Caprae Capital – AI Lead Recommendation Tool

Approach

This tool was built to help founder-led teams quickly discover companies they could realistically engage with for partnerships, sales, or acquisitions. The system works with just four inputs: country, domain, revenue, and growth rate. Based on these, it returns a short list of companies with similar operational and financial profiles. The goal was to keep the interface simple and the output commercially useful.

Model Selection

I used **KMeans clustering** because it works well when you don't have labels and need to group similar profiles. It's interpretable, efficient, and a great fit for lead segmentation. I selected 6 clusters to reflect different business sizes and growth profiles.

Data Preprocessing

The dataset included 1,000 companies (synthetic) with fields like Name, Country, Domain, Revenue_Cr, Growth_Rate, and Listed. I used a ColumnTransformer pipeline to:

- Scale numeric fields (Revenue, Growth Rate)

- One-hot encode categorical fields (Country, Domain)

Performance Evaluation

Since clustering is unsupervised, I validated results based on commercial logic. After clustering, I applied specific filters:

- Same **country**

- Same **domain**

- Revenue within **± ₹50 Cr** of input

- **Positive growth rate only**

This ensures the results are not just mathematically similar, but practically useful.

Rationale

This solution blends simple ML with real business filters. It's fast, easy to adapt, and relevant to how founders think. It reflects Caprae's mission to apply smart, real-world AI to drive better decisions across portfolio companies.