esperado por los analistas. Se procedió a normalizar el dato dividiéndolo por el *Revenue* estimado por los analistas (Zhou & Shon, 2012). De esta forma el dato resultante es una relación entre la sorpresa de *Revenue* y la estimación realizada por los analistas.

### **3.2.4. *Relación entre EPS y Revenue Surprises y el promedio de las sorpresas anteriores***

Estas dos variables se calculan dividiendo el EPS Surprise y Revenue Surprise del EA entre el promedio de los EPS Surprise históricos y el promedio de los Revenue Surprises históricos respectivamente menos uno (Ye & Schuller, 2021).

### **3.2.5. *Tasa de cambio trimestral de fundamentales***

Los fundamentales escogidos se tomaron de los 5 factores claves del modelo de *Extreme Gradient Boost* (XGBoost) usado por Ye & Schuller (2021) en su investigación de pronóstico de PEAD para empresas de Estados Unidos.

La tasa de cambio trimestral fue calculada dividiendo el valor reportado por el valor del trimestre anterior menos uno para hallar la tasa de cambio de un trimestre a otro para cada fundamental. Se realizó el cálculo sólo para las empresas que se tuviera sus fundamentales de forma consecutiva sin faltar datos desde 1992-04-21 (ó desde el inicio de los reportes de la empresa posterior a este día) hasta el último reporte previo a Noviembre 2024.

### **3.2.6. *Tasa de cambio anual de fundamentales***

La tasa de cambio anual fue calculada restando al valor reportado el valor del mismo fundamental del año anterior, dividido por el valor del año anterior menos 1 para hallar la tasa de cambio de un año a otro para cada fundamental. Esto con el fin de evaluar posible estacionalidad en los reportes de resultados cuando se comparan trimestre a trimestre (Ye & Schuller, 2021)

### **3.2.7. *Call Bid-Ask Spread normalizado por precio Spot***

Para identificar las acciones con mayor liquidez en el mercado de opciones se restó al precio ASK del contrato Call ATM su respectivo precio BID y se dividió entre el precio Spot del subyacente para normalizar la serie. Debido a que la liquidez de los contratos Call es mayor en mercados alcistas que la liquidez de los contratos PUT (Cao & Wei, 2008) se tomó el spread de los contratos Call ya que

para el rango de fechas bajo estudio (Enero 2017 a Noviembre 2024) el mercado estadounidense (más específicamente las 680 empresas analizadas) fue mayoritariamente alcista.

### **3.2.8. *Movimiento Esperado por el Mercado: Straddle ATM en T0***

Se toma como movimiento esperado por el mercado la suma del valor de las primas en Mid-Price del contrato Call y Put ATM con strike más cercano al precio spot de cierre de T0 y con fecha de expiración en T1. Normalizado por el precio de cierre en T0. El Mid-Price se obtiene sumando el precio en ASK y BID y dividido el resultado en dos.

### **3.2.9. *Movimiento Realizado entre T0 y T1 en Valor Absoluto***

Se obtiene cómo el valor absoluto de la resta entre el precio de cierre en T1 menos el precio de cierre en T0. Para su normalización se usó el precio de cierre en T0. Se tomó el valor absoluto para hacer viable su comparación con el movimiento esperado por el mercado.

### **3.2.10. *Booleano de Movimiento esperado por el Mercado versus Movimiento Realizado***

Para los *Earnings Announcements* cuyo movimiento esperado fue menor que el movimiento realizado (movimiento subestimado) la variable toma valor de 0 y cuando es igual o mayor (movimiento sobreestimado) la variable toma valor de 1. Esta variable actuará como *Label* o variable dependiente del modelo XGBoost.

## **3.3. Estadísticas Descriptivas**

A continuación, se muestra la descripción estadística de las variables del modelo.

### **3.3.1. *EPS y Revenue Surprises normalizados por precio de la acción***

La tabla 1 presenta los *EPS* y *Revenue Surprises* calculados. Los valores de *EPS Surprise* se muestran como la relación entre el *EPS Surprise* y el precio de cierre de la acción en T-1. Debido a que algunos precios son más cercanos a cero se tiene *outliers* debido al efecto del denominador. Para contrarrestar

dicho efecto se eliminaron los outliers recortando la cola izquierda por su percentil 0.01 y la cola derecha por el percentil 0.99. Obteniendo así los resultados mostrados a continuación:

**Tabla 1** *Estadística descriptiva EPS y Revenue surprises*

<b>Estadística</b>	<b>EPS Surprise</b>	<b>Revenue Surprise</b>
Registros	81,688	73,243
Media	0.000918232	0.115352
Desviación Est.	0.00355321	0.496235
Mínimo	-0.0221695	-0.777778
Percentil 10%	-0.00132076	-0.121212
Percentil 25%	0	-0.00203952
Percentil 50% (Mediana)	0.000505291	0.0335107
Percentil 75%	0.00168067	0.127636
Percentil 90%	0.0039899	0.324163
Máximo	0.0217096	6.16667

*Nota Tabla 1. La sorpresa del EPS es el resultado de restar el EPS reportado por la empresa menos el EPS estimado por los analistas dividido el precio de cierre que cotizaba en la fecha T-1. La sorpresa del Revenue es el resultado de restar el Revenue reportado por la empresa menos el Revenue estimado por los analistas y dividirlo por el valor estimado por los analistas. Se quitó los outliers menores al percentil 1 o mayores al percentil 99 con el fin de contrarrestar el efecto denominador.*

### **3.3.2. Relación entre EPS y Revenue Surprises y el promedio de sorpresas anteriores**

La tabla 2 muestra el cambio que tiene una sorpresa con respecto al promedio de las sorpresas anteriores de la misma empresa, eliminando outliers inferiores al percentil 0.01 o mayores al percentil 0.99. Para los EPS oscila entre -282% y el 378% para el 80% de los casos con una mediana de -27%. Mientras que para los *Revenue Surprise* este rango va de entre -25% a 20.5% para el 80% de los casos con mediana de -1.0%.

**Tabla 2** Estadística descriptiva EPS y Revenue surprises respecto al promedio de las sorpresas anteriores de cada acción.

<b>Estadística</b>	<b>EPS Surprise Backward Ave Diff</b>	<b>Revenue Surprise Backward Ave Diff</b>
Registros	81,611	73,243
Media	-0.567063	0.0199151
Desviación Est.	327.475	0.466162
Mínimo	-59153.1	-3.07876
Percentil 10%	-2.82907	-0.254712
Percentil 25%	-1	-0.0911097
Percentil 50% (Mediana)	-0.271555	-0.0101118
Percentil 75%	1.12571	0.0555496
Percentil 90%	3.78633	0.205288
Máximo	43484.1	6.01927

*Nota Tabla 2.* El dato se calcula tomando la diferencia entre EPS y Revenue Surprise y el promedio de los EPS y Revenue Surprise anteriores de la misma empresa. Se quitó los outliers menores al percentil 1 o mayores al percentil 99 con el fin de contrarrestar el efecto denominador.

### 3.3.3. Tasa de cambio trimestral de fundamentales

La Tabla 3 muestra las tasas de cambio trimestrales de los fundamentales escogidos para la investigación. Los fundamentales con mayor rango de variación de un trimestre a otro fueron el Free Cash Flow, el PE Ratio y el Net Income. Los de menor rango de variación fueron Total Liabilities, PS Ratio y PB Ratio.

**Tabla 3** Estadística descriptiva tasa de cambio trimestral de fundamentales

<b>Estadística</b>	<b>Operating Income</b>	<b>Net Income</b>	<b>Free Cash Flow</b>
Registros	102.095	98.118	95,973
Percentil 10%	-1.0528	-1.17132	-2.40581
Percentil 25%	-0.294962	-0.390162	-1.07161
Percentil 50% (Mediana)	0	-0.0073224	-0.565621
Percentil 75%	0.21549	0.273314	0.181818
Percentil 90%	0.768838	1.03635	1.54136
<b>Estadística</b>	<b>Total Liabilities</b>	<b>Gross Profit</b>	<b>Return On Assets</b>
Registros	97.002	101.591	95.159

Percentil 10%	-0.0476765	-0.234016	-1.15829
Percentil 25%	-0.00805492	-0.0607687	-0.417896
Percentil 50% (Mediana)	0	0.023744	-0.0403535
Percentil 75%	0.0318307	0.115183	0.234755
Percentil 90%	0.108549	0.304622	0.967599
<b>Estadística</b>	<b>PE Ratio</b>	<b>PS Ratio</b>	<b>PB Ratio</b>
Registros	97.702	101.519	96153
Percentil 10%	-1.15438	-0.256057	-0.242574
Percentil 25%	-0.428609	-0.117435	-0.0984384
Percentil 50% (Mediana)	-0.0618168	0.0038464	0.0063742
Percentil 75%	0.263894	0.139341	0.113203
Percentil 90%	1.07042	0.329065	0.25408

*Nota Tabla 3 Los datos se muestran restando al dato del trimestre su par correspondiente al trimestre anterior y dividiéndolo por el dato del trimestre anterior para obtener la tasa de cambio porcentual.*

### 3.3.4. Tasa de cambio anual de fundamentales

La Tabla 4 muestra las tasas de cambio anuales de los fundamentales escogidos para la investigación. Los fundamentales con mayor rango de variación de un año a otro fueron el Free Cash Flow, el PE Ratio y el Net Income. Los de menor rango de variación fueron Total Liabilities, PS Ratio y PB Ratio. **Tabla 4** *Estadística descriptiva tasa de cambio anual de fundamentales*

<b>Estadística</b>	<b>Operating Income</b>	<b>Net Income</b>	<b>Free Cash Flow</b>
Registros	99.103	95.126	93.585
Percentil 10%	-0.93668	-1.28013	-2.85467
Percentil 25%	-0.212931	-0.353846	-1.16327
Percentil 50% (Mediana)	0.0717452	0.0682688	-0.573584
Percentil 75%	0.314537	0.37739	0.204188
Percentil 90%	0.922894	1.2098	1.80441
<b>Estadística</b>	<b>Total Liabilities</b>	<b>Gross Profit</b>	<b>Return On Assets</b>
Registros	94.647	98.617	92.168
Percentil 10%	-0.0976825	-0.227577	-1.24626
Percentil 25%	-0.0228338	-0.0235147	-0.430769
Percentil 50% (Mediana)	0.0438413	0.0803474	-0.0398931
Percentil 75%	0.155817	0.217433	0.224762
Percentil 90%	0.406653	0.506452	0.930682
<b>Estadística</b>	<b>PE Ratio</b>	<b>PS Ratio</b>	<b>PB Ratio</b>
Registros	94.711	98.538	93.162

Percentil 10%	-1.23411	-0.35901	-0.401785
Percentil 25%	-0.429844	-0.164273	-0.182551
Percentil 50% (Mediana)	-0.0563529	0.0251103	0.0169064
Percentil 75%	0.273545	0.238969	0.232494
Percentil 90%	1.01643	0.540217	0.552659

*Nota Tabla 4 Los datos se muestran restando al dato del año su par correspondiente al año anterior y dividiéndolo por el dato del año anterior para obtener la tasa de cambio porcentual.*

### 3.3.5. Call Bid-Ask Spread normalizado por precio Spot

La Tabla 5 muestra los spreads encontrados para los contratos siguientes al *Earning Announcement* (en promedio 16 DTE). Se calculó para contratos Call y Put con strike más cercano al precio Spot del subyacente. La normalización se realizó por precio Spot.

**Tabla 5** Estadística descriptiva Call Bid-Ask Spread normalizado por precio Spot

Estadística	Bid-Ask Spread Call ATM	Bid-Ask Spread Put ATM
Registros	15,925	15,928
Promedio	0.00631217	0.00663777
Desv. Est.	0.00996644	0.0110706
Mínimo	9.65065e-05	5.36107e-05
Percentil 10%	0.000969237	0.000978282
Percentil 25%	0.00182482	0.00183981
Percentil 50% (Mediana)	0.00343926	0.00351406
Percentil 75%	0.00668896	0.00691772
Percentil 90%	0.0132101	0.01426
Máximo	0.228102	0.378984

*Nota Tabla 5. El dato se calcula tomando la resta entre precio Ask menos precio Bid para contratos Call y contratos Put con strike ATM y DTE más cercano al Earning Announcement, dividido por el precio spot.*

### 3.3.6. Movimiento Esperado, Movimiento Realizado y Diferencia

El movimiento esperado tiene menor desviación estándar (volatilidad) que la volatilidad de los movimientos realizados. Por otro lado, en promedio los movimientos esperados son mayores que lo

realizados cómo muestra la columna de diferencia entre ambos y el percentil 10, 25, 50 y cerca al 75 del movimiento esperado que son mayores movimientos mayores que los realizados (ver Tabla 6).

**Tabla 6** Estadística descriptiva Movimiento Esperado por el Mercado: Straddle ATM en T0 y Movimiento Realizado entre T0 y T1 en Valor Absoluto. Normalizados por Precio Cierre en T0

Estadística	Movimiento Esperado	Movimiento Realizado. Valor Abs	Diferencia Esp - Realizado
Registros	15,938	15,938	15,938
Promedio	0.0500514	0.0457774	0.004274
Desv. Est.	0.0233905	0.0494993	0.0451061
Mínimo	0.0062196	0	-0.817307
Percentil 10%	0.0267742	0.00573733	-0.0401953
Percentil 25%	0.0349548	0.0147367	-0.0123066
Percentil 50% (Mediana)	0.0458424	0.0330001	0.010079
Percentil 75%	0.0602512	0.0604296	0.0286705
Percentil 90%	0.0781359	0.0974369	0.0456292
Máximo	0.427193	1.08457	0.38517

*Nota Tabla 6.* Para el movimiento esperado se toma la suma del valor de un straddle ATM tomando el Mid-Price. Para el movimiento realizado se toma el valor absoluto del precio de cierre en T1 menos el precio de cierre en T0. Ambos se dividen por el precio de cierre en T0 hallar el movimiento porcentual.

### 3.3.7. Booleano de Movimiento esperado por el Mercado versus Movimiento Realizado

Como se muestra en la Tabla 7, el 63.25% de los 15,938 earnings analizados los movimientos estimados por el mercado fueron mayores que el movimiento que realmente tuvo la acción entre T0 y T1.

**Tabla 7** Booleano de Movimiento esperado por el Mercado versus Movimiento Realizado

Tipo de Movimiento	Registros	Peso
Subestimado (0)	5,858	36.75%
Sobreestimado (1)	10,080	63.25%

*Nota Tabla 7.* Si el movimiento esperado es menor al realizado se cuenta como subestimado y si es mayor o igual se cuenta como sobreestimado.

### 3.4. Modelo Extreme Gradient Boosting Classifier (XGBoostClassifier)

Extreme Gradient Boosting (XGBoost) es una librería de aprendizaje automático basada en la técnica de *boosting* de árboles de decisión. Dentro de ella cuenta con un módulo denominado

XGBoostClassifier el cuál es óptimo para la clasificación binaria, pertinente para la resolución del problema planteado por esta investigación.

**3.4.1. Descripción el Problema de Clasificación Binaria (Label)**

El modelo de aprendizaje automático de este estudio busca predecir si el movimiento esperado por el mercado el día del *Earnings Announcement* será mayor o menor al movimiento realizado por el subyacente antes del vencimiento de su próximo contrato regular de opciones. Esto constituye un problema de clasificación binaria en donde la variable dependiente del modelo ó *Label* se define cómo:

- 0: Movimiento subestimado (indica que un trade vega positivo será rentable)
- 1: Movimiento sobreestimado (indica que un trade vega negativo será rentable)

**3.4.2. Descripción de las Características (Features)**

La siguiente tabla presenta las características utilizadas en el modelo agrupadas en las mismas categorías que los datos recolectados:

<b>Categoría</b>	<b>Características (Features)</b>
Earning Announcement	<ul style="list-style-type: none"> <li>- EPS_Surprise</li> <li>- Revenue_Surprise</li> <li>- EPS_Surprise_Backward_Ave_Diff</li> <li>- Revenue_Surprise_Backward_Ave_Diff</li> </ul>
Fundamentales de la Empresa	<ul style="list-style-type: none"> <li style="width: 50%;">- Operating_Income_Q_Growth</li> <li style="width: 50%;">- Gross_Profit_Y_Growth</li> <li style="width: 50%;">- Operating_Income_Y_Growth</li> <li style="width: 50%;">- Return_On_Assets_Q_Growth</li> <li style="width: 50%;">- Net_Income_Q_Growth</li> <li style="width: 50%;">- Return_On_Assets_Y_Growth</li> <li style="width: 50%;">- Net_Income_Y_Growth</li> <li style="width: 50%;">- PE_Ratio_Q_Growth</li> <li style="width: 50%;">- Free_Cash_Flow_Q_Growth</li> <li style="width: 50%;">- PE_Ratio_Y_Growth</li> <li style="width: 50%;">- Free_Cash_Flow_Y_Growth</li> <li style="width: 50%;">- PS_Ratio_Q_Growth</li> <li style="width: 50%;">- Total_Liabilities_Q_Growth</li> <li style="width: 50%;">- PS_Ratio_Y_Growth</li> <li style="width: 50%;">- Total_Liabilities_Y_Growth</li> <li style="width: 50%;">- PB_Ratio_Q_Growth</li> <li style="width: 50%;">- Gross_Profit_Q_Growth</li> <li style="width: 50%;">- PB_Ratio_Y_Growth</li> </ul>

Se agruparon en cuatro grupos con los que se realizaron las iteraciones del modelo para encontrar el modelo con mejores métricas de precisión. Esta agrupación se hizo con el fin de evitar problemas de endogeneidad entre las mismas.

<b>Grupo Features A</b>	<b>Grupo Features B</b>
<ul style="list-style-type: none"> <li>- EPS_Surprise</li> <li>- Revenue_Surprise</li> </ul>	<ul style="list-style-type: none"> <li>- EPS_Surprise_Backward_Ave_Diff</li> <li>- Revenue_Surprise_Backward_Ave_Diff</li> </ul>
<b>Grupo Features C</b>	<b>Grupo Features D</b>
<ul style="list-style-type: none"> <li>- EPS_Surprise</li> <li>- Revenue_Surprise</li> <li>- Operating_Income_Q_Growth</li> <li>- Net_Income_Q_Growth</li> <li>- Free_Cash_Flow_Q_Growth</li> <li>- Total_Liabilities_Q_Growth</li> <li>- Gross_Profit_Q_Growth</li> <li>- Return_On_Assets_Q_Growth</li> <li>- PE_Ratio_Q_Growth</li> <li>- PS_Ratio_Q_Growth</li> <li>- PB_Ratio_Q_Growth</li> </ul>	<ul style="list-style-type: none"> <li>- EPS_Surprise</li> <li>- Revenue_Surprise</li> <li>- Operating_Income_Y_Growth</li> <li>- Net_Income_Y_Growth</li> <li>- Free_Cash_Flow_Y_Growth</li> <li>- Total_Liabilities_Y_Growth</li> <li>- Gross_Profit_Y_Growth</li> <li>- Return_On_Assets_Y_Growth</li> <li>- PE_Ratio_Y_Growth</li> <li>- PS_Ratio_Y_Growth</li> <li>- PB_Ratio_Y_Growth</li> </ul>

**3.4.3. Construcción del Modelo**

Al tratarse de un problema de clasificación binario el modelo seleccionado dentro todos los modelos disponibles de XGBoost fue XGBoost Classifier (XGBClassifier). La base de 15,938 registros que datan entre 2017-01-26 y 2024-09-11 se dividió en base de entrenamiento con fecha de corte 2021-12-31 día de inicio del retroceso en el SPX debido a la incertidumbre económica pos-COVID-19 y cercano al pico más alto de inflación y dos meses previo al inicio de incrementos en la tasa de interés de la FED. Los registros se dividieron en tres segmentos. Para cada segmento se evaluó cada uno de los cuatro grupos de características (Features) mencionados.

<b>Segmento A</b>	<b>Segmento B (Menos Líquidas)</b>	<b>Segmento C (Más Líquidas)</b>
Todas las Empresas	Empresas con Bid-Ask Spread $\geq$ Promedio Bid-Ask Spread	Empresas con Bid-Ask Spread $<$ Promedio Bid-Ask Spread

Los hiperparámetros modificados para optimización del modelo fueron: *max\_depth* (controla la profundidad del árbol), *learning\_rate* (utilizando menores valores para suavizar el aprendizaje) y *n\_estimators* (número de iteraciones del modelo) y *scale\_pos\_weight* (utilizado para compensar el

desbalance de los datos). Se utilizó el método *GridSearchCV* de la librería *sklearn* para iterar los hiperparámetros y encontrar los valores óptimos midiendo por F1-Score.

Hiperparámetro	Valores Probados	Valor Óptimo (Scoring F1)			
		Features A	Features B	Features C	Features D
max_depth	4, 5, 6, 7, 8, 9	7	9	4	7
learning_rate	0.005, 0.01, 0.04, 0.07, 0.1	0.005	0.005	0.005	0.04
n_estimators	30, 50, 100, 200	200	200	30	30

#### 3.4.4. Evaluación del Modelo

Para la medición y evaluación del modelo se usó las métricas: Precisión (% de predicciones positivas son acertadas), F1-Score (Promedio armónico entre precisión y *recall*, su rango va de 0.0 a 1.0 siendo 1.0 lo ideal) y AUC (Habilidad del modelo para clasificar siendo 0.5 misma habilidad que un modelo aleatorio y 1.0 ideal) . La métrica de F1-Score es especialmente útil para problemas con clases desbalanceadas como es el caso del presente estudio (ver balance en Tabla 7). Por esta misma razón se decidió utilizar la métrica de *Precision* en vez de *Accuracy*.

## 4. RESULTADOS

### 4.1. Descripción de las Variables

#### 4.1.1. Categorización de Earnings Surprises

La Tabla 6 muestra los EPS surprises agrupados según sí fueron una sorpresa positiva (EPS reportado > EPS estimado), una neutral (EPS reportado = EPS estimado) o una sorpresa negativa (EPS reportado < EPS estimado). De los 83,356 *Earnings Announcements* que contaban con EPS estimado y reportado, el 65% fueron positivos, el 21% negativos y en el 14% de los casos no hubo sorpresas.

**Tabla 8** Categorías de Earnings Surprises

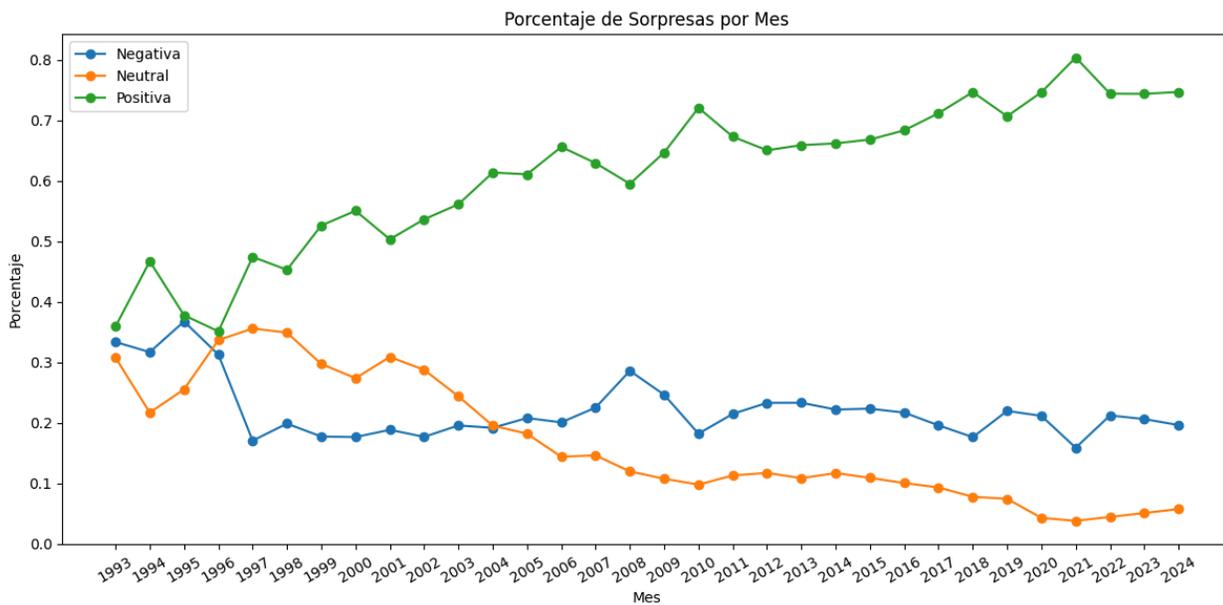
Tipo de Sorpresa	Registros	Peso
------------------	-----------	------

Positiva	53,843	64.59%
Negativa	17,437	20.92%
Neutral	12,076	14.49%

Nota Tabla 8. La sorpresa del EPS es la resta de lo estimado a lo reportado normalizado por el precio de la acción justo antes del reporte.

La gráfica 1 muestra que a pesar de que con el paso del tiempo se ha mejorado los métodos de estimación de los analistas de futuros EPS. Los Earnings Surprises siguen presentándose, siendo las sorpresas neutras las de menor presencia y las positivas las predominantes.

**Gráfico 1** Evolución en el tiempo de distribución de categorías de Earnings Surprises



Fuente: Elaboración propia

**4.1.2. Retornos en exceso: Post Earnings Announcement Drift a 7 días**

Los Earnings Announcements con EPS y Revenue Surprise Positivos tienen mejores retornos en el promedio que los que no tienen o tienen sorpresas negativas. Igualmente, los EA con sorpresas negativas tienen mayores pérdidas que los otros. Esto es coherente con la hipótesis de la actual presencia del PEAD en los mercados.

**Tabla 9** Estadística descriptiva PEAD.

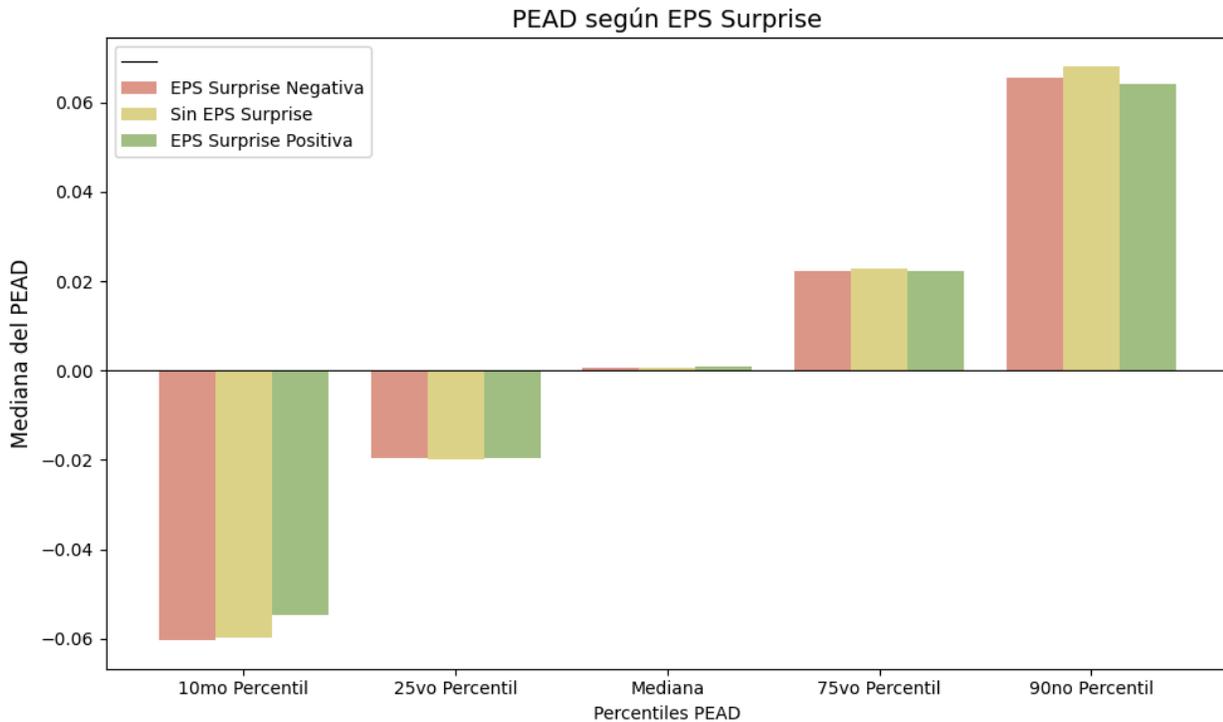
<b>Propiedad</b>	<b>PEAD de Todos los EA</b>	<b>PEAD EA con EPS y Revenue Surprise Positivas</b>	<b>PEAD EA con EPS y Revenue Surprise Negativas</b>
Registros	72,593	41,116	10,956
Promedio	0.003281	0.005765	-0.003387
Desv. Est.	0.061687	0.054553	0.061812
Min	-0.622526	-0.534697	-0.533978
Percentil 10%	-0.056746	-0.049609	-0.064485
Percentil 25%	-0.025797	-0.022412	-0.03098
Percentil 50% (Mediana)	0.000846	0.002674	-0.003638
Percentil 75%	0.029209	0.029796	0.02424
Percentil 90%	0.064929	0.063402	0.059084
Max	1.48801	1.48801	0.675525
<b>Propiedad</b>	<b>PEAD EA sin EPS Surprise o sin Revenue Surprises</b>	<b>PEAD EA con EPS y Revenue Surprise Percentil <math>\geq 90\%</math></b>	<b>PEAD EA con EPS y Revenue Surprise Percentil <math>\leq 10\%</math></b>
Registros	10,236	7,260	7,260
Promedio	0.001862	0.013493	-0.003251
Desv. Est.	0.066148	0.078852	0.087231
Min	-0.467503	-0.622526	-0.589009
Percentil 10%	-0.064441	-0.059016	-0.08341
Percentil 25%	-0.029493	-0.025354	-0.039624
Percentil 50% (Mediana)	-0.000443	0.006723	-0.00553
Percentil 75%	0.029978	0.044404	0.030332
Percentil 90%	0.069595	0.095262	0.07493
Max	0.990527	1.48801	1.48327

*Nota Tabla 9. El PEAD se calcula como retorno discreto simple de la resta del retorno continuo simple del precio de cierre en  $T_0$  a  $T_0+7$  del subyacente menos el retorno continuo simple para el mismo periodo del SPX tomando también precios de cierre en  $T_0$  a  $T_0+7$  para el índice (Shon & Zhou, 2011).*

El gráfico 2 soporta la necesidad de utilizar una estrategia de trading Delta neutral (bidireccional), dado que tanto para las sorpresas positivas como para las sorpresas negativas se pueden obtener retornos tanto positivos como negativos. Esto está en línea con Shon & Zhou (2011) quienes en

su investigación encontraron que sólo en el 60% de los casos el PEAD sigue la misma dirección que el signo de la sorpresa.

**Gráfico 2** Retornos en exceso a 7 días del EA medios según el tipo de EPS Surprise obtenido



Fuente: Elaboración propia. Metodología de cálculo de PEAD tomada (Shon & Zhou, 2011).

#### 4.2. Tablas de Resultados

La Tabla 10 muestra los resultados obtenidos para las doce iteraciones del modelo. Los cuatro grupos de Features para cada segmento. Las tres métricas de evaluación entre más cercanas a 1.0 sean mejor, por lo que el modelo cuya suma de las tres de más cercano a 3.0 será el modelo seleccionado como el mejor modelo de pronóstico.

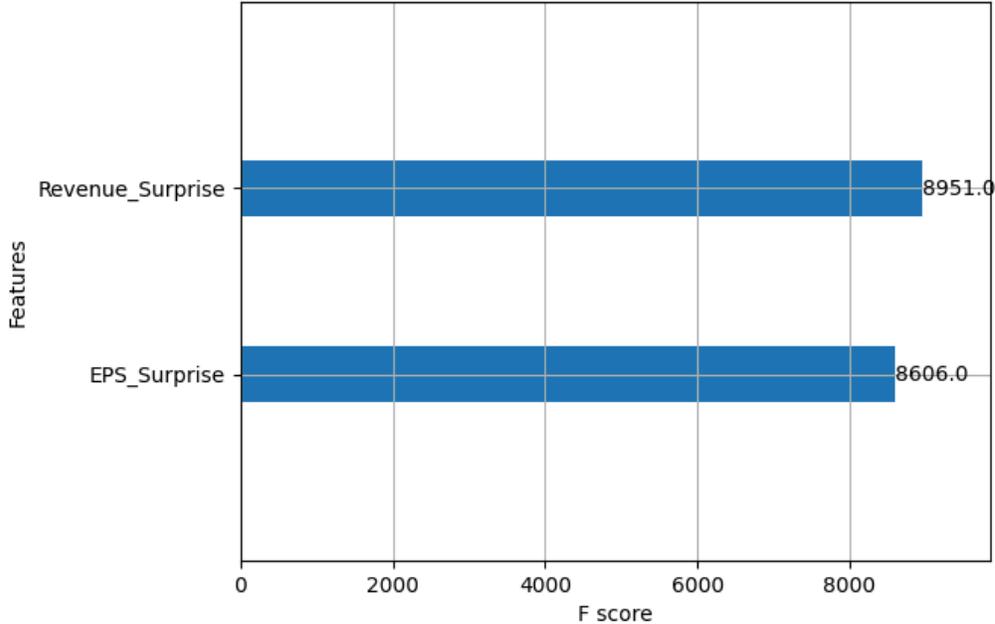
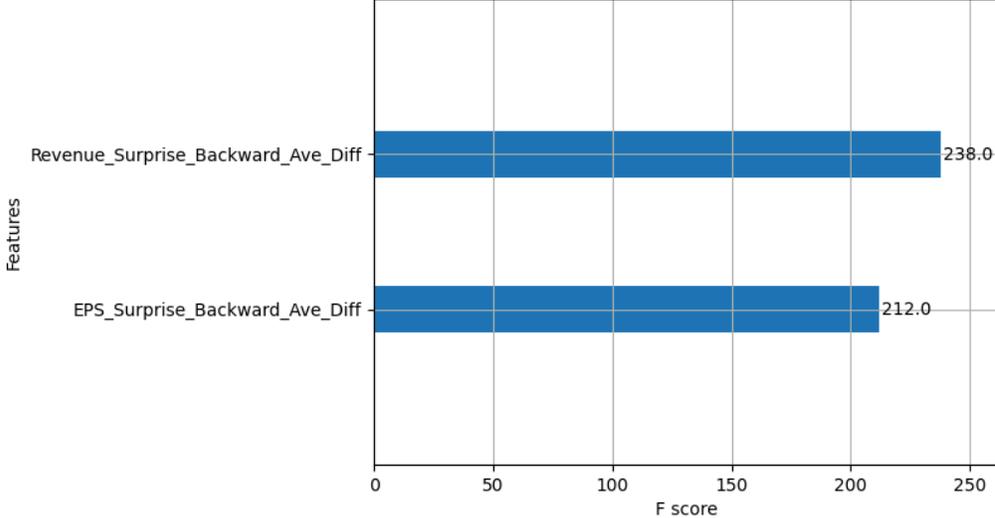
**Tabla 10** Métricas de evaluación iteraciones del modelo XGBoostClassifier

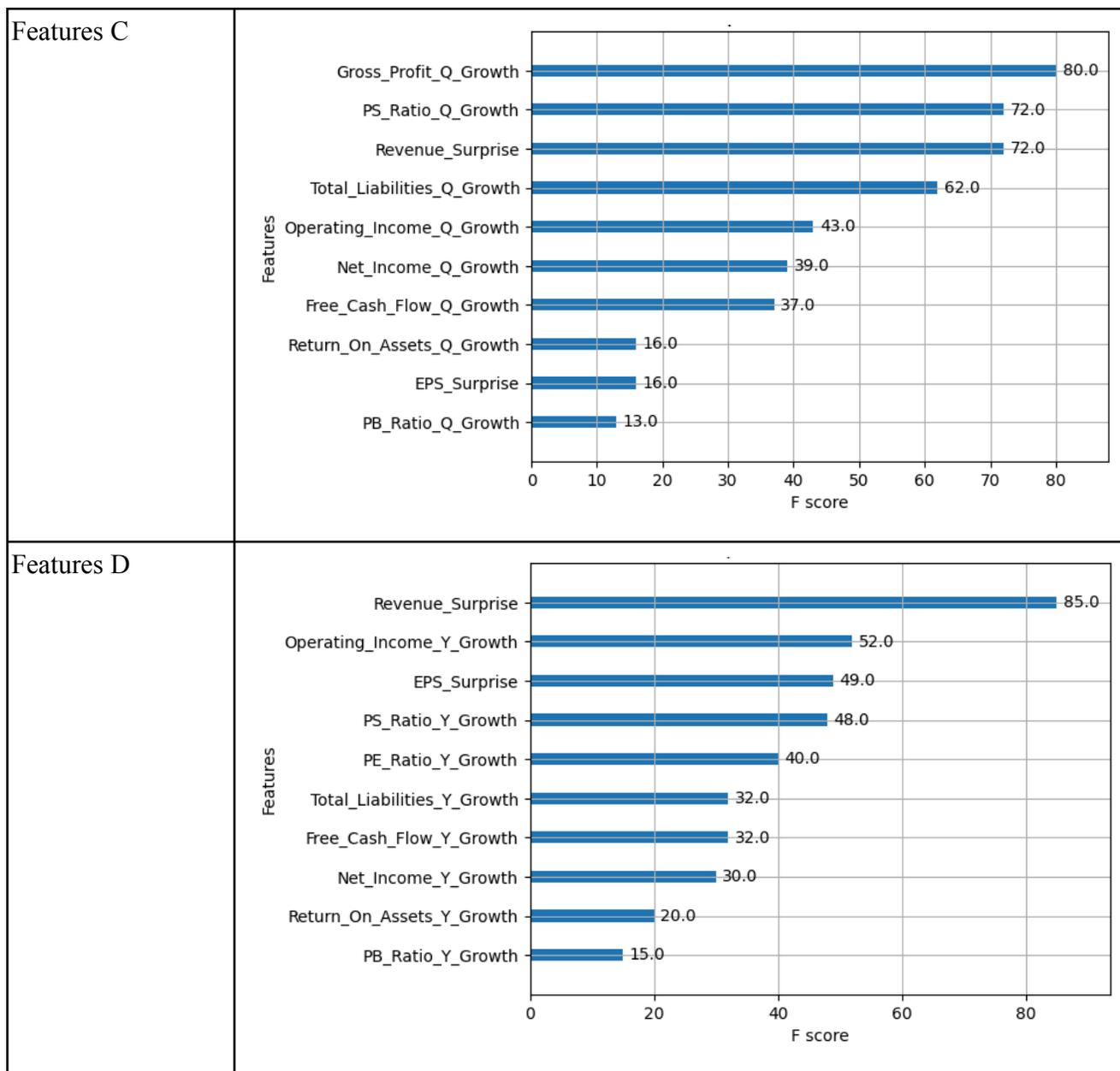
Grupo de Features	Resultados		
	Segmento A	Segmento B (Menos Líquidas)	Segmento C (Más Líquidas)
Features A	Entreno: 10,315 Prueba: 5,623 Precisión: 0.62 F1-Score: 0.77 AUC: 0.50	Entreno: 2,820 Prueba: 1,469 Precisión: 0.69 F1-Score: 0.81 AUC: 0.52	Entreno: 7,495 Prueba: 4,154 Precisión: 0.60 F1-Score: 0.75 AUC: 0.51
Features B	Entreno: 10,315 Prueba: 5,623 Precisión: 0.62 F1-Score: 0.77 AUC: 0.50	Entreno: 2,820 Prueba: 1,469 Precisión: 0.69 F1-Score: 0.81 AUC: 0.51	Entreno: 7,495 Prueba: 4,154 Precisión: 0.60 F1-Score: 0.75 AUC: 0.49
Features C	Entreno: 10,284 Prueba: 5,623 Precisión: 0.62 F1-Score: 0.77 AUC: 0.51	Entreno: 2,812 Prueba: 1,469 Precisión: 0.69 F1-Score: 0.81 AUC: 0.53	Entreno: 7,472 Prueba: 4,154 Precisión: 0.60 F1-Score: 0.75 AUC: 0.50
Features D	Entreno: 8,919 Prueba: 5,623 Precisión: 0.63 F1-Score: 0.77 AUC: 0.50	Entreno: 2,390 Prueba: 1,469 Precisión: 0.69 F1-Score: 0.81 AUC: 0.52	Entreno: 6,529 Prueba: 4,154 Precisión: 0.60 F1-Score: 0.75 AUC: 0.52

Nota Tabla 10: Métricas obtenidas con librería Sklearn. Los valores de Precisión, F1-Score y AUC entre más cercanos a 1.0 mejor.

Así mismo, se evaluó la importancia de las *features* tomando el F Score como la variable de selección, siendo un mayor F Score un indicador de mayor importancia de la *Feature* para el modelo y entre menor F Score menor relevancia tiene esa característica.

**Tabla 11** Importancia de Features según F Score para Segmento A

Grupo de Features	Importancia de Features (Según F Score)						
Features A	 <p>A horizontal bar chart titled 'Importancia de Features (Según F Score)' for 'Features A'. The y-axis is labeled 'Features' and lists 'Revenue_Surprise' and 'EPS_Surprise'. The x-axis is labeled 'F score' and ranges from 0 to 8000 with major ticks every 2000. The bar for 'Revenue_Surprise' has a value of 8951.0, and the bar for 'EPS_Surprise' has a value of 8606.0.</p> <table border="1"> <thead> <tr> <th>Feature</th> <th>F score</th> </tr> </thead> <tbody> <tr> <td>Revenue_Surprise</td> <td>8951.0</td> </tr> <tr> <td>EPS_Surprise</td> <td>8606.0</td> </tr> </tbody> </table>	Feature	F score	Revenue_Surprise	8951.0	EPS_Surprise	8606.0
Feature	F score						
Revenue_Surprise	8951.0						
EPS_Surprise	8606.0						
Features B	 <p>A horizontal bar chart titled 'Importancia de Features (Según F Score)' for 'Features B'. The y-axis is labeled 'Features' and lists 'Revenue_Surprise_Backward_Ave_Diff' and 'EPS_Surprise_Backward_Ave_Diff'. The x-axis is labeled 'F score' and ranges from 0 to 250 with major ticks every 50. The bar for 'Revenue_Surprise_Backward_Ave_Diff' has a value of 238.0, and the bar for 'EPS_Surprise_Backward_Ave_Diff' has a value of 212.0.</p> <table border="1"> <thead> <tr> <th>Feature</th> <th>F score</th> </tr> </thead> <tbody> <tr> <td>Revenue_Surprise_Backward_Ave_Diff</td> <td>238.0</td> </tr> <tr> <td>EPS_Surprise_Backward_Ave_Diff</td> <td>212.0</td> </tr> </tbody> </table>	Feature	F score	Revenue_Surprise_Backward_Ave_Diff	238.0	EPS_Surprise_Backward_Ave_Diff	212.0
Feature	F score						
Revenue_Surprise_Backward_Ave_Diff	238.0						
EPS_Surprise_Backward_Ave_Diff	212.0						



Notas Tabla 11: Cálculo realizado por librería xgboost. Entre mayor F-Score mejor.

### 4.3. Descripción de los Resultados

De los tres segmentos evaluados el segmento B de empresas de menor liquidez (Bid-Ask Spread mayor que el promedio) fue el que mejor resultado de precisión y F1-Score tuvo. Cabe aclarar que al tener poca liquidez, se requiere realizar operaciones con las empresas de este segmento que tengan suficiente *volumen* y *open interest* para buscar colocar posiciones cercanas al *Mid-Price*, como se utilizó

en la metodología de este estudio. Este segmento obtuvo un puntaje de precisión del 0.69, que en términos de la estrategia de trading de earnings con opciones sería un porcentaje de *Win-Rate* del 69%. Lo que indica una relación mínima de Riesgo-Retorno de 1:0.29 (por cada dólar arriesgado se espera obtener 29 centavos) para que las posiciones ganadoras compensen las perdedoras (en el largo plazo y con un alto número de *trades*, donde la ley de los grandes números entra en juego). Ahora bien, para obtener esta tasa de éxito implicaría que se mantendría la posición a vencimiento por lo que se descarta el uso de un Stop Loss o de un Take Profit y por ende, aunque no es parte de este estudio, la estrategia de rango óptima a utilizar sería una de riesgo limitado como por ejemplo un Short Iron Condor o un Short Iron Butterfly, así esto implique mayores costos de comisión y por ende menor rentabilidad.

Para el segmento ganador (segmento B: empresas menos líquidas) el grupo de características que tuvo las mejores métricas de desempeño fue el grupo de Features C, el cuál incluye EPS y Revenue Surprises y Tasas de cambio de Fundamentales Trimestrales. Los resultados del segmento A (todas las empresas) y C (empresas más líquidas) también indican que al agregar las tasas de cambio de los fundamentales el modelo tenía igual o mejor desempeño aunque no en la misma magnitud que para las empresas menos líquidas. Esto se puede deber a que las empresas menos líquidas tienen menos analistas informados publicando estimaciones de EPS y Revenue para futuros Earnings Announcements.

De todos los grupos de Features cabe destacar que el Revenue Surprise fue la característica con mayor relevancia en los 4 grupos lo que indica que son las sorpresas en Revenue son la mejor señal para el modelo predecir movimientos sobreestimados o subestimados excepto para el grupo de Features C que el *Gross\_Profit\_Q\_Growth* y el *PS\_Ratio\_Q\_Growth* la superaron. Para el grupo de fundamentales los relacionados al *Gross Profit*, *PS Ratio*, *Operating Income* y *Total Liabilities* son los que mayor importancia tienen entre los analizados. Cabe destacar que el cambio de medición de la sorpresa expresado en términos de su tamaño respecto al promedio de sorpresas anteriores (*EPS* y *Revenue*

*Surprise Backward Ave Diff*) obtuvo los mismos resultados que las sorpresas sin tomar en cuenta sus sorpresas pasadas (*EPS* y *Revenue Surprise*).

Más aún, debido a que el modelo de clasificación utilizado identifica los escenarios donde el movimiento es sobreestimado (1) y al ser este el escenario con mayor porcentaje de precisión en su estimación, la estrategia de trading planteada será de venta de contratos de opciones con estrategias de rango donde se beneficie de la caída en la volatilidad al mantenerse el precio dentro del rango, del paso del tiempo al ser contratos cercanos a su expiración y delta neutral que no asumen direccionalidad del subyacente. Adicionalmente, se sugiere utilizar como parte de manejo del riesgo del portafolio, Earnings que ayuden al Beta Hedging del portafolio para que se mantenga el delta neutral a nivel del portafolio.

## 5. DISCUSIÓN

La pertinencia de este estudio y la justificación detrás del modelo predictivo es la existencia actual de la anomalía del PEAD, ya que, de otra forma, sí se apoya la idea de que el mercado es eficiente en su sentido semifuerte (Fama, 1991), la información revelada el día del *Earnings Announcement* (EA) sería asimilada de inmediato por el precio y no permitiría en el largo plazo lograr retornos mayores que el mercado. La demostración de la existencia del PEAD, realizada inicialmente por Bernard & Thomas en 1989, ha sido aceptada y ampliada por autores como Griffin et al. (2010), Meursault et al. (2023) y Ye & Schuller (2021) entre otros. Sin embargo, también ha sido debatida por otros autores que argumentan que su magnitud ha venido disminuyendo a niveles insignificantes (Chordia et al., 2014; Johnson and Schwartz, 2000; Martineau, 2019) lo que indicaría que el mercado está asimilando eficientemente la información publicada en el EA. Al respecto, este documento encontró que el modelo con mejores

métricas de predicción fue el que incorporaba fundamentales y adicionalmente documentó efectos de PEAD de magnitud considerable hasta noviembre del 2024.

Por otro lado, este estudio se enfocó en los días posteriores al anuncio con el fin de estudiar la mejor estrategia en el corto plazo. Esto es apoyado por autores que demuestran que el mayor efecto del PEAD se da en el trimestre siguiente al anuncio (e.g., Cox, 2020; Li et al., 2020; Martineau, 2019). Sin embargo, otros investigadores argumentan que este efecto se da en su mayor parte previo al *Earning Announcement* debido a inversionistas con información privilegiada que se filtra en el mercado antes del día del anuncio (Milian, 2015). Aunque, este documento no estudió el efecto del PEAD los días previos al anuncio, esto le permitió centrarse en otorgar información dirigida a inversionistas que no cuentan con acceso a información privilegiada y solo conocen los la nueva información el día del anuncio.

## 6. CONCLUSIONES

El objetivo de este documento era desarrollar un modelo de aprendizaje automático que pronosticara si el movimiento esperado por el mercado será mayor o menor que el movimiento que realizará la acción después de su *Earnings Announcement*. Monetizando el efecto del PEAD en el mercado de opciones con estrategias de rango para generar Alfa (retorno en exceso) consistentemente en el tiempo. Se encontró que segmentando las empresas pertenecientes al Russel 1000 con menor liquidez por tener un Bid-Ask Spread menor que el promedio, y utilizando tanto las medidas de sorpresa del EPS y Revenue y las tasas de cambio trimestrales de los fundamentales más claves como el *Gross Profit*, *PB Ratio* y *Total Liabilities*, se logra obtener un modelo predictivo, entrenado con datos (2017 al 2021), una precisión del 69% en los datos de prueba (2022 al 2024).

El modelo identifica los escenarios donde el movimiento esperado por el contrato de opciones con fecha de expiración más cercana al *Earning Announcement* está sobreestimando el movimiento a vencimiento que tendrá el subyacente por lo que es indicador de que se puede colocar un trade de rango de venta de contratos (vega negativo, theta positivo, delta neutro) que se beneficiará del paso del tiempo, la caída de la volatilidad y no asume riesgo direccional para ser rentable. En la medida que el precio se mantenga dentro del rango estimado por el Straddle ATM en T0 y pase el tiempo la probabilidad de quedar fuera del rango será menor por lo que disminuirá el valor de la prima y por ende será rentable.

## 7. RECOMENDACIONES Y LIMITACIONES

Debido a que la base de datos histórica de opciones se consiguió con periodicidad diaria tomando el precio minutos antes del cierre del mercado, esto limitó el análisis de este estudio en términos de poder evaluar la fluctuación de los contratos durante el día y no permitió hacer una mejor selección de precios para los contratos BMO (*Before Market Opens*). Por otro lado, la serie de datos de EPS y Revenue Estimados que se usó fue únicamente el promedio de los reportes de todos los analistas suscritos a la empresa por lo que no permitió explorar lo encontrado por (Shon & Zhou, 2011) que tomar la estimación máxima y mínima de todos los analistas podría servir como mejor indicador de estimación de EPS y Revenue del Mercado.

Este estudio no contaba con todo el histórico de precios de los contratos de opciones entre T0 y T1 por lo que se limitó únicamente a analizar los precios de los contratos en T0 más no la evolución de los mismos para haber examinado a profundidad qué valor de TP (*Take Profit*) y SL (*Stop Loss*) hubieran sido óptimos para aumentar el porcentaje de Trades exitosos. Por último, este documento se limitó a realizar el cálculo de Alfa (retorno en exceso) bajo la misma metodología de (Shon & Zhou,

2011) y a usar *XG Boost Classifier* por lo encontrado por Ye & Schuller (2021) por lo que para futura investigación se recomienda evaluar otros modelos de clasificación binaria y de cálculo de Alfa.

Respecto al modelo de XG Boost Classifier, se recomienda continuar la investigación incorporando Features relacionados a los rezagos de retornos de pasados Earnings, indicadores técnicos y indicadores macroeconómicos para buscar mejorar la métrica de AUC que fue la métrica de evaluación más baja de las tres. Por último, se recomienda profundizar en el análisis particular de las diferentes estrategias de rango con opciones como el *Straddle*, *Strangle*, *Iron Condor*, entre otros para determinar con herramientas de backtesting cual de ellas da mayor rentabilidad.

## 8. BIBLIOGRAFÍA

- Ball, R., & Brown, P. (1968). An Empirical Evaluation of Accounting Income. In *Source: Journal of Accounting Research* (Vol. 6, Issue 2).
- Bernard, V. L., & Thomas, J. K. (1989). Post-Earnings-Announcement Drift: Delayed Price Response or Risk Premium? In *Current Studies on The Information Content of Accounting Earnings* (Vol. 27).
- Brav, A., & Heaton, J. B. (2002). Competing Theories of Financial Anomalies. *Review of Financial Studies*, 15(2 SPEC.), 575–606. <https://doi.org/10.1093/rfs/15.2.575>
- Bruce Johnson, W., & Schwartz Jr, W. C. (2000). Evidence that Capital Markets Learn from Academic Research: Earnings Surprises and the Persistence of Post-Announcement Drift.
- Cao, M., & Wei, J. Z. (2008). Option Market Liquidity: Commonality and Other Characteristics. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1082642>
- Chordia, T., Subrahmanyam, A., & Tong, Q. (2014). Have capital market anomalies attenuated in the recent era of high liquidity and trading activity? *Journal of Accounting and Economics*, 58(1), 41–58. <https://doi.org/https://doi.org/10.1016/j.jacceco.2014.06.001>

- Cox, J. (2020). Market fragmentation and post-earnings announcement drift. *Journal of Economics and Finance*, 44(3), 587-610.
- Fama, E. F. (1991). Efficient Capital Markets: II. In *Source: The Journal of Finance* (Vol. 46, Issue 5).
- Foster, G., Olsen, C., & Shevlin, T. (1984). Earnings Releases, Anomalies, and the Behavior of Security Returns. In *Source: The Accounting Review* (Vol. 59, Issue 4).
- Fink, J. (2021). A review of the Post-Earnings-Announcement Drift. In *Journal of Behavioral and Experimental Finance* (Vol. 29). Elsevier B.V. <https://doi.org/10.1016/j.jbef.2020.100446>
- Griffin, J. M., Kelly, P. J., & Nardari, F. (2010). Do Market Efficiency Measures Yield Correct Inferences? A Comparison of Developed and Emerging Markets. *The Review of Financial Studies*, 23(8), 3225–3277. <https://doi.org/10.1093/rfs/hhq044>
- Landsman, W., & Maydew, E. (2002). Has the Information Content of Quarterly Earnings Announcements Declined in the Past Three Decades? *Journal of Accounting Research*, 40, 797–808. <https://doi.org/10.1111/1475-679X.00071>
- Lev, B. (1989). On the Usefulness of Earnings and Earnings Research: Lessons and Directions from Two Decades of Empirical Research. *Journal of Accounting Research*, 27, 153–192. <https://doi.org/10.2307/2491070>
- Li, Y., Nekrasov, A., & Teoh, S. H. (2020). Opportunity knocks but once: delayed disclosure of financial items in earnings announcements and neglect of earnings news. *Review of Accounting Studies*, 25, 159-200.
- Milian, J. A. (2015). Unsophisticated arbitrageurs and market efficiency: Overreacting to a history of underreaction? *Journal of Accounting Research*, 53(1), 175–220. <https://doi.org/10.1111/1475-679X.12070>

- Martineau, C. (2019). The Evolution of Financial Market Efficiency: Evidence from Earnings Announcements. <https://www.charlesmartineau.com>.
- Meursault, V., Liang, P. J., Routledge, B. R., & Scanlon, M. M. (2023). PEAD.txt: Post-Earnings-Announcement Drift Using Text. *Journal of Financial and Quantitative Analysis*, 58(6), 2299–2326. <https://doi.org/10.1017/S0022109022001181>
- Roll, R., Schwartz, E., & Subrahmanyam, A. (2009). Options trading activity and firm valuation. *Journal of Financial Economics*, 94(3), 345–360. <https://doi.org/https://doi.org/10.1016/j.jfineco.2009.02.002>
- Shon, J., & Zhou, P. (2011). *Trading on Corporate Earnings News* (Pearson Education LTD, Ed.).
- Ye, Z. J., & Schuller, B. W. (2021). Capturing dynamics of post-earnings-announcement drift using a genetic algorithm-optimized XGBoost. *Expert Systems with Applications*, 177. <https://doi.org/10.1016/j.eswa.2021.114892>
- Zhou, P., & Shon, J. (2012). *Option Strategies for Earnings Announcements*. Pearson Education, Inc.