



GRADO EN ADMINISTRACIÓN Y DIRECCIÓN DE EMPRESAS  
(ADE)

TRABAJO FIN DE GRADO (ADE)

El valor de ser un líder en sostenibilidad: Estudio de impacto de la  
economía sostenible en la empresa

Autor: Jorge Eugenio Prieto Rivero

Director: José Carlos Romero Mora

Madrid

Noviembre de 2023

Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título:  
El valor de ser un líder en sostenibilidad: Estudio de impacto de la economía sostenible en la  
empresa  
en la ETS de Ingeniería - ICAI de la Universidad Pontificia Comillas en el  
curso académico 2023-2024 es de mi autoría, original e inédito y  
no ha sido presentado con anterioridad a otros efectos. El Proyecto no es plagio de otro, ni  
total ni parcialmente y la información que ha sido tomada  
de otros documentos está debidamente referenciada.


Fdo.: Jorge Prieto Rivero Fecha: 26/ 11/ 2023



Autorizada la entrega del proyecto

EL DIRECTOR DEL PROYECTO

Fdo.: José Carlos Romero Mora Fecha: 26/ 11/ 2023





GRADO EN ADMINISTRACIÓN Y DIRECCIÓN DE EMPRESAS  
(ADE)

TRABAJO FIN DE GRADO (ADE)

El valor de ser un líder en sostenibilidad: Estudio de impacto de la  
economía sostenible en la empresa

Autor: Jorge Eugenio Prieto Rivero

Director: José Carlos Romero Mora

Madrid

Noviembre de 2023

## Contenido

Resumen Ejecutivo.....	5
Alineación con los Objetivos de Desarrollo Sostenible (ODS).....	7
Limitaciones de las variables estudiadas.....	8
Capítulo 1.- Introducción y planteamiento del proyecto.....	10
1.1.- Introducción.....	10
1.2.- Planteamiento del proyecto.....	11
Capítulo 2.- Descripción de las tecnologías (estado de la técnica) .....	13
Capítulo 3.- Modelo desarrollado.....	15
3.1.- Variables.....	15
3.2.- Estudio de Relaciones.....	17
3.2.1.- Caso Base.....	21
3.2.2.- Caso Ratios Financieros.....	22
3.2.3.- Caso Ratios ESG.....	24
3.2.4.- Caso Deltas Financieras.....	25
3.2.5.- Caso Deltas ESG.....	28
3.2.6.- Caso ESG Previo.....	29
3.2.7.- Caso CAGR.....	31
3.2.8.- Caso CAGR – Otros Algoritmos.....	32
3.3.- Revisión de los Datos.....	35
3.4.- Análisis Predictivo.....	38
3.4.1.- Predicción por regiones.....	39
3.4.1.- Predicción por sector.....	43
Capítulo 4.- Conclusiones.....	47
4.1.- Conclusiones sobre la metodología.....	47
4.2.- Conclusiones sobre los resultados.....	48
4.3.- Recomendaciones para futuros estudios.....	51
Capítulo 5.- Bibliografía.....	53
Capítulo 6.- Apéndices.....	55
6.1.- Apéndice A: Visualización de los datos.....	55

# Resumen Ejecutivo

El valor de ser un líder en sostenibilidad: Estudio de impacto de la economía sostenible en la empresa

**Autor: Prieto Rivero, Jorge Eugenio.**

Director: Romero Mora, José Carlos.

Entidad Colaboradora: Kearney

## RESUMEN DEL PROYECTO

La generación de valor por parte de empresas alineadas con objetivos sostenibles es innegable. La transformación esencial del modelo productivo de las sociedades a nivel internacional, cambios regulatorios y desarrollos tecnológicos han cambiado de gran manera el funcionamiento de las empresas y los mecanismos que las recompensan.

El sector energético es especialmente estudiado en este aspecto por la escala de los efectos sociales y medioambientales que tiene su funcionamiento. El sector ha sido históricamente uno de los mayores pesos económicos y de empleo a nivel mundial y ha llevado a cabo una gran transformación hacia fuentes de energía más eficientes y limpias que permiten una transición sostenible para los países a nivel internacional.

Su rol dentro de la transición sostenible hace que una justificación objetiva de la relevancia e importancia de las variables de sostenibilidad sea importante no solo para aumentar la consciencia de este factor sino también para favorecer su comprensión. Este trabajo por tanto se posiciona para hacer un análisis objetivo estadístico que permita establecer los vínculos aparentes entre la métricas de sostenibilidad y las económicas.

La selección de las variables a estudiar se ha llevado a cabo de manera razonada y meticulosa por su gran importancia a la hora de determinar las conclusiones que son posibles conseguir. Su recogida se ha llevado a cabo mediante Bloomberg y CapitalIQ por su facilidad de acceso a datos y las bases de datos de información histórica que tienen.

Para el desarrollo del estudio se han realizado durante este trabajo una serie de dinámicas de análisis estadístico modernas en forma de correlación y algoritmos de árbol de decisión, bosques aleatorios, Naïve Bayes, perceptrón multi capa, k-means y “State Of The Art” (SOTA).

El estudio masivo de estas empresas ha demostrado de manera estadística y objetiva que la primera hipótesis de este trabajo sobre la existencia de una relación entre variables de sostenibilidad y financieras dentro de las empresas. Además, el estudio de los diferentes casos ha permitido interpretar relaciones más profundas entre algunas de las variables interpretando patrones de ciclos, efectos económicos en años específicos o mecanismos de recompensa regulatorios en el mercado.

La segunda hipótesis de establecer un modelo claro de predicción para conectar la rentabilidad de la inversión en una empresa a través de sus variables de sostenibilidad ha conseguido resultados mixtos. Por un lado, se ha demostrado que con una variedad de algoritmos y con grupos de variables adecuados se pueden conseguir éxitos de predicción muy elevados. Sin embargo, también se establece que no existe un modelo único y específico para realizar este análisis. Es por tanto recomendable que el análisis sea lo más rico posible e incluyendo variables de sostenibilidad además de financieras para poder llevar a cabo una mejor inversión y favorecer el cambio sostenible con una distribución de capital más madura.

# Alineación con los Objetivos de Desarrollo Sostenible (ODS)

El estudio realizado en este trabajo sigue un alineamiento con los Objetivos de Desarrollo Sostenible polifacético. Esto se debe al estudio de diferentes variables de la sostenibilidad y su impacto en las variables financieras, así como el estudio de un patrón de conducta que pueda servir para explicar decisiones de inversión alineadas con los ODS y su empuje a futuro.



Figura 1: Objetivos Desarrollo Sostenible a los que afecta el proyecto.

Su vínculo principal se encuentra con el ODS9 que representa la “Industria, Innovación e Infraestructura”. Según la propia página de la UNPD “con más de la mitad de la población mundial viviendo en ciudades, el transporte masivo y la energía renovable son cada vez más importantes” para ello “formas importantes para facilitar el desarrollo sostenible son la promoción de industrias sostenibles” (UNPD23).

Dentro de esta labor se encuentra el eje principal del trabajo que consiste en conseguir identificar mecanismos y la existencia general de vínculos entre las variables que definen la sostenibilidad de una empresa, sus variables financieras y el transcurso de su valoración durante los años. Con esta función troncal el trabajo promueve una mejor comprensión del desarrollo sostenible entre las compañías energéticas, población de estudio del trabajo.

Por su naturaleza estadística e investigadora durante el trabajo se establece una disciplina de estudio. Esta disciplina de estudio, como se explicará en más profundidad en el planteamiento del proyecto, abarca el estudio de las variables a utilizar, la identificación de limitaciones y de maneras eficientes de obtener esos datos, el uso de métodos

estadísticos para la exploración de cómo abordar la hipótesis, determinar los posibles estudios cualitativos y algorítmicos que emplear y desarrollar conclusiones y desafíos con los que poder replicar y avanzar el estudio a futuro. Esta disciplina de estudio dentro del trabajo busca encontrar un patrón estandarizable que ayude a proyectos posteriores y permita ayudar en la consecución de los ODS.

De manera importante, pero menos con menos incidencia necesariamente, se hace una alineación fuerte con los ODS 5, 6, 7, 10, 12 y 13.

El ODS5 “Igualdad de Género” expone que “aunque hay más mujeres que nunca en el mercado laboral, todavía hay grandes desigualdades en algunas regiones, y sistemáticamente a las mujeres se les niegan los mismos derechos laborales que tienen los hombres” (UNPD23). Durante la realización del trabajo las variables vinculadas a la incorporación de la mujer a las empresas de energía se han tenido en cuenta. La representación a través de los años en los diferentes niveles de dirección de la empresa permite de una manera sencilla vincular este objetivo a las hipótesis y resultados del estudio. El vínculo con el ODS10 “Reducción de las Desigualdades” viene representado a través de variables adicionales de convicción social de la empresa que vienen detalladas en apartados posteriores.

Los ODS6 “Agua Limpia y saneamiento”, ODS7 “Energía Asequible y No Contaminante”, ODS12 Producción y Consumo Responsables” y ODS13 “Acción por el Clima” vienen representadas a través de una serie de variables de producción y consumo de las empresas. El consumo de agua, residuos generados, consumo de energía renovable, residuos reciclados y emisiones de gases invernadero en diferentes perímetros son algunas de las variables que se estudian y que han permitido durante el transcurso del trabajo estudiar el impacto en las variables financieras y en la inversión en la empresa.

## Limitaciones de las variables estudiadas

Antes de la introducción del trabajo es relevante resaltar que, siempre intentando abarcar el máximo posible de variables, este trabajo no pretende descartar aspectos de la sostenibilidad como menos importantes ni trata de establecer algunos como principales.

La primera de las razones por la que no se puede atribuir esa intención al trabajo es por la imposibilidad de encontrar contenido estandarizado, reportado y transparente durante los años estudiados para las empresas escogidas de muchos aspectos muy relevantes para la transición sostenible de nuestras sociedades a futuro.

Aspectos como la pobreza energética e impactos sociales de las instalaciones de transmisión y generación, por nombrar algunos, tienen una gran importancia dentro del desarrollo de las empresas de energía de cara a la sostenibilidad, pero por su estudio reciente, falta de mecanismos claros para reportar estos impactos y por su falta de estandarización no se han encontrado variables adaptables al proyecto con las que representarlos.

La segunda es que la transición global se encuentra en un proceso de definición a nivel global. Nuevas iniciativas para la promoción de energía verde y procesos novedosos transformadores en la generación y transporte de energía en Estados Unidos como el



empujado con el “Inflation Reduction Act” y cambios en el marco regulatorio de la financiación sostenible en la Unión Europea redefinen globalmente la aplicación de variables y criterios que establezcan una guía sobre la que medir el impacto de las empresas (EC23).

# Capítulo 1.- Introducción y planteamiento del proyecto

## 1.1.- Introducción

El impacto de las variables ESG en el tejido financiero y económico empresarial es indudable. El acceso a financiación ESG, además de las regulaciones a favor de la transición a medidas de sostenibilidad y justicia social, está acelerando en los últimos años. Desde el sector de la energía, diferentes medidas de compensación de los precios en la unión europea y el 'IRA' en estados unidos han reforzado la posición preponderante de la sostenibilidad en la energía por lo que analizar un número amplio de variables ESG en las variables financieras de las empresas es necesario para los inversores.

Además, desde organizaciones internacionales, nacionales y de origen privado como la "WEF", "Bank of England" y el "Glasgow Alliance for Net Zero" se organizan a agentes privados y públicos para organizar medidas de ayuda al medioambiente de manera más eficiente, compartir los recientes avances y empujar la percepción de los problemas de la sostenibilidad como acuciantes. El propio presidente del Banco de Inglaterra en 2015, Mark Carney, explicaba en su discurso sobre la tragedia en el horizonte que "cada vez hay más pruebas del papel del hombre en el cambio climático. Se considera muy probable que los factores humanos sean la causa principal del calentamiento global desde mediados del siglo XX. Aunque las fluctuaciones naturales pueden enmascararlo temporalmente, la tendencia subyacente de calentamiento inducido por el hombre de dos décimas de grado por década no ha disminuido desde la década de 1970" (BOE15).

A esta coyuntura positiva para la transición, se le contraponen acciones legales por parte de la política estadounidense buscando limitar la capacidad de la sostenibilidad como variable para la inversión. Además, relevantes fondos y bancos de inversión están limitando sus iniciativas de sostenibilidad y actuando contra sus accionistas climáticos para preservar la rentabilidad de sus inversiones. De acuerdo con Moody's "aunque muchos gestores de activos consideran que el enfoque medioambiental de las inversiones es clave para la rentabilidad a largo plazo, las autoridades de un número cada vez mayor de estados de Estados Unidos (27, representando el 35% de los activos de fondos de pensiones nacionales) consideran que este planteamiento es incompatible con la responsabilidad fiduciaria de los gestores de activos y afirman que puede violar las leyes antimonopolio" (IE23). Dentro de los principales gestores de la inversión a nivel mundial, JPMorgan recomendó a sus accionistas a oponerse a todas las medidas que forzarían una política de descarbonización estricta a la institución de cara a reunión anual de mayo 2023 (STUTTS23) y BlackRock redujo el porcentaje de propuestas ESG aprobadas del 47% en 2021 al 7% en 2023 (MASTERS23).

Con una oposición creciente en focos institucionales financieros de gran importancia es más relevante que nunca fundamentar el valor de la inversión sostenible desde una multipolaridad de factores para poder evidenciar la necesidad de continuar este tipo de inversión.

El origen del proyecto se ubica en gran parte sobre las ideas y generosa ayuda de Celine Bak, Caballero de “l'Ordre National du Mérite” por su labor en la cumbre de Paris, sobre continuar la labor realizada por Kearney con la elaboración de un índice de sostenibilidad para poder ubicar que sectores eran más beneficiados por los mercados de capital por sus políticas ESG (BAK21).

Además, durante la realización del proyecto José Ruíz López y Álvaro de Palacio, de Kearney, han llevado a cabo un seguimiento del trabajo con el objetivo de facilitar su transcurso y orientarlo hacía resultados más eficaces además de muy generosa ayuda durante la realización de la fase de recolección de datos de las empresas, fase detallada en el planteamiento.

## 1.2.- Planteamiento del proyecto

El proyecto se basa en la consecución de dos hipótesis la primera es averiguar si existe una relación entre las variables financieras de una empresa y las de sostenibilidad y la segunda es si estas variables pueden establecer un patrón claro y predecible para la evolución en el valor capitalizable de las empresas a través de los años.

Con el objetivo de abordar esas hipótesis al plantear el proyecto se tuvo que decidir cuáles son las variables más adecuadas de carácter financiero para representar a las empresas. Esta primera criba se ha realizado principalmente por contraste con especialistas y por la experiencia adquirida durante la realización de mis prácticas en una gran gestora de activos (simultáneas a esta fase del proyecto). Este proceso iterativo llevo a la limitación de las variables financieras a Deuda Neta, Deuda Total, Ingresos, EBITDA, EBIT, EBT, Beneficios Netos, inversión en inmovilizado (CAPEX), Capitalización Bursátil, valor de la empresa (EV), EV/EBITDA, Intereses Minoritarios, efectivo de las operaciones y gastos provenientes de intereses sobre la deuda.

Para la obtención de estas variables financieras se tuvo que prever algún método automatizado o semi automatizado con el que conseguir las previamente mencionadas variables y es por ello por lo que se optó por la plataforma de S&P “CapitalIQ”. La plataforma es parte de la rama de inteligencia de mercado de la compañía y permite a sus usuarios acceder a la información pública de las empresas de una manera masiva y contrastable.

El proceso de selección de las variables financieras se tuvo que llevar a cabo de manera iterativa para contrastar no solo cuáles de las variables se encontraban representadas en la plataforma (con el fin de tener una única fuente del dato) sino también cuáles se encontraban suficientemente pobladas para todas las empresas y el rango de años seleccionados.

La selección de las variables de sostenibilidad se realizó a la inversa por limitación de fuentes automatizadas de las que poder obtener los datos de las empresas a nivel masivo.

Esta limitación se ha considerado suficientemente relevante como para incluir en el apartado de alineación con las ODS.

La aplicación que se consideró de mayor idoneidad por su cantidad de variables y población de los datos ha sido la plataforma de “Bloomberg”. Al igual que “CapitalIQ” la plataforma de “Bloomberg” obtiene sus datos de forma masiva de la información pública de las empresas.

La filtración de las variables de sostenibilidad por tanto pasa por un primer filtro de las variables disponibles en la aplicación de Bloomberg. Sobre estas se ha realizado un filtro académico a través de diferentes especialistas en diferentes aspectos de la sostenibilidad hasta llegar a la siguiente lista de variables: Emisiones de gases de efecto invernadero de Scope 1 - 2 (Por localización) – 2 (Por Mercado) – 3, carbono emitido por unidad de producción, Energía Total Consumida, Energía Renovable Consumida, Residuos Totales, Residuos Reciclados, Consumo Total de Agua, Porcentaje de Mujeres en la compañía, Porcentaje de Mujeres en Posiciones de Mando Medio, Porcentaje de Mujeres en Alta Dirección, Número de Incidentes Laborables Recogidos, Número de Empleados y las puntuaciones de Bloomberg de ESG, Medio Ambiente, Social y Gobernanza.

La siguiente fase del proyecto abordó la decisión sobre la longitud temporal en la que los datos de las empresas se enmarcan. Con intención de ofrecer el mayor horizonte temporal de los datos posible y enmarcarlo en el estudio de 2021 de Kearney se ha escogido el intervalo entre los años 2015 y 2022. Para la fecha de 2022 se encuentran ciertas limitaciones en la población de las variables de las empresas por lo que para algunos de los estudios se utiliza 2021 como último año.

Por último, antes de aplicar modelos y empezar el estudio sobre los datos, se partió a la recolección de los datos. Esta recolección se llevó a cabo a través de un equipo especializado de Kearney con el objetivo de descargar las bases de datos con el menos coste monetario y temporal posible. También por la inmensidad de variables y de empresas que se listaron el proyecto se ha planteado para incluir 113 empresas del sector eléctrico a nivel mundial organizadas por valor de compañía.

# Capítulo 2.- Descripción de las tecnologías (estado de la técnica)

El trabajo se encuentra en un estado de la técnica desarrollado, aunque se presenta como diferencial respecto a los trabajos previos encontrados por su profundidad y su variedad de algoritmos empleados.

Por el lado de la interacción entre las variables financieras existen numerosos índices de sostenibilidad que tratan de agrupar las variables de ESG para mayor facilidad de interpretación. En el trabajo se utilizan las variables de índice de sostenibilidad de Bloomberg para estudiar el vínculo que estas tienen con las financieras y como manera de enriquecer las variables.

Ejemplos de índices de sostenibilidad se pueden encontrar en los índices de S&P para financiación verde, financiación social, inversión de impacto (S&P23) y el “Dow Jones Sustainability Index” que se enfoca en agregar a las 2,500 compañías más grandes del “Dow Jones Global Total Stock Market Index” y seleccionar al top 10% en materia de variables ESG (BBVA22).

A su vez, se pueden destacar publicaciones previas estudiando el vínculo entre las variables ESG y las variables económicas de empresas. El trabajo realizado por Kearney, comentado anteriormente, para identificar la posición de las compañías en diferentes sectores respecto a variables medioambientales centro su estudio sobre las emisiones por lo que se amplía, en un número menor de sectores, el número de variables de estudio (BAK21). Una tesis realizada en 2019 por la Universidad de Lausanne buscando una relación entre ESG y el ROE encontró vínculos positivos de las variables ESG aunque su estudio se aplica con uso considerablemente reducido de algoritmos y estudios estadísticos (RADIC19). Por último, NYU llevó a cabo un estudio sobre 1,000 estudios previos sobre la materia que llegó a la conclusión de que el 58% de los estudios corporativos llegaban a la misma conclusión de que existía un impacto positivo, especialmente en el ROE, ROA y el precio de la acción (WHELAN21).

Sobre las técnicas de análisis estadístico y algorítmico empleadas cabe comentar que todas ellas son técnicas con pasado académico establecido para todos los usos empleados. El uso para este tipo de estudio, revisando publicaciones anteriores, resulta novedoso salvo por el uso de correlaciones, las cuáles son un método común de encontrar patrones sencillos que vinculen el comportamiento de variables. El resto de los algoritmos, por su uso para confirmar modelos predictivos resultan innovadores respecto al estado de la técnica actual para este tipo de trabajo.

Las técnicas se han aplicado primero con un filtrado de los datos recibidos por parte de Kearney directamente sobre Excel para tener un formato estandarizado y digerible por las aplicaciones de estudio de datos.

Posteriormente los datos se han procesado mediante la aplicación de datos masivos KNIME. KNIME permite la aplicación de algoritmos basados en código de Python de manera masiva lo cual permite una labor extensiva de estudio sobre bases de datos masivas como la requerida en este estudio. A lo largo del trabajo se describirán los diferentes bloques utilizados, así como su uso cuando sean novedosos. Cuando los bloques de código sean reutilizados en los diferentes casos no se hará el mismo nivel de inciso en ellos con el objetivo de evitar repetitividad y extensión innecesarias.

# Capítulo 3.- Modelo desarrollado

Como se ha mencionado anteriormente, la aplicación seleccionada para la realización de los modelos es KNIME. Para comprender el modelo desarrollado es importante introducir las variables seleccionadas para los datos. Una revisión extensiva de los datos mediante correlaciones, árboles de decisión y bosques aleatorios permite entender las relaciones entre variables más fuertes. Tras ello una revisión de los datos a través de dos variables estudiadas como relevantes para visualizar los datos de las empresas estudiadas. Por último, se ha realizado un análisis predictivo de una variable relevante de inversión a través de las variables de ESG con el objetivo de determinar su eficacia a la hora de realizar predicciones, su influencia y si existe algún patrón o algoritmo de preferencia para los clústeres seleccionados.

## 3.1.- Variables

El estudio de las variables financieras es necesario para poder comprender la salud general de la empresa y la percepción que pueden tener los accionistas de ella. Con la intención de representar el máximo posible de variables necesarias se han escogido las siguientes entre las disponibles en “CapitalIQ”:

- Deuda Neta: Representa la parte de la deuda de la empresa no compensada por activos líquidos.
- Deuda Total: Representa el montante de deuda de la empresa.
- Ingresos: Representa los ingresos brutos de la empresa.
- EBITDA: Beneficio restante antes de aplicar impuestos, intereses y depreciaciones.
- EBIT: Beneficio restante antes de aplicar intereses e impuestos.
- EBT: Beneficio restante antes de aplicar impuestos sobre los beneficios.
- Beneficios Netos: Valor obtenido por la empresa tras aplicar costes e impuestos.
- Inversión en inmovilizado (CAPEX): Inversión de la empresa en sus activos capitalizables y su mantenimiento.
- Capitalización Bursátil: Valor total de las acciones de la empresa. El valor de las acciones es el reportado en las cuentas anuales como último del año contable en cuestión.
- Valor de la empresa (EV): Valor total de la empresa al añadir su capitalización, deuda neta y sus intereses minoritarios.
- EV/EBITDA: Ratio utilizado para medir el valor de una empresa respecto a su EBITDA. Suele utilizarse para ver de manera sencilla si una empresa se encuentra sobrevalorada o infravalorada respecto a empresas de su sector.
- Intereses Minoritarios: Representa el valor minoritario que controla la empresa en otras.
- Efectivo de las operaciones: La cantidad de caja obtenida de las operaciones regulares de la empresa.

- Gastos provenientes de intereses sobre la deuda: Intereses pagados sobre la deuda de la empresa.

Además de esas variables directamente obtenidas de “CapitalIQ” se han creado una serie de variables financieras derivadas a lo largo del trabajo para poder profundizar el estudio. Las variables financieras derivadas son:

- EBITDA Ratio: La ratio de EBITDA representa la parte de los ingresos brutos que se convierten en EBITDA tras los costes devengados de vender los productos (COGS).
- Coste de la Deuda: Representa el porcentaje pagado de la deuda en intereses. El coste se ha calculado dividiendo los intereses entre la deuda bruta. Este indicador no representa el coste de deuda exacto de la empresa ya que los valores de deuda bruta y los valores de intereses reportados por “CapitalIQ” no incluyen gran parte de la complejidad de productos de la deuda de una empresa. Aun así el coste calculado es una aproximación muy cercana y permite enriquecer de gran manera el estudio.
- Profit Ratio: Representa la ratio de ingresos brutos que se convierten en beneficio neto.
- Ratio de Caja de Operaciones: Porcentaje de los ingresos que se convierten en la caja.
- EBITDA/Interés: Indicador utilizado para reflejar el coste que representan los intereses sobre el EBITDA y por tanto la capacidad que tiene la empresa para pagarlos.
- ROE: Ingresos netos de la compañía divididos entre la capitalización de la empresa. Es un indicador de rentabilidad de la empresa para los accionistas.

Por la parte de las variables medioambientales se han obtenido a través de “Bloomberg” con el objetivo de dar una representación lo más diversa posible de los esfuerzos en sostenibilidad de las empresas. Las variables ESG son las siguientes:

- Emisiones de gases de efecto invernadero de Scope 1
- Emisiones de gases de efecto invernadero de Scope 2 (Por localización)
- Emisiones de gases de efecto invernadero de Scope 2 (Por Mercado)
- Emisiones de gases de efecto invernadero de Scope 3
- Carbono emitido por unidad de producción
- Energía Total Consumida
- Energía Renovable Consumida
- Residuos Totales
- Residuos Reciclados
- Consumo Total de Agua
- Porcentaje de Mujeres en la compañía
- Porcentaje de Mujeres en Posiciones de Mando Medio
- Porcentaje de Mujeres en Alta Dirección
- Número de Incidentes Laborables Recogidos
- Número de Empleados



Respecto a variables derivadas las primeras que merecen mención son los propios índices directamente extraídos de Bloomberg y que son las puntuaciones de Bloomberg de ESG, Medio Ambiente, Social y Gobernanza. Además, durante el trabajo se ha creado una subvariable de “Residuos Reciclados” como porcentaje de la división entre “Residuos Reciclados” y “Residuos Totales”.

A su vez, con el objetivo de evitar un análisis con mucha sensibilidad respecto a la escala se han introducido subvariables de ESG dividiéndolas entre los ingresos respectivos de cada empresa. La razón por la que se han escogido los ingresos de las empresas como estándar de la escala es por ser el indicador financiero menos distorsionado por las operaciones particulares de cada empresa y por ser un dato poblado al completo para todas las empresas en todos los años de estudio.

### 3.2.- Estudio de Relaciones

Tras obtener las bases de datos deseadas es necesario adecuarlas al funcionamiento del programa, crear las subvariables requeridas y aplicar la programación necesaria para obtener los resultados deseados. En este estudio de relaciones entre las variables se llevan a cabo dos etapas estandarizadas para diferentes agrupaciones de los datos incluyendo un estudio de correlaciones de todas las variables contra el resto y un estudio de variables de ESG predictoras de variables financieras a través de algoritmos de árboles de decisión y de bosques aleatorios.

Debido a que es la primera vez que se explica el uso del modelo se va a proceder a la explicación del funcionamiento de los diferentes bloques usados. Como se ha mencionado antes esta explicación no se repetirá durante el trabajo y solo se dedicará a aquellos bloques que sean novedosos respecto a apartados anteriores del proyecto.

Para la creación de un modelo sobre datos masivos en KNIME se deben realizar una serie de etapas para completar la gestión de los datos masivos.

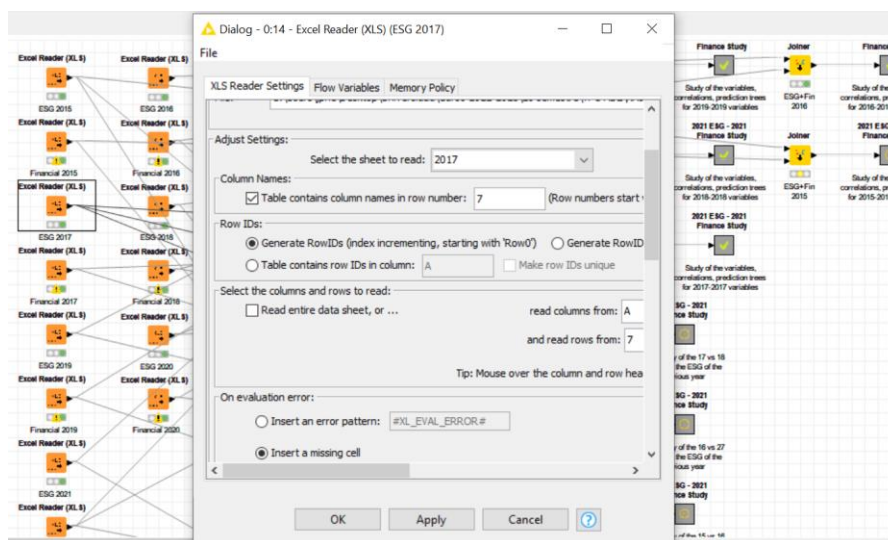


Figura 2: Menú de configuración del bloque lector

La primera de estas etapas es el acceso a los datos que en KNIME se lleva a cabo a través de uno de los diferentes tipos de “Reader”. Debido a que nuestros datos se encuentran en un formato de Excel se utiliza el “Excel Reader”. En la configuración del bloque se selecciona la dirección de los datos, las celdas que se utilizan como referencia y la generación de columnas y filas identificadoras o usar ya existentes. Por último, el bloque permite una demostración del output obtenido para acelerar el proceso de acceso a los datos.

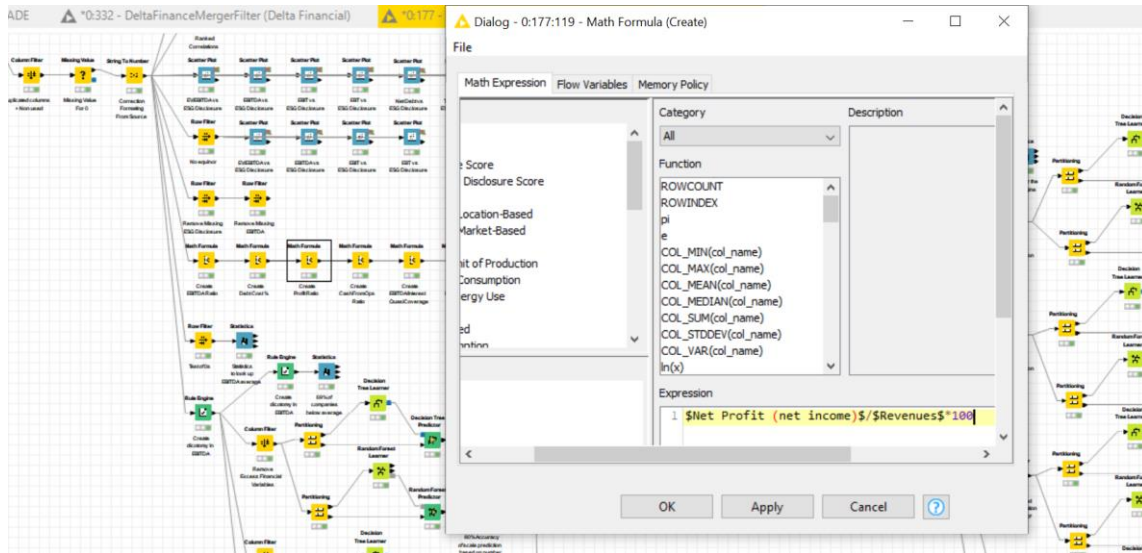


Figura 3: Menú de configuración de fórmulas matemáticas

La segunda de estas etapas es la transformación de los datos. La transformación de los datos es la fase más importante del modelo ya que implica la adaptación de los datos del mundo real a un formato digerible por los diferentes algoritmos que apliquemos. Para nuestro estudio de correlaciones y predictivo debemos adecuar los datos introducidos a los algoritmos de bosques aleatorios y árbol de decisión. Esto necesariamente por formatear la variable a predecir en una variable binaria.

Los bloques que se incluyen dentro de la transformación son bloques de filtrado de columnas y filas, valores perdidos, conversión de formatos, renombrar, particiones, fórmulas matemáticas, bloques de reglas, juntar bases de datos y normalizadores. Los bloques vienen identificados con el color amarillo dentro del programa.

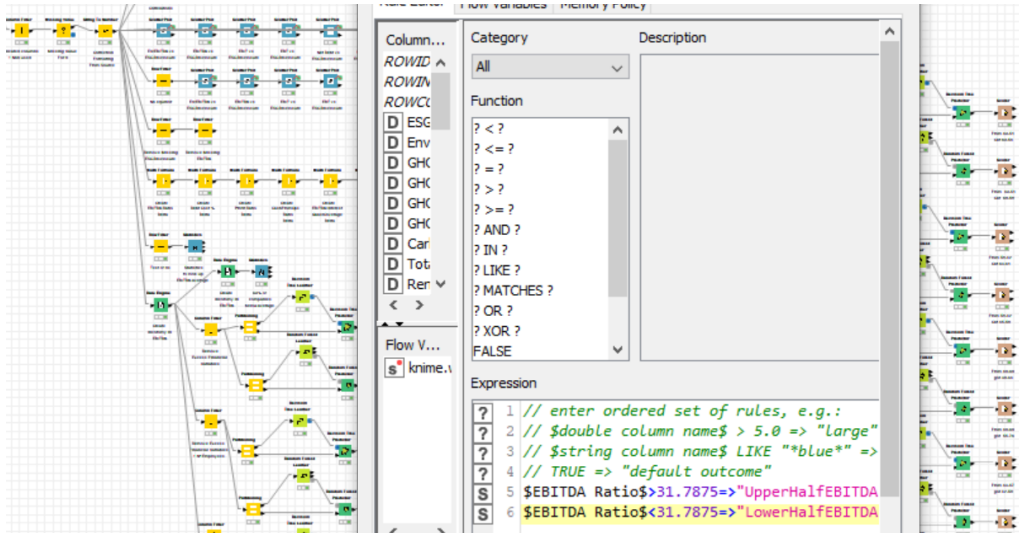


Figura 4: Menú de configuración de un bloque de reglas

Durante la transformación de datos, la conversión de variables financieras para la predicción se ha llevado a cabo con un proceso iterativo de transformación para determinar una mediana aceptable de los datos con la que establecer dos grupos con una población de empresas que superan a la media en esa variable financiera y otra que se encuentra por debajo. En casos en los que el baile de datos era muy diferente a una campana de gauss centrada se ha optado por establecer la media como punto de ruptura entre los grupos y destacar el porcentaje de las muestras que se encuentran en cada uno.

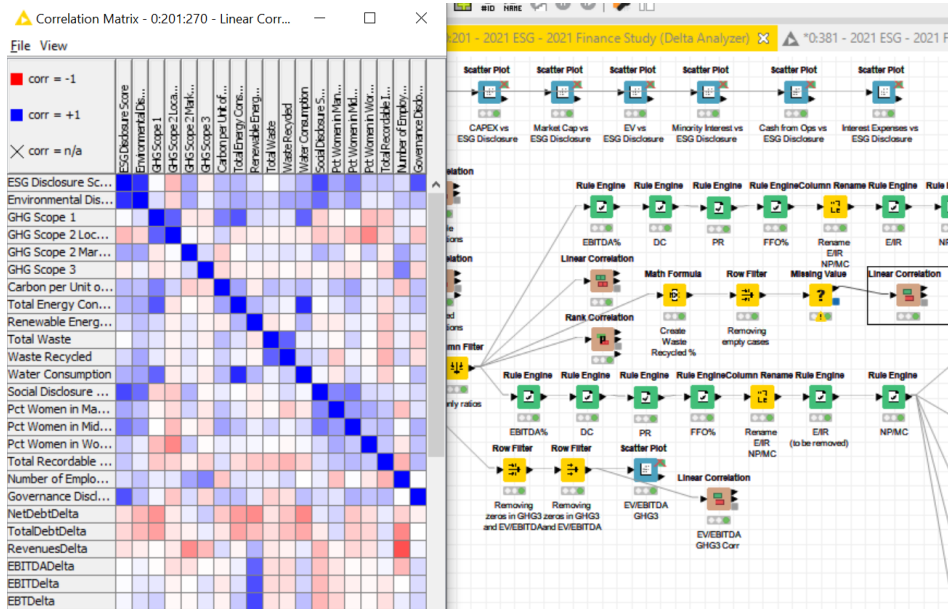


Figura 5: Matriz de correlaciones

La correlación forma parte de la tercera fase de un modelo, el análisis y la minería de datos. Para realizar esto el programa cuenta con un bloque de Python que permite una visualización fácil de estos datos mediante una tabla de colores. A su vez para mayor facilidad de extracción de conclusiones se han realizado tablas con los vínculos de

correlación más destacables y los cuáles se comentarán para cada uno de los casos realizados.

Tras el análisis de las correlaciones más fuertes entre variables se lleva a cabo un análisis con algoritmos de árbol de decisión y de bosques aleatorios. La razón principal por la que el uso de estos algoritmos es tan potente se debe a que permiten una visualización de las variables escogidas por el algoritmo como mejores predictores de las variables financieras. En casos con porcentajes de predicción muy superiores a la media nos permite hacernos a la idea de posibles vínculos entre ciertas variables que no se detectan con tanta facilidad con la realización de estudios puramente basados en la correlación.

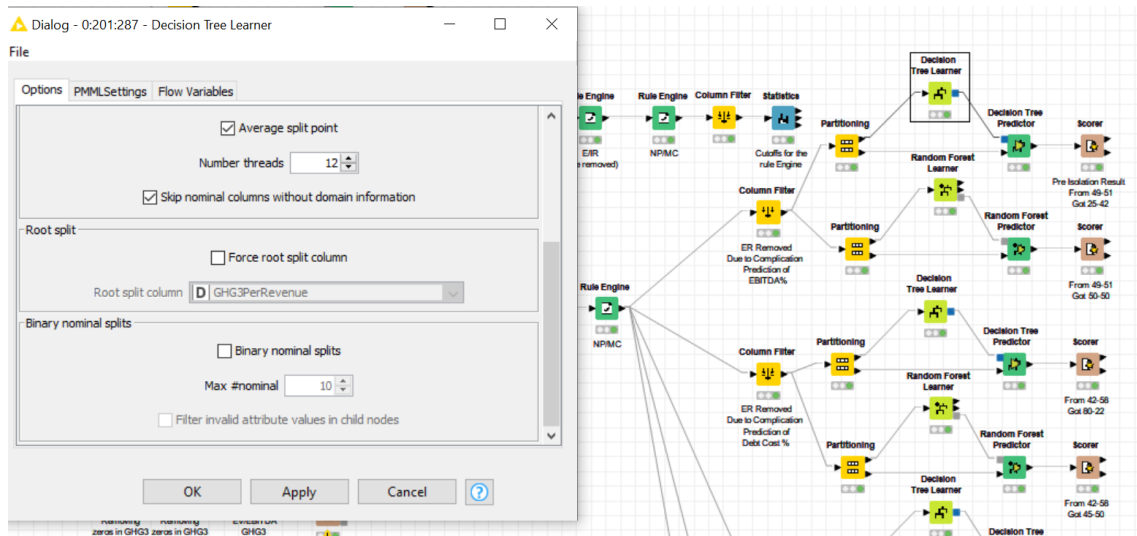


Figura 6: Menú de configuración del bloque de aprendizaje

Para la realización de un análisis de predicción con estos algoritmos se debe primero adecuar las variables a predecir y realizar una partición aleatoria de las compañías entre el grupo de entrenamiento y el grupo de práctica.

Debido a que la muestra poblacional de empresas (n=113) no es demasiado grande se ha optado por una separación entre los dos grupos de 80-20. Esta separación es óptima ya que permite una gran representación de datos para el entrenamiento sin comprometer demasiado la representatividad de las predicciones ya que si el grupo de práctica fuera demasiado pequeño la aleatoriedad sería muy determinante en el éxito de las predicciones. En las predicciones realizadas durante el proyecto se revisa la kappa de Cohen como indicador para analizar la replicabilidad de las predicciones de los estudios. En muchos de los casos en los que se han obtenido kappas de cohen pequeñas se ha reintroducido la predicción hasta conseguir un resultado suficientemente estable.

Por último, es relevante destacar que en las tablas en las que se plasma el éxito de la predicción de los algoritmos se incluye el éxito de predicción para los dos grupos con el objetivo de minimizar el error en ambos y poder ver casos en los que se mejore la predicción de uno de los grupos y no del otro.

### 3.2.1.- Caso Base

El caso base explora el estudio de las correlaciones de variables y predictivos de todas las variables financieras del año 2021 sin realizar modificaciones contra las variables de ESG del año 2021 sin realizar modificaciones. Este estudio considera aceptable que vínculos que se puedan presentar tengan en cuenta no solo comportamientos comunes a otros años sino también únicos a este año.

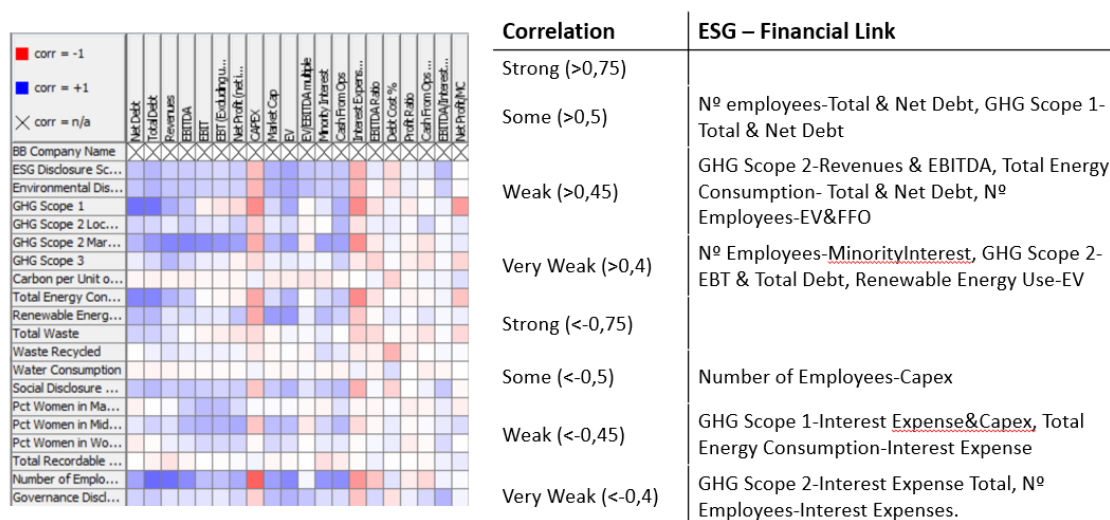


Figura 7: Resultados de correlación del caso base

En la figura 7 se pueden ver los resultados de la correlación del caso base. Para este y otros casos se excluye destacar los casos en los que la correlación, tanto positiva como negativa, es muy débil. En el caso base se puede ver que no existen correlaciones fuertes entre variables de sostenibilidad y financieras. Aun así, sí que existe alguna correlación entre variables y en muchos casos correlaciones débiles o muy débiles.

En todos los casos las correlaciones tienen sentido y demuestran un vínculo lógico entre las variables de ESG y las financieras. Es merecido destacar que las variables de gases de efecto invernadero demuestran un vínculo débil con a la facturación y el EBITDA que estaría alineado con el alza de precios dramáticos durante la crisis provocada por el conflicto en Ucrania.

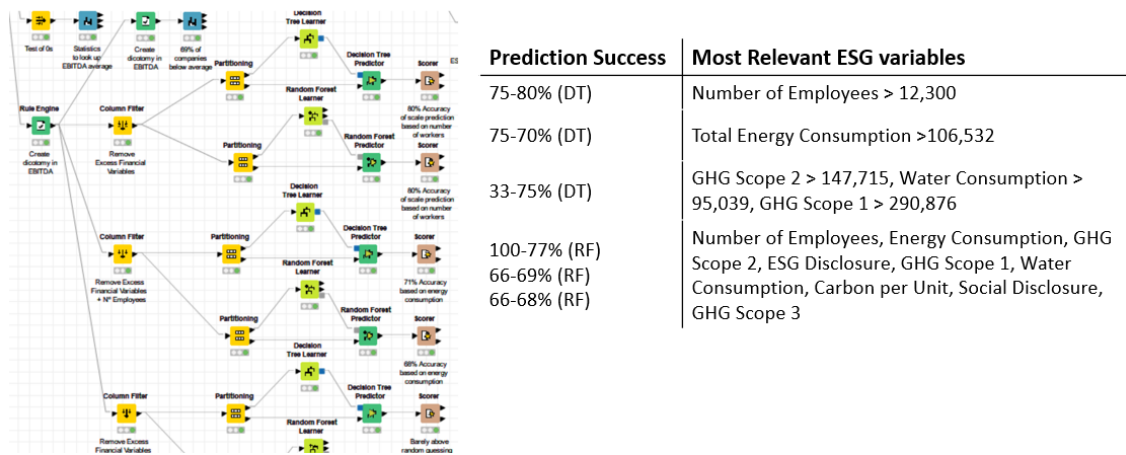


Figura 8: Resultados de predicción con DT y RF en el caso base

En la figura 8 se pueden ver los resultados de predicción de los algoritmos de árbol de decisión (DT) y de bosques aleatorios (RF) para el caso base. En este primer caso base se destacan las variables de ESG más relevantes, aunque en los posteriores se destacará principalmente las ramas primarias.

Para este caso base se ha empleado una dicotomía respecto al EBITDA generado por las empresas. Este razonamiento de análisis se ha empleado ya que permite conocer las empresas más eficientes operativamente y su vinculación con el ESG. Cabe destacar que a priori se puede entender que estas variables pueden indicar a subgrupos dentro de las empresas de energía basadas en su operativa o su localización.

La distribución de los datos en esta dicotomía es 31-69 para el grupo por encima de la media y por debajo. De los bloques en los que se obtiene una predicción mayor se puede ver que el número de empleados y la energía consumida mejoran los resultados de predicción para ambos grupos. Estas variables medioambientales tienen un carácter muy vinculado a la escala y carecen de una perspectiva de crecimiento temporal por lo que se requieren más casos.

### 3.2.2.- Caso Ratios Financieros

El siguiente caso representa una actualización respecto al caso base para tratar de compensar la influencia de la escala en las variables de ESG. Para ello se escogen los ingresos como variable de referencia de la dimensión de las empresas y se utiliza para compensar las variables financieras que serán utilizadas para la correlación y la predicción.



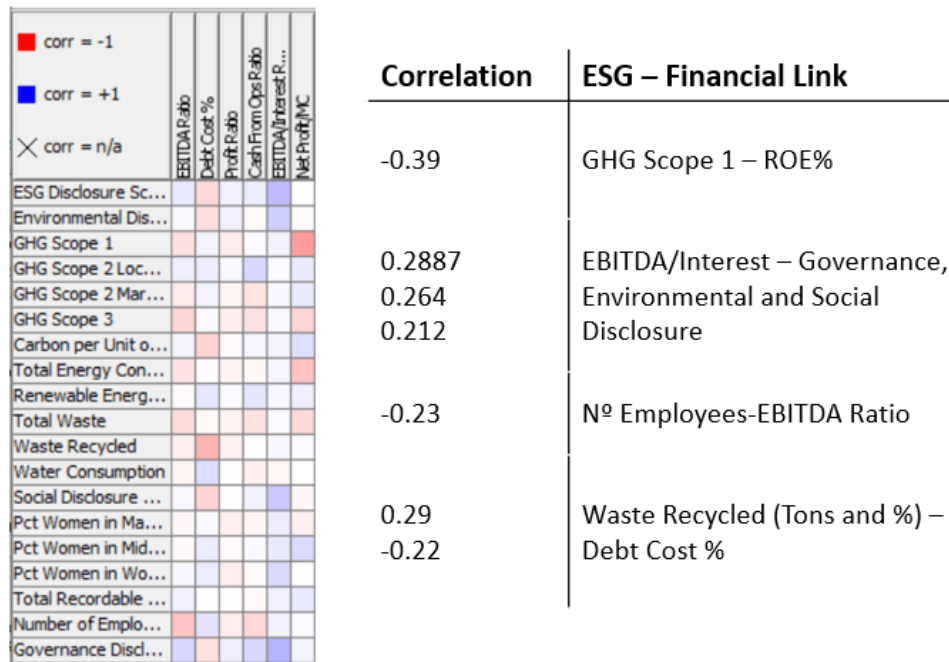


Figura 9: Resultados de correlación para el caso de ratios financieros

De los resultados de correlación para este caso cabe destacar primero que las correlaciones son por lo general más débiles. Aun así se puede ver que el emplear estos ratios financieros ya nos permite sacar conclusiones sobre la naturaleza de la relación entre las variables de ESG y las financieras. La relación indirecta entre las emisiones de primer alcance y la rentabilidad de la inversión demuestran que existe un vínculo positivo que recompensa a las empresas menos contaminantes en su generación de valor al eliminar el factor de la escala. Otras conclusiones, con mayor debilidad por su débil correlación, son la mejora de las condiciones de pago de intereses respecto al EBITDA para empresas que se alinean con los índices de Bloomberg y menores costes de deuda en empresas que reciclan más en porcentaje, pero no en total.

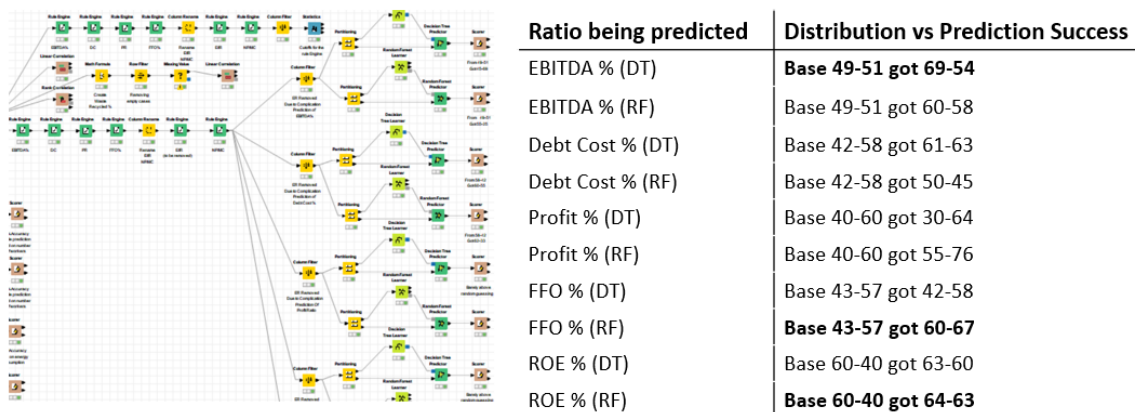


Figura 10: Resultados de predicción con DT y RF en el caso de ratios financieros

En preparación de realizar la predicción con muchas más variables hay que agregar reglas lógicas para cada uno con el bloque de fórmulas matemáticas y escoger un punto de corte

diferente para cada variable para conseguir una distribución adecuada entre los grupos para las predicciones de los algoritmos.

Para todos los algoritmos y en todos los ratios se puede ver una mejora de la predicción respecto a una distribución media estrictamente aleatoria. Aun así siguen, como se ha podido ver en la correlación, es difícil ver la influencia de manera adecuada debido al efecto de la escala sobre las variables de ESG que llevan a casos extraños como el del reciclaje bruto y porcentual.

### 3.2.3.- Caso Ratios ESG

En el siguiente caso se trata de compensar la influencia de la escala de las empresas en relación con las variables de sostenibilidad de las empresas. Para ello se escogen todas las variables con valores brutos y las que ya son variables que no dependen de la escala.

	Energ...	Energ...	Total...	Waste...	Water...	Empl...	GHG...	GHG...	GHG...	Correlation	ESG – Financial Link
Net Debt										0.66	GHG Scope 1&2 per Revenue – Profit Ratio
Total Debt											
Revenues										0.5607	Energy Consumption per Revenue – Profit Ratio
EBITDA											
EBIT											
EBT (Excluding un...)											
Net Profit (net inc...)											
CAPEX										0.34	Total Waste per Revenue – Debt Cost % & Profit Ratio
Market Cap										0.33	
EV											
EV/EBITDA multiple											
Minority Interest											
Cash From Ops										0.32	Debt Cost % - Nº Employees per Revenue
Interest Expense...											
EBITDA Ratio											
Debt Cost %											
Profit Ratio										-0.14	Waste per Revenue – Debt Cost %
Cash From Ops R...											
EBITDA/Interest ...											
Net Profit/MC											

Figura 11: Resultados de correlación para el caso de ratios ESG

Solo con la aplicación de los ratios ESG la correlación entre las variables mejora razonablemente y permite ver más razonamientos que vinculan las variables de las empresas. La relación para el año 2021 de las variables de gases de efecto invernadero de alcance 1 y 2 emitidos partidos de los ingresos con la conversión a beneficios puede verse altamente influenciada por el drástico incremento de los precios de las materias primas en el mundo a consecuencia del conflicto en Ucrania. Aunque el vínculo no sea el habitual, derivado de la eficiencia y reclamo de políticas sostenibles, este tipo de vínculo con la geopolítica también es signo de la influencia de las variables no económicas sobre las económicas.



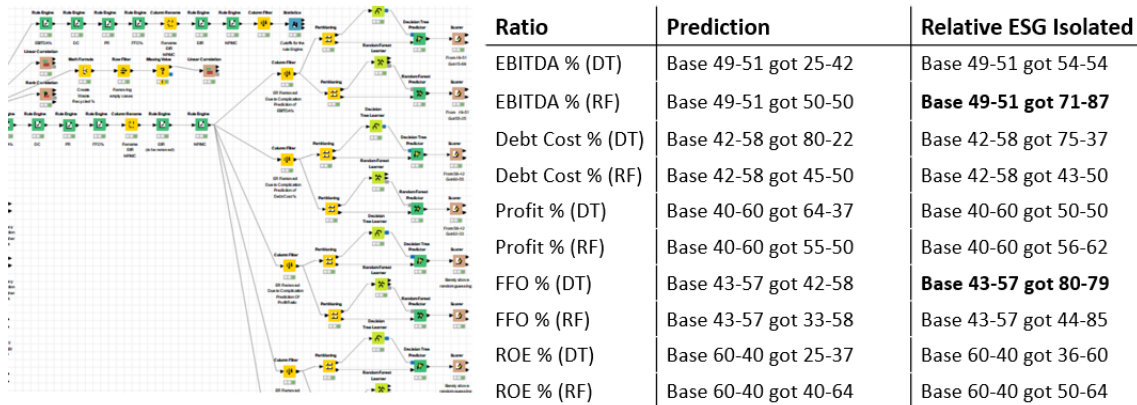


Figura 12: Resultados de predicción con DT y RF en el caso de ratios ESG

La predicción de las variables financieras a través de los algoritmos de árbol de decisión y bosques aleatorios se polariza respecto al caso anterior al incrementar a niveles altos para algunas variables y empeora considerablemente para otros con algunos casos encontrándose alineado con una muestra aleatoria.

### 3.2.4.- Caso Deltas Financieras

El caso anterior, aunque destaca las capacidades de predicción y de correlación de las variables de ESG sobre las financieras, deja algunas conclusiones que se pueden afinar con un análisis más profundo.

Por un lado, se ve como las compañías con escala a través del número de empleados e ingresos, entre otras variables, tienen altas correlaciones positivas con variables de ESG. Esto se puede ver a través de la escalabilidad que permite un mayor empleo de la eficiencia que permiten los beneficios de una política ESG avanzada en la financiación y en la nueva inversión en proyectos de gran escala necesarios para la transformación de las economías a nivel mundial en el sector energético.

A su vez, en las predicciones de los casos, las variables preferidas por los algoritmos de predicción de árbol de decisión y de bosques aleatorios suelen ser de escala (número de empleados, consumo de energía, emisiones y consumo de agua). Es por ello por lo que la siguiente fase del estudio va a buscar encontrar correlaciones y predicciones a través de la diferencia de los valores económicos respecto al año anterior. Esta delta se aplica sobre las variables financieras brutas y los ratios ya que la evolución de ambas resulta de interés para ver la correlación con las variables de ESG. Para poder realizar comparaciones con la correlación en las variables de 2021 se aplica la delta entre los años 2020 y 2021.

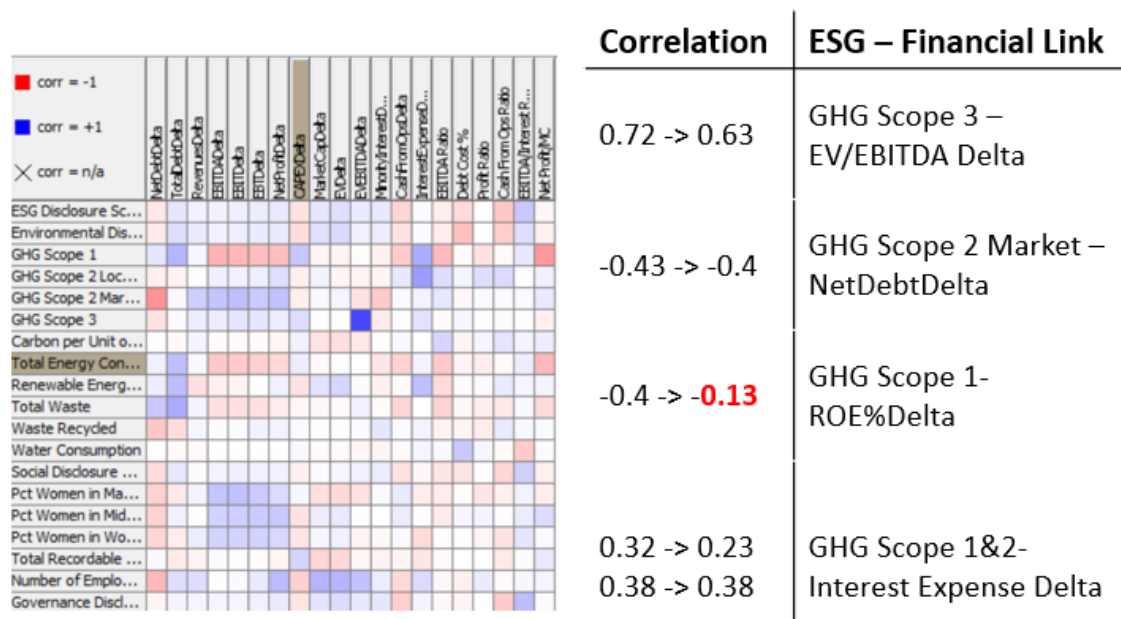


Figura 13: Resultados de correlación para el caso de ratios financieros con variables de ESG brutas

En la figura 13 se puede ver el resultado de la aplicación de las deltas de variables financieras y compararlas con las correlaciones obtenidas con esas mismas variables financieras solo contabilizadas para el año 2021. Las correlaciones entre los cambios de las variables financieras se encuentran en línea con las de los casos anteriores.

Los cambios de valoración entre los años 2020 y 2021 se encuentran muy influidos por las emisiones de gases de efecto invernadero a la vez que sus cambios en deuda neta. La rentabilidad de sus operaciones ha permitido acumular grandes cantidades de caja a empresas emisoras lo que les ha servido para reducir la deuda neta, aunque su deuda bruta se mantenga a niveles similares a otros años. Los costes por intereses sin embargo siguen una relación positiva con las emisiones con lo que se podría razonar un efecto de los cambios en la financiación ofrecida por los grandes bancos.

Es también destacable como la relación con la rentabilidad que existía para el año 2021 se debilita considerablemente al aplicar la delta respecto al año anterior.

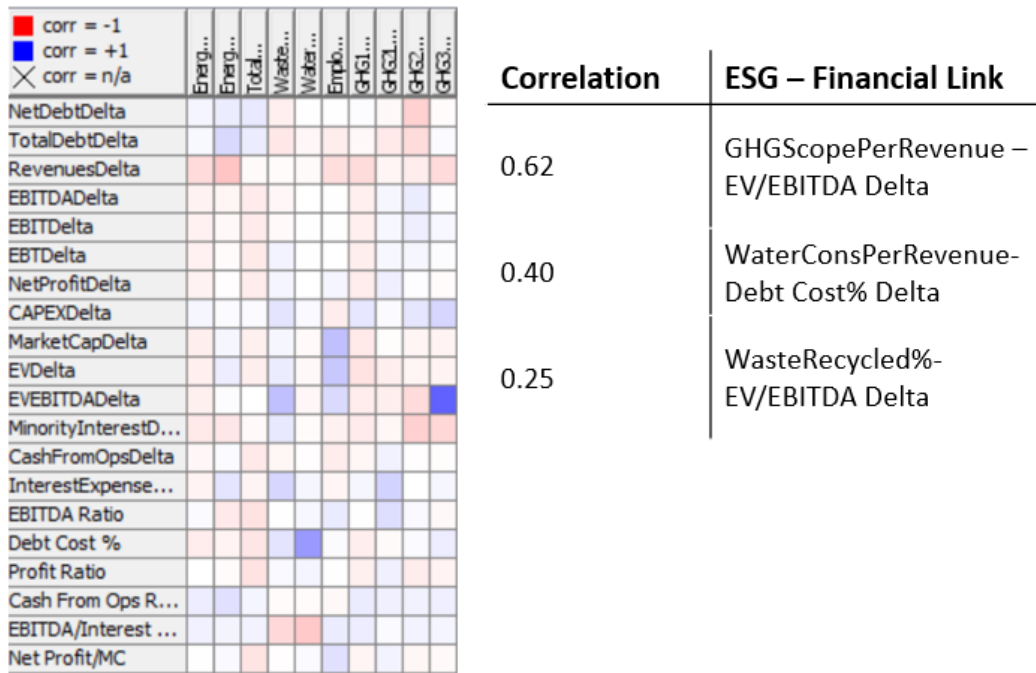


Figura 14: Resultados de correlación para el caso de ratios financieros con variables de ESG divididas entre ingresos

Al realizar el análisis con las variables de ESG con el efecto de la escala compensado nos encontramos la misma correlación en la emisión de gases de efecto invernadero además de una correlación débil entre el consumo de agua alto y el coste de deuda y entre la valoración y los residuos reciclados.

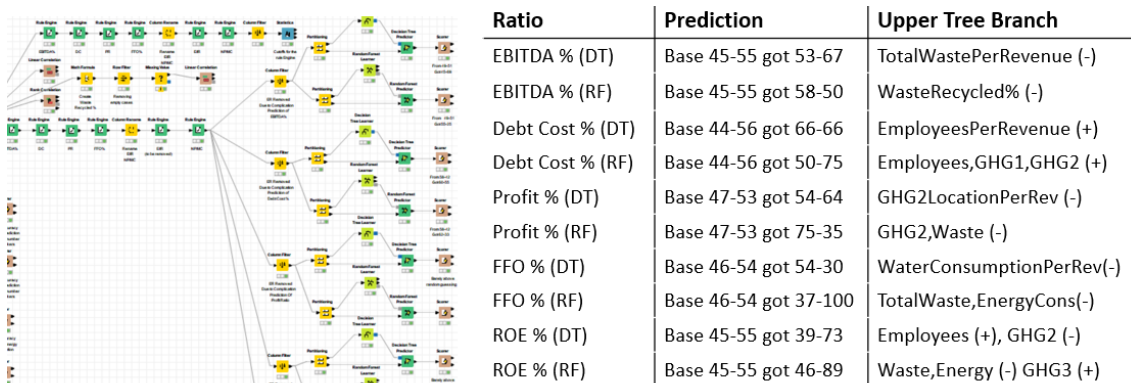


Figura 15: Resultados de predicción con DT y RF en el caso de deltas financieras

Al realizar la predicción se ve que se mejora ligeramente respecto a la media. Las variables que los algoritmos de DT y RF priorizan suelen tener efecto de la escala lo que puede indicar un importante papel en la relación entre las variables de ESG y financieras.

El signo en la figura 15, en la sección que describe las variables que conforman la primera selección dicotómica, define la relación con la variable a predecir. Por ejemplo, en el primer caso del EBITDA % con árbol de decisión la variable ESG más relevante es el residuo generado total entre ingresos. La relación es negativa por lo que a menor residuos

generados por ingresos el algoritmo predice que la empresa tiene mayor porcentaje de EBITDA %.

Otro aspecto que cabe destacar en los resultados de esta predicción es como para algunas de las variables la predicción general es mejor que el aleatorio, pero para uno de los dos grupos es peor que la media. Esta situación se presenta ya que el algoritmo prioriza crear un árbol de decisión que maximice el existo de predicción combinado. Es también importante monitorizar esta casuística ya que el algoritmo se puede sobre entrenar para predecir uno de los dos grupos o tratar de predecir todos los grupos de datos como si perteneciesen a un grupo únicamente.

### 3.2.5.- Caso Deltas ESG

De la misma manera que no solo las variables reportadas en el año, sino que su progresión importan para las variables financieras este mismo estudio se ha realizado con las variables de ESG. Para esta correlación se incluyen tanto las variables financieras brutas como los ratios.

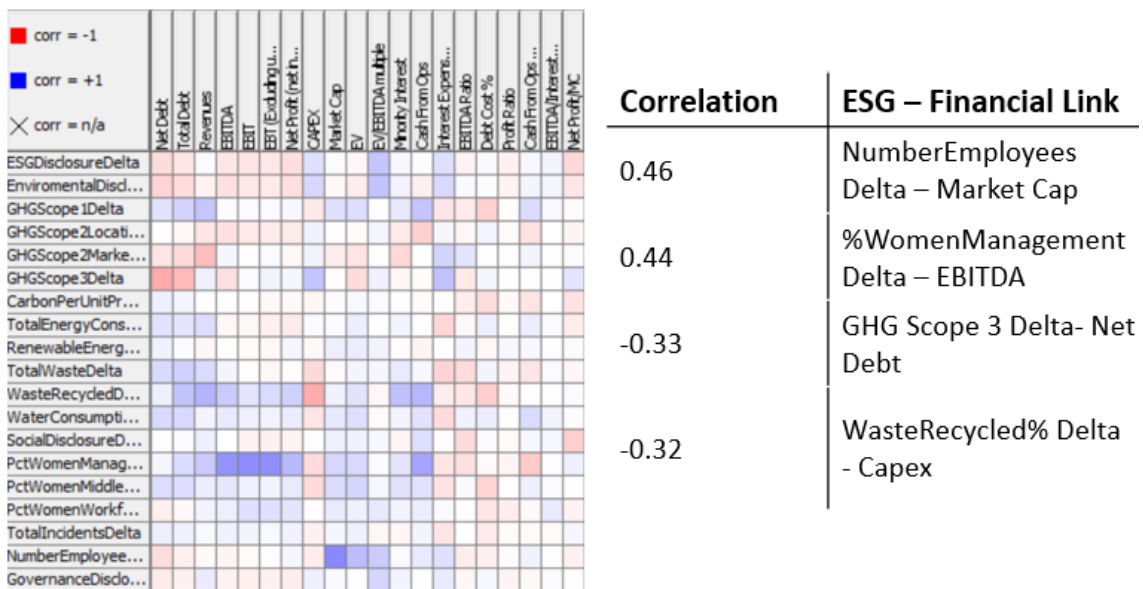


Figura 16: Resultados de correlación para el caso de deltas ESG

Del resultado de la correlación se puede ver correlación débil entre variables de ESG y financieras. La relación entre el incremento de empleados y el valor de la empresa tiene lógica, aunque la relación se puede entender como de doble dirección, la empresa se encuentra con buena salud y en expansión por lo que contrata a más empleados y esto a su vez le reporta más inversión. El incremento del porcentaje de mujeres en las empresas con mayor EBITDA, la deuda neta alta en empresas que reducen sus emisiones de alcance 3 o alto CAPEX (la relación es positiva ya que el valor se contabiliza en los datos como negativo) con el incremento del reciclaje en porcentaje tienen también correlaciones débiles y posibles razonamientos.

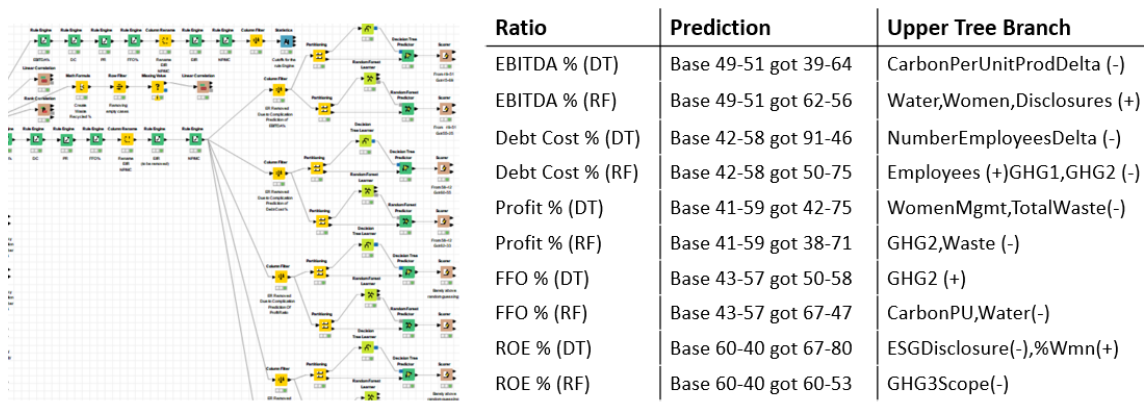


Figura 17: Resultados de predicción con DT y RF en el caso de deltas ESG

Con los resultados de la predicción se puede ver una mejoría general respecto a los resultados aleatorios, aunque no para todos los ratios financieros. A su vez, es meritorio destacar como no en todas las predicciones la variable de clasificación y su interacción con el ratio a predecir puede resultar lógico. La rentabilidad de la inversión tiene una relación positiva con el porcentaje de mujeres en altas posiciones de dirección, pero negativa con el índice de ESG realizado con Bloomberg. Por otro lado, la reducción del carbón emitido por unidad de producción como predictor de mayor EBITDA % puede resultar lógico por una generación de valor más eficiente, aunque la mejora de predicción no es muy alta.

### 3.2.6.- Caso ESG Previo

Uno de los problemas en el análisis de los casos que aparece es el estudio de que año de ESG afecta a que año de información financiera. En este análisis se pueden diferenciar dos maneras de atacar este caso.

Por un lado, existe el procedimiento estándar. Según este procedimiento la información publicada en el año 2021 se encuentra vinculada a la información financiera del mismo año. Esto en efecto significa que los proyectos anunciados, las noticias de la empresa y la información publicada trimestralmente vincula las variables de ESG y financieras dentro de un mismo año.

Por otro lado, se puede llevar a cabo un análisis reputacional. La información de ESG en los reportes de sostenibilidad suele publicarse con meses de diferencia respecto a la información financiera por lo que la única información de ESG que un inversor o acreedor tiene de la empresa durante el año de 2021, a ciencia cierta, es la del año 2020 que se publica normalmente en torno a mediados del año 2021. Esto conlleva que se puede argumentar que el casado de las variables de ESG se debería hacer con el año anterior por lo que se realizará ese caso de estudio.

Con el objetivo de analizar esta casuística se realiza el consiguiente análisis, incluyendo ratios financieros y de ESG, y comparando las correlaciones con las del estudio para el mismo año.

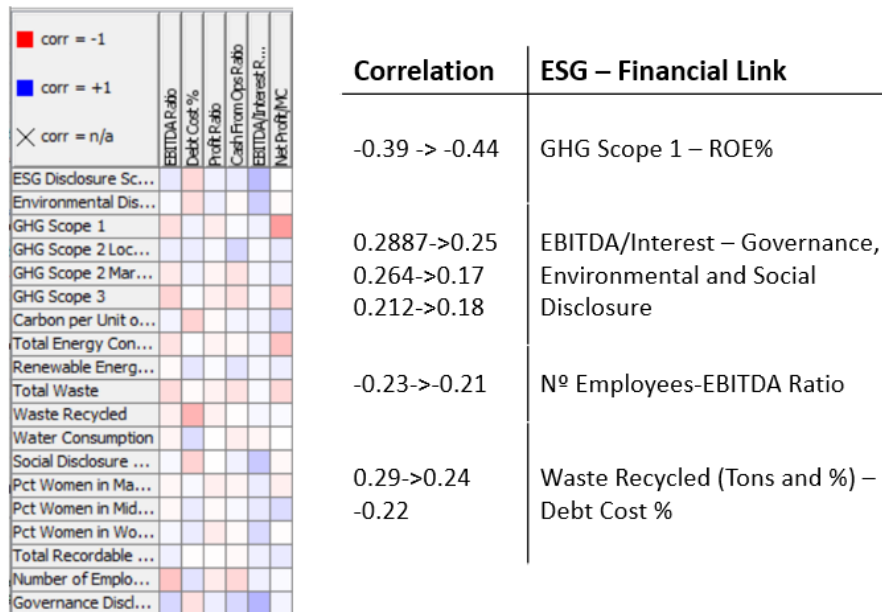


Figura 18: Resultados de correlación para el caso de ESG Previo

Los resultados corroboran que las correlaciones son similares, con algunas intensificándose y otras debilitándose. Las interacciones entre las variables no cambian por lo que a su vez demuestra que los análisis de los casos anteriores no sufren de una inestabilidad grande por el cambio entre años.

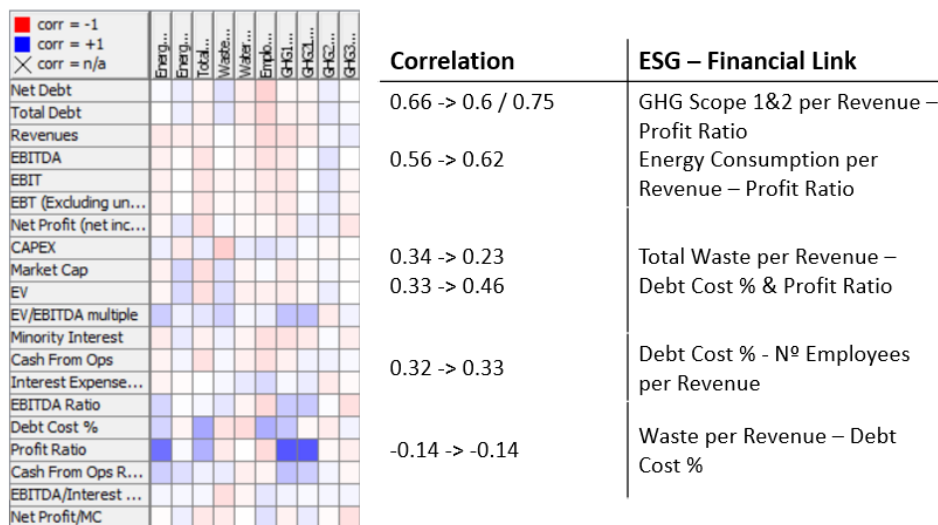
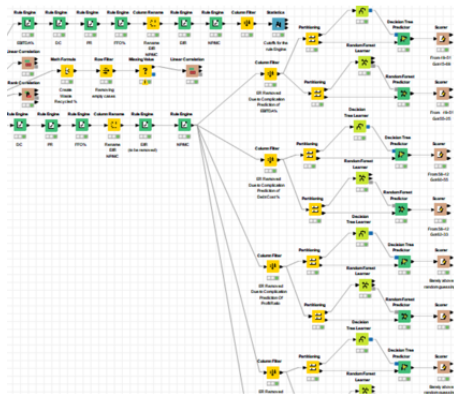


Figura 19: Resultados de correlación para el caso de ESG Previo con variables de ESG relativas

En la correlación usando variables de ESG relativas la correlación entre los gases de efecto invernadero se refuerza, aunque para el resto de las variables se mantiene en línea con el análisis para el mismo año de variables ESG.





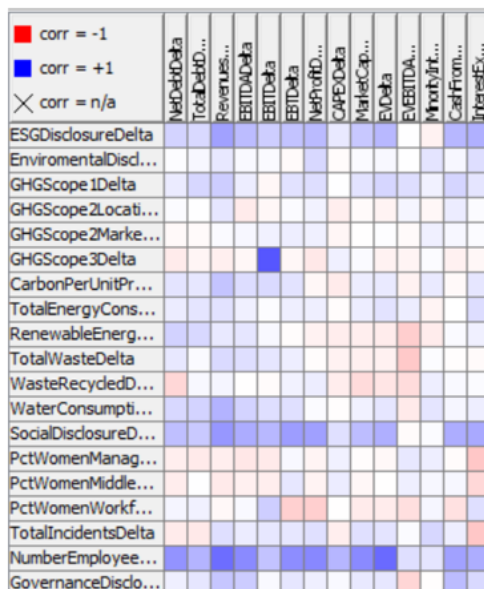
Ratio	Prediction	Relative ESG Isolated
EBITDA % (DT)	Base 49-51 got 25-42	Base 49-51 got 54-54
EBITDA % (RF)	Base 49-51 got 50-50	<b>Base 49-51 got 71-87</b>
Debt Cost % (DT)	Base 42-58 got 80-22	Base 42-58 got 75-37
Debt Cost % (RF)	Base 42-58 got 45-50	Base 42-58 got 43-50
Profit % (DT)	Base 40-60 got 64-37	Base 40-60 got 50-50
Profit % (RF)	Base 40-60 got 55-50	Base 40-60 got 56-62
FFO % (DT)	Base 43-57 got 42-58	<b>Base 43-57 got 80-79</b>
FFO % (RF)	Base 43-57 got 33-58	Base 43-57 got 44-85
ROE % (DT)	Base 60-40 got 25-37	Base 60-40 got 36-60
ROE % (RF)	Base 60-40 got 40-64	Base 60-40 got 50-64

Figura 20: Resultados de predicción con DT y RF en el caso de ESG previo

El análisis de predicción resulta de gran valor para este caso en el que se puede ver la gran diferencia en el éxito de predicción cuando se aíslan las variables de ESG relativas en vez de introducir todas las variables de sostenibilidad juntando brutas y relativas.

### 3.2.7.- Caso CAGR

Para analizar la evolución de las variables y su correlación un ejercicio que puede aportar resultados de interés es evaluar la correlación entre el CAGR (crecimiento acumulado compuesto) desde el año 2015 hasta el año 2021.



Correlation	ESG – Financial Link
0.66	GHG Scope 3 CAGR – EBIT CAGR
0.58	Number of Employees CAGR –
0.57	Revenues CAGR, EV CAGR, MC
0.46	CAGR
0.4	Social Disclosure CAGR – Revenues CAGR

Figura 21: Resultados de correlación para el caso de CAGR

Los resultados del análisis del caso CAGR se encuentran alineados con los de otros casos con correlaciones medias o débiles, aunque destaca que el comportamiento está relacionado con la emisión de gases de efecto invernadero como indicador de escala por ingresos a la vez que el efecto que tiene el crecimiento de empleados como un indicador del crecimiento de una empresa entre los años 2015 y 2021.

### 3.2.8.- Caso CAGR – Otros Algoritmos

En los casos previos se ha analizado utilizando algoritmos de árbol de decisión y bosques aleatorios. Estos algoritmos son útiles para el análisis debido a su construcción lógica y utilidad para predecir variables dicotómicas a través de una serie de variables continuas o dicotómicas.

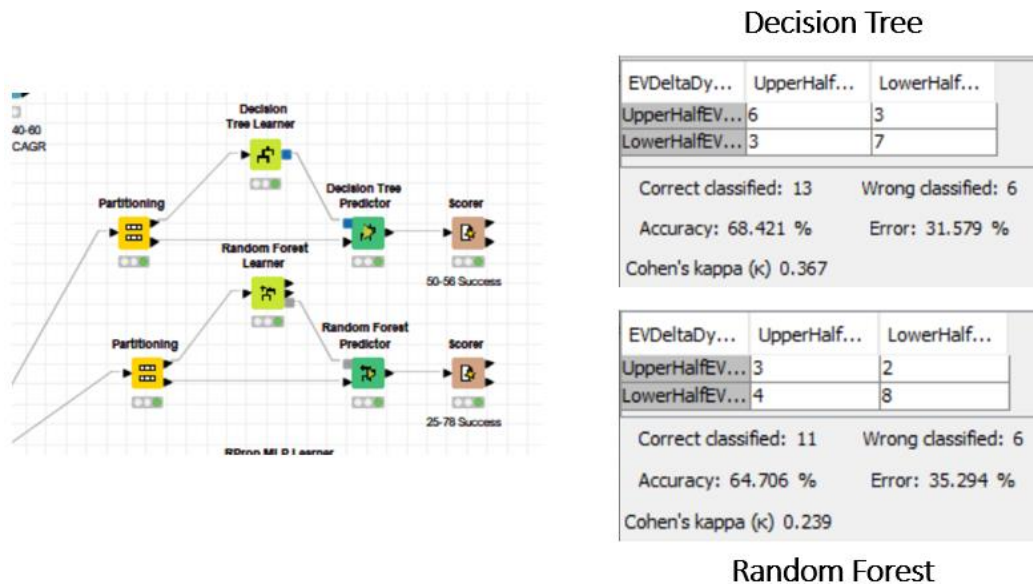


Figura 22: Resultados de predicción en el caso CAGR con DT y RF

El uso de estos algoritmos para el caso CAGR deja una precisión no muy alta y sobre todo tiene un bajo valor de replicabilidad debido a su índice de coincidencia Kappa que se encuentra en valores débiles de entre 0.21 y 0.4. En este análisis, sin embargo, se pueden utilizar una serie de algoritmos que se presentarán a continuación ya que serán utilizados en gran cantidad para todos los años en la sección del análisis predictivo.

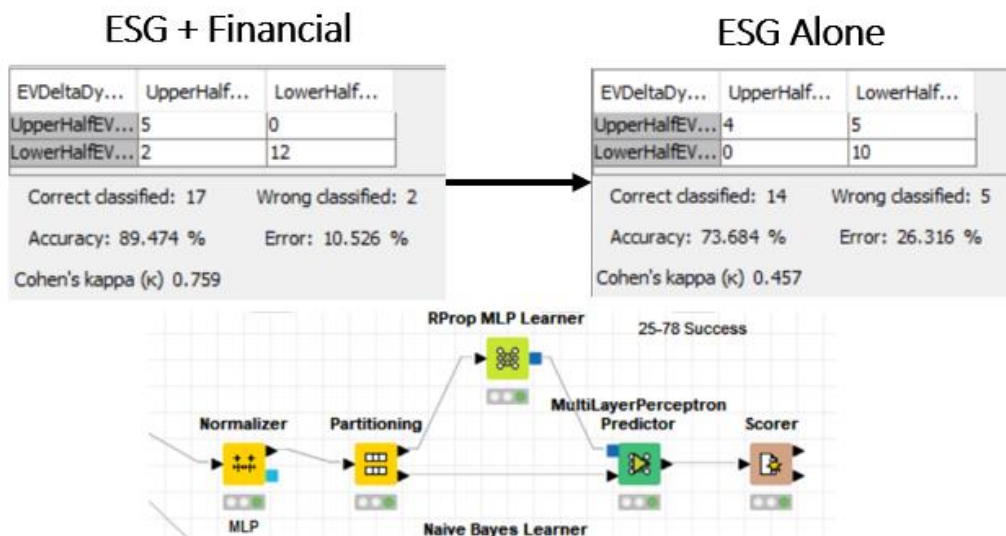


Figura 23: Resultados predicción caso CAGR con MLP



El siguiente algoritmo utilizado es el perceptrón multicapa (MLP). El perceptrón multicapa es un algoritmo que utiliza una red neuronal artificial formada por n capas de entrada cuyas salidas desembocan en m capas ocultas antes de pasar a la salida. Este algoritmo permite obtener soluciones a problemas no linealmente separables y si se entrena bien puede obtener salidas muy precisas.

Para fundamentar el efecto de la predicción a través de las variables de ESG en la figura 23 se puede ver el éxito de predicción alto para el EV CAGR de las compañías entre 2015 y 2021 en dos grupos. Al utilizar las variables financieras y de ESG previsiblemente se obtiene un éxito muy alto cercano al 90% y un índice de coincidencia Kappa bueno cercano a 0.8. Sin embargo, al realizar la misma predicción solo utilizando las variables de ESG se sigue consiguiendo una precisión elevada del 73.7% y un índice de coincidencia moderado de 0.46.

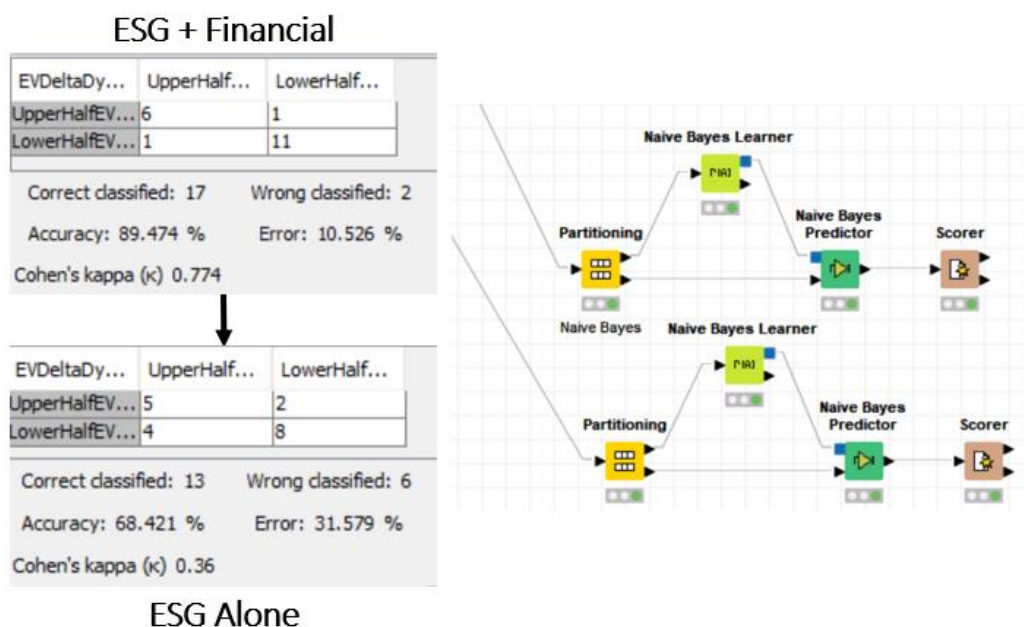


Figura 24: Resultados predicción caso CAGR con Naive Bayes

El siguiente algoritmo que encaja con la predicción deseada es el clasificador bayesiano ingenuo (Naive Bayes). Mediante el clasificador Naive Bayes se puede obtener una predicción de precisión similar a la del MLP utilizando las variables de ESG y financieras. Aun así, al utilizar solo las variables de ESG la predicción empeora en mayor medida que con el anterior algoritmo, aunque sigue siendo una buena predicción con una precisión cercana al 70% aunque con un índice de coincidencia débil que hace que oscile más la precisión en cada iteración del algoritmo.

El clasificador Naive Bayes es también de gran utilidad en la realización de modelos ya que permite aplicar pesos ponderados a las variables con facilidad. Esta aplicación sin embargo no es parte de las hipótesis y propósito de este trabajo por lo que quedaría como una posible ampliación futura.

### 2 clusters

Cluster \ P...	cluster_0	cluster_1
cluster_0	17	2
cluster_1	0	0

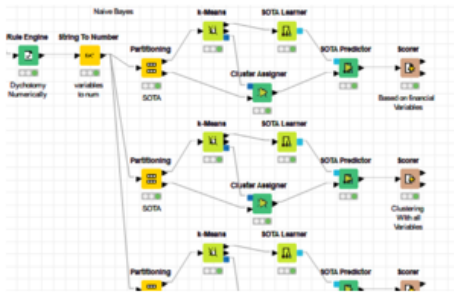
Correct classified: 17    Wrong classified: 2  
 Accuracy: 89.474 %    Error: 10.526 %  
 Cohen's kappa ( $\kappa$ ) 0

### 3 clusters

Cluster \ P...	cluster_2	cluster_1	cluster_0
cluster_2	15	1	0
cluster_1	0	2	0
cluster_0	1	0	0

Correct classified: 17    Wrong classified: 2  
 Accuracy: 89.474 %    Error: 10.526 %  
 Cohen's kappa ( $\kappa$ ) 0.616

### 4 clusters



Cluster \ P...	cluster_0	cluster_1	cluster_3
cluster_0	5	1	2
cluster_1	1	1	1
cluster_3	2	1	5

Correct classified: 11    Wrong classified: 8  
 Accuracy: 57.895 %    Error: 42.105 %  
 Cohen's kappa ( $\kappa$ ) 0.321

Figura 25: Clusterización mediante k-medias y SOTA

Por último, se utilizan algoritmos para lograr dividir las variables entre clústeres utilizables más adelante en el análisis predictivo. La formación de clústeres es muy útil ya que permite separar la población de datos entre sectores con un comportamiento similar y con compañías entre las que es más fácil identificar patrones.

Esta labor de determinar clústeres se ha realizado mediante una combinación de uso de k-medias y SOTA. K-medias permite la división de los datos entre k clústeres que se dividen por centroides óptimos para los diferentes grupos de datos. Tras identificar cada set de datos con su centroide más cercano se utiliza un algoritmo de “State-Of-The-Art” (SOTA) que permite entrenar y predecir sobre los clústeres determinados anteriormente mediante k-medias.

En la figura 25 se puede ver la limitación de la división en 2, 3 y 4 clústeres. Al dividir los datos en 2 o 3 clústeres óptimos matemáticamente se obtiene una sobrepoblación de uno de los clústeres que hace que la predicción opte por predecir todos los datos como pertenecientes a este grupo. Al incrementar el número de clústeres a 4 los datos se distribuyen mejor pero el algoritmo es entonces deficiente al determinar adecuadamente que datos corresponden a cada centroide. La precisión queda en un valor débil menor al 60% y la coincidencia se sitúa en un valor débil menor a 0.4.

Esta incapacidad de creación de clústeres claros para nuestra base de datos se debe a dos principales razones. La primera es que debido a la gran cantidad de variables los centroides no se escogen adecuadamente. La segunda razón es que debido a la limitación de compañías (n=113) y la población más ligera de los datos de ESG en los periodos cercanos a 2015 dificulta la tarea de dividir los datos en más de 3 clústeres.

Por estas razones en los casos posteriores se ha optado por realizar clústeres de dos tipos en torno a categorías como la región y el tipo de especialización de la empresa dentro del sector de la energía.

### 3.3.- Revisión de los Datos

Para realizar el análisis predictivo sobre todos los años y ver la viabilidad de realizar una predicción de un valor financiero de inversión es útil llevar a cabo dos primeros pasos. El primero es la decisión sobre los clústeres

Previa a la realización del análisis predictivo procedemos con la revisión de datos para decidir los clústeres. Esta revisión de datos cualitativos que permitan realizar la clusterización es de gran ayuda cuando los métodos de modelado no encuentran centroides que definan grupos subyacentes.

Row Labels	Count of Main Continent of Action (pregrouping)	
EastAsia,MiddleEast,India	30	26.55%
Europe	27	23.89%
North America	48	42.48%
South America	8	7.08%
(blank)		
<b>Grand Total</b>	<b>113</b>	<b>100.00%</b>

Figura 26: Distribución de las empresas de acuerdo con región

Al realizar la clasificación de las empresas escogidas de acuerdo con su región de operación mayoritaria la distribución se encuentra razonablemente separada entre Asia, Europa y América del Norte. América del Sur tiene una representación inferior por lo que a la hora de analizar por clústeres se ha decidido agregarla a las muestras de América del Norte por mayor similitud que con el resto de las regiones.

Row Labels	Count of Sector? Majority (pregrouping)	
Batteries	5	4.42%
Coal, Oil and Natural Gas	40	35.40%
Distribution/Transmission	28	24.78%
Nuclear	3	2.65%
Renewables	37	32.74%
(blank)		
<b>Grand Total</b>	<b>113</b>	<b>1</b>

Figura 27: Distribución de las empresas por especialización

Dentro del sector de la energía otra manera cualitativa de gran utilidad para segmentar la influencia de las variables de ESG sobre las financieras es la especialización particular mayoritaria de las empresas seleccionadas.

En el caso de las empresas de energía en gran parte de las empresas más grandes su origen como compañía nacional de energía implica un mixto de ingresos de varias de las categorías. Por ello las empresas se han categorizado por su fuente de ingresos mayoritaria.

Las categorías tienen tres grupos mayoritarios en la distribución y transmisión, la generación con energía intensa en gases de efectos invernadero y empresas renovables. La nuclear y las de baterías se han excluido del análisis para no distorsionar los análisis por especialización y ya que su representación es muy pequeña para hacer análisis separados sobre las mismas.

Además de los clústeres para poder realizar un análisis predictivo adecuado se debe establecer una serie de variables clasificatorias para las empresas desde la que poder estudiar los datos en busca de patrones de cara a escoger los algoritmos de aplicación.

Para el caso de las empresas de energía, y tratando de confirmar la segunda hipótesis de este trabajo de poder “establecer un patrón claro y predecible para la evolución en el valor capitalizable de las empresas a través de los años”, se han escogido las siguientes variables:

- **Ingresos:** Descrito anteriormente en el apartado de variables, representa los ingresos brutos de la empresa. Su aplicación como variable clasificatoria es ideal debido a ser una variable de escala previa a filtros de funcionamiento de la empresa que podrían generar distorsiones temporales en la percepción de empresas más y menos grandes. Se mide en millones de euros.
- **Delta capitalización:** Se estima como el porcentaje de incremento entre la capitalización de la empresa el año anterior y el año de los ingresos. Es ideal para medir el impacto de la percepción de mercado de la empresa y de sus actuaciones ESG. Sus principales limitaciones son que el dato se obtiene a día 31 de diciembre por lo que se puede ver alterado por factores específicos de la fecha y ajenos a la gestión del resto del año.

Una vez establecida las variables sobre las que representar las empresas se pueden organizar en gráficas para observar la relación entre las dos para todos los años y separadas por regiones y por especializaciones. Las gráficas se pueden encontrar en el apéndice A en el capítulo 6.1 para su observación detenida, aunque se comentarán a continuación.

Para el año 2022 el impacto de la guerra ruso-ucraniana y el consecuente impacto en los precios de los principales combustibles fósiles consigue que las mayores empresas de generación con estos combustibles se enriquezcan tanto por las operaciones que algunos de estas operan como por las coberturas previas sobre combustibles para generación.

Esta relación se puede ver en menor medida en empresas americanas grandes del sector y en mayor medida en las empresas asiáticas que se han beneficiado de flujos de crudo y gas ruso a precios reducidos por los bloqueos en occidente.

Las renovables y las empresas de transmisión, así como las empresas en Europa, no presentan relaciones visibles.

Para el año 2021 la mayoría de las empresas en todas las diferentes categorías presentan un comportamiento similar al del año anterior en términos de capitalización. Es merecedor destacar que este año cuenta con un mayor número de valores atípicos. Los valores atípicos en la diferencia de capitalización entre 2020 y 2021, generalmente de

tendencia positiva, puede ser por empresas que redujeran su capitalización considerablemente durante la pandemia Covid-19 y que al recuperarse su expectativas de generación de caja con la desescalada de las restricciones mejorara sus resultados.

Para el año 2020 se ve un comportamiento de pérdidas capitalización en los combustibles fósiles y de ganancias de capitalización en las renovables. Este comportamiento se puede justificar a través del cambio considerable en la demanda energética a la baja en las economías mundiales del 10-30% (WEF20). En Europa especialmente se ha aprovechado para acelerar la transición energética evitando que los efectos de la reducción de demanda afecten a la generación renovable en detrimento de la generación con combustibles fósiles (EU21).

Para el año 2019 la variación de la capitalización es positiva para las empresas americanas, las empresas asiáticas registran progresiones de su capitalización más negativas en las más grandes y las europeas encuentran una dispersión equilibrada en torno al estancamiento de la capitalización. Para la especialización los combustibles fósiles están equilibrados entre crecientes y decrecientes mientras que renovables y transmisión tuvieron más crecimiento, aunque centrado en la media de ingresos. En comparación con años interiores la falta de claros impactos macroeconómicos y geopolíticos dificulta la identificación de una causa raíz para el crecimiento de las empresas americanas o las renovables y de transmisión.

2018 fue un año de crecimiento para empresas europeas y 2017 un año de crecimiento para todas las empresas en todas las especializaciones y regiones.

2016 acusa una pronunciada contracción en las capitalizaciones de las empresas energéticas europeas. En este aspecto podrían existir varias causas como la caída pronunciada en el precio de los combustibles fósiles entre 2014 y 2016 o el tercer plan de rescate organizado para Grecia en 2015. Ninguna de ellas parece explicar de manera concluyente este efecto tan pronunciado solo en las empresas europeas.

A pesar de la complicación de establecer causas raíz para los comportamientos alineados de empresas en clusters de regiones o especialización se puede ver la potencia de estos análisis en el alineamiento de empresas en torno a ciertos años. Esto refuerza la justificación de realizar los análisis predictivos sobre estos grupos ya que demuestran tener características únicas que pueden ayudar a encontrar patrones de relación entre variables ESG y la capitalización para esos mercados específicos dentro de las empresas mundiales de energía.

## 3.4.- Análisis Predictivo

El análisis predictivo de la evolución de la capitalización inter anualmente es la última fase de desempeño de los modelos realizados.

En esta fase primero se debe establecer cuáles son las diferentes segmentaciones de datos a introducir para determinar la eficacia de unas variables contra otras. Este prefiltro es necesario para la mayoría de los algoritmos antes descritos ya que muchos de ellos constituyen cajas negras en las que no se pueden identificar los patrones o relaciones subyacentes. Por otro lado, la aplicación estandarizada de las variables permite el poder representar de manera masiva los resultados lo que permite simplificar su presentación de manera visual.

A su vez se ha optado por no introducir los índices de Bloomberg en estas predicciones ya que podrían tener un efecto distorsionador en las predicciones al ser otorgados a posteriori sobre empresas que ya han reportado su ESG y no ser una variable que emane naturalmente por el desempeño de la empresa durante el transcurso de sus operaciones.

Las agrupaciones de variables escogidas son las siguientes:

- Todas las variables: Aplicar todas las variables de ESG obtenidas como inputs de los modelos nos permite ver cuál es el impacto en el algoritmo de todas ellas además de poder ver el efecto de ruido al reducir estas variables. Cuando una predicción con todas las variables resulta buena, pero mejora en la aplicación de un segmento específico se debe a que la sobre especificación de añadir demasiadas variables afecta a la capacidad de los algoritmos de realizar estos procesos.
- Solo variables suficientemente pobladas: Algunas de las variables no tienen el mismo nivel de riqueza reportada que otras. La no inclusión de variables menos pobladas (GHG2 Mercado, Consumo de Agua, Residuos reciclados, Consumo de Energía Renovable, % de Mujeres en Alta Dirección y el carbono emitido por unidad de producción) ayuda a comprender el impacto de que las empresas reporten más información y además permite a los algoritmos utilizar variables que afectan a un mayor número de empresas.
- Variables con mayor representación en los análisis anteriores: Algunas variables durante los estudios realizados en los diferentes casos previos han demostrado tener una mayor relación tanto en las correlaciones con variables financieras como en las predicciones de variables financieras mediante algoritmos (Alcances GHG, Consumo de Agua, Residuos Reciclados, Residuos Totales, Número de Empleados y Consumos de Energía). Aplicar estas variables a los algoritmos permite el estudio y perfeccionamiento continuo desde los casos anteriores además de tener ya información contrastable de los vínculos posibles.
- Variables relacionadas con las emisiones de gases de efecto invernadero: Ver la relevancia en el peso de la predicción del crecimiento de la capitalización de las empresas a través de las emisiones permite ver si esta es una de las agrupaciones de variables prioritarias en la valoración de una empresa (GHG 1, 2, 3 y carbono por unidad de producción).

- Variables sociales: La representación de la mujer en los distintos niveles de dirección de las empresas, los niveles de accidentes y el número de empleados como indicadores sociales del progreso de las empresas permiten ver la relevancia de la dimensión social en un nivel simplificado contra la predicción de la valoración de las empresas.
- Variables relacionadas con el consumo: Consumo de energía, energía renovable, agua y generación de residuos y su reciclaje son variables relacionadas con la eficiencia en el consumo de las empresas y que pueden tener una relación separada con la valoración tanto por su carácter para la sostenibilidad como por su relación con la eficiencia en las cadenas de suministro y en el funcionamiento de la empresa.

A	All variables
B	Only well reported variables
C	Successful variables from the initial analysis
D	Emission related variables
E	Social Variables
F	Consumption behavior variables

Figura 28: Clasificación de las agrupaciones de variables

Para facilitar la visualización de los resultados de la predicción se utilizará el código con letras para representar a las diferentes segmentaciones de las variables que se puede ver en la figura 28.

### 3.4.1.- Predicción por regiones

En la realización de todas las predicciones se aplican los algoritmos detallados en el apartado 3.2.8. Es por ello por lo que cada uno de estos resultados requiere de entre diez y cincuenta bloques de programación (para preparación de los datos y aplicación de los algoritmos). Esto resalta la capacidad de simplificación de KNIME como programa de análisis de datos masivos ya que los miles de bloques necesarios para este trabajo en líneas de código serían una gran limitación a la hora de personalizar los análisis y realizar revisiones en caso de errores.

Para todos los casos los algoritmos deben predecir si los datos de práctica se encuentran en la distribución superior o inferior al valor más cercano a la media de crecimiento de capitalización.

Como se ha demostrado en el análisis de los datos el crecimiento de la capitalización de las empresas varía considerablemente dependiendo de la región. Aún así se ha escogido la media global para realizar el corte de las bases de datos y así realizar un análisis común pero que resalte las diferencias de predicción entre las regiones.

Success chance in prediction %	Asia						America						Europe					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
Decision Tree	16	50	66	33	50	<b>83</b>	<b>70</b>	40	40	40	<b>70</b>	40	60	<b>100</b>	80	80	40	80
Random Forest	50	16	66	66	66	50	30	60	<b>70</b>	40	40	40	<b>100</b>	80	80	80	<b>80</b>	60
Multi Layer	50	66	50	50	50	<b>83</b>	60	40	50	40	50	30	40	40	60	80	<b>100</b>	60
Naïve Bayes	50	16	50	50	33	33	50	40	40	50	60	60	0	40	40	60	80	60

Figura 29: Éxito en la predicción por región para el delta de capitalización 2021-2022

Para la predicción por región entre los años 2021 y 2022 con las variables de ESG del año 2021 se ha utilizado un crecimiento de capitalización corte del 10%. La predicción devuelve valores óptimos en Asia al introducir las variables de consumo y en todas las regiones a través del árbol de decisión.

Las predicciones en Europa tienen resultados mejores mientras que en América la media de las predicciones es peor. Aun así, es posible predecir con éxito alto para todas las geografías en diferentes algoritmos y sets de variables.

Success chance in prediction %	Asia						America						Europe					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
Decision Tree	66	<b>83</b>	50	33	66	33	18	45	36	55	<b>63</b>	55	<b>80</b>	40	40	20	<b>80</b>	20
Random Forest	50	50	33	50	<b>83</b>	17	27	45	54	45	33	36	60	40	<b>80</b>	20	<b>80</b>	60
Multi Layer	50	66	66	50	50	33	45	55	<b>66</b>	36	27	27	0	<b>80</b>	40	20	0	40
Naïve Bayes	50	66	<b>83</b>	66	66	17	45	36	54	45	55	<b>63</b>	40	60	40	60	40	40

Figura 30: Éxito en la predicción por región para el delta de capitalización 2020-2021

Para la predicción entre los años 2020 a 2021 con las variables de ESG del año 2020 el incremento de la capitalización de corte es el 10%. Las variables sociales y las variables obtenidas en los casos previos de este trabajo son las mejores predictoras de un incremento en la capitalización.

Al observar los algoritmos el árbol de decisión consigue altas probabilidades de éxito. Las predicciones en Europa y en Asia consiguen probabilidades altas de éxito mientras que en Estados Unidos las predicciones más altas no sobrepasan el 70% de éxito. Este efecto es comprensible desde la información previa del análisis de los datos ya que los crecimientos de capitalización de los años 2020 a 2021 son generalmente negativos para la región con altas volatilidades positivas en muchos casos.



Success chance in prediction %	Asia						America						Europe					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
Decision Tree	50	67	17	83	50	83	40	40	50	60	50	30	60	60	60	60	80	20
Random Forest	67	33	67	67	33	<b>100</b>	50	30	<b>80</b>	40	60	50	20	80	80	40	80	80
Multi Layer	50	67	67	50	67	50	<b>70</b>	50	<b>70</b>	30	60	40	80	80	60	40	80	60
Naïve Bayes	33	83	<b>100</b>	83	67	33	50	50	40	30	60	50	40	60	80	20	<b>100</b>	20

Figura 31: Éxito en la predicción por región para el delta de capitalización 2019-2020

Para la predicción por región entre los años 2019 y 2020 con las variables de ESG del año 2019 se ha utilizado un crecimiento de la capitalización corte nula del 0%. Para este año no hay ningún algoritmo o set de datos que coincida con el éxito máximo de las predicciones para las diferentes regiones. Aun así, las predicciones a través de las variables seleccionadas por casos anteriores dan buenos resultados para todas las regiones.

Success chance in prediction %	Asia						America						Europe					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
Decision Tree	83	67	50	67	50	<b>100</b>	60	90	80	70	80	90	20	40	<b>80</b>	40	60	60
Random Forest	83	83	67	67	67	<b>100</b>	90	<b>100</b>	<b>100</b>	70	70	<b>100</b>	20	40	20	20	60	<b>80</b>
Multi Layer	50	67	83	50	67	67	70	80	70	70	80	80	20	40	40	60	40	20
Naïve Bayes	67	50	50	50	50	33	50	50	60	40	80	20	0	40	60	40	<b>80</b>	40

Figura 32: Éxito en la predicción por región para el delta de capitalización 2018-2019

Para la predicción por región entre los años 2018 y 2019 con las variables de ESG del año 2018 se ha utilizado un crecimiento de la capitalización del 10%. Las variables sociales y la predicción por bosques aleatorios dan resultados óptimos de predicción para todas las regiones. Además, se puede ver que las predicciones son notablemente peores para Europa que para Asia a pesar de que la diferencia en la posición de las empresas respecto al punto de corte es similar para ambas regiones. Se puede ver también con facilidad la idoneidad de ciertos algoritmos respecto a otros por la caída considerable de precisión para todos los sets de variables en la región americana al aplicar los algoritmos de predicción por Naïve Bayes.

Success chance in prediction %	Asia						America						Europe					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
Decision Tree	<b>80</b>	60	60	60	60	40	70	50	30	50	60	60	20	60	<b>100</b>	40	<b>100</b>	80
Random Forest	<b>80</b>	20	<b>80</b>	<b>80</b>	60	<b>80</b>	50	<b>100</b>	40	50	60	50	60	80	80	<b>100</b>	<b>100</b>	80
Multi Layer	<b>80</b>	40	40	40	60	<b>80</b>	60	30	50	50	60	50	80	60	80	60	80	80
Naïve Bayes	<b>80</b>	40	20	40	40	40	70	60	10	30	60	40	60	80	80	80	<b>100</b>	60

Figura 33: Éxito en la predicción por región para el delta de capitalización 2017-2018

Para la predicción por región entre los años 2017 y 2018 con las variables de ESG del año 2017 se ha utilizado un crecimiento de la capitalización corte nula del 0%. Para este año la predicción por bosques aleatorios da resultados por encima de la media para casi todos los sets de variables además de representar muchos de los mejores resultados para todas

las regiones. Es interesante la importancia para algunas regiones de la selección de sets de variables ESG ya que en Asia al incluir todas se obtienen resultados buenos para todos los algoritmos y en Europa al incluir variables relacionadas con el consumo de la empresa.

Success chance in prediction %	Asia						America						Europe					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
Decision Tree	40	40	40	20	<b>100</b>	40	<b>100</b>	60	70	60	60	60	<b>80</b>	20	20	40	60	60
Random Forest	60	60	0	20	60	40	50	50	50	50	40	70	20	20	40	20	60	20
Multi Layer	<b>80</b>	0	0	40	60	40	70	50	<b>80</b>	<b>80</b>	30	80	60	40	60	<b>80</b>	60	0
Naïve Bayes	60	0	0	60	40	60	<b>80</b>	70	50	<b>80</b>	50	50	0	20	40	<b>80</b>	60	60

Figura 34: Éxito en la predicción por región para el delta de capitalización 2016-2017

Para la predicción por región entre los años 2016 y 2017 con las variables de ESG del año 2016 se ha utilizado un crecimiento de la capitalización del 15%. Para este año destacan los algoritmos por encima de las variables al demostrarse los bosques aleatorios como peor predictor mientras que los bosques aleatorios, el perceptrón multi capa y Naïve Bayes devolviendo precisiones altas para todas las regiones. La media general de la precisión de las predicciones para este año es generalmente baja, aunque los resultados con predicciones acertadas tienen alta precisión. Esta alta dispersión en los resultados es mucho más acentuada que en otros años especialmente en Asia donde los algoritmos obtienen precisiones nulas en muchos de los algoritmos para las variables de casos anteriores y para las variables más reportadas.

Success chance in prediction %	Asia						America						Europe					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
Decision Tree	20	40	20	20	20	60	<b>80</b>	50	<b>80</b>	50	70	60	75	75	75	<b>100</b>	<b>100</b>	75
Random Forest	<b>80</b>	40	40	20	20	<b>80</b>	50	50	60	50	<b>80</b>	40	75	50	50	75	<b>100</b>	<b>100</b>
Multi Layer	<b>80</b>	20	60	20	40	40	40	50	<b>80</b>	50	60	30	<b>100</b>	50	75	<b>100</b>	50	50
Naïve Bayes	40	60	60	60	60	40	70	70	60	40	20	40	<b>100</b>	75	75	<b>100</b>	75	50

Figura 35: Éxito en la predicción por región para el delta de capitalización 2015-2016

Por último, para la predicción por región entre los años 2015 y 2016 con las variables de ESG del año 2015 se ha utilizado un crecimiento de la capitalización del 5%. La región europea tiene alta precisión en sus predicciones apoyada principalmente por la distribución específica de la región para este año con resultados muy superiores a las de otras regiones en crecimiento de capitalización. Aun así, los algoritmos de perceptrón multicapa y de bosques aleatorios son capaces de encontrar precisiones altas. La región asiática tiene precisiones en la predicción peores que la región americana, a pesar de su distribución similar en este año, pudiendo indicar a un vínculo más débil en la valoración de las empresas asiáticas respecto a sus valores de ESG en comparación con las americanas para el año 2015.

### 3.4.1.- Predicción por sector

El mismo análisis de predicción del crecimiento de la capitalización para las diferentes regiones se ha realizado sobre los clústeres por sector. Para estos casos las distribuciones de las empresas en los diferentes años son diferentes por lo que también la precisión de las predicciones será ideal bajo diferentes sets de variables y algoritmos para los diferentes años.

Los puntos de corte son los mismos que en el caso anterior para determinar el grupo de datos que supera la media de crecimiento de capitalización y por tanto en la que se desearía invertir.

Success chance in prediction %	Coal, Oil & Gas						Transmission & Distribution						Renewables					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
Decision Tree	25	38	63	25	<b>75</b>	38	40	40	60	40	80	60	43	<b>86</b>	43	57	29	71
Random Forest	50	38	63	25	<b>75</b>	13	<b>80</b>	<b>80</b>	40	40	60	20	57	86	57	57	71	14
Multi Layer	38	38	<b>75</b>	50	38	63	40	<b>80</b>	40	60	40	40	71	43	57	57	57	29
Naïve Bayes	25	25	63	63	38	38	80	60	40	40	80	40	29	57	<b>86</b>	57	43	29

Figura 36: Éxito en la predicción por sector para el delta de capitalización 2021-2022

Para la predicción por sector entre los años 2021 y 2022 con las variables de ESG del año 2021 se ha utilizado un crecimiento de capitalización corte del 10%. Para este año no hay ningún algoritmo ni ningún set de variables aplicable a todos los sectores para obtener resultados de predicción óptimos. Además, al escoger cualquier set de variables o algoritmo la media de las predicciones resultantes se aproxima al 50% lo que indica que los algoritmos no consiguen interpretar patrones que sean aplicables con la clusterización por sectores.

Success chance in prediction %	Coal, Oil & Gas						Transmission & Distribution						Renewables					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
Decision Tree	63	13	63	50	63	38	33	17	0	33	50	17	29	71	29	71	75	57
Random Forest	38	25	50	63	38	50	17	33	50	50	67	17	71	57	71	43	<b>81</b>	43
Multi Layer	50	63	25	63	63	25	<b>83</b>	67	33	50	33	33	29	43	<b>100</b>	43	76	43
Naïve Bayes	50	50	<b>88</b>	50	25	38	33	50	50	50	50	50	43	<b>86</b>	57	71	43	57

Figura 37: Éxito en la predicción por sector para el delta de capitalización 2020-2021

Para la predicción entre los años 2020 a 2021 con las variables de ESG del año 2020 el incremento de la capitalización de corte es el 10%. Al igual que en el año anterior es difícil encontrar algoritmos que mejoren considerablemente la predicción respecto a la media aleatoria. Aunque para alguno de los sectores específicos se obtengan con algoritmos aislados precisiones altas no se puede extrapolar al resto de sectores. Esto puede interpretarse de manera que la división por sectores para los años 2021 y 2020 devuelve grupos tan heterogéneos que aplicar las variables de sostenibilidad no permite detectar cuales son ideales para la inversión como si se podía hacer mejor con las regiones.

Success chance in prediction %	Coal, Oil & Gas						Transmission & Distribution						Renewables					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
Decision Tree	38	38	50	63	25	<b>75</b>	50	33	33	50	<b>83</b>	33	57	86	86	57	86	86
Random Forest	63	63	38	<b>75</b>	63	50	50	67	50	50	67	50	86	86	71	57	86	71
Multi Layer	<b>75</b>	63	<b>75</b>	63	25	50	50	50	17	33	50	67	57	71	57	57	57	71
Naïve Bayes	<b>75</b>	<b>75</b>	50	38	63	50	50	50	50	17	33	0	71	43	43	86	86	<b>100</b>

Figura 38: Éxito en la predicción por sector para el delta de capitalización 2019-2020

Para la predicción por sector entre los años 2019 y 2020 con las variables de ESG del año 2019 se ha utilizado un crecimiento de la capitalización corte nula del 0%. Para este año hay algoritmos y sets de variables que tienen precisión en la predicción alta para sectores específicos, pero al igual que en los anteriores con dificultad de encontrar algoritmos o sets de variables que permitan la predicción para todas. Para este caso y para los anteriores las variables más reportadas tienen una precisión alta para casi todos los algoritmos lo que puede indicar que se valoran más las empresas de generación renovable con enfoque en esas variables. En el caso de la generación con combustibles fósiles parece que la predicción en torno a las variables escogidas como relevantes en los casos anteriores devuelven precisiones mayores.

Success chance in prediction %	Coal, Oil & Gas						Transmission & Distribution						Renewables					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
Decision Tree	75	63	75	63	<b>88</b>	50	67	50	67	50	67	50	<b>86</b>	57	71	57	71	57
Random Forest	<b>88</b>	75	75	75	<b>88</b>	63	67	67	67	67	67	50	57	57	29	71	43	<b>100</b>
Multi Layer	<b>88</b>	63	<b>88</b>	75	63	75	<b>100</b>	67	67	67	67	<b>100</b>	71	14	71	43	43	43
Naïve Bayes	63	<b>88</b>	75	38	<b>88</b>	63	17	50	50	50	83	50	29	29	71	43	71	43

Figura 39: Éxito en la predicción por sector para el delta de capitalización 2018-2019

Para la predicción por sector entre los años 2018 y 2019 con las variables de ESG del año 2018 se ha utilizado un crecimiento de la capitalización del 10%. Para este año las predicciones son mejores con el set con todas las variables y el set con las variables escogidas en casos anteriores dando buenos resultados para todos los sectores. A su vez todos los algoritmos salvo el de Naïve Bayes dan resultados considerablemente mejores a la media en casi todos los casos.

Este cambio respecto a los años posteriores puede indicar a un cambio en la percepción del impacto de las variables de ESG sobre la valorización de la empresa y consecuentes inversiones que alteren la capitalización desde el covid-2019. Los cambios generados en la demanda de energía y en la capacidad de aprovisionamiento de combustibles fósiles puede haber alterado este comportamiento y hacer que las predicciones de los algoritmos tengan menor éxito con variables que al clusterizar a nivel regional se han mantenido relevantes.

Success chance in prediction %	Coal, Oil & Gas						Transmission & Distribution						Renewables					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
Decision Tree	50	63	75	<b>88</b>	50	<b>88</b>	50	83	50	50	33	67	67	50	50	67	50	50
Random Forest	50	63	63	63	38	50	<b>100</b>	67	50	<b>100</b>	83	83	67	67	<b>83</b>	67	67	67
Multi Layer	75	75	25	63	50	38	67	67	50	83	67	<b>83</b>	<b>83</b>	<b>83</b>	50	33	67	<b>83</b>
Naïve Bayes	63	75	<b>88</b>	63	63	38	67	67	67	83	50	33	17	67	67	83	67	17

Figura 40: Éxito en la predicción por sector para el delta de capitalización 2017-2018

Para la predicción por sector entre los años 2017 y 2018 con las variables de ESG del año 2017 se ha utilizado un crecimiento de la capitalización corte nula del 0%. Las predicciones realizadas con las variables relacionadas con las emisiones y con el algoritmo de bosques aleatorios son particularmente precisas para todos los sets de variables.

La dinámica de precisiones elevadas previa a la pandemia continúa también este año con los ejemplos mencionados de algoritmos con resultados de predicción precisos y con sets de variables que para cada sector obtienen resultados buenos.

Success chance in prediction %	Coal, Oil & Gas						Transmission & Distribution						Renewables					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
Decision Tree	50	63	<b>88</b>	50	63	50	50	67	50	67	50	83	50	50	50	50	<b>83</b>	<b>83</b>
Random Forest	38	38	25	25	63	50	67	<b>83</b>	33	33	50	50	67	67	50	<b>83</b>	33	33
Multi Layer	50	<b>75</b>	63	25	<b>75</b>	38	33	50	17	50	33	50	50	50	83	67	33	67
Naïve Bayes	63	25	25	38	50	63	33	33	33	67	17	67	67	50	<b>83</b>	17	67	<b>83</b>

Figura 41: Éxito en la predicción por sector para el delta de capitalización 2016-2017

Para la predicción por sector entre los años 2016 y 2017 con las variables de ESG del año 2016 se ha utilizado un crecimiento de la capitalización del 15%. Para este año las predicciones realizadas con el árbol de decisión son para todos los sets de variables más precisas que la media. Algoritmos como Naïve Bayes tienen el efecto contrario, consiguiendo predicciones muy dispersas con resultados muy superiores a la media y muy inferiores.

Las variables vinculadas al consumo también tienen una precisión más elevada para todos los grupos. A su vez en todos los sectores un set de variables destaca por precisiones elevadas para todos los algoritmos lo que destaca el carácter separado pero eficiente de la clusterización por sectores para el año 2016.

Success chance in prediction %	Coal, Oil & Gas						Transmission & Distribution						Renewables					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
Decision Tree	29	57	86	43	57	<b>100</b>	50	67	50	33	33	66	83	67	<b>100</b>	50	33	<b>100</b>
Random Forest	86	43	57	14	71	86	33	50	33	<b>83</b>	50	50	50	67	<b>100</b>	67	83	50
Multi Layer	71	71	<b>100</b>	14	57	71	67	50	17	33	33	66	50	67	83	83	33	67
Naïve Bayes	29	71	57	71	71	57	67	<b>83</b>	50	33	50	33	67	67	67	66	33	83

Figura 42: Éxito en la predicción por sector para el delta de capitalización 2015-2016

Por último, para la predicción por sector entre los años 2015 y 2016 con las variables de ESG del año 2015 se ha utilizado un crecimiento de la capitalización del 5%. Para este año el set de variables de consumo da buenos resultados para todos los indicadores. Sin embargo, a nivel de algoritmos ninguno es mucho más preciso que los otros aunque todos obtienen resultados de precisión que en conjunto son mejores que la media de distribución.

# Capítulo 4.- Conclusiones

Durante la realización del proyecto y la conclusión de los diferentes casos se obtiene información relevante para la interpretación de las hipótesis y que se recoge a continuación para mayor facilidad de acceso a la misma.

La primera hipótesis del proyecto cuenta con una complejidad especial por la gran cantidad de componentes del ESG además de la gran cantidad de efectos que influyen en una empresa a raíz de estas variables, tanto de manera positiva como negativa.

La segunda hipótesis contiene mucha complejidad ya que el valor capitalizable no solo responde a dinámicas de las empresas del sector energético o de las variables ESG que pueden afectarlas, sino que es parte del permanentemente cambiante sector financiero. Además, como se ha comentado durante el transcurso del trabajo, la variable de la capitalización de las empresas cuenta con la limitación de la temporalidad del valor capturado a final de año que se puede haber visto influenciado por circunstancias exógenas al funcionamiento de esta o condiciones transitorias que provocaron pánico/euforia en los mercados.

## 4.1.- Conclusiones sobre la metodología

Pese a la complejidad de las hipótesis, la metodología seguida favorece un proceso metódico basado en un entendimiento progresivo de las variables utilizadas y de su conexión entre ellas.

Las primeras conclusiones sobre la metodología provienen de la decisión de las variables a escoger. La dificultad de escoger variables de ESG que resaltar en trabajos de análisis de datos proviene principalmente del limitado registro histórico que existe. La falta de estándares de reportado en muchas de las geografías, la falta de acuerdo en la naturaleza de las variables más resaltables y el interés de las empresas por acentuar su labor en unas variables sobre otras permiten que la información sea de una riqueza muy asimétrica y que los estudios como el realizado tengan dificultad para abarcar un número grande de aspectos de la sostenibilidad.

Las plataformas empleadas para la descarga de los datos resultan adecuadas para aligerar la carga de trabajo de poco valor añadido que representa la búsqueda masiva de datos para todas las empresas y años seleccionados durante el estudio. Estas plataformas no son infalibles pero la capacidad de agregación de datos y de descarga masiva representan una mejora de los tiempos de desarrollo de ejercicios como el realizado.

El programa seleccionado para la realización del trabajo (KNIME) demuestra su gran utilidad ya que con herramientas no dedicadas al análisis de datos masivos este estudio no habría sido posible. La capacidad de simplificación del análisis de datos con código basado en Python de una manera fácilmente revisable y con gran facilidad para extracción de las figuras con las que explicar las conclusiones de análisis y predicciones.

El uso de esta herramienta, u otras herramientas de análisis de datos, es encarecidamente recomendable sobre herramientas clásicas de modelaje por la gran profundidad que permiten y la eficiencia en la capacidad requerida de los recursos computacionales.

Respecto a la parte de la metodología realizada mediante el uso de KNIME, se puede separar en el estudio de los datos mediante correlaciones y algoritmos fácilmente interpretables y la predicción incluyendo la revisión de datos, clústeres y uso de algoritmos.

El uso metódico de la correlación para interpretar la conexión entre las variables es de gran utilidad y recomendado para trabajos posteriores. Además de permitir una gran visibilidad de las conexiones entre las variables permite un análisis rico por la inclusión de la comprobación de las conexiones mediante los algoritmos de árbol de decisión y de bosques aleatorios.

Ha sido adicionalmente de gran utilidad para la identificación de las diferentes problemáticas de la interpretación de los datos el uso de la correlación combinada con algoritmos. Prevalencia parcial de la escala de las empresas como consideración creadora de ruido, el estudio de las variables como progreso respecto a años anteriores, la visión de estas con el crecimiento compuesto anualizado y

## 4.2.- Conclusiones sobre los resultados

Al estudiar el caso base en la primera correlación se puede ver el vínculo que existe entre las variables financieras y las de sostenibilidad a través de una correlación de valor medio entre ambas. Cabe destacar que este vínculo es lógico ya que refleja comportamientos de las empresas que se vinculan con su estructura productiva y por tanto con su capacidad de aprovecharse de ciclos económicos y de las diferentes oportunidades de mercado.

La correlación entre las emisiones y los ingresos son razonables debido a la escala de las empresas históricamente emisoras, las cuáles, aunque con problemas de proyección a largo plazo se mantienen como algunas de las empresas más rentables a nivel internacional.

La correlación entre menores emisiones y menores intereses o menores empleados y menores intereses o menos CAPEX, alude a consideraciones principalmente de escala ya que, aunque para ambos casos podría ser resultado de políticas particulares de la empresa son también variables características de la escala de la empresa.

La predicción es increíblemente acertada para encontrar las empresas con más ebitda enfocándose solo en variables de escala como los empleados y la energía consumida. Debido a este enfoque en la escala de las empresas se valora la importancia de encontrar variables específicas que no permitan este tipo de influencias.

Al quitar el factor de la escala de las variables financieras se obtienen resultados interesantes. Una relación entre menores emisiones y mayor rentabilidad de la inversión encaja con las empresas novedosas tanto de fabricación de baterías como renovables que



destacan por tener valoraciones muy atractivas y altos resultados. Menores empleados y un ratio de EBITDA mayor. Mayor porcentaje de reciclado con un menor coste de deuda. El ratio entre el interés y el EBITDA con los Disclosures de Bloomberg. Todas estas relaciones dan resultados super interesantes sobre cuál es el comportamiento de la empresa, que tipo de empresas son y como se determinan factores como los de los intereses.

Las predicciones al utilizar los ratios financieros se dificultan, aunque se mejora la predicción sobre EBITDA, FFO y ROE lo cual demuestra que las correlaciones indican a variables que al parametrizar pueden predecir el resultado de las empresas.

Al quitar el factor de la escala tanto de las variables ESG como de las financieras las correlaciones ascienden a valores medios y casi fuertes. Relación similar entre las emisiones relativas a ingresos y la generación de beneficios demuestran comportamientos subyacentes de las empresas que han materializado los ingresos como las de gas y petróleo durante 2021. El consumo de energía y los beneficios. Del lado de la predicción con los algoritmos se consiguen ratios de predicción muy buenos para predecir EBITDA y FFO demostrando que para este año las variables ESG fueron muy relevantes cuando las variables que se introducen tienen sentido.

Al aplicar el estudio de correlaciones se ha podido demostrar como la mayoría de las relaciones entre las variables se mantienen. Esto es de especial interés ya que refuerza la realidad de que se crean clústeres de empresas con sistemas productivos similares. En el caso de las variables de sostenibilidad provoca el efecto deseado de convertirse en variables esenciales para definir las prácticas de la empresa y fundamentar que estos clústeres con prácticas sostenibles similares obtengan beneficios alineados para ciertas variables financieras.

Al realizar el mismo análisis de las deltas financieras aplicando además las deltas de ESG los resultados que se obtienen relaciones débiles entre las variables de ESG y financieras que dan información añadida sobre el comportamiento de estas. El incremento de empleados tiene una relación positiva con el valor capitalizable probablemente debido al mayor acceso a capitales que reporta una mejor valoración en bolsa y por tanto un crecimiento de la plantilla mayor en años de recuperación como 2021.

La correlación positiva entre el aumento de mujeres en posiciones de alta dirección con altos niveles de EBITDA puede indicar el efecto positivo de la diversidad como generador de riqueza. La correlación entre una deuda neta mayor y emisiones menores puede comprenderse dentro de la proyección a futuro que permite una empresa con menores emisiones tener posiciones más endeudadas para generar proyectos. Empresas con mayores porcentajes de reciclaje y mayores niveles de CAPEX puede demostrar más capacidad de inversión en empresas que mejoran sus indicadores ESG.

La predicción mejora en general respecto al caso anterior lo que indica una mayor capacidad de replicabilidad de los resultados al tener un análisis más profundo basado en deltas de variables ESG y financieras. A su vez, al mirar las variables con las que se consiguen los mejores resultados de predicción se puede ver que la rentabilidad de las inversiones tiene relaciones positivas con el porcentaje de mujeres y la generación de mayor EBITDA en porcentaje con el carbono emitido por unidad de producción.

Al realizar el análisis del año previo se puede ver que las correlaciones entre variables para algunos casos incrementan y en otros se debilitan, pero se mantienen en línea al realizar el análisis para el mismo año. Esto es un resultado de gran riqueza porque por un lado confirma que las correlaciones no tienen un comportamiento errático y que las empresas tienen valores de variables ESG similares para años contiguos. Por otro lado, se puede ver como las correlaciones cambian ligeramente, pero en ambas direcciones por lo que el resultado por correlación es inconcluso.

Para las predicciones sin embargo se mejora a porcentajes muy elevados de éxito para la predicción del EBITDA% y del FFO%. Estos porcentajes elevados pueden indicar que la caja de operaciones y el EBITDA se pueden predecir con mayor éxito analizándolos desde una perspectiva sostenible de año previo más que por el procedimiento estándar. Para el resto de las variables la predicción está alineada con casos anteriores.

De la correlación del caso CAGR destaca el crecimiento de las emisiones de alcance 3 con el crecimiento de EBIT. Esta relación resulta curiosa por no extenderse al CAGR de EBITDA y por tanto puede indicar que los efectos de la amortización en empresas que requieren de mayores niveles de inversión para realizar su transición no son irrisorios. El CAGR del crecimiento del número de empleados encuentra altos niveles de correlación con Ingresos, EV y valor capitalizable de la empresa. Esta relación tiene mucho sentido ya que representa la variable de ESG más relacionada con la escala de la empresa respecto a variables financieras muy íntimamente relacionadas con la escala para un periodo de crecimiento grande entre 2015 y 2021.

Sobre el caso CAGR se presenta el resto de los algoritmos a desarrollar en la fase de análisis predictivo. Estos son la técnica de clustering mediante algoritmos de k means y SOTA la predicción mediante algoritmos de árbol de decisión, bosques aleatorios, Naïve Bayes y perceptrón multicapa.

Los resultados de los algoritmos aplicados en el resto de los casos anteriores, árbol de decisión y bosques aleatorios, dan resultados algo bajos de precisión y demuestran una replicabilidad débil a través del coeficiente de kappa.

Al utilizar el algoritmo de perceptrón multicapa se puede ver la relación más clara entre variables financieras comparada con el uso de únicamente variables ESG. Aun así el éxito de la predicción se mantiene alto con un coeficiente de kappa moderado por lo que las variables ESG permiten identificar patrones de las empresas con mejores resultados.

El algoritmo de Naïve Bayes da resultados ligeramente peores que los del perceptrón multicapa, pero también permite ver que, aunque menos clara que la relación con variables financieras y de ESG, se pueden dibujar suficientemente mejor respecto a una predicción aleatoria los patrones de las empresas con mejores resultados.

Se han utilizado algoritmos SOTA para dividir las variables entre clústeres utilizables durante el análisis predictivo. El uso combinado de SOTA y k-means permite crear estos grupos eficientes y a su vez realizar la predicción con un grupo de testeo y uno de entrenamiento para ver si esas diferencias son claramente predecibles.

En esta labor de división por grupos eficientes se obtiene que el algoritmo prioriza sobrepoblar algunos de los clústeres sobre otros por lo que, aunque puede que este

identificando grupos diferenciados estos son demasiado específicos para llevar a cabo un análisis posterior de las variables separadas por esos grupos. Por otro lado, al incrementar el número de clústeres deseados hasta 4 clústeres los algoritmos tienen errores considerables para predecir qué población de testeo pertenece a cada grupo y por tanto pone en duda su diferenciación como extrapolable a todas las empresas. Esta limitación para crear clústeres se debe a la gran cantidad de empresas diferentes que se encuentran en el análisis, así como la población variable de todas las variables para estas empresas.

Al crear clústeres basados en la geografía y diferencias sectoriales dentro de la población de empresas energéticas y realizar un análisis a través de la visualización de los datos para las variables de ingresos y del delta de capitalización se pueden ver relaciones claras entre empresas dependiendo de los años. La guerra en ucrania, la pandemia y ciclos económicos dentro de las regiones y los sectores se han podido ver de manera rápida a través de este análisis.

Al realizar el análisis completo de para todos los años, con los diferentes clústeres de variables y los algoritmos de árbol de decisión, bosques aleatorios, perceptrón multicapa y Naïve Bayes se puede ver claramente la complejidad de la modelización.

Los resultados de predicción altos para todos los años en todas las geografías y sectores, pero utilizando diferentes algoritmos y grupos de variables define que un sistema complejo y diverso de predicción es el más adecuado para poder predecir las empresas con mejores resultados económicos solo mediante variables de sostenibilidad.

A su vez, en el desarrollo del trabajo se ha podido ver que, aunque no ha sido posible establecer una única manera de realizar la predicción del rendimiento de las empresas a través de variables de sostenibilidad si que queda claro un vínculo fuerte entre las variables de sostenibilidad y las financieras.

Este análisis además ha sido de mucha utilidad no solo por demostrar de manera estadística ese vínculo sino por también permitir entender dinámicas más profundas de estas relaciones al verse que variables específicas se vinculan unas con otras.

## 4.3.- Recomendaciones para futuros estudios

Durante la realización del trabajo ha quedado definida la riqueza y necesidad de reajustar los escenarios y casos de análisis a medida que se alcanzan resultados. Por ello es recomendable para trabajos posteriores seguir esta misma metodología de tratar de nulificar influencia de factores de escala, temporales o de agrupación de variables para tener una imagen más rica del estado de la cuestión y para poder obtener de manera más fácil conclusiones.

El programa utilizado para realizar las predicciones a su vez ha demostrado su utilidad para poder realizar estudios estadísticos masivos y por tanto es recomendable para estudios futuros.

La metodología de estudio con correlación y predicción es estándar para estudios estadísticos, aunque su aplicación sucesiva e iterativa para los diferentes casos además del análisis de como determinar los clústeres tanto con algoritmos como con razonamiento son muy recomendables para estudios futuros.

La mayor limitación sobre este tipo de estudios estadísticos es la riqueza de las variables, su veracidad, su interpretación y su variedad por lo que es muy recomendable, como se ha realizado en este estudio, dedicar una parte importante del estudio a encontrar variables adecuadas y fuentes que permitan una riqueza de estas.

# Capítulo 5.- Bibliografía

- [UNPD23] Objetivos de desarrollo sostenible | Programa de las Naciones Unidas para el Desarrollo. (2023). UNDP. <https://www.undp.org/es/sustainable-development-goals>
- [EC23] European Commission - European Commission. [https://ec.europa.eu/commission/presscorner/detail/en/ip\\_23\\_3192](https://ec.europa.eu/commission/presscorner/detail/en/ip_23_3192)
- [BOE15] Breaking the Tragedy of the Horizon - Climate Change and Financial Stability - Speech by Mark Carney. (2015). Bank of England. <https://www.bankofengland.co.uk/speech/2015/breaking-the-tragedy-of-the-horizon-climate-change-and-financial-stability>
- [MASTERS23] Masters, B., & Temple-West, P. (2023, 23 agosto). BlackRock's support for climate and social resolutions falls sharply. Financial Times. <https://www.ft.com/content/06fb1b85-56ba-48cd-b6f6-75f8b8eee7e1>
- [IE23] Investment Executive. (2023). When it comes to ESG, fund managers damned if they do and damned if they don't. Investment Executive. <https://www.investmentexecutive.com/news/research-and-markets/when-it-comes-to-esg-fund-managers-damned-if-they-do-and-damned-if-they-dont/#:~:text=Intensifying%20political%20polarization%20over%20the%20role%20of%20ESG,challenge%20to%20asset%20managers%20caught%20in%20the%20middle>
- [STUTTS23] Stutts, J. (2023, 5 abril). JPMorgan opposes climate shareholder proposals ahead of proxy vote. American Banker. <https://www.americanbanker.com/news/jpmorgan-opposes-climate-shareholder-proposals-ahead-of-proxy-vote>
- [BAK21] Bak, C., Szapáry, P. M., Ibanez, E. P., & González, J. (2023). The ESG Value and Leadership Index. Kearney. <https://www.kearney.com/service/sustainability/article/-/insights/the-esg-value-and-leadership-index>
- [S&P23] ESG and sustainable indices | S&P Global. (s. f.). <https://www.spglobal.com/spdji/en/indices/products/esg-sustainable-indices.html>
- [BBVA22] BBVA Communications. (2022, 18 mayo). ¿Sabes qué es el índice Dow Jones Sustainability Index? BBVA NOTICIAS. <https://www.bbva.com/es/sostenibilidad/sabes-que-es-el-indice-dow-jones-sustainability-index/>
- [RAMIC19] Ramić, H. (2019). Relationship between ESG performance and Financial Performance of Companies: An overview of the issue. ResearchGate. <https://doi.org/10.13140/RG.2.2.34837.68322>
- [WHELAN21] Whelan, T., Atz, U., Van Holt, T. and Clark, C. (2021). Uncovering the Relationship by Aggregating Evidence from 1,000 Plus Studies Published between 2015 – 2020. NYU Stern Center. [https://www.stern.nyu.edu/sites/default/files/assets/documents/NYU-RAM\\_ESG-Paper\\_2021%20Rev\\_0.pdf](https://www.stern.nyu.edu/sites/default/files/assets/documents/NYU-RAM_ESG-Paper_2021%20Rev_0.pdf)
- [WEF20] These 3 charts show what COVID-19 has done to global energy demand. (2020, 3 Agosto). World Economic Forum.

<https://www.weforum.org/agenda/2020/08/covid19-change-energy-electricity-use-lockdowns-falling-demand/>

- [EU21] Eurostat. (2021, junio 29). Renewable energy largely spared from pandemic effects. Eurostat. <https://ec.europa.eu/eurostat/web/products-eurostat-news/-/DDN-20210629-1>

Todas las figuras y tablas realizadas son de realización propia.

# Capítulo 6.- Apéndices

## 6.1.- Apéndice A: Visualización de los datos

Las gráficas expuestas en este apéndice corresponden a la representación de los datos en torno a las dos variables escogidas: Ingresos (€m, eje horizontal) y el incremento de la capitalización de la empresa como porcentaje respecto al año anterior.

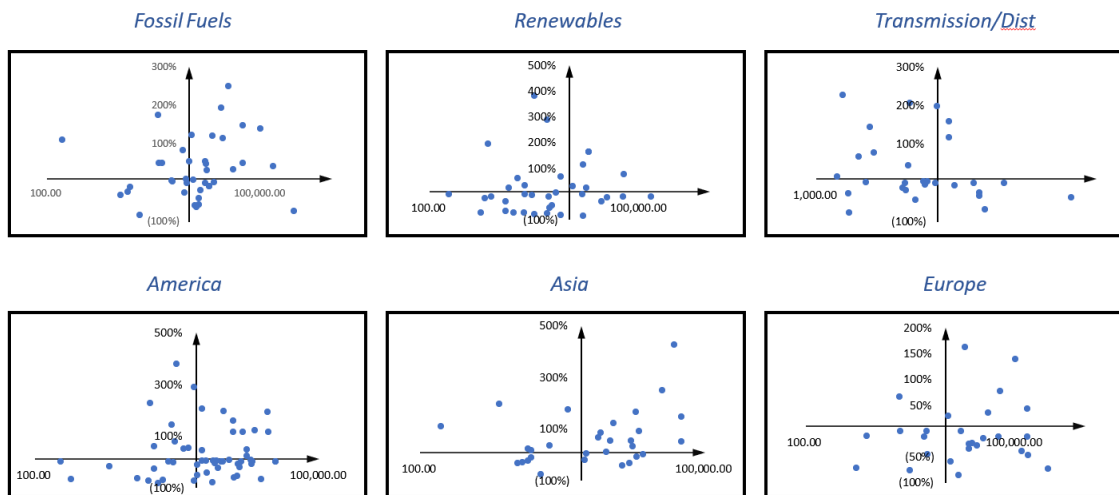


Figura 43: Empresas año 2022 Ingresos – Delta Capitalización (21-22)

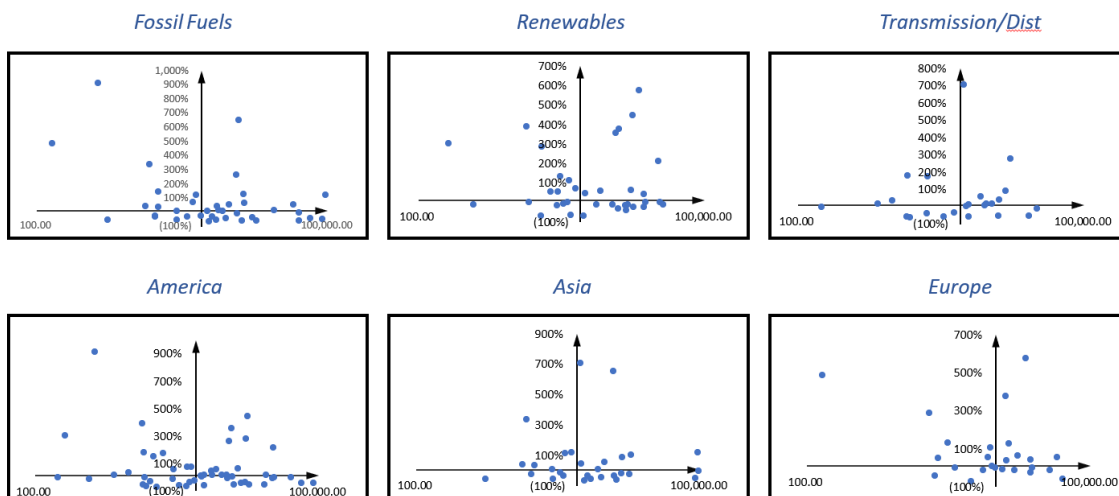


Figura 44: Empresas año 2021 Ingresos – Delta Capitalización (20-21)

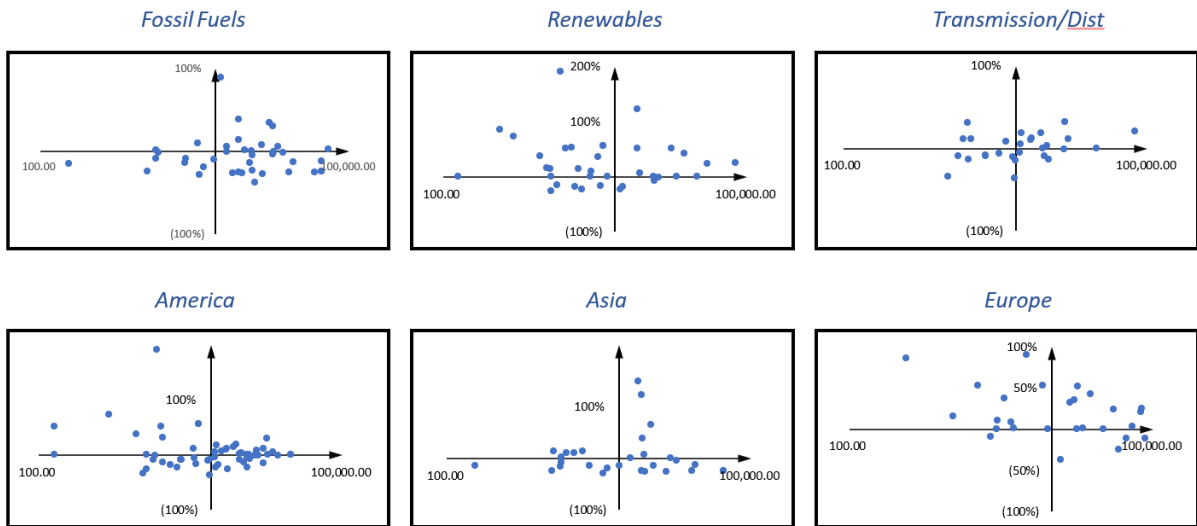


Figura 45: Empresas año 2020 Ingresos – Delta Capitalización (19-20)

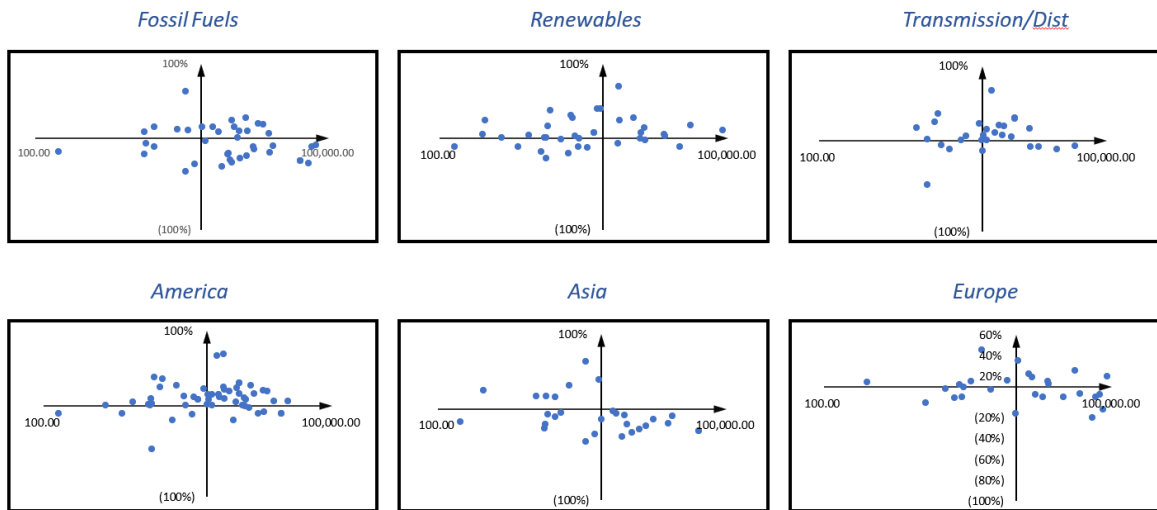


Figura 46: Empresas año 2019 Ingresos – Delta Capitalización (18-19)



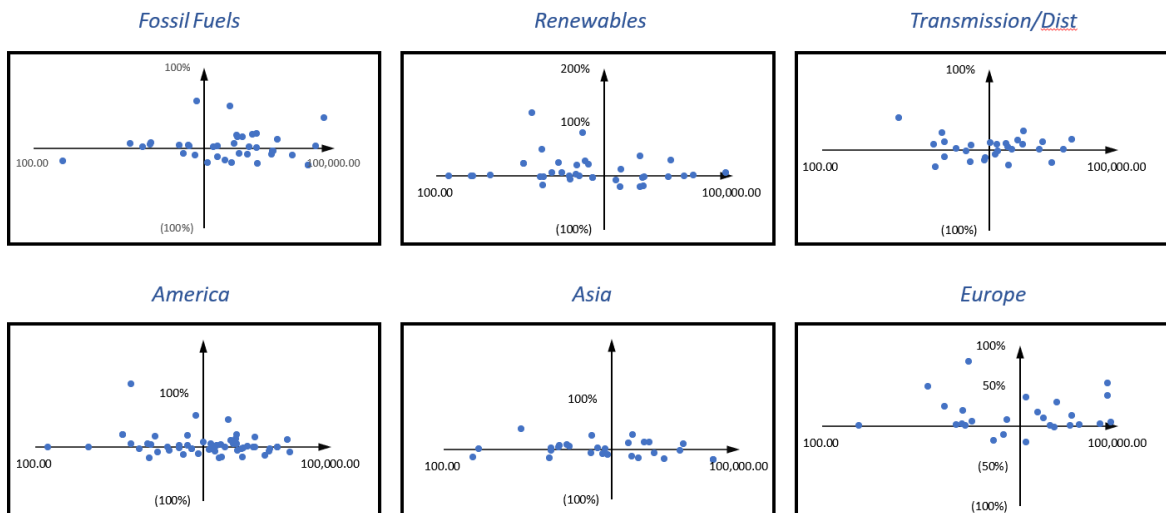


Figura 47: Empresas año 2018 Ingresos – Delta Capitalización (17-18)

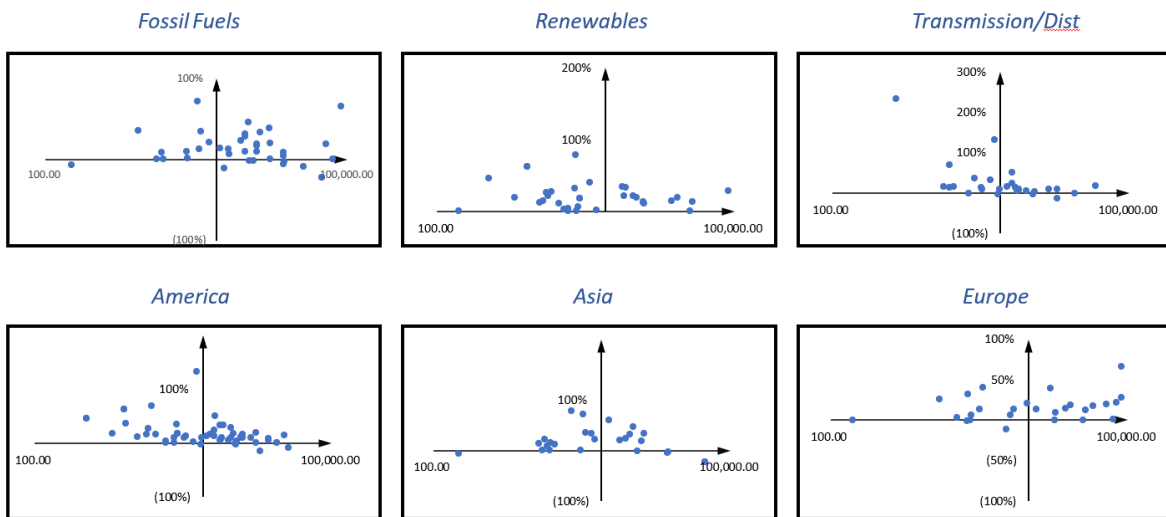


Figura 48: Empresas año 2017 Ingresos – Delta Capitalización (16-17)

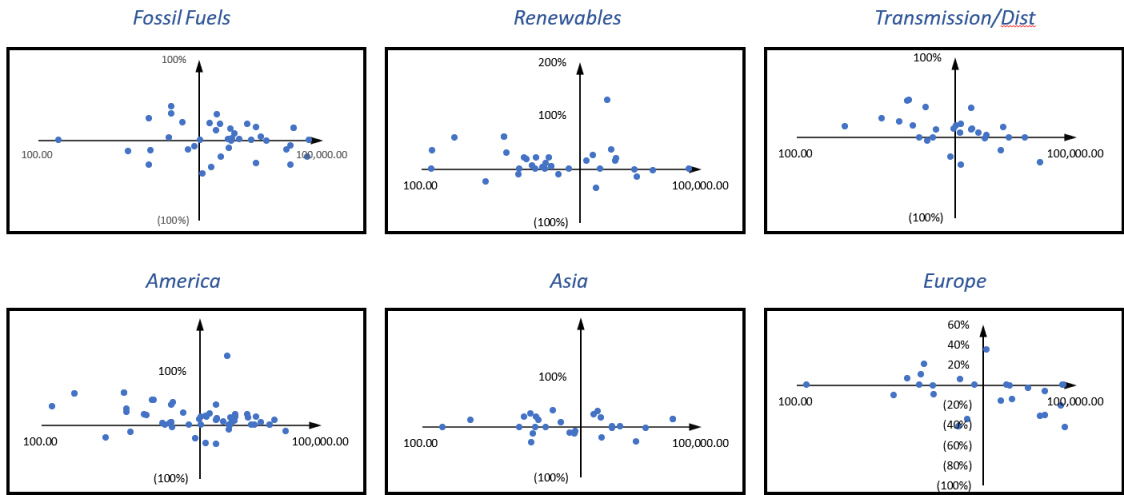


Figura 49: Empresas año 2016 Ingresos – Delta Capitalización (15-16)