

Entrega Semana 7

Integrantes: Alexis Abreu, Maria Paula Cortés, Ricardo Pretelt, Edgar Balaguera

¿Cómo determinar la estructura del mercado de valores en términos de la correlación condicional entre instrumentos financieros para la ayuda en la toma de decisiones de inversión?

Resumen

Se implementará un algoritmo de clustering que a partir de analizar las correlaciones condicionadas entre los precios de cierre de las acciones seleccionadas en el período de 1 de enero de 2018 a 30 de junio de 2022 se logra encontrar patrones similares en los movimientos del precio de las acciones y determinar a través de estas similitudes que acciones tienen los mismos comportamientos para diversificar las inversiones.

Se parte de la agrupación de las acciones por el comportamiento en su precio en el mercado, analizar su rendimiento y así poder establecer una medida de comparación con las demás. Una vez se establecen las medidas de comparación se realizará un proceso de clustering con GraphicalLassoCV. Para que el clustering de resultado sea graficado se utiliza un embedding de 2 dimensiones.

El objetivo del proyecto es identificar los instrumentos financieros con comportamientos similares e identificar, por ende, las posiciones de mayor rentabilidad al momento de tomar una decisión de inversión. Estos agrupamientos permiten aumentar la posibilidad de rentabilidad positiva a partir del entendimiento de más de 1 cluster, es decir diversificar el portafolio de inversión.

En el proceso de clustering estos grupos con comportamientos similares deben ser mutuamente excluyentes; es decir, cada instrumento debe pertenecer a un sólo grupo y cada grupo debe estar formado por empresas en el que su comportamiento sea lo más parecido posible y a la vez, con comportamientos diferentes entre grupos. En la práctica, se busca encontrar grupos con elementos lo más homogéneos posibles pero heterogéneos entre ellos.

Introducción

Históricamente las personas con intenciones de inversión han tenido que acercarse a una comisionista o a un inversionista profesional que establezca un vínculo entre las oportunidades de inversión del mercado y el perfil de riesgo del cliente. El deber de asesoría obliga al asesor a mantener las características de apetito de riesgo del cliente alineadas con el portafolio de inversión.

Recientemente se han popularizado las páginas de trading independiente en donde las personas tienen la libertad de invertir en los instrumentos financieros que resultan atractivos. Sin embargo, son pocas las personas con alto Financial Literacy y por ende dependen de las asesorías y consejos de personas que se consideran expertas en el campo.

Un desarrollo de este tipo, habilita la toma de decisiones con base en afinidad entre los instrumentos financieros objeto de estudio o seleccionados por cada persona. Es por esto, que pretendemos implementar un algoritmo de clustering que permite encontrar grupos de interés basados en en los

precios de cierre de las acciones en un corte transversal del tiempo, analizar correlaciones condicionadas entre los precios de cierre de cada una de las acciones seleccionadas y así lograr encontrar patrones similares en los movimientos del precio de las acciones y determinar a través de estas similitudes que acciones tienen los mismos comportamientos para diversificar las inversiones.

Se parte de la agrupación de las acciones por el comportamiento en su precio en el mercado, analizar su rendimiento y así poder establecer una medida de comparación con las demás. Una vez establezcamos estas medidas de comparación se realizará un proceso de clustering con GraphicalLassoCV.

El clustering, en otras palabras, permite identificar los instrumentos financieros con comportamientos similares e identificar, por ende, las posiciones de mayor rentabilidad al momento de tomar una decisión de inversión. Estos agrupamientos permiten aumentar la posibilidad de rentabilidad positiva a partir del entendimiento de más de 1 cluster, es decir diversificar el portafolio de inversión.

En el proceso de clustering estos grupos con comportamientos similares deben ser mutuamente excluyentes; es decir, cada instrumento debe pertenecer a un sólo grupo y cada grupo debe estar formado por empresas en el que su comportamiento sea lo más parecido posible y a la vez, con comportamientos diferentes entre grupos. En la práctica, se busca encontrar grupos con elementos lo más homogéneos posibles pero heterogéneos entre ellos.

Materiales y métodos

En cuanto a materiales se usan los datos históricos de valor de acciones en cierre ajustado para empresas de Standard & Poor's 500 (S&P 500) desde 1/01/2018 hasta 30/06/2022, teniendo en cuenta que la bolsa sólo abre en días laborales. Se obtuvieron los datos gracias a la librería de python Yfinance[1]. También de Yahoo Finance se obtiene información importante como categorías de las empresas, rentabilidad de dividendos y la descripción de la empresa.

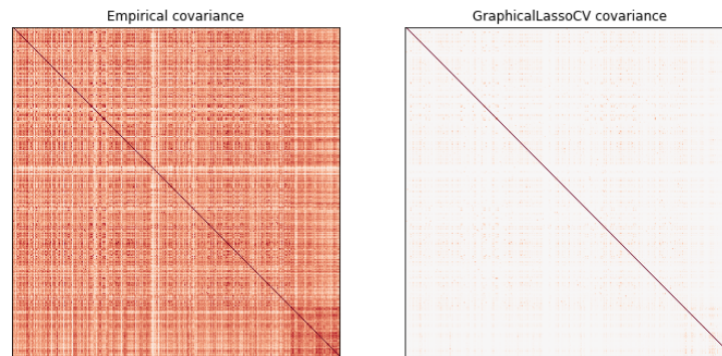
Para el desarrollo de este proyecto, se usa GraphicalLassoCV: Luego de obtener la variación porcentual para los precios de cierre de los activos durante el periodo analizado, lo que se busca es encontrar que precios de cierre están correlacionados unos con otros de forma condicional, para esto se hará uso de la matriz inversa de covarianza dispersa (sparse inverse covariance estimation), esto nos dará como resultado una lista de símbolos y conexiones. Finalmente se realizará la clusterización y se realizará el gráfico.

Usaremos la sparse inverse covariance matrix para encontrar que variaciones de precio están relacionadas condicionalmente con otras. Primero comparamos resultados para elección de GraphicalLassoCV.

La matriz inversa de covarianza es la que será nuestro insumo para realizar el siguiente paso, la clusterización. Lo anterior debido a que con esta matriz tenemos el cálculo de los precios que se comportan de forma similar. Después de identificar los clúster se puede graficar en 2D anidado.

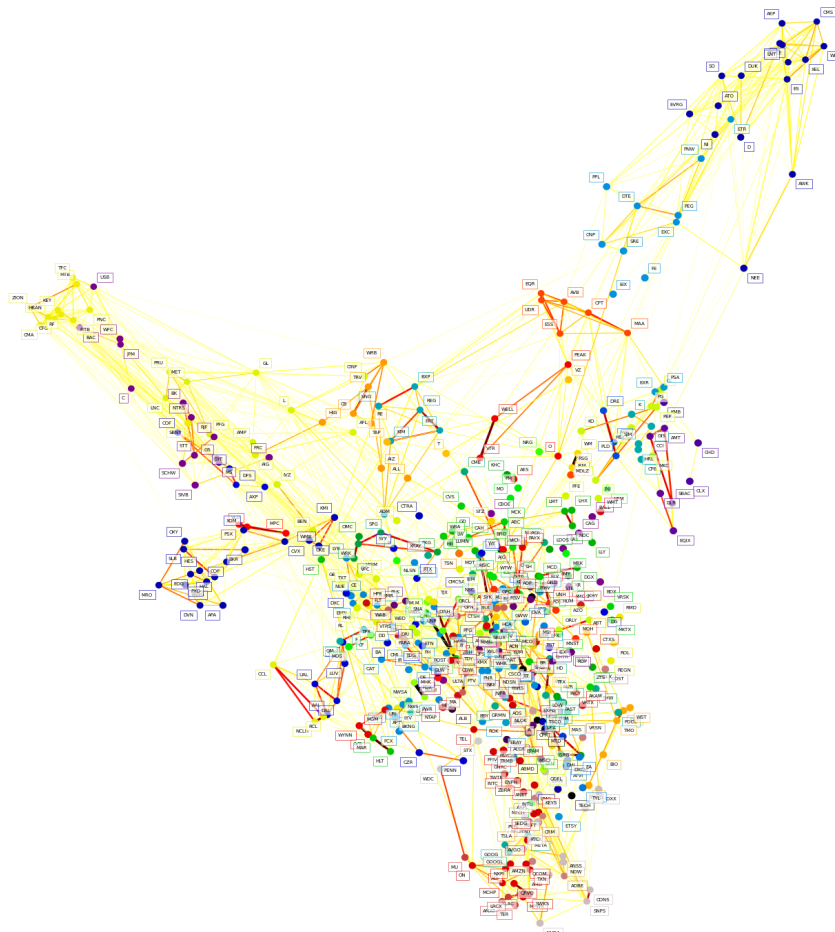
Resultados y discusión

De acuerdo a lo planteado metodológicamente se realiza el modelo de GraphicalLassoCV con los datos y se procede a encontrar las covarianzas empíricas y covarianza GraphicalLassoCV. Se observa que GraphicalLassoCV muestra resultados suavizados.

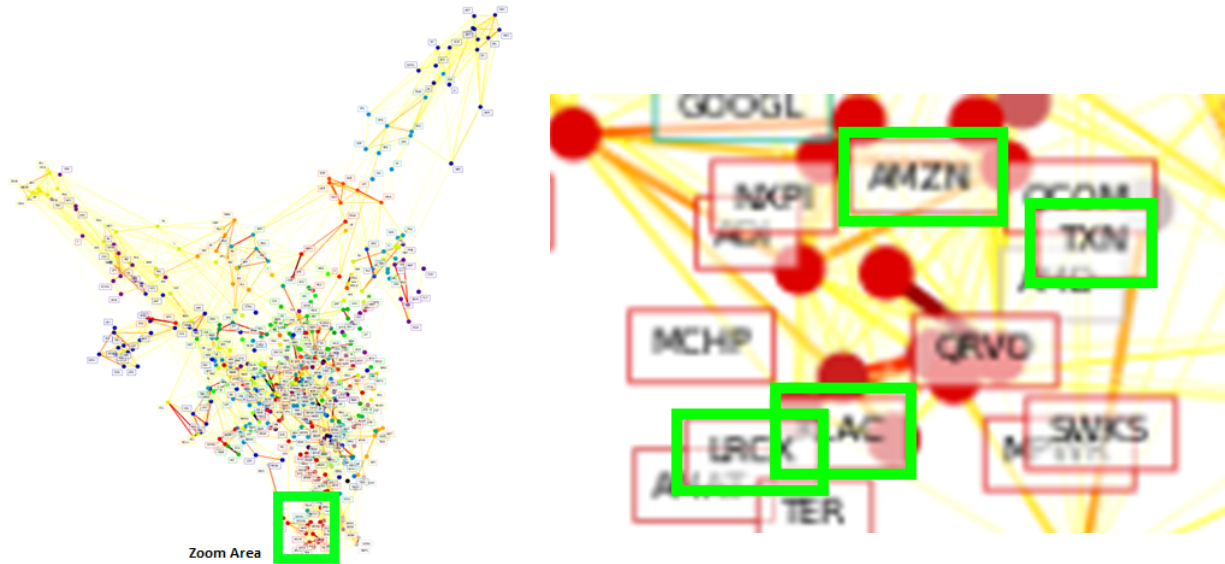


En resumen, las dos gráficas anteriores nos muestran cómo difiere la representación de la matriz inversa de covarianza, mostrando que el algoritmo GraphLassoCV hace un buen trabajo reduciendo el ruido entre las relaciones de los símbolos permitiendo un mejor insumo para el algoritmo de clusterización.

Para realizar la gráfica de clúster se procede a hacer una embedding de 2 dimensiones y se obtiene la siguiente gráfica:



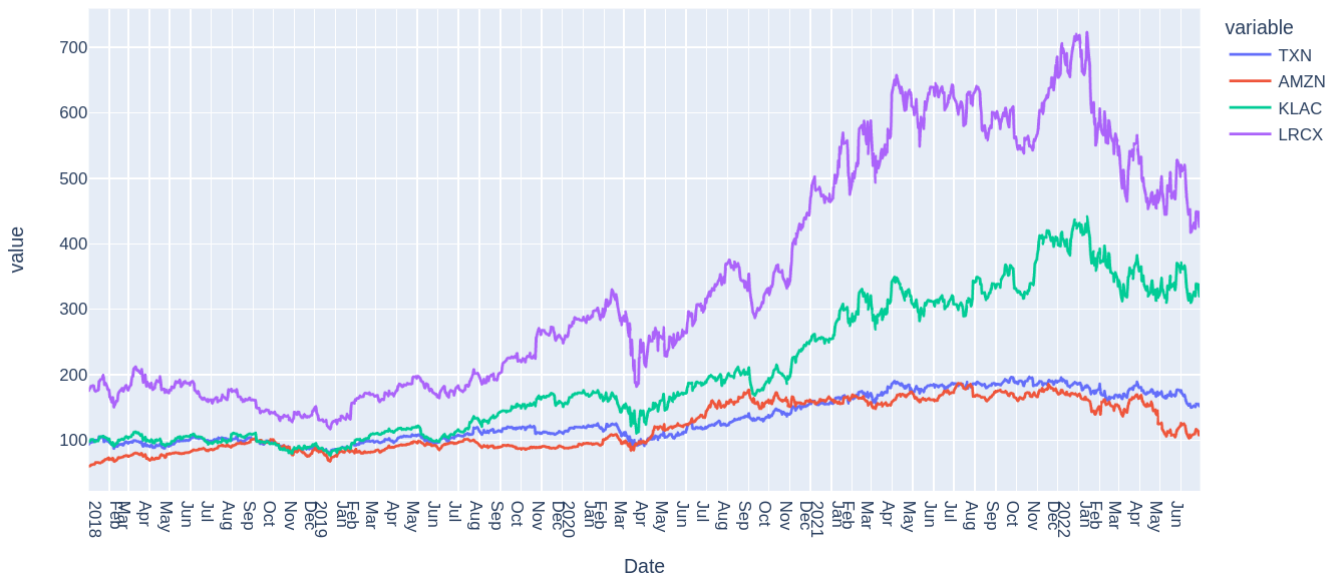
Con la última gráfica se procede a analizar los resultados. Para este caso se toman 4 empresas que están relacionadas entre sí, primero TXN y AMZN que pertenecen al mismo cluster y están próximos en el espacio, y no muy lejos KLAC y LRCX.



Interesantemente, 3 de las 4 compañías pertenecen al sector de tecnología y AMZN al de consumo. Lógicamente empresas de sectores similares tienden a tener comportamientos similares en los precios de sus acciones por lo que se agrupan.

Symbol	Description	Category2	Category3	GICS Sector	Market cap	Dividend yield	Price to TTM earnings	Price to TTM sales	Price to book value	Action
TXN	Texas Instrument Inc	Common stocks	Large cap	Information Technology	\$167,537,382,446	2.48%	0.0	0.0	0.0	Analyze
AMZN	Amazon.Com Inc.	Common stocks	Large cap	Consumer Discretionary	\$1,474,954,186,550	0.00%	0.0	0.0	0.0	Analyze
KLAC	KLA-Tencor Corporation	Common stocks	Large cap	Information Technology	\$53,988,951,636	1.16%	0.0	0.0	0.0	Analyze
LRCX	Lam Research Corp	Common stocks	Large cap	Information Technology	\$70,794,601,180	1.17%	0.0	0.0	0.0	Analyze

Comportamiento de compañías seleccionadas



De acuerdo a la gráfica se observa que el comportamiento de las compañías son similares o siguen algunos patrones. Por las fechas se observa que son compañías que se beneficiaron de la virtualidad por las cuarentenas de pandemia que hicieron incrementar su valor entre 2020 y finales de 2021. Al reducirse las restricciones las acciones empezaron a perder valor. En su mayoría son compañías del sector de tecnologías de la información como era de esperarse, pero también hay empresas de consumo de productos, debido a la alta demanda de productos de primera necesidad en tiempos de cuarentena.

Conclusiones

Para el proyecto usando un dataset de histórico de precios de acciones se mostró que es posible agrupar haciendo clustering comportamientos similares entre empresas. Empresas similares o por lo menos del mismo sector productivo tienden a estar en el mismo clúster. Estas empresas a pesar de estar en el mismo clúster pueden tener incrementos o decrecimientos de su valor más fuertes entre sí, pero con una misma tendencia. Para invertir, es un buen comienzo si se tiene un interés en una empresa y se buscan similares ya que se reduce el grupo de estudio y se puede encontrar relaciones entre empresas, que no se tenían en cuenta previamente.

Referencias

[1] Documentación Yfinance en python <https://pypi.org/project/yfinance/>