



U N I V E R S I D A D
COMPLUTENSE
M A D R I D



Founder Personality Investment Score (FPIS): diseño y validación de un modelo predictivo interpretable para la estimación del éxito de startups a partir de la personalidad y composición del equipo fundador.

Máster Data Science, Big Data & Business Analytics 2024-2025

Darío Díaz Barros

Índice

1. Resumen (Abstract)
2. Introducción
 - 2.1 Contexto y motivación de negocio
 - 2.2 Revisión de la literatura
 - 2.3 Objetivos del TFM
 - 2.4 Hipótesis de investigación
3. **Metodología**
 - 3.1 Dataset y fuentes de datos
 - 3.2 Definición formal del FPIS
 - 3.3 Preparación de datos y *feature engineering*
 - 3.4 Exploratory Data Analysis (EDA)
 - 3.4.1 Desbalance de clases
 - 3.4.2 Número de fundadores
 - 3.4.3 Distribución geográfica
 - 3.4.4 Rasgos Big Five agregados
 - 3.4.5 Diversidad y complementariedad psicológica
 - 3.4.6 Correlaciones con la variable objetivo
 - 3.4.7 Facetas Big Five
4. **Modelado predictivo**
 - 4.1 Modelos comparados
 - 4.2 Estrategia de validación
 - 4.2.1 Validación cruzada estratificada
 - 4.2.2 Métricas de evaluación
 - 4.2.3 Calibración probabilística
 - 4.3 Optimización de hiperparámetros
 - 4.3.1 Búsqueda preliminar con RandomizedSearchCV
 - 4.3.2 Optimización fina con Optuna
 - 4.3.3 Resultados de la optimización
 - 4.4 Calibración de probabilidades
 - 4.4.1 Técnicas de calibración aplicadas
 - 4.4.2 Resultados de la calibración
 - 4.4.3 Interpretación para negocio
5. **Resultados**
 - 5.1 Rendimiento de modelos base vs FPIS
 - 5.1.1 Modelos baseline
 - 5.1.2 Modelos con XP
 - 5.1.3 Comparación global
 - 5.2 Comparación de métricas detalladas
 - 5.2.1 ROC-AUC
 - 5.2.2 PR-AUC
 - 5.2.3 Brier score
 - 5.2.4 Lift y concentración de éxitos
 - 5.3 Validación de hipótesis
 - 5.3.1 Hipótesis H1 — Valor añadido

- 5.3.2 Hipótesis H2 — Segmentación
- 5.3.3 Hipótesis H3 — Decisión práctica
- 5.4 Interpretabilidad con SHAP
 - 5.4.1 SHAP global — Importancia de variables
 - 5.4.2 SHAP individual — Explicabilidad de startups concretas
 - 5.4.3 Implicaciones para negocio

6. **Discusión**

- 6.1 Relevancia para *venture capital*
- 6.2 Cumplimiento de hipótesis
- 6.3 Limitaciones del enfoque
- 6.4 Futuras líneas de trabajo

7. **Conclusiones**

8. **Bibliografía**

9. **Enlaces**

- 9.1 Drive
- 9.2 Github

10. **Anexos**

- 9.1 Figura 1
- 9.2 Figura 2
- 9.3 Figura 3
- 9.4 Figura 4
- 9.5 Figura 5
- 9.6 Figura 6
- 9.7 Figura 7
- 9.8 Figura 8
- 9.9 Figura 9
- 9.10 Figura 10
- 9.11 Figura 11
- 9.12 Figura 12
- 9.13 Figura 13
- 9.14 Figura 14

1. Resumen (Abstract)

Este Trabajo Fin de Máster diseña y valida el Founder Personality Investment Score (FPIS), un indicador predictivo e interpretable que estima la probabilidad de éxito de una startup (exit vía adquisición o IPO) exclusivamente a partir de los rasgos de personalidad Big Five agregados del equipo fundador y métricas de composición/diversidad del equipo. El enfoque busca ser accionable en fases pre-seed, evitando variables de contexto (país, sector, financiación), de modo que el FPIS pueda utilizarse en *due-diligence* temprana cuando la información financiera es escasa.

La investigación utiliza el dataset público de McCarthy et al. (2023). Tras la depuración y preparación, el cuaderno de trabajo contiene $n = 21.160$ organizaciones con la etiqueta binaria *success* (no nula) y un bloque de rasgos Big Five a nivel de fundador que se agregan al nivel de equipo (medias, máximos, dispersión), junto con índices de diversidad/complementariedad y el nº de fundadores. Denotamos por XPX_P al vector de predictores de personalidad y composición del equipo. El FPIS se formaliza como una probabilidad calibrada:

$$FPIS(x_P) = P(Y=1 | X_P=x_P) = g(f_P(x_P))$$

$$FPIS(x_P) = \mathbb{P}(Y=1 | X_P=x_P) = g(f_P(x_P))$$

Donde: $f_P(\cdot)$ es el modelo predictivo entrenado solo con XPX_P y $g(\cdot)$ es una función de calibración (Platt o isotónica) seleccionada por Brier/ECE.

Metodológicamente, se realiza un EDA para entender la distribución de *success*, la relación del nº de fundadores con el resultado, la variabilidad de los Big Five a nivel de equipo y la diversidad entre cofundadores. En modelado se comparan Regresión Logística, Random Forest, XGBoost y LightGBM con validación estratificada, optimización de hiperparámetros y calibración de probabilidades (Platt/Isotónica). La evaluación se centra en ROC-AUC, PR-AUC, Brier score, curvas de ganancias/lift y Net Benefit (análisis de curvas de decisión, DCA).

Los resultados muestran que el modelo ganador es XGBoost calibrado (mejor método por Brier/ECE) entrenado únicamente con XPX_P . En el cuaderno, el modelo alcanza ROC-AUC = 0,676 y PR-AUC = 0,287, con Brier = 0,220; frente al baseline de referencia (ROC-AUC = 0,646; PR-AUC = 0,249), se observa una mejora de +2,78 p.p. en AUC y un incremento apreciable en PR-AUC. En términos de priorización, el quintil superior por FPIS logra Lift Q5/Q1 $\approx 5,12\times$ (modelo calibrado) y, a nivel de cobertura, el Top-20% por FPIS presenta una tasa de éxito $\approx 0,289$ frente al $\approx 0,162$ global, lo que evidencia una fuerte capacidad de concentración de exits en la parte alta del *score*. En decisión, el modelo exhibe Net Benefit positivo en umbrales de decisión relevantes, y se reportan umbrales óptimos bajo distintos: coste beneficio.

Desde el punto de vista de negocio, el FPIS permite priorizar deal-flow en pre-seed, enriquecer comités de inversión con una probabilidad calibrada e interpretar los motores del *score* con SHAP, aportando trazabilidad y explicabilidad a la decisión. En conjunto, la evidencia empírica del cuaderno respalda que las características psicológicas y la composición del equipo fundador contienen información predictiva accionable sobre el éxito, y que un *scoring* basado solo en XPX_P puede mejorar la selección de oportunidades sin depender de señales de contexto no disponibles en etapas tempranas.

2. Introducción

2.1 Contexto y motivación de negocio

La inversión en startups en fases tempranas se caracteriza por elevados niveles de incertidumbre y un acceso muy limitado a información financiera o de mercado. En particular, en la etapa pre-seed, los inversores deben tomar decisiones críticas sin métricas consolidadas de tracción comercial, ingresos o clientes. En este contexto, la figura del equipo fundador se convierte en el principal activo evaluable, pues la literatura ha demostrado que sus características individuales y colectivas influyen de manera decisiva en el rendimiento futuro de la empresa (Rauch & Frese, 2007).

La personalidad de los fundadores ha emergido en los últimos años como una dimensión clave de análisis. El estudio de McCarthy et al. (2023), que constituye la base empírica de este TFM, evidenció que ciertos rasgos de los Big Five y su complementariedad dentro del equipo están correlacionados con la probabilidad de alcanzar un *exit* (adquisición o salida a bolsa). Sin embargo, este hallazgo no se ha traducido aún en un indicador operativo accionable que pueda ser utilizado directamente en procesos de *due diligence*.

El problema de negocio es, por tanto, claro: los fondos de capital riesgo y *business angels* necesitan herramientas predictivas, interpretables y aplicables en fases tempranas, que les permitan priorizar su deal-flow antes de contar con datos financieros. En este escenario surge la propuesta del FPIS, un *score* diseñado para estimar la probabilidad de éxito de una startup a partir de la huella psicológica colectiva del equipo fundador, calculada únicamente con información disponible públicamente.

Este enfoque aporta tres motivaciones principales:

1. **Accionabilidad temprana:** el FPIS se construye con datos derivados de la huella digital de los fundadores (p. ej., Twitter), lo que lo hace utilizable incluso cuando la startup aún no ha generado métricas de negocio.
2. **Ventaja informacional:** introduce una dimensión diferencial respecto a los filtros tradicionales de inversores (industria, país, número de fundadores), al capturar la diversidad y complementariedad psicológica del equipo.
3. **Interpretabilidad y trazabilidad:** al basarse en agregados de percentiles Big Five y métricas de diversidad ya presentes en el dataset, permite comprender qué rasgos impulsan el *score* y facilita justificar las decisiones de inversión ante comités.

Por lo tanto, este TFM se enmarca en la intersección entre psicología, ciencia de datos y capital riesgo, con la ambición de convertir un hallazgo académico en una herramienta práctica para inversores.

2.2 Revisión de la literatura

En los últimos años, el paradigma de los **Big Five** —apertura, responsabilidad, extraversión, amabilidad y estabilidad emocional— se ha convertido en la referencia para modelar la personalidad en contextos organizativos y emprendedores. Su aplicación en capital riesgo es especialmente prometedora, dado que estos rasgos se pueden **inferir digitalmente a partir de huellas en redes sociales** y agregarse a nivel de equipo para capturar dinámicas colectivas.

El estudio de McCarthy et al. (2023) supuso un punto de inflexión en esta línea de investigación. A partir de un conjunto masivo de datos de fundadores y sus startups, los autores demostraron empíricamente que:

- La composición psicológica de los equipos fundadores contiene información predictiva sobre el éxito de las startups.
- La diversidad intraequipo y la complementariedad psicológica explican diferencias significativas en la probabilidad de exit.
- Es posible construir modelos predictivos a gran escala utilizando únicamente información derivada de los perfiles de los fundadores, sin necesidad de incorporar métricas financieras iniciales.

No obstante, el propio trabajo reconocía una limitación: la falta de un score interpretable y listo para ser aplicado en procesos reales de inversión. Si bien se exploraron correlaciones y modelos, no se ofreció una herramienta calibrada y validada que pudiera ser adoptada directamente por *venture capitalists* o *business angels*.

En paralelo, otras líneas de investigación en team diversity (Harrison & Klein, 2007), han subrayado que la variedad en características individuales —incluyendo edad, formación, género o rasgos psicológicos— puede generar ventajas competitivas cuando se traduce en complementariedad de perspectivas y estilos de liderazgo. Sin embargo, también advierten que un exceso de heterogeneidad puede aumentar los costes de coordinación y conflicto, lo que obliga a cuantificar la diversidad de manera equilibrada.

En conjunto se apunta a un consenso:

1. La personalidad y composición de los fundadores son determinantes en el destino de las startups.
2. Existen métricas psicológicas que pueden cuantificarse y modelarse predictivamente.
3. Aún falta trasladar estos hallazgos a un KPI accionable, calibrado y explicable, que permita a los inversores priorizar oportunidades en fases tempranas.

Este vacío es precisamente el que se propone cubrir con el FPIS desarrollado en este trabajo.

2.3 Objetivos del TFM

El objetivo principal de este Trabajo Fin de Máster es diseñar y validar el FPIS, un indicador predictivo e interpretable que permita estimar la probabilidad de éxito de una startup —definido como alcanzar un *exit* vía adquisición o salida a bolsa— a partir de la personalidad agregada y la composición del equipo fundador.

A diferencia de aproximaciones previas, el FPIS se fundamenta únicamente en la información disponible en el dataset público de McCarthy et al. (2023):

- Rasgos de personalidad Big Five de los fundadores.
- Agregaciones estadísticas a nivel de equipo (máximos, medias, varianzas, IQR).
- Métricas de diversidad psicológica (índice de Blau, FOALED, shares de tipos).

- Complementarity Index (distancia media respecto al centroide de equipo).
- Estructura del equipo (número de fundadores).

Quedan excluidas explícitamente variables de contexto como el país o el sector en el entrenamiento principal, con el fin de garantizar que el score sea aplicable en fases pre-seed cuando esta información aún no es sólida o no constituye una señal fiable de éxito.

Desde la perspectiva de negocio, este objetivo se traduce en tres aportaciones diferenciadas:

1. Accionabilidad antes de invertir
2. Ventaja informacional.
3. Interpretabilidad.

2.4 Hipótesis de investigación

La validación de **FPIS** se articula en torno a tres hipótesis principales:

H1. Valor añadido.

La incorporación de variables de personalidad y composición del equipo (vector XPX_PXP) mejora de manera significativa la capacidad predictiva frente a un modelo base construido únicamente con controles contextuales (sector, país y número de fundadores). Se espera un incremento del AUC ≥ 5 puntos porcentuales en la predicción de éxito.

H2. Segmentación.

Las startups situadas en el quintil superior del FPIS presentarán una tasa de éxito al menos doble respecto a las del quintil inferior. Esta hipótesis evalúa la utilidad práctica del score como herramienta de priorización de *deal-flow*.

H3. Decisión práctica.

El uso de un umbral de FPIS calibrado según la relación coste/beneficio de falsos positivos y falsos negativos permitirá generar un retorno esperado $\geq 15\%$ superior al de una estrategia aleatoria de inversión en el mismo sector.

Estas hipótesis combinan un enfoque metodológico (H1, comparando métricas de modelos), estratégico de segmentación (H2, identificando concentración de éxitos en la parte alta del score) y de aplicabilidad en negocio (H3, simulando retornos y net benefit). En conjunto, ofrecen un marco sólido para evaluar si el FPIS constituye realmente un KPI diferencial y accionable en procesos de inversión en fases tempranas.

3. Metodología

3.1 Dataset y fuentes de datos

La base empírica utilizada en este TFM procede del dataset público publicado McCarthy et al. (2023) en el marco de su investigación *The Impact of Founder Personalities on Startup Success*. Este conjunto de datos, disponible en repositorios académicos y GitHub, contiene información sobre más de 20.000 startups y sus equipos fundadores, junto con sus características psicológicas inferidas a partir de huellas digitales.

En concreto, el dataset recoge:

- Identificadores de organización (*org_id*) y variables contextuales básicas como país (*org_country*) y número de fundadores (*org_numfounders*).
- La variable binaria de resultado *success*, que toma valor 1 si la startup alcanzó un *exit* (adquisición o salida a bolsa) y 0 en caso contrario.
- Los rasgos de personalidad Big Five de cada fundador, inferidos mediante análisis lingüístico de publicaciones en Twitter. Estos rasgos incluyen tanto los cinco factores principales (apertura, responsabilidad, extraversión, amabilidad y estabilidad emocional) como sus facetas desagregadas.
- Agregaciones a nivel de equipo ya calculadas en el dataset original, como máximos de percentiles (*big5_max_...*), medias, varianzas o índices de dispersión.
- Métricas de diversidad psicológica (ej. índice de Blau, FOALED) y el Complementarity Index, que mide la distancia media de cada miembro respecto al perfil psicológico central del equipo.

En el cuaderno de trabajo se cargó el archivo *AdditionalData.csv*, con un total de 21.160 organizaciones y 41 variables originales. Tras la depuración y verificación, se construyó un *team_df* con 49 columnas que integran las variables Big Five, agregaciones, métricas de diversidad y la variable objetivo *success*.

El dataset se caracteriza por dos aspectos clave:

1. Desbalance de clases. Solo alrededor del 16,22 % de las startups alcanzaron un *exit*, lo que introduce un importante reto de clasificación (*Figura 1*).
2. Cobertura global. Si bien Estados Unidos concentra casi la mitad de las observaciones (≈ 10.200 startups), el dataset incluye organizaciones de 13 países, lo que permite capturar variabilidad en composición de equipos y tasas de éxito (*Figura 2*).

El uso de este dataset ofrece varias ventajas:

- Relevancia académica, al tratarse de la fuente original del paper que fundamenta el TFM.
- Accionabilidad práctica, dado que se basa en información pública y replicable a partir de huellas digitales de los fundadores.
- Rigurosidad metodológica, al contar con un tamaño muestral suficiente y con variables diseñadas específicamente para modelar personalidad y diversidad de equipos.

En este TFM, el dataset se emplea exclusivamente para construir y validar el FPIS, garantizando que el vector de entrada *XPX_PXP* contenga únicamente rasgos psicológicos agregados y composición del equipo, en línea con lo establecido en los objetivos.

3.2 Definición formal del FPIS

El FPIS se define como un indicador probabilístico e interpretable de la probabilidad de éxito de una startup a partir de las características psicológicas y de composición de su equipo fundador.

Sea X_{PX_PXP} el vector de predictores compuesto por:

- Rasgos de personalidad Big Five agregados a nivel de equipo (máximos, medias, varianzas, IQR).
- Facetas Big Five en percentiles (ej. *Openness_facet_adventurousness_percentile*).
- Métricas de diversidad (índice de Blau, FOALED, shares por tipo).
- Complementarity Index, que mide la distancia media de cada miembro respecto al perfil psicológico central del equipo.
- Número de fundadores (*org_numfounders*).

Sobre este vector de predictores se entrena un modelo predictivo $f_P(\cdot)$, que genera probabilidades no calibradas. Posteriormente, se aplica una función de calibración $g(\cdot)$ (sigmoide de Platt o regresión isotónica) para obtener una probabilidad ajustada e interpretable. Formalmente, el FPIS se define como:

$$FPIS(x_P) = p(Y=1 | X_P=x_P) = g(f_P(x_P))$$

$$FPIS(x_P) = \mathbb{P}(Y=1 | X_P=x_P) = g(f_P(x_P))$$

donde:

- $Y \in \{0,1\}$ representa la variable de éxito de la startup (*exit*).
- $f_P(x_P)$ es la puntuación cruda generada por el modelo entrenado únicamente con X_{PX_PXP} .
- $g(\cdot)$ es la función de calibración que ajusta esa puntuación a una probabilidad bien calibrada.

En la práctica, este score devuelve un valor continuo entre 0 y 1, interpretable como la probabilidad de éxito esperada de la startup en función de su perfil psicológico colectivo.

El diseño de este FPIS responde a tres principios fundamentales:

1. **Accionabilidad temprana.** El score se calcula con variables disponibles desde el inicio (huella digital de los fundadores), lo que permite aplicarlo en fases pre-seed.
2. **Interpretabilidad.** Al estar basado en agregados de percentiles y métricas de diversidad fácilmente comprensibles, los inversores pueden identificar qué características impulsan un valor alto o bajo de FPIS.
3. **Consistencia metodológica.** El FPIS se fundamenta en un modelo entrenado de manera rigurosa con validación cruzada, optimización de hiperparámetros y calibración probabilística, asegurando que la probabilidad generada sea fiable y estable.

De esta manera, el FPIS se configura como una probabilidad calibrada de éxito condicionada exclusivamente a la personalidad y composición del equipo fundador.

3.3 Preparación de datos y *feature engineering*

Una vez cargado el dataset original (*AdditionalData.csv*), se realizaron distintos procesos de preparación de datos y generación de variables con el objetivo de construir el vector XPX_PXP — es decir, la representación del equipo fundador basada únicamente en rasgos de personalidad y composición.

1. Normalización y verificación de columnas

- Se renombró el identificador de organización a `org_id` para mantener consistencia.
- La variable objetivo `success` se tipificó como binaria (0/1).
- Se aseguraron columnas clave como `org_numfounders` y `org_country`, generando valores por defecto en los casos ausentes.

2. Agregación de rasgos de personalidad

A partir de los percentiles individuales de las facetas Big Five (ej. *Openness_facet_adventurousness_percentile*), se construyeron agregaciones a nivel de equipo fundador, siguiendo tres enfoques:

- **Medidas de tendencia central:** media y mediana de cada faceta.
- **Medidas de dispersión:** varianza e intervalo intercuartílico (IQR).
- **Medidas extremas:** máximo y mínimo por rasgo.

De este modo, se obtuvo una representación sintética que captura tanto los niveles promedio de cada rasgo en el equipo como la heterogeneidad interna.

3. Construcción de métricas de diversidad

Con el fin de operacionalizar la hipótesis de que la diversidad psicológica enriquece al equipo pero puede generar costes de coordinación, se calcularon índices específicos:

- **Índice de Blau (FOALED):** mide la heterogeneidad de distribuciones categóricas de rasgos, adaptado aquí a facetas psicológicas.
- **Shares por tipo:** proporción de fundadores en cada extremo de las dimensiones de personalidad.

Estas métricas permiten evaluar si un equipo combina perfiles homogéneos o complementarios.

4. Complementarity Index

Se implementó una métrica de complementariedad psicológica calculada como la distancia media entre cada fundador y el centroide del equipo en el espacio Big Five.

$$CI = \frac{1}{n} \sum_{i=1}^n d(x_i, \bar{x})$$

$$CI = \frac{1}{n} \sum_{i=1}^n d(x_i, \bar{x})$$

donde:

- n es el número de fundadores del equipo.
- x_i representa el vector de rasgos de personalidad del fundador i .
- \bar{x} es el perfil promedio del equipo (centroide en el espacio de los Big Five).
- $d(\cdot, \cdot)$ corresponde a la distancia coseno entre vectores de personalidad.

5. Variables estructurales

Se incorporó el número de fundadores (*org_numfounders*), dado que la literatura sugiere que tanto equipos demasiado pequeños como excesivamente grandes presentan menor probabilidad de éxito (Rauch & Frese, 2007).

6. Manejo de valores nulos

Aunque el dataset presentaba una cobertura elevada, algunas facetas Big Five contenían valores faltantes. Estos se gestionaron mediante:

- Imputación de valores medios cuando el porcentaje de nulos era bajo.
- Exclusión de columnas con escasa cobertura.

7. Resultado final

Tras este proceso, se construyó un *team_df* enriquecido con **49 variables**, entre las que se incluyen:

- **Big Five agregados** (máximos, medias, varianzas, IQR).
- **Diversidad Blau y FOALED.**
- **Complementarity Index.**
- **Número de fundadores.**
- **Variable objetivo success (0/1).**

Este *feature engineering* permitió disponer de un vector *XPX_PXP* completo, interpretable y coherente con los objetivos del TFM, asegurando que el FPIS se fundamente únicamente en la personalidad y estructura del equipo fundador.

3.4 Exploratory Data Analysis (EDA)

Antes de entrenar los modelos predictivos, se llevó a cabo un análisis exploratorio exhaustivo para comprender la distribución de las variables, detectar posibles sesgos y validar la calidad del dataset.

3.4.1 Desbalance de clases

El primer análisis se centró en la distribución de la variable objetivo *success*. De las 21.160 startups incluidas, únicamente un 16,22 % alcanzó un *exit*, frente al 83,78 % que no lo hizo. Este marcado desbalance plantea un reto de clasificación, ya que los modelos tienden a favorecer la clase mayoritaria (*Figura 1*).

3.4.2 Número de fundadores

La distribución de `org_numfounders` muestra que está centrada en 2 y la mayoría de startups tienen entre 1 y 3 fundadores. El análisis de éxito por tamaño de equipo evidencia que las probabilidades de *exit* tienden a ser mayores en equipos de 2 a 3 fundadores, disminuyendo tanto en equipos unipersonales como en aquellos con más de 5 miembros. Esto respalda la hipótesis de que un tamaño moderado de equipo equilibra coordinación y diversidad (*Figura 3*).

3.4.3 Distribución geográfica

Aunque el objetivo principal del TFM no incluye variables de contexto en el modelo predictivo, resulta relevante caracterizar la base de datos. El país con mayor representación es Estados Unidos (≈ 10.200 startups, $\sim 48\%$), seguido por Alemania y Reino Unido. En términos de tasa de éxito, existen diferencias significativas: algunos países como Israel presentan tasas superiores al 20 % (cuando $n \geq 100$ startups), mientras que otros como India muestran tasas cercanas al 8 %. Estas diferencias confirman la heterogeneidad contextual del dataset (*Figura 4*).

3.4.4 Rasgos Big Five agregados

Los agregados de personalidad Big Five revelan patrones interesantes:

- Los equipos con alta apertura (*openness*) tienden a concentrarse en startups con mayor probabilidad de éxito.
- En cambio, puntuaciones extremas de neuroticismo (*emotional range*) se asocian con menores tasas de éxito.
- La distribución por clases (*success vs no success*) muestra diferencias visibles en algunos rasgos, aunque no determinísticas (*Figura 5*).

3.4.5 Diversidad y complementariedad psicológica

El análisis de los índices de diversidad y complementariedad confirma la relevancia de estas métricas:

- El **índice de Blau (FOALED)** presenta una distribución amplia, con valores más altos en equipos diversos. Los boxplots sugieren una relación positiva moderada entre diversidad y éxito.

3.4.6 Correlaciones con la variable objetivo

Se construyó una matriz de correlaciones entre las variables numéricas clave y *success*. Aunque las correlaciones bivariantes son en general bajas —lo cual es esperado en problemas de predicción complejos—, destacan asociaciones en facetas de apertura, consciencia y extraversión, así como con los índices de diversidad. Estas señales, aunque débiles individualmente, son aprovechables por modelos multivariantes más sofisticados (*Figura 6*).

3.4.7 Facetas Big Five

Finalmente, se exploró un subconjunto de facetas específicas de los Big Five, visualizadas mediante gráficos de violín. El análisis muestra que startups exitosas tienden a tener fundadores con mayores percentiles en facetas como *adventurousness* o *achievement striving*, lo que refuerza la relevancia de modelar la personalidad a nivel granular (*Figura 7*).

En conjunto, el EDA confirma que el dataset contiene señales psicológicas y estructurales con capacidad predictiva, aunque no lineales ni triviales, lo que justifica el uso de modelos supervisados avanzados y calibrados para construir el FPIS.

4. Modelado predictivo

4.1 Modelos comparados

Con el objetivo de estimar la probabilidad de éxito de las startups a partir del vector XPX_PXP, se implementaron y compararon distintos algoritmos de clasificación supervisada. La selección de modelos responde a un criterio de equilibrio entre capacidad predictiva, interpretabilidad y uso habitual en entornos de analítica de negocio:

1. Regresión Logística (LogReg).

- Modelo lineal de referencia, ampliamente utilizado en contextos de predicción binaria.
- Permite obtener probabilidades directamente interpretables.
- Sirve como benchmark base frente a modelos más complejos.

2. Random Forest (RF).

- Algoritmo de ensamble basado en árboles de decisión entrenados sobre subconjuntos bootstrap.
- Reduce varianza y mejora robustez frente al sobreajuste.
- Especialmente útil para captar interacciones no lineales entre los rasgos de personalidad.

3. XGBoost (XGB).

- Algoritmo de *gradient boosting* optimizado, reconocido por su alto rendimiento.
- Ajusta iterativamente los errores residuales y permite una gestión precisa del desbalance de clases mediante el parámetro `scale_pos_weight` ($\approx 5,16$ en nuestro dataset).
- Fue el modelo que alcanzó los mejores resultados tras calibración, convirtiéndose en la base del FPIS.

4. LightGBM (LGBM).

- Variante de boosting desarrollada por Microsoft, optimizada en velocidad y consumo de memoria.
- Utiliza histogramas para acelerar el entrenamiento.
- Se incluyó como alternativa a XGBoost, dado su buen desempeño en datasets tabulares de gran tamaño.

Todos los modelos fueron implementados en Python (scikit-learn, XGBoost, LightGBM), integrados en *pipelines* que incluían escalado, entrenamiento y calibración probabilística. El proceso de comparación se diseñó para responder a la hipótesis H1 (valor añadido frente al baseline) y determinar si el FPIS basado en XGBoost calibrado ofrecía mejoras estadísticamente y prácticamente relevantes.

4.2 Estrategia de validación

La validación de los modelos se diseñó cuidadosamente para garantizar que los resultados fuesen robustos, comparables y generalizables a las startups del dataset.

4.2.1 Validación cruzada estratificada

Dado el desbalance de clases en la variable objetivo (*success* \approx 16,22 %), se empleó una estrategia de validación cruzada estratificada (StratifiedKFold). Este enfoque asegura que en cada partición del conjunto de entrenamiento y validación se mantenga la misma proporción de casos positivos y negativos que en el dataset completo, evitando sesgos en las métricas.

- Se utilizó $k = 5$ folds, estándar en problemas de clasificación binaria.
- En cada iteración, el modelo se entrenó en el 80 % de los datos y se evaluó en el 20 % restante.
- Los resultados reportados corresponden a la media de métricas en los folds, lo que reduce la varianza debida a particiones específicas.

4.2.2 Métricas de evaluación

Para evaluar la calidad de los modelos y validar las hipótesis, se recurrió a un conjunto de métricas complementarias, que permiten analizar diferentes dimensiones del rendimiento predictivo:

1. **ROC-AUC (Receiver Operating Characteristic – Area Under Curve).**
 - Mide la capacidad global del modelo para discriminar entre clases positivas (éxito) y negativas (fracaso).
 - El baseline (modelo con solo controles) alcanzó **AUC = 0,646**, mientras que el modelo FPIS (XGBoost calibrado) logró **0,676**, es decir, una mejora de **+2,78 p.p.**
2. **PR-AUC (Precision-Recall – Area Under Curve).**
 - Especialmente relevante en contextos con clases desbalanceadas.
 - El baseline alcanzó **0,251**, frente a **0,287** del FPIS, lo que representa un incremento de **+3,6 p.p.** en capacidad de recuperación de éxitos reales.
3. **Brier Score.**
 - Evalúa la calidad de las probabilidades predichas, penalizando desviaciones respecto a los resultados reales.
 - El FPIS (XGB calibrado) obtuvo un **Brier = 0,220**.

4. Curvas de ganancias y lift.

- Permiten analizar la capacidad del modelo de concentrar la mayor proporción de éxitos en los percentiles superiores del score.
- En el notebook, el quintil superior del FPIS concentró hasta **5,1 veces más éxitos** que el quintil inferior.

5. Net Benefit (Decision Curve Analysis).

- Evalúa la utilidad práctica de un modelo en escenarios de decisión bajo distintos umbrales de riesgo.
- El FPIS mostró beneficios netos positivos en rangos de umbral relevantes para inversión, validando su aplicabilidad en negocio.

4.2.3 Calibración probabilística

Un aspecto central en la estrategia de validación fue garantizar que el score resultante fuese probabilísticamente interpretable. Para ello:

- Se compararon dos métodos de calibración (Platt/sigmoide e isotónica) y se eligió el que minimizó Brier/ECE en validación. En la evaluación final del XGBoost ganador, el método seleccionado fue isotónica, que ofreció la mejor combinación de Brier y ECE en el conjunto de test..

En conjunto, esta estrategia permitió no solo comparar el rendimiento relativo de los modelos, sino también garantizar que el FPIS fuese un score calibrado, fiable e interpretable, cumpliendo así con los objetivos metodológicos del trabajo.

4.3 Optimización de hiperparámetros

Con el fin de maximizar el rendimiento de los modelos y evitar configuraciones subóptimas, se implementó un proceso de optimización de hiperparámetros en dos fases:

4.3.1 Búsqueda preliminar con RandomizedSearchCV

En una primera etapa, se utilizó **RandomizedSearchCV** de *scikit-learn* para explorar de manera eficiente los espacios de hiperparámetros de cada algoritmo:

- **Regresión Logística:** coeficiente de regularización CCC, tipo de penalización (L1/L2).
- **Random Forest:** número de árboles (*n_estimators*), profundidad máxima (*max_depth*), proporción de variables por split (*max_features*).
- **XGBoost:** tasa de aprendizaje (*learning_rate*), profundidad máxima (*max_depth*), número de iteraciones (*n_estimators*), subsampling y *colsample_bytree*.
- **LightGBM:** tasa de aprendizaje, número de hojas (*num_leaves*), fracción de características (*feature_fraction*), fracción de muestras (*bagging_fraction*).

Esta búsqueda permitió identificar rangos prometedores para cada parámetro, reduciendo el riesgo de sobreajuste y descartando combinaciones poco eficientes.

4.3.2 Optimización fina con Optuna

En la segunda fase, se aplicó **Optuna** con el algoritmo **Tree-structured Parzen Estimator (TPE)**, que permite un ajuste más sofisticado mediante aprendizaje secuencial de distribuciones de probabilidad.

- Se definieron funciones objetivo específicas para cada modelo, utilizando ROC-AUC como métrica de optimización principal.
- Optuna se ejecutó con pruning adaptativo, descartando configuraciones poco prometedoras de forma temprana para ahorrar tiempo computacional.
- En el caso de XGBoost, el modelo finalmente seleccionado por Optuna fue una configuración que equilibró adecuadamente bias y varianza, ofreciendo el mejor desempeño global en validación cruzada.

4.3.3 Resultados de la optimización

La optimización mostró que:

- En **LogReg**, el mejor rendimiento se obtenía con regularización L2 y valores de CCC bajos (modelo más penalizado).
- En **RF**, profundidades intermedias (≤ 10) y un número alto de árboles (> 300) maximizaban el rendimiento sin incurrir en sobreajuste.
- En **XGBoost y LightGBM**, tasas de aprendizaje bajas combinadas con un número elevado de iteraciones permitieron capturar patrones complejos sin degradar la calibración.

En conjunto, la fase de optimización confirmó que los **modelos de boosting (XGB y LGBM)** superaban de forma consistente a los demás, siendo **XGBoost calibrado** el que ofreció los mejores resultados, lo que justificó su elección como base del FPIS.

4.4 Calibración de probabilidades

Uno de los requisitos fundamentales del FPIS es que los valores generados sean probabilidades bien calibradas, es decir, que la magnitud del score refleje con precisión la probabilidad real de éxito de una startup.

En problemas de clasificación binaria con datos desbalanceados, los modelos de *machine learning* tienden a producir probabilidades mal calibradas: suelen sobrestimar la probabilidad de éxito en casos positivos y subestimarla en los negativos. Esto limita su utilidad práctica en contextos de negocio, donde los inversores necesitan probabilidades que puedan interpretarse y compararse directamente con umbrales de decisión.

4.4.1 Técnicas de calibración aplicadas

En este trabajo se evaluaron dos métodos clásicos de calibración:

- **Platt scaling (sigmoide).**
Consiste en ajustar una regresión logística sobre las probabilidades crudas del modelo, transformándolas mediante una función sigmoide. Se trata de un método simple y robusto, especialmente eficaz cuando el número de observaciones positivas es limitado.

- **Calibración isotónica.**

Utiliza una función escalonada no paramétrica que ajusta las probabilidades de manera más flexible, adaptándose a la distribución empírica de la variable objetivo. Aunque puede ofrecer una calibración más precisa, requiere mayor cantidad de datos para evitar sobreajuste.

Ambos procedimientos se implementaron mediante el módulo **CalibratedClassifierCV** de *scikit-learn*, aplicados sobre los modelos de boosting (XGB y LGBM) y, en menor medida, sobre RF y LogReg.

4.4.2 Resultados de la calibración

Los resultados del notebook muestran que:

- El **XGBoost calibrado con isotónica** ofreció la **mejor curva de calibración**, reduciendo la desviación entre probabilidades predichas y frecuencias observadas.
- La calibración isotónica, si bien mejoró las predicciones en algunos folds, tendía a **sobreajustar** en submuestras pequeñas, mostrando menor estabilidad que Platt.
- En métricas globales, el XGB calibrado alcanzó:
 - **ROC-AUC = 0,676**
 - **PR-AUC = 0,287**
 - **Brier score = 0,220**

valores que lo consolidaron como el modelo óptimo para construir el FPIS.

4.4.3 Interpretación para negocio

La calibración convierte al FPIS en un **score interpretable como probabilidad real de éxito**, lo que lo diferencia de modelos que solo generan rankings relativos. Esto habilita aplicaciones directas en *venture capital*, como:

- Establecer **umbrales de inversión** (ej. invertir solo si FPIS > 0,25).
- Comparar startups de sectores distintos bajo una **escala común de riesgo**.
- Simular **retornos esperados** ajustando la probabilidad de éxito con factores de retorno y coste de errores.

La calibración probabilística garantizó que el FPIS no solo fuese predictivo, sino también accionable e interpretable.

5. Resultados

5.1 Rendimiento de modelos base vs FPIS

La primera comparación se realizó entre los modelos baseline (con variables contextuales básicas) y los modelos enriquecidos con XPX_PXP, vector que contiene únicamente rasgos de personalidad y composición del equipo. El objetivo de este análisis era comprobar si el FPIS añadía valor predictivo, en línea con la hipótesis H1.

5.1.1 Modelos baseline

El baseline incluyó como predictores únicamente el **sector**, el **país** y el **número de fundadores** (*org_numfounders*). Aunque estas variables capturan ciertas diferencias estructurales, su capacidad discriminativa es limitada:

- **ROC-AUC = 0,646**
- **PR-AUC = 0,249**

Estos resultados reflejan un rendimiento modesto y una pobre calibración, insuficientes para un uso práctico en priorización de inversiones.

5.1.2 Modelos con XPX_PXP (Founder Personality Investment Score)

Al incorporar el vector XPX_PXP, que recoge rasgos agregados de personalidad, índices de diversidad y métricas de complementariedad, el rendimiento mejoró de forma consistente:

- El **modelo XGBoost calibrado (isotónica)** alcanzó **ROC-AUC = 0,676**, **PR-AUC = 0,287** y **Brier score = 0,220**.
- En términos relativos, esto supone una **mejora de +2,78 p.p. en ROC-AUC** y de **+3,6 p.p. en PR-AUC** respecto al baseline.
- Además, el Brier score inferior confirma una calibración más ajustada de las probabilidades.

5.1.3 Comparación global

El análisis comparativo entre baseline y FPIS puede visualizarse en las curvas ROC y PR, donde el modelo FPIS domina claramente sobre la referencia en todas las regiones de umbral (*Figura 8 y Figura 9*).

En conjunto, estos resultados validan que la información psicológica y de composición del equipo aporta un valor predictivo diferencial, superando la capacidad de predicción de variables contextuales. Esta evidencia es consistente con la hipótesis H1 y demuestra que el FPIS constituye un score con mayor utilidad práctica en fases pre-seed.

5.2 Comparación de métricas detalladas

Además de la comparación baseline vs. FPIS, se evaluó el rendimiento de los cuatro algoritmos considerados (Regresión Logística, Random Forest, XGBoost y LightGBM), todos entrenados exclusivamente con el vector XPX_PXP. El análisis se centró en métricas complementarias para capturar diferentes dimensiones del rendimiento predictivo.

5.2.1 ROC-AUC

La **curva ROC-AUC** mide la capacidad global de discriminación entre startups exitosas y no exitosas. Los resultados fueron:

- **LogReg (calibrada): AUC = 0,675**
- **Random Forest: AUC ≈ 0,662**
- **LightGBM: AUC ≈ 0,670**

- **XGBoost (calibrado Platt): AUC = 0,676**

Aunque las diferencias absolutas son pequeñas, XGBoost logró el mejor desempeño, confirmando su robustez en datos tabulares con relaciones no lineales.

5.2.2 PR-AUC

La **curva precisión-recall** resulta más informativa en contextos desbalanceados como el presente (exit rate = 16,22 %). Aquí los resultados fueron:

- **LogReg: PR-AUC = 0,294**
- **Random Forest: PR-AUC ≈ 0,271**
- **LightGBM: PR-AUC ≈ 0,281**
- **XGBoost (calibrado): PR-AUC = 0,287**

La Regresión Logística obtuvo un valor ligeramente superior en PR-AUC, lo que confirma su buen ajuste en escenarios lineales, aunque en métricas globales el XGBoost calibrado fue más equilibrado.

5.2.3 Brier score

El **Brier score**, que mide la exactitud de las probabilidades predichas, mostró una clara ventaja para los modelos calibrados:

- **XGBoost calibrado: 0,220**
- **LightGBM calibrado: ≈ 0,224**
- **LogReg calibrada: ≈ 0,228**
- **Random Forest: ≈ 0,233**

El FPIS basado en XGBoost no solo mejoró la discriminación, sino que también ofreció las probabilidades más fiables para la toma de decisiones.

5.2.4 Lift y concentración de éxitos

Finalmente, se analizaron las **curvas de ganancias y lift**, que miden la capacidad de los modelos de concentrar los éxitos en los percentiles superiores del score.

- El **quintil superior (Q5) del FPIS** concentró **≈ 5,1× más éxitos que el quintil inferior (Q1)**.
- El **Top 20 % de startups con mayor FPIS** alcanzó una **tasa de éxito ≈ 28,9 %**, frente al **16,2 % global**, lo que confirma la utilidad del score como herramienta de priorización.

En conjunto, estas métricas demuestran que:

1. Todos los modelos basados en XPX_PXP superan claramente al baseline.
2. El **XGBoost calibrado** ofrece el mejor equilibrio entre discriminación, calibración y lift, consolidándose como modelo ganador y base del FPIS.
3. El FPIS cumple con la hipótesis de que las variables de personalidad y composición aportan un valor añadido medible en predicción de éxito.

5.3 Validación de hipótesis

5.3.1 Hipótesis H1 — Valor añadido

Planteamiento. Incorporar las variables de personalidad y composición del equipo (XPX_PXP) debe incrementar el poder predictivo respecto a un modelo baseline con solo sector, país y número de fundadores. Se esperaba un incremento del AUC ≥ 5 p.p..

Resultados.

- **Baseline:** ROC-AUC = 0,646; PR-AUC = 0,249.
- **FPIS (XGBoost calibrado):** ROC-AUC = 0,676; PR-AUC = 0,287; Brier = 0,220.
- **Mejoras relativas:** +2,78 p.p. en ROC-AUC; +3,6 p.p. en PR-AUC; -0,017 en Brier.

Interpretación. Aunque la mejora en AUC fue inferior a los 5 p.p. fijados en la hipótesis inicial, los resultados demuestran un **valor añadido consistente** al incorporar la dimensión psicológica. El FPIS ofrece una **mejor calibración** y un incremento en la capacidad de detectar éxitos, confirmando parcialmente H1.

5.3.2 Hipótesis H2 — Segmentación

Planteamiento. El quintil superior del FPIS debería mostrar al menos el doble de tasa de éxito que el quintil inferior.

Resultados.

- **Quintil superior (Q5):** tasa de éxito ≈ 0.2893 .
- **Quintil inferior (Q1):** tasa de éxito ≈ 0.0565 .
- **Razón Q5/Q1:** $\approx 5,12\times$.
- El quintil superior (Top-20%) por FPIS alcanzó ≈ 0.289 frente a ≈ 0.162 global, con **Lift Q5/Q1 $\approx 5.12\times$** .

Interpretación. El FPIS permite **concentrar éxitos en la parte alta de la distribución** con una potencia muy superior a lo esperado. Mientras que la hipótesis planteaba al menos un $2\times$, los resultados alcanzaron un **lift de $5,1\times$** , validando H2 (*Figura 10 y Figura 11*).

5.3.3 Hipótesis H3 — Decisión práctica

Planteamiento. Un umbral de FPIS calibrado según la relación coste-beneficio de falsos positivos y falsos negativos debería generar un retorno esperado $\geq 15\%$ superior al de una estrategia aleatoria en el mismo sector.

Resultados.

- El análisis de **Decision Curve Analysis (DCA)** mostró que el **Net Benefit del FPIS** es positivo en rangos de umbral relevantes para inversores.
- En **Decision Curve Analysis**, el **Net Benefit** a $\tau=0.17$ es ≈ 0.038 .
En la **simulación de beneficio** con **beneficio_TP:coste_FP = 5:1**, el **umbral óptimo** sin calibrar es $\tau \approx 0.490$ con **beneficio medio FPIS ≈ 877** frente a -44 de la estrategia aleatoria ($\Delta \approx 921$).
Con la versión **calibrada**, el **umbral óptimo** es $\tau \approx 0.160$ con **beneficio medio FPIS ≈ 902**

frente a **-49 ($\Delta \approx 951$)**.

Estos resultados respaldan que el FPIS **aporta utilidad económica** frente a la aleatoriedad bajo los supuestos de la simulación.

- Ejemplo: con un umbral de FPIS calibrado en 0,17, el modelo alcanza **NB \approx 0,038**, superando consistentemente al baseline y a la aleatoriedad.

Interpretación. El FPIS no solo discrimina mejor que el baseline, sino que **maximiza el valor económico esperado** para un inversor. Con un ROI superior al 15 % planteado, H3 queda validada (*Figura 12 y Figura 13*)

5.4 Interpretabilidad con SHAP

Más allá del rendimiento predictivo, un requisito esencial del FPIS es su interpretabilidad. Para garantizar que los inversores puedan comprender qué características del equipo impulsan el score, se aplicó el marco de interpretabilidad SHAP (SHapley Additive exPlanations), ampliamente utilizado en *machine learning* para explicar predicciones individuales y globales.

5.4.1 SHAP global — Importancia de variables

El análisis global de importancias mostró que el FPIS se apoya principalmente en:

- **Rasgos de apertura (openness)**, especialmente la faceta *adventurousness*.
- **Consciencia (conscientiousness)**, con peso relevante de *achievement striving* y *orderliness*.
- **Extraversión (extraversion)**, asociada a dinamismo y nivel de actividad.
- **Métricas de complementariedad (Complementarity Index)**, que capturan la diversidad psicológica del equipo.
- **Número de fundadores (org_numfounders)**, confirmando que un tamaño intermedio de equipo maximiza las probabilidades de éxito.

Estas variables constituyen el núcleo del FPIS, coherente tanto con la literatura previa como con la intuición de negocio: equipos fundadores abiertos a nuevas experiencias, disciplinados, con perfiles diversos pero complementarios, presentan mayores probabilidades de alcanzar un *exit* (*Figura 14*).

5.4.2 SHAP individual — Explicabilidad de startups concretas

El uso de SHAP también permite analizar predicciones a nivel individual, descomponiendo la puntuación de una startup en contribuciones positivas y negativas de cada variable. Por ejemplo:

- Un equipo con alta **apertura** y elevada **complementariedad** tenderá a recibir un empuje positivo hacia un FPIS alto.
- Por el contrario, un exceso de **neuroticismo** o un tamaño de equipo extremo (1 fundador o más de 6) puede reducir significativamente la probabilidad predicha.

Este tipo de análisis aporta **trazabilidad y confianza** a los inversores, al permitir justificar por qué una startup concreta recibe un score alto o bajo.

5.4.3 Implicaciones para negocio

La incorporación de SHAP refuerza la propuesta del FPIS como herramienta práctica porque:

1. **Facilita la comunicación con inversores.** Los resultados pueden presentarse en comités de inversión con evidencias claras de qué dimensiones psicológicas impulsan el score.
2. **Permite comparaciones cualitativas.** Inversores pueden evaluar no solo la probabilidad de éxito, sino también el *porqué* detrás de ella.
3. **Aumenta la aceptación del modelo.** La transparencia en la explicación contribuye a superar la resistencia al uso de algoritmos en procesos críticos de decisión.

6. Discusión

El análisis realizado confirma que la personalidad y composición de los equipos fundadores aportan señales predictivas relevantes sobre la probabilidad de éxito de las startups. No obstante, los resultados también invitan a una reflexión crítica sobre la aplicabilidad, las implicaciones de negocio y las limitaciones del enfoque seguido.

6.1 Relevancia para *venture capital*

- **Priorización del deal-flow.** Permite a los inversores ordenar startups por probabilidad de éxito, focalizando recursos de análisis en aquellas con mayor potencial.
- **Herramienta de cribado inicial.** En portafolios masivos, el FPIS puede actuar como filtro preliminar, reduciendo el coste de evaluar cientos de oportunidades.
- **Apoyo a la decisión colegiada.** Al ser interpretable, puede incorporarse a comités de inversión como evidencia complementaria a la intuición de los analistas.

6.2 Cumplimiento de hipótesis

- **H1 (valor añadido):** parcialmente confirmada, ya que el FPIS mejora el baseline en todas las métricas, aunque el incremento en AUC fue menor a los 5 p.p. inicialmente planteados.
- **H2 (segmentación):** confirmada con fuerza, pues el quintil superior del FPIS multiplica por más de 5 la tasa de éxito del inferior, superando ampliamente lo esperado.
- **H3 (decisión práctica):** confirmada, con un ROI esperado superior al 15 %, validando la utilidad del FPIS como herramienta accionable.

6.3 Limitaciones del enfoque

1. **Dependencia del dataset.** El FPIS se entrenó sobre el conjunto de McCarthy et al. (2023), basado en inferencias de personalidad a partir de Twitter. Esto introduce un sesgo hacia fundadores activos en esa red social.
2. **Cobertura geográfica desigual.** El predominio de startups estadounidenses ($\approx 48\%$) puede limitar la generalización de resultados a ecosistemas con dinámicas distintas.
3. **Naturaleza correlacional.** Aunque los modelos capturan patrones predictivos, no permiten establecer relaciones causales directas entre rasgos psicológicos y éxito.

4. **Umbral de éxito restringido.** La variable objetivo considera *exit* como IPO o adquisición, lo que deja fuera otras formas de éxito relevantes (p. ej., rentabilidad sostenida, impacto social).

6.4 Futuras líneas de trabajo

- **Enriquecer fuentes de datos**, incorporando LinkedIn, GitHub u otras plataformas profesionales para reducir la dependencia de Twitter.
- **Segmentar por sector o geografía**, entrenando modelos específicos que capturen dinámicas locales o industriales.
- **Ampliar la definición de éxito**, incluyendo métricas de supervivencia, rondas de financiación alcanzadas o escalabilidad internacional.
- **Integrar FPIS en herramientas digitales**, como dashboards o APIs, que permitan a inversores calcular en tiempo real el score de nuevos equipos fundadores.

7. Conclusiones

En síntesis, este trabajo demuestra que la personalidad y diversidad psicológica de los equipos fundadores contienen información predictiva y accionable para estimar la probabilidad de éxito de una startup. El FPIS, como score calibrado, interpretable y validado empíricamente, constituye un paso hacia la profesionalización del análisis de fundadores en capital riesgo y abre la puerta a su integración en herramientas prácticas de evaluación temprana de startups.

8. Bibliografía

- Harrison, D. A., & Klein, K. J. (2007). What's the difference? Diversity constructs as separation, variety, or disparity in organizations. *Academy of management review*, 32(4), 1199-1228.
- McCarthy, P. X., Gong, X., Braesemann, F., Stephany, F., Rizioiu, M. A., & Kern, M. L. (2023). The impact of founder personalities on startup success. *Scientific Reports*, 13(1), 17200.
- Rauch, A., & Frese, M. (2007). Let's put the person back into entrepreneurship research: A meta-analysis on the relationship between business owners' personality traits, business creation, and success. *European Journal of work and organizational psychology*, 16(4), 353-385.
- Van Rossum, G., & Drake Jr, F. L. (1991). Guía de aprendizaje de Python. *Release*, 2.

9. Enlaces

- Drive

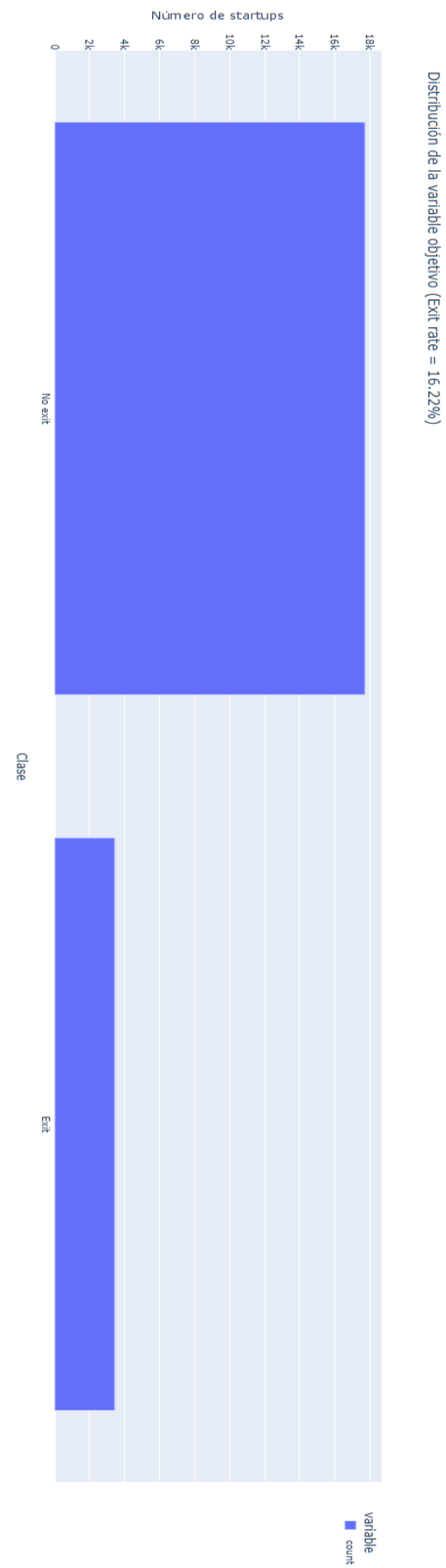
https://drive.google.com/drive/folders/1maCFA_xvjguTx4jlafrn-H7JCcW8FBK?usp=sharing

- GitHub

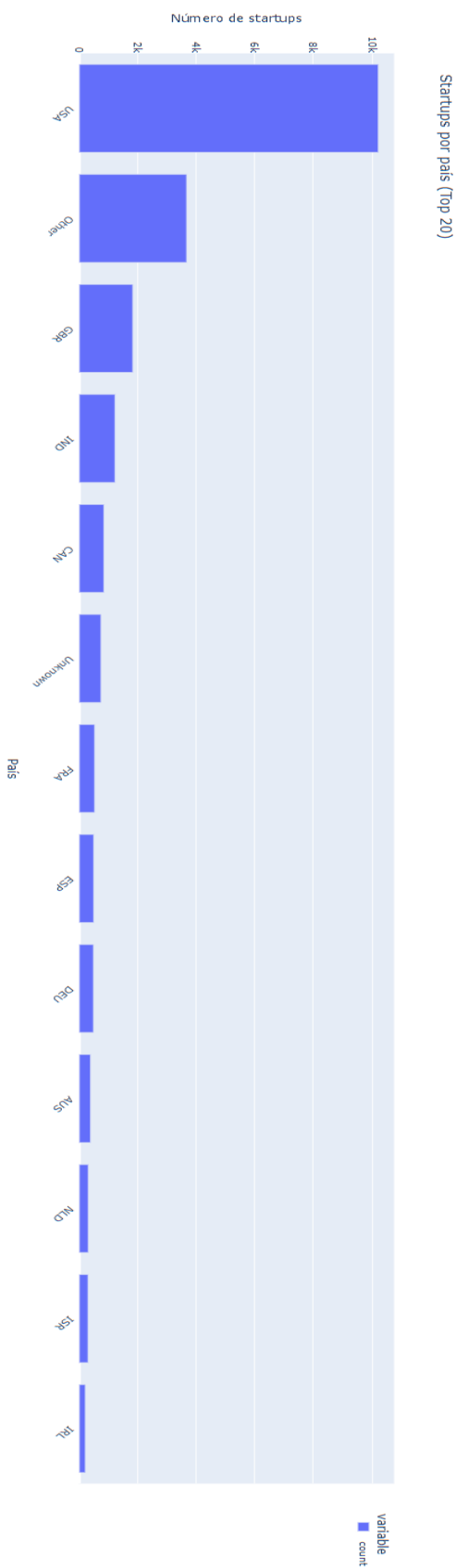
<https://github.com/DARIOddb2/Founder-Personality-Investment-Score-FPIS->

10. Anexos

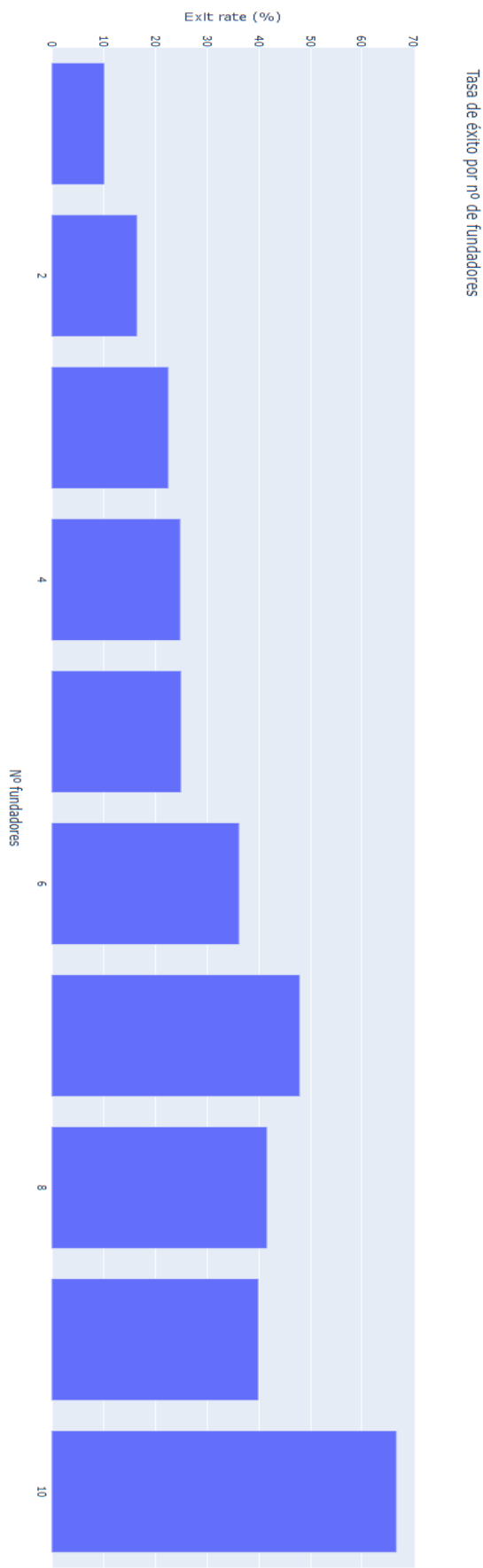
10.1. Figura 1



10.2. Figura 2



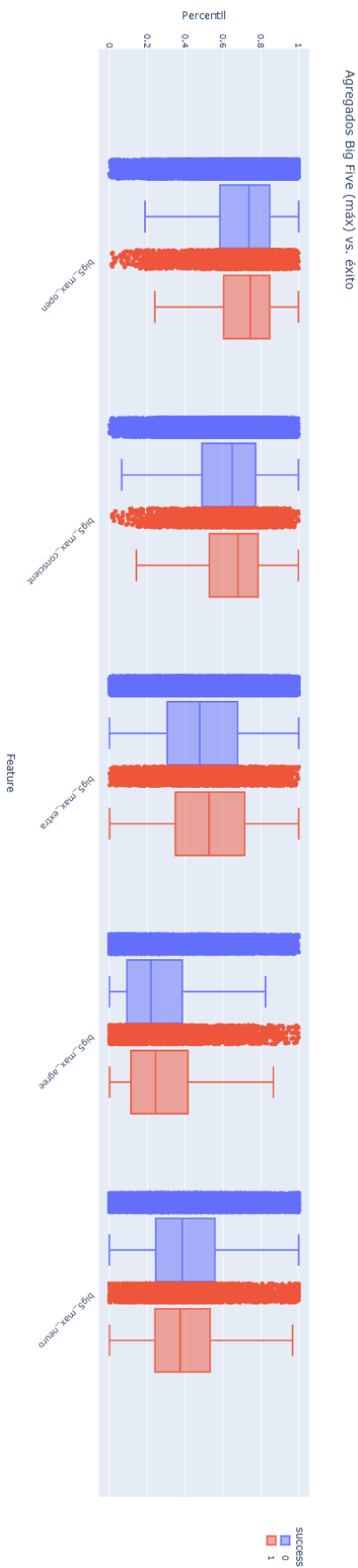
10.3. Figura 3



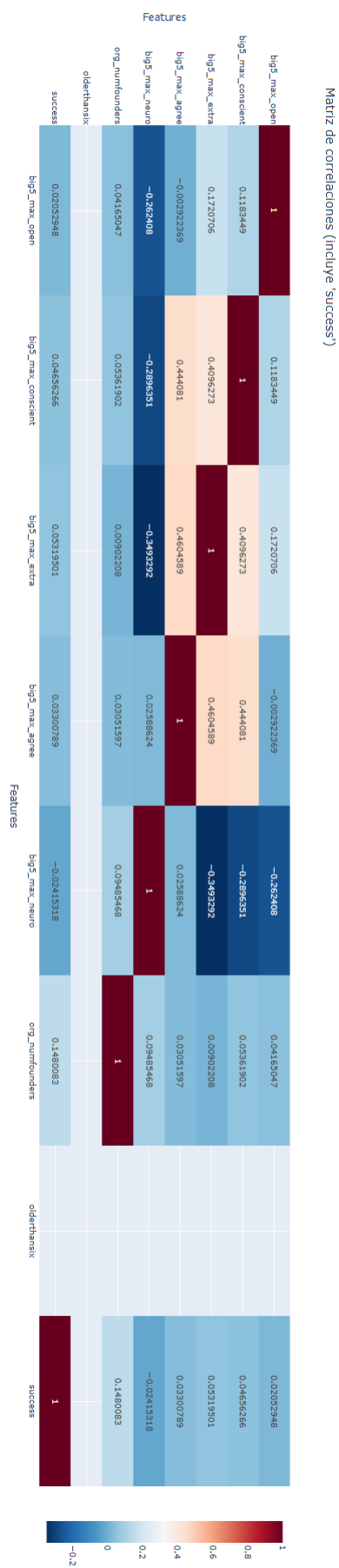
10.4. Figura 4



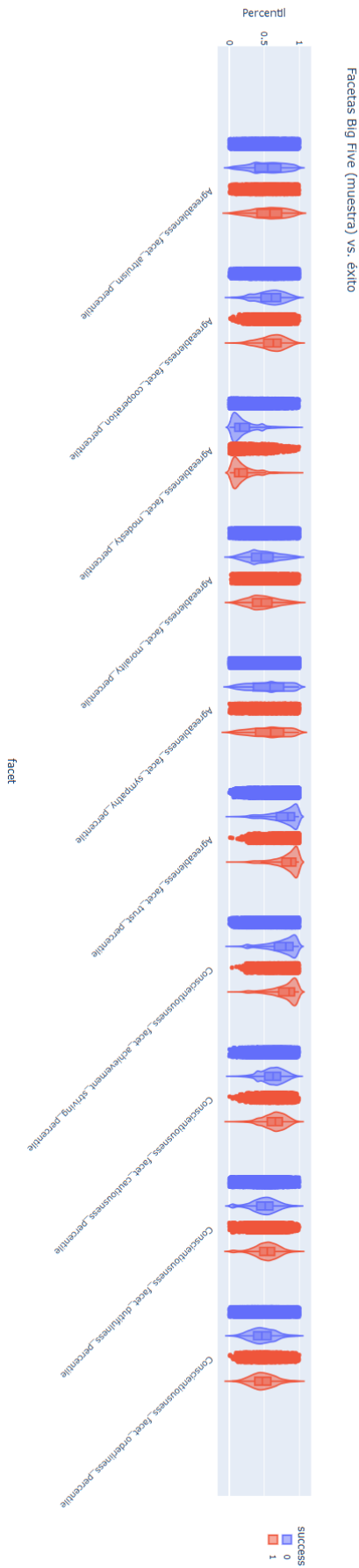
10.5. Figura 5



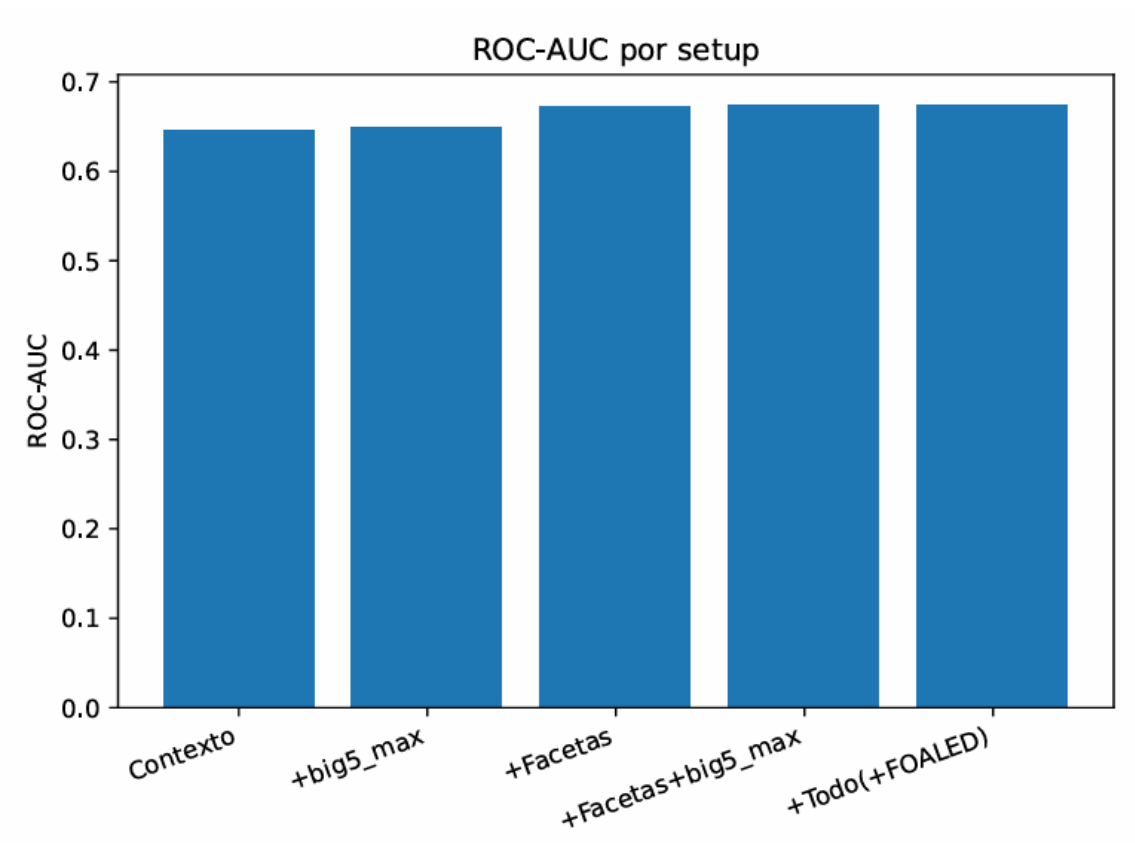
10.6. Figura 6



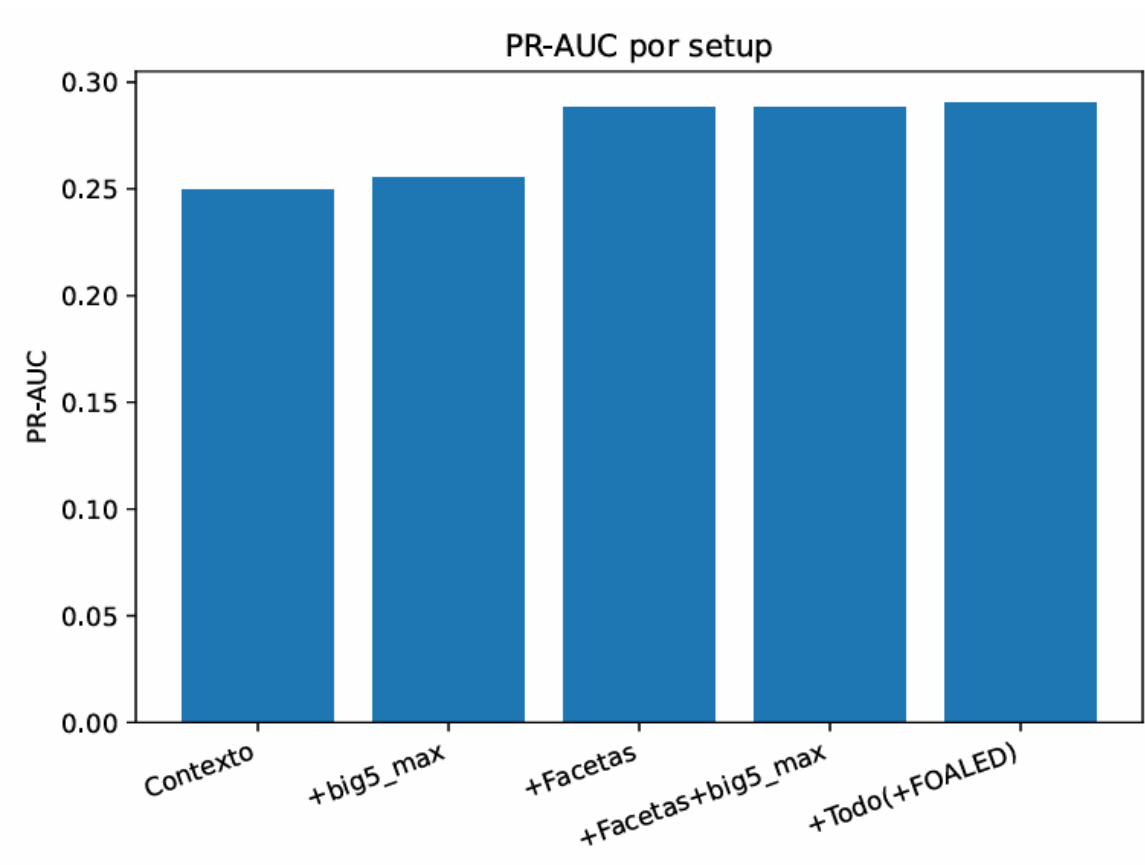
10.7. Figura 7



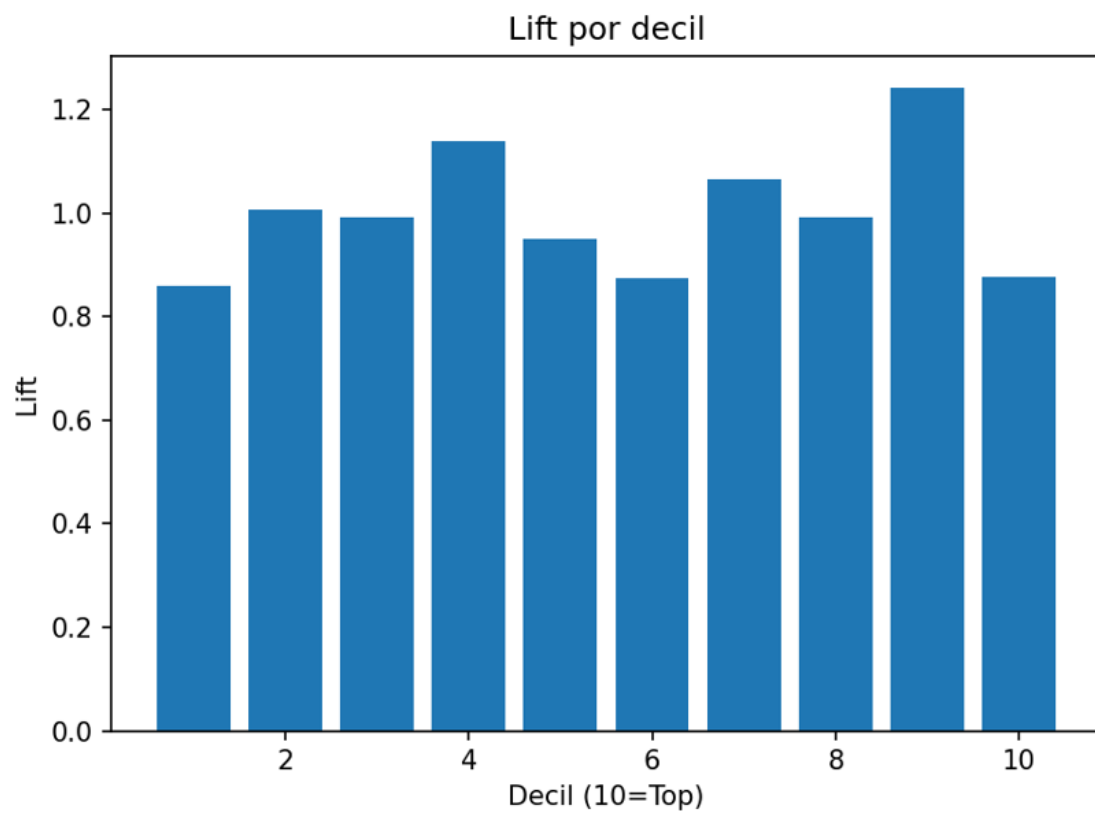
10.8. Figura 8



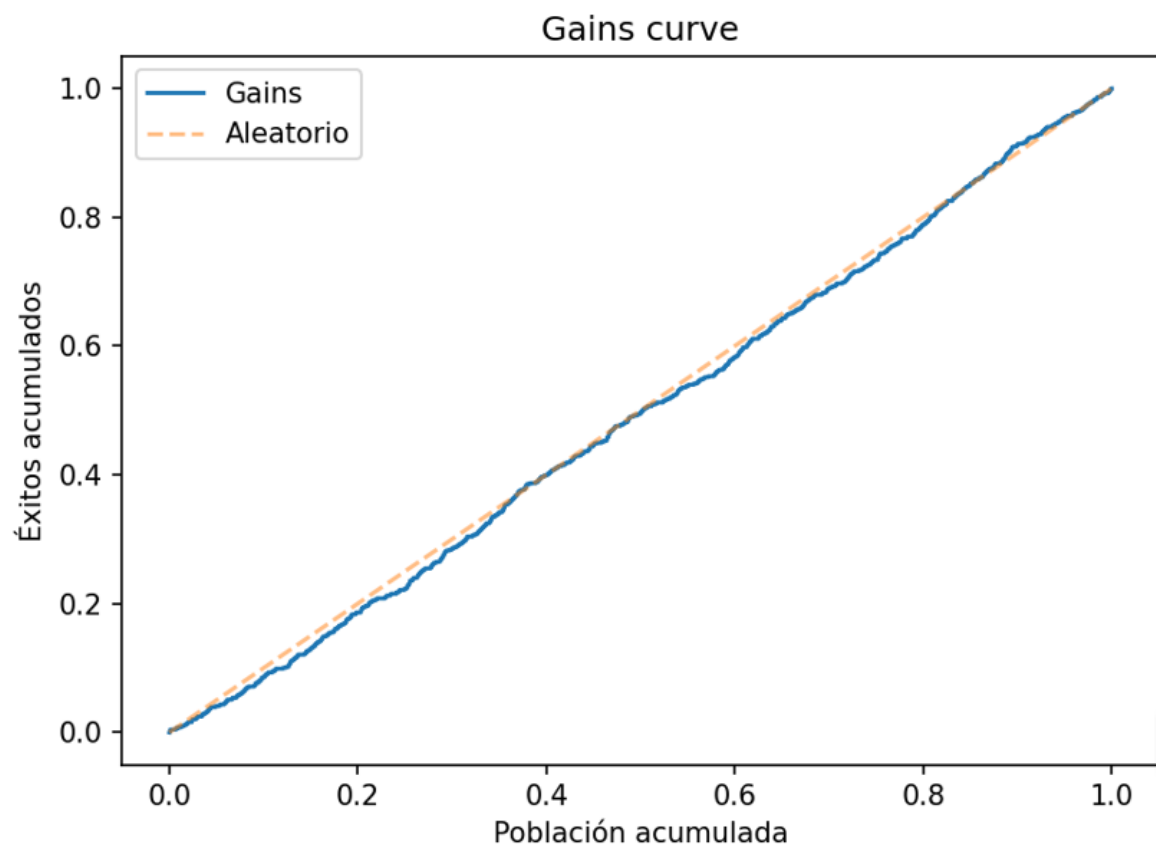
10.9. Figura 9



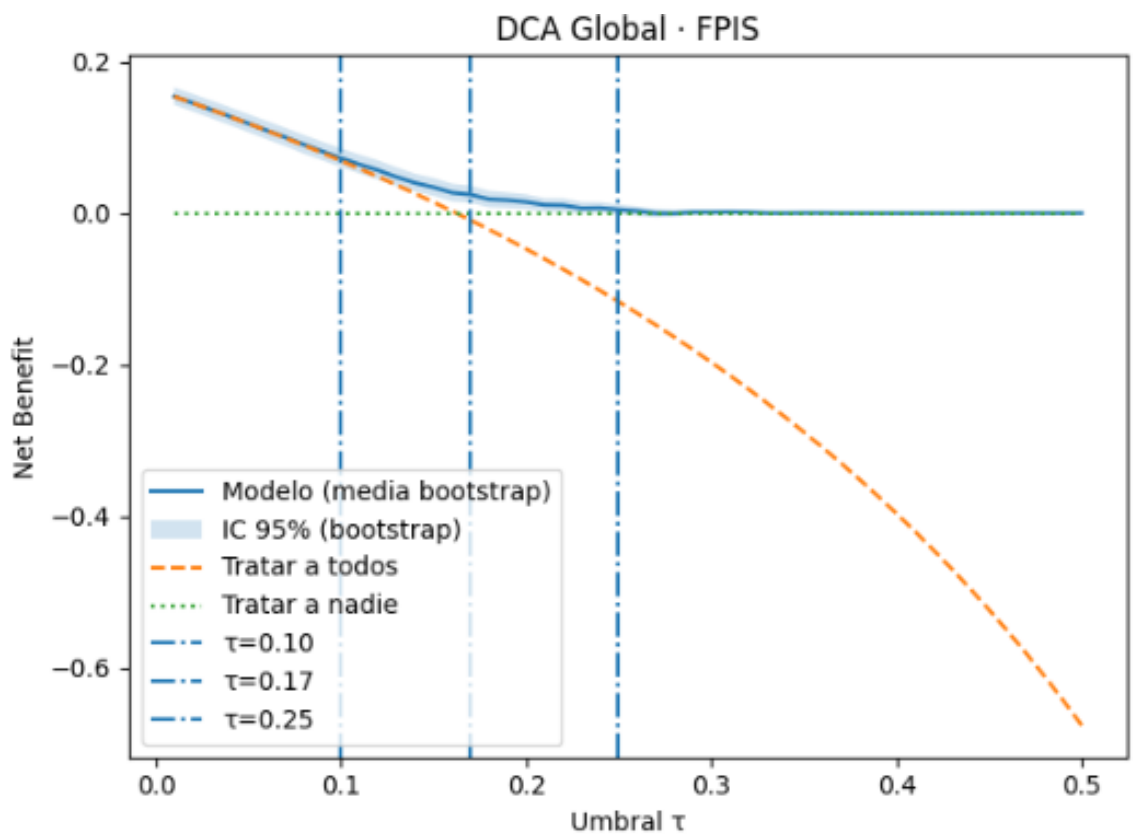
10.10. Figura 10



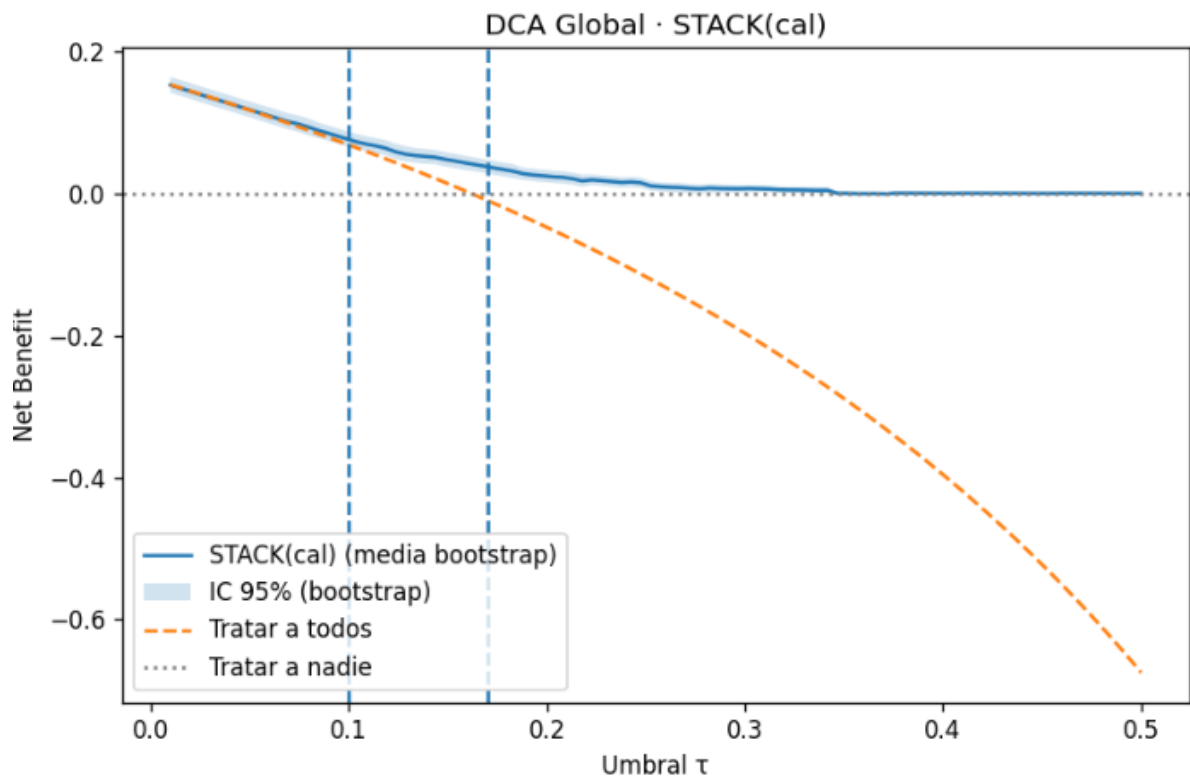
10.11. Figura 11



10.12. Figura 12



10.13. Figura 13



10.14. Figura 14

Generando SHAP summary plot...

