

Báo Cáo Phân Tích Cơ Hội Đầu Tư (Investment Opportunity Analysis EDA)

1. Giới Thiệu Đề Tài

Trong bối cảnh thị trường đầu tư toàn cầu ngày càng cạnh tranh, việc xác định các cơ hội đầu tư tiềm năng là yếu tố then chốt giúp các nhà đầu tư và doanh nghiệp đưa ra quyết định chính xác. Bài báo cáo này tập trung vào việc phân tích dữ liệu đầu tư để xác định các lĩnh vực, quốc gia và loại hình đầu tư phù hợp nhất, dựa trên chiến lược "đầu tư vào nơi những người khác đang đầu tư". Đặc biệt, đề tài yêu cầu tìm ra các phân tích trong đó khoảng vốn có thể huy động của doanh nghiệp nằm trong phạm vi **5-15 triệu USD** để đưa ra quyết định đầu tư tốt nhất. Bằng cách sử dụng các kỹ thuật Phân Tích Dữ Liệu Khám Phá (EDA), chúng tôi sẽ đưa ra các kết luận và khuyến nghị chi tiết.

2. Giới Thiệu Tập Dữ Liệu

Tập dữ liệu được sử dụng trong bài báo cáo này bao gồm thông tin về các khoản đầu tư vào các công ty khởi nghiệp và doanh nghiệp trên toàn cầu. Các cột chính trong tập dữ liệu bao gồm:

- **permalink** và **company_permalink**: Định danh công ty.
- **funding_round_type**: Loại hình đầu tư (ví dụ: seed, venture, private equity).
- **raised_amount_usd**: Số tiền đầu tư (USD).
- **country_code**: Mã quốc gia.
- **category_list**: Danh mục ngành nghề của công ty.

Các cột khác như **homepage_url**, **founded_at**, **state_code**, **region**, **city** đã được loại bỏ do không phù hợp với mục tiêu phân tích.

3. Thu Thập và Tiền Xử Lý Dữ Liệu

3.1. Hợp Nhất và Làm Sạch Dữ Liệu

Hợp nhất dữ liệu: Do hai cột **permalink** và **company_permalink** giống nhau, một trong hai cột đã được xóa để tránh trùng lặp.

Loại bỏ cột không cần thiết: Các cột như **homepage_url**, **founded_at**, **state_code**, **region**, **city** và **funding_round_code** (với 73% giá trị bị thiếu) đã được loại bỏ.

Xử lý giá trị thiếu:

- Cột **raised_amount_usd** (số tiền đầu tư) là cột quan trọng, do đó các hàng có giá trị thiếu trong cột này đã được loại bỏ.
- Cột **country_code** (mã quốc gia) có khoảng 6% giá trị thiếu. Các hàng này cũng đã được loại bỏ để đảm bảo tính chính xác của phân tích.
- Cột **category_list** có tỷ lệ giá trị thiếu thấp (0,65%), các hàng này cũng đã được loại bỏ.

3.2. Xử Lý Outliers và Phân Tích Sơ Bộ

Số tiền đầu tư: Giá trị trung bình của `raised_amount_usd` là khoảng 10 triệu USD, trong khi giá trị trung vị chỉ 1,68 triệu USD. Sự chênh lệch lớn này cho thấy sự phân bố không đồng đều của dữ liệu, với nhiều giá trị ngoại lệ (outliers).

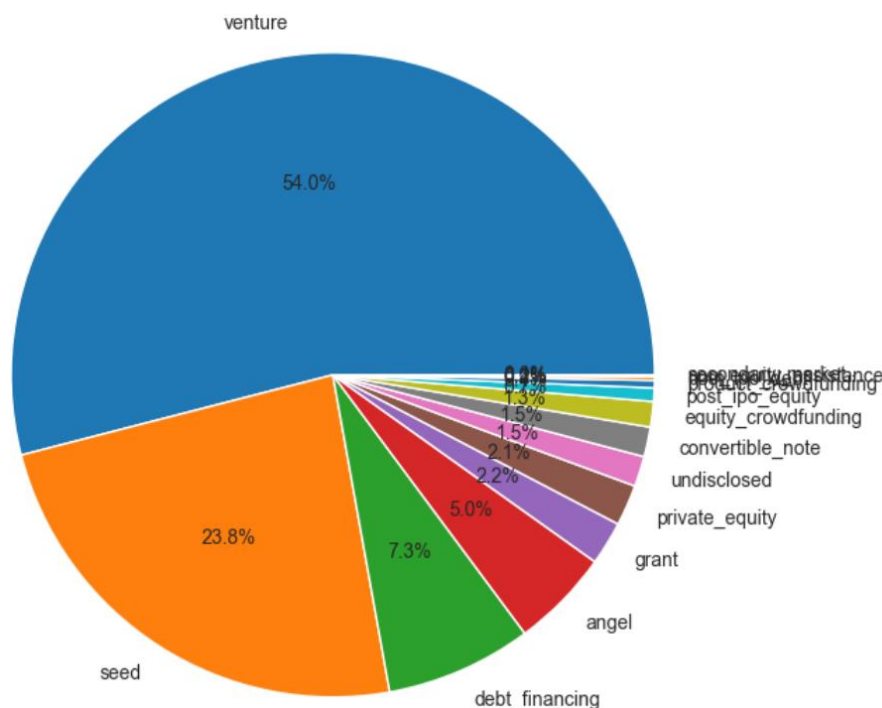
Xử lý outliers: Để đảm bảo tính chính xác của phân tích, các giá trị ngoại lệ trong cột `raised_amount_usd` đã được xử lý bằng cách loại bỏ các hàng có giá trị quá cao hoặc quá thấp so với phạm vi đầu tư mục tiêu (5-15 triệu USD).

4. Phân Tích Dữ Liệu

4.1. Phân Tích Loại Hình Đầu Tư

Dữ liệu bao gồm 14 loại hình đầu tư đa dạng. Thể hiện sự phong phú của cơ hội đầu tư trong dữ liệu, từ các startup mới đến các doanh nghiệp lớn và có thị phần ổn định. Điều này cung cấp một cơ sở vững chắc cho việc phân tích và lựa chọn loại hình đầu tư phù hợp với mục tiêu đầu tư và chiến lược của nhà đầu tư.

Biểu đồ phân phối số lượng giao dịch đầu tư của từng loại hình



Kết quả phân tích:

- Venture và Seed đang chiếm vị trí dẫn đầu về số lượng giao dịch (54% và 23.8%), vượt xa so với các loại hình đầu tư khác. Sự chênh lệch lớn trong số lượng giao dịch giữa Venture, Seed và các loại hình đầu tư khác cho thấy rằng Venture và Seed là những lựa chọn phổ biến và được ưa chuộng từ các nhà đầu tư trong dữ liệu này. Bảng số tiền đầu tư trung bình của các loại hình đầu tư

Bảng giá trị trung đầu tư giữa các loại hình đầu tư

```
[48]: funding_round_type
      secondary_market      45.850
      private_equity      20.000
      post_ipo_debt       19.900
      post_ipo_equity     12.265
      venture             5.000
      debt_financing       1.100
      undisclosed         1.100
      angel               0.415
      convertible_note     0.300
      seed                0.300
      grant               0.220
      product_crowdfunding 0.210
      equity_crowdfunding  0.080
      non_equity_assistance 0.060
      Name: raised_amount_usd, dtype: float64
```

- Loại hình đầu tư Private Equity có sự chênh lệch lớn hơn so với các loại hình đầu tư khác. Do thường thực hiện ở giai đoạn sau khi doanh nghiệp đã có sự phát triển và thị phần ổn định, ta có thể thấy được sự quan tâm của các nhà đầu tư là các doanh nghiệp đã ổn định và có tiềm năng tăng trưởng cao
- Venture tuy có số tiền đầu tư ít hơn so với Private Equity nhưng vẫn chiếm vị trí quan trọng trong bảng xếp hạng loại hình đầu tư. Điều này cho thấy các nhà đầu tư cũng có quan tâm đến các doanh nghiệp mới nổi và trên đà phát triển nhanh
- Angel và Seed tuy số tiền đầu tư ít hơn hẳn so với 2 loại hình trước nhưng mà vẫn cân nhắc đầu tư vào các doanh nghiệp ở giai đoạn đầu sự phát triển, tuy tiềm năng rủi ro khá cao.
- Các loại hình đầu tư khác không đáng kể so với Private Equity, Venture, Angel và Seed do có ít số lượng giao dịch.

Kết quả phân tích:

- **Private Equity** có số tiền đầu tư trung bình cao nhất (khoảng 20 triệu USD), vượt quá phạm vi chênh lệch giữa giá trị trung bình và trung vị (10,5 1,8 triệu USD). Tuy nhiên, giá trị này nằm ngoài phạm vi đầu tư mục tiêu (5-15 triệu USD).
- **Venture** có số tiền đầu tư trung bình là 5 triệu USD, nằm trong phạm vi đầu tư mục tiêu (5-15 triệu USD). Đây là loại hình đầu tư phù hợp với các doanh nghiệp mới nổi và đang phát triển nhanh.
- **Angel** và **Seed** có số tiền đầu tư trung bình thấp hơn (0,415 triệu USD và 0,3 triệu USD), nằm ngoài phạm vi đầu tư mục tiêu.

Kết luận: Loại hình **Venture** là phù hợp nhất để đầu tư, vì nó chiếm 54% tổng và nằm trong phạm vi 5-15 triệu USD và phù hợp với chiến lược đầu tư vào các doanh nghiệp đang phát triển .

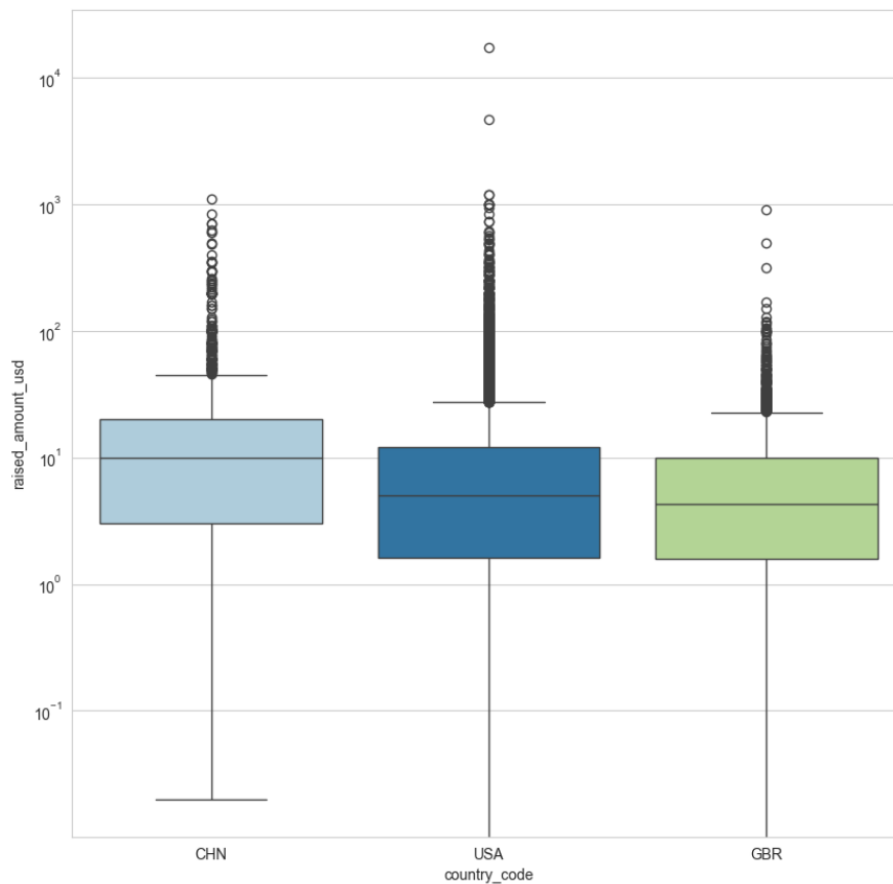
4.2. Phân Tích Quốc Gia

Bảng số liệu về số tiền đầu tư hay số tiền vốn huy động được dựa trên các quốc gia

```
[49]: country_code
      USA    420067.65
      CHN    39338.79
      GBR    20072.77
      IND   14261.56
      CAN    9482.20
      FRA    7226.81
      ISR    6854.34
      DEU    6307.00
      JPN    3167.67
      SWE    3145.88
      Name: raised_amount_usd, dtype: float64
```

Trong số 10 quốc gia dẫn đầu, Mỹ, Trung Quốc và Anh là ba quốc gia đứng đầu với số tiền đầu tư nhiều nhất.

Biểu đồ phân bố số tiền đầu tư từ các quốc gia



- Mỹ là quốc gia nhận được số tiền đầu tư cao nhất trong quá khứ, hộp màu xanh đậm lớn có nhiều điểm ngoại lai điều này cho thấy quốc gia này có sự biến động lớn về số tiền đầu tư.
- Trung Quốc có vùng màu xanh nhạt và ngắn hơn so với Mỹ, ít điểm ngoại lai hơn, cho thấy ở đây ít biến động hơn về số tiền đầu tư.
- Anh Quốc tuy đứng thứ 3 về số lượng tiền đầu tư vào, cũng như là phạm vi đầu tư nhỏ hơn nhưng điều này cho thấy ở đây có sự ổn định và nhất quán trong hoạt động đầu tư. Không như Mỹ và Trung Quốc vì 2 quốc gia này có sự biến động lớn trong số tiền đầu tư, cho thấy có sự không ổn định và rủi ro trong việc đầu tư tại đây.

Kết quả phân tích:

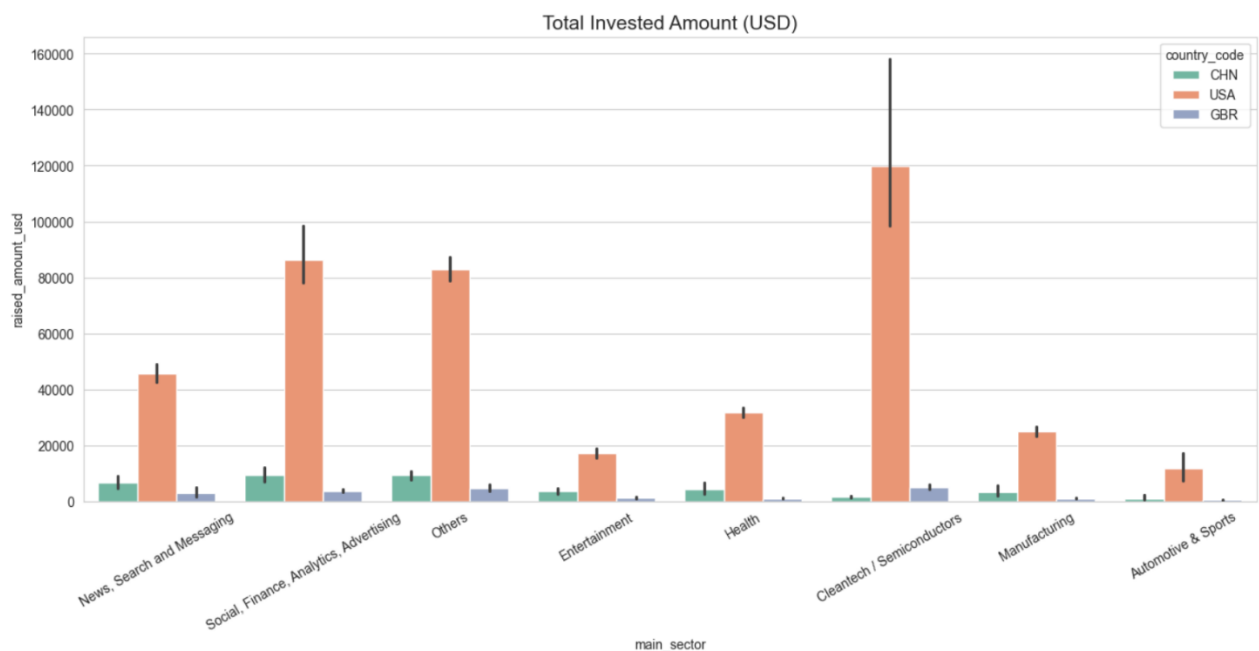
- **Mỹ (USA)** là quốc gia dẫn đầu về cả số lượng và tổng số tiền đầu tư, chiếm phần lớn các giao dịch đầu tư.
- Các quốc gia khác như Trung Quốc (CHN), Anh (GBR) và Ấn Độ (IND) cũng có số lượng đầu tư đáng kể nhưng thấp hơn nhiều so với Mỹ.
- Biểu đồ phân bố đầu tư theo quốc gia cho thấy Mỹ chiếm ưu thế rõ rệt, tiếp theo là các quốc gia có nền kinh tế phát triển và đang phát triển.

Kết luận: Mỹ là quốc gia phù hợp nhất để đầu tư dựa trên số lượng và tổng số tiền đầu tư.

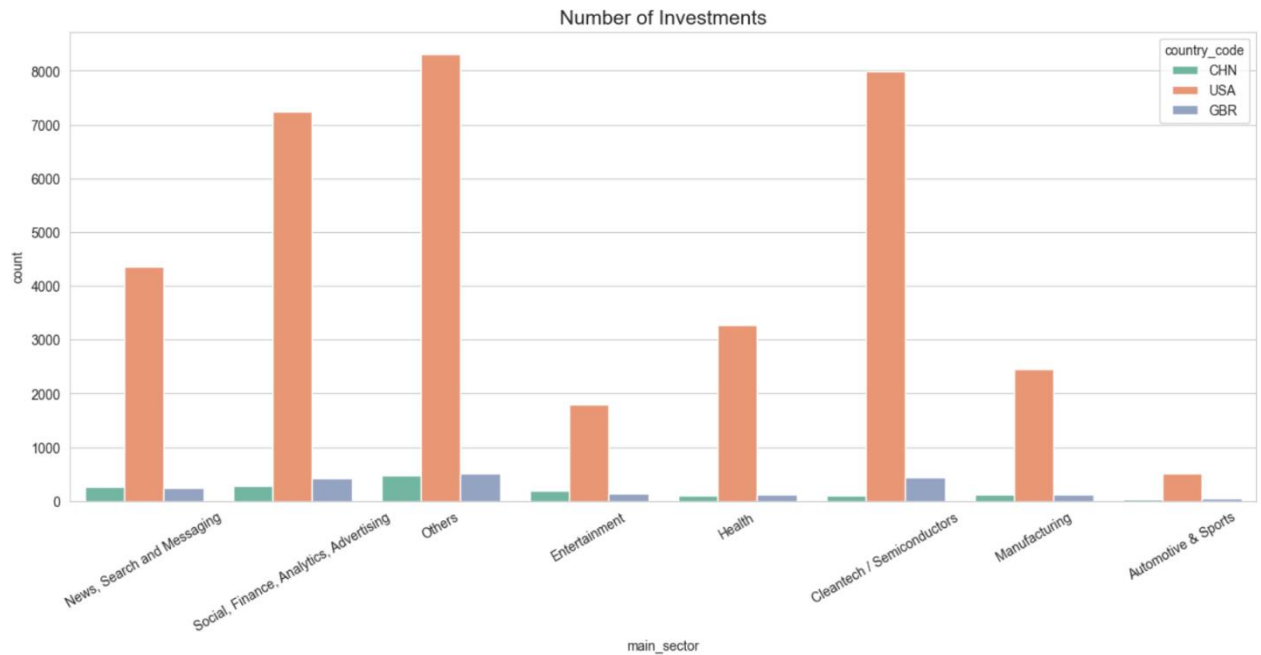
4.3. Phân Tích Lĩnh Vực Đầu Tư

Tập dữ liệu được chọn lọc theo ba quốc gia hàng đầu là Mỹ, Trung Quốc, Anh với loại hình đầu tư phù hợp nhất là Venture có số tiền đầu tư và số lượng đầu tư cao nhất.

Biểu đồ tổng số tiền đầu tư theo lĩnh vực của ba quốc gia



Biểu đồ tổng số lượng đầu tư theo lĩnh vực của ba quốc gia



Kết quả phân tích:

Các lĩnh vực được đầu tư nhiều nhất bao gồm:

- Others (các lĩnh vực khác).
- Cleantech/Semiconductors (công nghệ sạch và chất bán dẫn).
- Social, Finance, Analytics, Advertising (mạng xã hội, tài chính, phân tích và quảng cáo).

Nếu loại bỏ lĩnh vực **Others**, lĩnh vực **News, Search and Messaging** (tin tức, tìm kiếm và nhắn tin) là lựa chọn tốt tiếp theo.

Biểu đồ phân bố đầu tư theo lĩnh vực cho thấy sự tập trung cao vào các ngành công nghệ và tài chính.

Kết luận: Các lĩnh vực **Cleantech/Semiconductors** và **Social, Finance, Analytics, Advertising** là những lĩnh vực tiềm năng để đầu tư.

5. Kết Luận

Qua phân tích dữ liệu, chúng tôi đưa ra các kết luận sau:

- **Loại hình đầu tư phù hợp nhất: Venture** (với số tiền đầu tư trung bình 5 triệu USD) là tăng trưởng cao. Đây cũng là loại hình đầu tư nằm trong phạm vi vốn huy động mục tiêu (5-15 triệu USD).
- **Quốc gia đầu tư tiềm năng nhất: Mỹ (USA)** là thị trường hàng đầu với số lượng và tổng số tiền đầu tư lớn nhất, tạo điều kiện thuận lợi cho các nhà đầu tư.

- **Lĩnh vực đầu tư tiềm năng:** Các lĩnh vực **Cleantech/Semiconductors, Social, Finance, Analytics, Advertising**, và **News, Search and Messaging** là những lĩnh vực thu hút nhiều vốn đầu tư và có tiềm năng tăng trưởng cao.

6. Bài Học Rút Ra

Xử lý dữ liệu thiếu và outliers: Việc xử lý cẩn thận các giá trị thiếu và outliers (đặc biệt là trong các cột quan trọng như `raised_amount_usd` và `country_code`) là yếu tố then chốt để đảm bảo tính chính xác của phân tích.

Lựa chọn thống kê phù hợp: Trong trường hợp dữ liệu có nhiều giá trị ngoại lệ, giá trị trung vị thường phản ánh chính xác hơn so với giá trị trung bình.

Chiến lược đầu tư theo xu hướng: Đầu tư vào các lĩnh vực và quốc gia mà các nhà đầu tư khác đang tập trung là một chiến lược hiệu quả để giảm thiểu rủi ro và tối đa hóa lợi nhuận.