

PRIVATE EQUITY ADVISORY AI TOOL

Predicting the most relevant Private Equity investors for target companies

Authors:

- Aviv Rabinowitz
- Blanca Burgaleta García
- Maha Alkaabi
- Noureldin Sewilam
- Omer Althwaini
- Pablo Gallegos



INTRODUCTION

The current process of matching Private Equity (PE) firms to investment targets is fragmented and manual, relying on historical intuition and scattered data sources. Our objective was to design an AI-powered tool that automates PE-to-company match prediction using machine learning and enriched data, helping EY consultants quickly identify relevant investors for any target company.

METHOD

We developed an end-to-end ML pipeline to predict the most relevant PE investors for a given company:

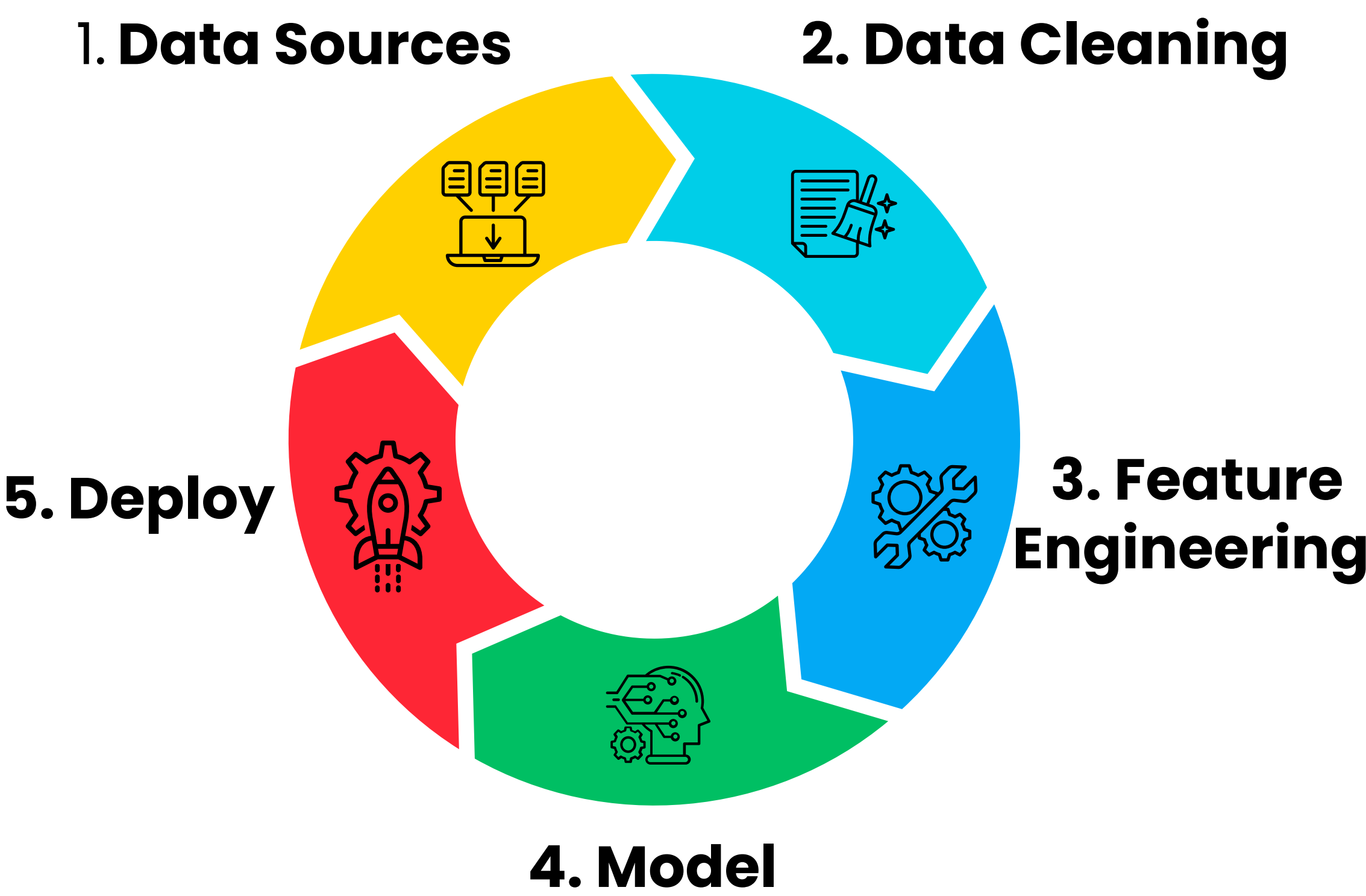


Figure: End-to-End Machine Learning Pipeline – from raw data ingestion to investor match prediction and application deployment

- 1

EY’s internal PE and company datasets enriched with websites, sectors, and financials
- 2

Standardized sector and geography fields, removed low-quality entries, and ensured data consistency for modeling.
- 3

Engineered **200+** features capturing sector alignment, geographic fit, historical investment patterns, and deal size compatibility.
- 4

Trained an **XGBoost** classifier to predict likely PE investors for a company. Prioritized interpretability, performance, and scalability.
- 5

Deployed the model in a **Streamlit** app featuring interactive dashboards for investor targeting, portfolio exploration, and Spanish market insights.

DATA ANALYSIS (EDA)

We performed exploratory data analysis (EDA) to uncover investment patterns and dynamics that shaped our model design and PE firm targeting logic.

1. PE Landscape

- **169 PE firms** analyzed with HQs mapped globally
- Highest Density in US, UK, France, & Spain
- Sector breakdown highlighted dominant investment verticals

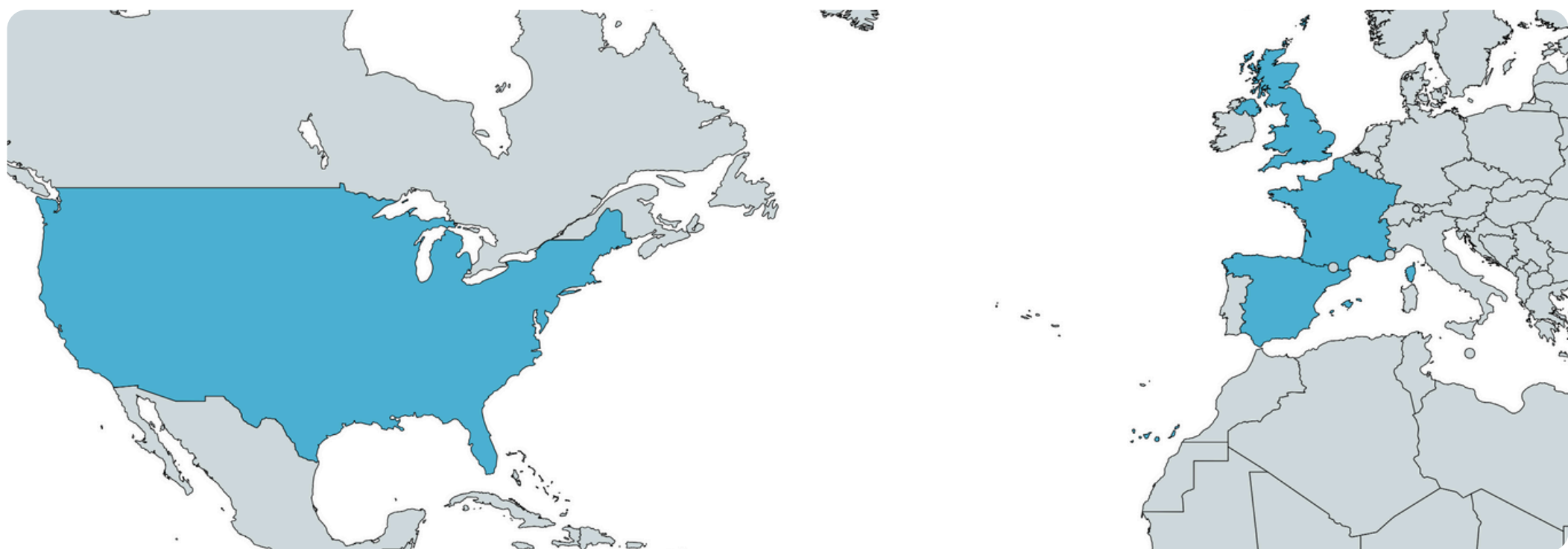


Figure: Geographic distribution of 169 PE firms in the dataset. Majority are headquartered in the US, UK, France, and Spain, reflecting key hubs of investment activity.

2. Portfolio Behavior

- Mapped co-investment patterns and sector overlap.
- Tracked investment volume trends over time.
- Identified firms’ historical preferences and diversification levels

3. Spanish Market exploration

- Explored CPR & Education sectors (**1,800+ companies**)
- Clustered by **EBITDA**, **CAGR**, and **ownership type**
- Highlighted high-growth, high-margin segments for deal sourcing

MODEL EVALUATION

We evaluated our model using classification metrics and business-driven interpretation techniques to ensure performance, transparency, and usability for PE advisors

1. Problem Framing

Will this PE firm invest in the selected company?

2. Final Model

- Best performance on tabular data.
- Handles class imbalance
- Provides explainable feature importances.



3. Evaluation Metrics

Accuracy	Precision	Recall	F1 Score
84%	81%	89%	85%

Table: Performance metrics of the final XGBoost model on the test set, demonstrating strong predictive accuracy and balance across key classification measures.

STREAMLIT DEPLOYMENT

We deployed our model into a Streamlit web app to enable real-time PE-investor matching and market exploration through interactive dashboards.

Key Features

1. Investor Match Tool

Input a company to get ranked PE investors based on alignment scores.

2. PE Explorer

View global HQs, sector focus, and portfolio patterns.

3. Spanish Market Deep Dive

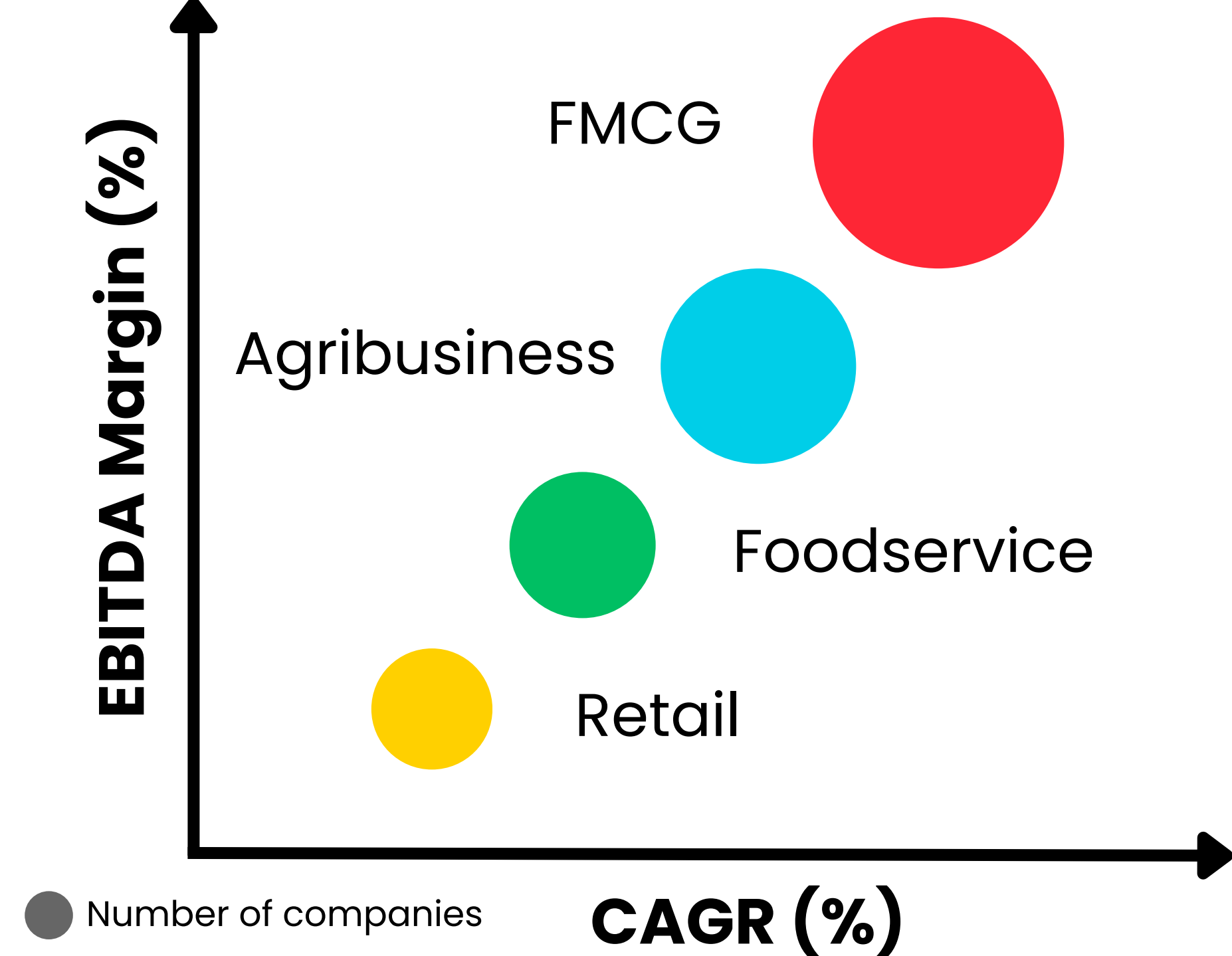
Visualize CPR/Education sectors by CAGR, EBITDA margin, and ownership type.

4. Company Profile & Thesis

Generate LLM-powered company profiles and structured diligence questions in one-click

5. Admin Panel

Upload new datasets and trigger the data pipeline.



Spanish CPR subsectors mapped by 3-Year CAGR (X) and EBITDA Margin (Y). FMCG emerges as a standout with the strongest growth and profitability profile, signaling prime PE investment potential.

CONCLUSION

We delivered a scalable AI tool that predicts the most relevant PE investors for any company. By combining machine learning with rich data and interactive dashboards, our platform streamlines investor targeting and empowers EY consultants with faster, smarter, insight-driven decisions.

FUTURE WORK

1. Data Governance

Centralize and automate data cleaning using a structured SQL pipeline to boost reliability and scale.

2. Cloud Deployment

Host on Azure for secure, real-time access.

