|  |
| --- |
| detail of persons hands with scissors, markers, workingBANK MARKETING ANALYSIS REPORT |

|  |  |  |
| --- | --- | --- |
| GROUP 1 – DATA BROS |  |  |
| Lê Hương Giang  Phan An Giang  Vũ Việt Tùng |  |  |

TABLE OF CONTENTS

[1. Tổng quan Dữ liệu 3](#_Toc180358695)

[2. Khám phá dữ liệu 4](#_Toc180358696)

[a. Phân tích đơn biến 4](#_Toc180358697)

[b. Phân tích đa biến 14](#_Toc180358698)

[c. Ma trận tương quan và biểu đồ heatmap 23](#_Toc180358699)

[d. Các biến quan trọng nhất ảnh hưởng đến quyết định đăng ký tiền gửi có kỳ hạn. 24](#_Toc180358700)

[e. Kết luận chung 27](#_Toc180358701)

[3. Đề xuất chiến lược 27](#_Toc180358702)

[4. Đánh giá tác động: 29](#_Toc180358703)

[KẾT LUẬN 30](#_Toc180358704)

## Báo cáo này phân tích kết quả chiến dịch tiếp thị của ngân hàng thông qua các cuộc gọi điện thoại nhằm dự đoán hành vi của khách hàng trong việc đăng ký gửi tiền có kỳ hạn. Dữ liệu từ hơn 11.000 khách hàng đã được thu thập và phân tích để hiểu rõ những yếu tố ảnh hưởng đến quyết định tài chính của họ, bao gồm các đặc điểm nhân khẩu học, lịch sử tài chính và tần suất tương tác trong chiến dịch. Báo cáo cũng tập trung vào việc khám phá các chiến lược tiếp cận khách hàng hiệu quả, xác định phân khúc khách hàng tiềm năng và đưa ra các đề xuất tối ưu hóa chiến lược tiếp thị để nâng cao tỷ lệ thành công.

## 1. Tổng quan Dữ liệu

Dataset được cung cấp cung cấp thông tin toàn diện về tương tác của khách hàng trong chiến dịch tiếp thị dựa trên điện thoại, bao gồm biến định lượng và biến phân loại.

Dữ liệu phân tích gồm **11,162 hàng** và **16 cột**, bao gồm 15 biến đầu vào, trong đó có 7 biến định lượng (day, age, duration, balance, campaign, pdays, previous) and 8 biến phân loại (education, job, marital, default, contact, housing, loan, poutcome). Dữ liệu được thu thập từ các chiến dịch tiếp thị qua điện thoại của ngân hàng, với mục tiêu dự đoán xem khách hàng có đăng ký gửi tiền có kỳ hạn hay không. Biến mục tiêu là biến phân loại nhị phân, cho biết liệu khách hàng có gửi tiền không (yes/no).

**Một số thống kê chính về tuổi khách hàng, số dư và thời lượng cuộc gọi là nền tảng cho việc phân tích sâu, thực hiện mô hình và đánh giá:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Biến** | **Trung bình (Mean)** | **Min** | **Max** |
| **Tuổi** | 41.23 | 18 | 95 |
| **Số dư** | 1529 | -6847 | 81204 |
| **Thời gian cuộc gọi** | 372 | 2 | 3881 |

## 2. Khám phá dữ liệu

### a. Phân tích đơn biến

#### Phân tích biến phân loại

A graph of different colored bars

Description automatically generated

##### Biến Nghề nghiệp (Job):

Phân tích tần suất cho biến **job** cho thấy sự phân bổ khách hàng theo nghề nghiệp. Cụ thể:

* **"Blue-collar"** và **"Management"** là hai nhóm nghề nghiệp có số lượng khách hàng cao nhất, chiếm tỷ trọng lớn trong tổng số khách hàng.
* Một số nhóm nghề nghiệp khác như **"technician"** và **"services"** cũng chiếm phần đáng kể, tuy nhiên tỷ lệ khách hàng trong các nhóm này ít hơn.
* **Những khách hàng thuộc nhóm nghề "student"** (sinh viên) và **"retired"** (người về hưu) có tần suất thấp hơn.

A pie chart with text

Description automatically generated

##### Biến Tình trạng hôn nhân (Marital status):

Biểu đồ tần suất cho biến **tình trạng hôn nhân** cho thấy:

* **Khách hàng đã kết hôn** là nhóm chiếm đa số trong tổng thể khách hàng.
* **Khách hàng độc thân** chiếm số lượng ít hơn so với khách hàng đã kết hôn, nhưng lại có tỷ lệ gửi tiền cao hơn.
* **Khách hàng đã ly hôn** chiếm tần suất nhỏ nhất trong dữ liệu, và cũng là nhóm có tỷ lệ gửi tiền thấp hơn.

A graph of a number of people

Description automatically generated with medium confidence

##### 

##### Biến Trình độ học vấn (Education):

Về trình độ học vấn, phân tích tần suất cho thấy:

* Phần lớn khách hàng có trình độ học vấn trung học ("secondary"), điều này phản ánh phần đông khách hàng có học vấn phổ thông.
* Trình độ đại học ("tertiary") và tiểu học ("primary") là hai nhóm tiếp theo cũng có tần suất xuất hiện tương đối.

A blue circle with red and black text

Description automatically generated

##### Biến Tình trạng nợ xấu (Default):

Phân tích tần suất cho biến **default** chỉ ra rằng:

* Hầu hết khách hàng **không có nợ xấu** ("no"), chiếm phần lớn trong dữ liệu.
* **Rất ít khách hàng có nợ xấu** ("yes"), điều này cho thấy hầu hết các khách hàng trong tập dữ liệu là những người có tín dụng tốt.

A diagram of a loan

Description automatically generated with medium confidence

##### Biến Vay nợ nhà ở (Housing loan):

Phân tích tần suất cho biến **housing loan** cho thấy số lượng khách hàng có vay nợ nhà ở và không vay nợ khá tương đương nhau:

* **Nhóm khách hàng không có khoản vay nhà ở** chiếm tỷ trọng nhỉnh hơn, cho thấy họ có khả năng tài chính tốt hơn để gửi tiền.
* **Nhóm có khoản vay nhà ở** có tỷ lệ gửi tiền thấp hơn, vì họ có thể có gánh nặng tài chính lớn hơn.

##### Biến Vay cá nhân (Loan):

Phân tích tần suất cho biến **loan** cho thấy:

* **Phần lớn khách hàng không có khoản vay cá nhân**, điều này có thể lý giải rằng họ có điều kiện tài chính tốt hơn để gửi tiền.
* **Nhóm khách hàng có khoản vay cá nhân** chiếm tần suất nhỏ hơn, và điều này phản ánh rằng khách hàng có khoản vay cá nhân thường gặp khó khăn trong việc tham gia các khoản gửi tiền có kỳ hạn.

A pie chart with different colored circles

Description automatically generated

#### 

##### 

##### Biểu đồ tần suất cho biến "Contact" (Phương thức liên lạc):

* **Phương thức liên lạc qua điện thoại di động ("cellular")** chiếm tỷ lệ cao nhất với **8,642** khách hàng, thể hiện đây là phương thức liên lạc chủ yếu mà ngân hàng sử dụng để tiếp cận khách hàng trong chiến dịch.
* **Liên lạc qua điện thoại cố định ("telephone")** được sử dụng ít hơn nhiều với **2,346** khách hàng, cho thấy phương thức này không phổ biến trong chiến dịch này.
* **Phương thức liên lạc không xác định ("unknown")** chỉ có **774** trường hợp, phản ánh một tỷ lệ nhỏ khách hàng không xác định rõ ràng về phương thức liên lạc.

Vì vậy, **Phương thức liên lạc qua điện thoại di động** là lựa chọn hiệu quả nhất, chiếm phần lớn các cuộc tiếp cận với khách hàng.

A graph of different colored squares

Description automatically generated

##### Biểu đồ tần suất cho biến "Month" (Tháng):

* **Tháng 5 (May)** có số lượng liên hệ cao nhất với **2,824** khách hàng, chiếm phần lớn trong tổng số liên hệ, cho thấy đây là tháng có hoạt động tiếp cận mạnh mẽ nhất trong chiến dịch.
* Các tháng **tháng 8 (August)**, **tháng 7 (July)** và **tháng 6 (June)** cũng có số lượng liên hệ khá cao, dao động từ **1,500 - 1,700** khách hàng, cho thấy hoạt động tiếp cận vẫn diễn ra khá đều trong các tháng mùa hè.
* **Tháng 11 (November)** và **tháng 12 (December)** có số lượng liên hệ thấp nhất với **276** và **110** khách hàng, cho thấy đây là giai đoạn hoạt động tiếp cận yếu nhất, có thể do ngân hàng giảm thiểu chiến dịch vào những tháng này.

Từ đó, ta có thể rút ra kết luận:

* **Chiến dịch cao điểm** diễn ra vào **tháng 5**, có thể là thời gian ngân hàng đẩy mạnh các chiến dịch marketing hoặc thực hiện các chương trình khuyến mãi để thu hút khách hàng.
* **Sự sụt giảm mạnh vào cuối năm** có thể cho thấy chiến lược tiếp cận cần được điều chỉnh để duy trì hiệu quả vào các tháng thấp điểm.

#### A screenshot of a computer screen Description automatically generated

##### Biểu đồ tần suất cho biến "Poutcome" (Kết quả trước đó):

* **"unknown"** là nhóm lớn nhất với **8,326** khách hàng, thể hiện rằng phần lớn khách hàng không có thông tin về kết quả chiến dịch trước đó. Điều này có thể do khách hàng chưa từng tham gia các chiến dịch trước hoặc dữ liệu chưa được ghi lại đầy đủ.
* Nhóm **"failure"** có **1,228** khách hàng, biểu thị những trường hợp mà chiến dịch trước đó không thành công trong việc thuyết phục khách hàng gửi tiền.
* **"success"** có **1,071** khách hàng, phản ánh những khách hàng đã có kết quả tích cực từ chiến dịch trước đó, cho thấy rằng chiến dịch trước đó đã đạt được kết quả nhất định đối với nhóm này.
* Nhóm **"other"** nhỏ nhất với **537** khách hàng, thể hiện các kết quả khác không thuộc các nhóm trên, có thể bao gồm các kết quả không rõ ràng hoặc không theo tiêu chuẩn.

#### A diagram of a pie chart Description automatically generated

##### Biểu đồ tần suất cho biến "Deposit" (Tiền gửi):

* **Nhóm không gửi tiền ("no")** chiếm tỷ lệ lớn với **5,873** khách hàng, cho thấy phần lớn khách hàng không đăng ký tiền gửi sau khi tham gia chiến dịch. Điều này có thể xuất phát từ nhiều nguyên nhân như chiến dịch chưa đủ sức hấp dẫn hoặc khách hàng chưa đủ điều kiện tài chính.
* **Nhóm gửi tiền ("yes")** chiếm **5,289** khách hàng, thể hiện rằng chiến dịch đã đạt được một mức độ thành công nhất định với gần một nửa số khách hàng đã tham gia gửi tiền.

***Kết luận chung:***

Các biểu đồ này cung cấp cái nhìn tổng quan về phân phối của các biến phân loại trong dữ liệu. Từ đó, có thể nhận thấy một số nhóm khách hàng đặc trưng như những người **có nghề nghiệp trong lĩnh vực "technician" hoặc "management", phần lớn không có nợ xấu và không có khoản vay cá nhân. Phương thức liên lạc qua điện thoại di động là phương thức chính mà ngân hàng sử dụng để tiếp cận khách hàng trong các chiến dịch**.

#### Phân tích biến định lượng

A graph of age distribution

Description automatically generated

##### Biến Age (tuổi):

* + Phân phối cho thấy hầu hết khách hàng nằm trong độ tuổi từ 20 đến 60, với đỉnh cao ở khoảng 30–40 tuổi.
  + Số lượng khách hàng giảm dần sau 60 tuổi.

##### Biến Balance (số dư tài khoản):

* + Phần lớn khách hàng có số dư tài khoản nằm trong khoảng từ 0 đến 2000 euro.
  + Có một số ít khách hàng có số dư tài khoản âm (thể hiện khoản nợ), nhưng đây không phải là số lượng lớn.

A green line graph with white text

Description automatically generated

##### Biến Duration (thời gian cuộc gọi):

* + Phần lớn các cuộc gọi kéo dài dưới 500 giây (dưới 10 phút).
  + Một số ít cuộc gọi kéo dài rất lâu (từ 1000 giây trở lên).

##### Biến Campaign (số lần khách hàng được liên hệ trong chiến dịch hiện tại):

* + Phần lớn khách hàng được liên hệ 1-3 lần trong chiến dịch.
  + Số lần liên hệ giảm mạnh sau 5 lần.

##### Biến Pdays (số ngày kể từ khi khách hàng được liên hệ lần cuối trong chiến dịch trước đó):

* + Phần lớn khách hàng có giá trị Pdays cao hoặc chưa được liên hệ trước đó (-1), thể hiện qua tần suất cao ở các giá trị lớn hoặc bằng 999 (giá trị đại diện cho "chưa liên hệ").

***Kết luận chung:***

Biểu đồ này cung cấp thông tin về cách các biến định lượng trong dữ liệu được phân bố. Các kết quả cho thấy phần lớn khách hàng nằm trong độ tuổi trung niên, có số dư tài khoản dương và chỉ được liên hệ một vài lần trong chiến dịch hiện tại. Thời gian cuộc gọi cũng có xu hướng ngắn. Những thông tin này có thể giúp xác định chiến lược tiếp cận khách hàng hiệu quả hơn trong các chiến dịch tiếp theo.

### Phân tích đa biến

#### Mối quan hệ giữa các biến độc lập với biến phụ thuộc

A graph of a graph

Description automatically generated with medium confidence

Biểu đồ box plot mà bạn cung cấp mô tả mối quan hệ giữa **"thời gian của lần liên hệ cuối cùng" (giây)** (trục y) và **"trạng thái gửi tiền"** (trục x), với hai nhóm: **"có gửi tiền"** (yes) và **"không gửi tiền"** (no).

* Nhóm khách hàng gửi tiền ("yes") có thời gian liên hệ trung bình và sự phân bố thời gian cao hơn so với nhóm không gửi tiền ("no").
* Những lần liên hệ dài hơn thường gắn liền với việc khách hàng gửi tiền, trong khi các cuộc gọi ngắn hơn phổ biến hơn ở nhóm không gửi tiền.
* Một số cuộc liên hệ kéo dài rất lâu (trên 2500 giây) có khả năng cao đã dẫn đến quyết định gửi tiền.

Vì vậy, Thời gian liên hệ dài hơn (tính bằng giây) có mối liên hệ tích cực với khả năng khách hàng gửi tiền. Nhóm "yes" có thời gian trung vị dài hơn và phạm vi thời gian liên hệ rộng hơn, điều này có thể cho thấy các cuộc gọi dài hơn và chi tiết hơn có khả năng thúc đẩy khách hàng gửi tiền.

A red and green circular objects with black text

Description automatically generated

Biểu đồ trên cho thấy mối quan hệ giữa việc có vay nhà và quyết định gửi tiết kiệm kỳ hạn.

* Khách hàng có vay nhà:
  + Tỷ lệ khách hàng có vay nhà và gửi tiết kiệm kỳ hạn là 36,6%. Điều này cho thấy một bộ phận không nhỏ khách hàng vẫn có thể cân đối giữa việc trả nợ và tiết kiệm.
  + Tỷ lệ khách hàng có vay nhà nhưng không gửi tiết kiệm kỳ hạn là 63,4%. Con số này khá cao, cho thấy áp lực tài chính từ khoản vay nhà có thể ảnh hưởng đến khả năng tiết kiệm của một số người.
* Khách hàng không có vay nhà:
  + Tỷ lệ khách hàng không có vay nhà và gửi tiết kiệm kỳ hạn là 57%. Điều này cho thấy khách hàng không có gánh nặng nợ nhà thường có khả năng tiết kiệm cao hơn.
  + Tỷ lệ khách hàng không có vay nhà nhưng không gửi tiết kiệm kỳ hạn là 43%. Con số này cho thấy vẫn có một số người không có vay nhà nhưng không có thói quen tiết kiệm.

=> Vay nhà ảnh hưởng đến hành vi tiết kiệm: Khách hàng có vay nhà thường có xu hướng tiết kiệm ít hơn so với những người không có vay nhà. Điều này có thể do các khoản chi trả hàng tháng cho khoản vay nhà đã chiếm một phần lớn ngân sách của họ.

A red and green circular objects with text

Description automatically generated

Biểu đồ trên chia khách hàng thành hai nhóm chính: những người có vay tiêu dùng và những người không có vay tiêu dùng

Khách hàng có vay tiêu dùng:

* Tỷ lệ khách hàng có vay tiêu dùng và gửi tiết kiệm kỳ hạn là 33.2%. Con số này cho thấy một bộ phận khách hàng vẫn có thể cân đối giữa việc trả nợ và tiết kiệm.
* Tuy nhiên, tỷ lệ khách hàng có vay tiêu dùng nhưng không gửi tiết kiệm kỳ hạn lại chiếm tới 66.8%. Điều này cho thấy áp lực tài chính từ khoản vay tiêu dùng có thể ảnh hưởng đáng kể đến khả năng tiết kiệm của nhiều người.

Khách hàng không có vay tiêu dùng:

* Tỷ lệ khách hàng không có vay tiêu dùng và gửi tiết kiệm kỳ hạn là 49.5%. Con số này cho thấy những người không có gánh nặng nợ tiêu dùng thường có khả năng tiết kiệm cao hơn.
* Mặc dù vậy, vẫn có 50.5% khách hàng không có vay tiêu dùng nhưng cũng không gửi tiết kiệm kỳ hạn. Điều này có thể do nhiều yếu tố khác nhau, chẳng hạn như thói quen chi tiêu, mục tiêu tài chính cá nhân, hoặc các yếu tố kinh tế khác.

A blue and orange pie chart

Description automatically generated

Biểu đồ trên chia khách hàng thành ba nhóm chính dựa trên tình trạng hôn nhân: đã ly hôn, đã kết hôn và độc thân. Mỗi nhóm lại được chia nhỏ theo hai tiêu chí: có gửi tiết kiệm kỳ hạn hay không.

Người đã ly hôn:

* Tỷ lệ người đã ly hôn và gửi tiết kiệm kỳ hạn là 48.1%. Điều này cho thấy một bộ phận người đã ly hôn vẫn có thói quen tiết kiệm.
* Tuy nhiên, tỷ lệ người đã ly hôn nhưng không gửi tiết kiệm kỳ hạn lại chiếm tới 51.9%. Điều này có thể do nhiều yếu tố, như áp lực tài chính sau ly hôn, hoặc ưu tiên chi tiêu cho các nhu cầu khác.

Người đã kết hôn:

* Tỷ lệ người đã kết hôn và gửi tiết kiệm kỳ hạn là 56.6%. Con số này cho thấy những người đã kết hôn thường có xu hướng tiết kiệm cao hơn. Điều này có thể do nhu cầu tích lũy cho tương lai, như nuôi con, mua nhà,...
* Tuy nhiên, vẫn có 43.4% người đã kết hôn nhưng không gửi tiết kiệm kỳ hạn. Điều này có thể do chi tiêu sinh hoạt gia đình cao, hoặc các yếu tố cá nhân khác.

Người độc thân:

* Tỷ lệ người độc thân và gửi tiết kiệm kỳ hạn là 45.7%. Con số này cho thấy người độc thân cũng có thói quen tiết kiệm, nhưng tỷ lệ này thấp hơn so với người đã kết hôn.
* Mặc dù vậy, vẫn có 54.3% người độc thân không gửi tiết kiệm kỳ hạn. Điều này có thể do thu nhập không ổn định, hoặc xu hướng chi tiêu cho các nhu cầu cá nhân.

A graph of blue and green bars

Description automatically generated

Biểu đồ cột so sánh việc đăng ký tiền gửi có kỳ hạn giữa các nhóm nghề nghiệp khác nhau. Trục hoành biểu thị các loại nghề nghiệp, trong khi trục tung hiển thị số lượng người, được phân loại thành những người có đăng ký ("yes" - màu xanh lá) và không đăng ký ("no" - màu xanh dương) tiền gửi có kỳ hạn.

1. Các nhóm nghề có tỷ lệ đăng ký tiền gửi cao ("Yes"):
   * Sinh viên có tỷ lệ đăng ký cao nhất, với 74.7% lựa chọn tiền gửi có kỳ hạn.
   * Người đã nghỉ hưu có tỷ lệ đăng ký cao là 66.3%.
   * Nhân viên quản lý và người thất nghiệp cũng có tỷ lệ đăng ký tương đối cao, lần lượt là 50.7% và 56.6%.
2. Các nhóm nghề có tỷ lệ đăng ký tiền gửi thấp ("No"):
   * Công nhân lao động phổ thông có tỷ lệ không đăng ký cao nhất, lên đến 63.6%.
   * Doanh nhân cũng không mặn mà với tiền gửi có kỳ hạn, với tỷ lệ không đăng ký là 62.5%.
   * Kỹ thuật viên có nhiều người không đăng ký hơn (53.9%) so với số người đăng ký.
3. Các nhóm nghề có tỷ lệ cân bằng:
   * Những người có nghề nghiệp không rõ có phân bố gần như cân bằng giữa đăng ký (48.6%) và không đăng ký (51.4%).
   * Nhân viên hành chính cũng khá cân bằng, mặc dù hơi nghiêng về phía không đăng ký với 52.7%.

***Nhận xét:***

* Những nghề như sinh viên, người đã nghỉ hưu và người thất nghiệp có xu hướng đăng ký tiền gửi có kỳ hạn nhiều hơn.
* Ngược lại, những nghề có thể liên quan đến lao động tay chân hoặc hoạt động doanh nghiệp, như công nhân lao động phổ thông và doanh nhân, có xu hướng không đăng ký nhiều hơn.

A graph of red and green bars

Description automatically generated

Biểu đồ trên cho thấy mối quan hệ giữa thời điểm liên hệ cuối cùng với khách hàng và quyết định gửi tiết kiệm kỳ hạn của họ.

* Sự khác biệt rõ rệt giữa các tháng: Tỷ lệ khách hàng gửi tiết kiệm kỳ hạn khác nhau đáng kể tùy thuộc vào tháng liên hệ cuối cùng, cho thấy thời điểm liên hệ có thể ảnh hưởng đến quyết định của khách hàng.
* Tháng 5 (May) có tỷ lệ gửi tiết kiệm cao nhất: Tháng 5 là tháng có tỷ lệ khách hàng gửi tiết kiệm cao nhất, cho thấy chiến dịch tiếp thị hoặc các hoạt động liên hệ với khách hàng trong tháng này đã rất hiệu quả.
* Một số tháng có tỷ lệ gửi tiết kiệm thấp: Các tháng như tháng 1 (Jan), tháng 3 (Mar), tháng 10 (Oct) có tỷ lệ khách hàng gửi tiết kiệm tương đối thấp, cho thấy các chiến dịch tiếp thị trong những tháng này có thể chưa đạt hiệu quả như mong đợi.
* Xu hướng chung: Không có xu hướng tăng giảm rõ ràng theo thời gian. Tỷ lệ gửi tiết kiệm có sự biến động khá lớn giữa các tháng.

A graph of blue and green bars

Description automatically generated

Biểu đồ trên so sánh việc đăng ký tiền gửi có kỳ hạn theo các cấp độ giáo dục khác nhau. Trục hoành biểu thị các cấp độ giáo dục (primary - tiểu học, secondary - trung học, tertiary - đại học và unknown - không rõ), trong khi trục tung hiển thị số lượng người, được phân chia thành nhóm có đăng ký ("yes" - màu xanh lá) và không đăng ký ("no" - màu xanh dương).

* Những người có trình độ học vấn cao hơn, cụ thể là cấp đại học (tertiary), có xu hướng đăng ký tiền gửi có kỳ hạn nhiều hơn so với những người chỉ có trình độ tiểu học hoặc trung học.
* Nhóm có trình độ tiểu học và trung học phần lớn không đăng ký tiền gửi có kỳ hạn, cho thấy có thể có mối liên hệ giữa mức độ giáo dục và khả năng hoặc xu hướng tài chính của một người.

A graph of a number of blue and purple bars

Description automatically generated

Biểu đồ trên so sánh việc đăng ký tiền gửi có kỳ hạn dựa trên tình trạng hôn nhân và nhóm tuổi. Trục hoành biểu thị tình trạng hôn nhân (đã ly hôn - divorced, đã kết hôn - married, độc thân - single) và trục tung biểu thị số lượng người đăng ký. Các nhóm tuổi bao gồm Minor (vị thành niên), Young Adult (người trưởng thành trẻ), Adult (người trưởng thành), Middle-Aged (trung niên) và Senior (người cao tuổi.

* Nhìn chung, tỷ lệ không đăng ký tiền gửi có kỳ hạn cao hơn ở các nhóm người đã kết hôn và ly hôn, đặc biệt là ở nhóm tuổi trẻ và trưởng thành.
* Người lớn tuổi và người đã ly hôn có xu hướng đăng ký tiền gửi nhiều hơn so với các nhóm khác, đặc biệt trong nhóm người cao tuổi, có thể do nhu cầu tài chính và tiết kiệm ở giai đoạn này tăng cao.

A graph of a diagram

Description automatically generated with medium confidence

**Biểu đồ khu vực xếp chồng trên thể hiện mối quan hệ giữa nhóm tuổi, tình trạng hôn nhân, và quyết định đăng ký tiền gửi có kỳ hạn.**

Xu hướng theo nhóm tuổi

* Khách hàng trẻ tuổi hơn (Young Adult, Adult) chiếm tỉ lệ đông đảo, nhưng tỷ lệ đăng ký tiền gửi không cao.
* Nhóm tuổi trưởng thành (Adult) có lượng khách hàng lớn nhất, nhưng cũng như các nhóm khác, đa phần họ không đăng ký.
* Nhóm tuổi trung niên và người cao tuổi (Middle-Aged, Senior) chiếm tỷ lệ nhỏ và tỷ lệ đăng ký tiền gửi cũng thấp hơn so với nhóm trẻ hơn.

Các nhóm cần chú ý

* Nhóm khách hàng độc thân ở độ tuổi Adult và Young Adult mặc dù chiếm tỷ lệ lớn nhưng phần lớn không đăng ký, vì vậy nhóm này có thể cần được chú trọng hơn trong các chiến dịch marketing.
* Nhóm khách hàng đã kết hôn mặc dù có tỷ lệ đăng ký cao hơn một chút nhưng vẫn là phần nhỏ, điều này có thể do họ có những ưu tiên tài chính khác.

### Ma trận tương quan và biểu đồ heatmap

A screenshot of a graph

Description automatically generated

* **Các đặc trưng có tương quan mạnh với biến mục tiêu ("deposit"):**
* Kết quả chiến dịch trước ("poutcome"): Có tương quan dương tương đối mạnh (~0.5) với "deposit". Điều này cho thấy kết quả của chiến dịch tiếp thị trước đó có ảnh hưởng lớn đến việc khách hàng có gửi tiền hay không.
* Thời gian cuộc gọi ("duration"): Cũng có tương quan dương đáng kể (~0.5) với "deposit". Điều này có thể ám chỉ rằng những cuộc gọi tiếp thị kéo dài có thể dẫn đến khả năng gửi tiền cao hơn từ phía khách hàng.
* Số liên hệ trong chiến dịch ("campaign"): Có tương quan âm nhẹ với "deposit", cho thấy số lần liên hệ không ảnh hưởng tích cực đến quyết định gửi tiền.
* **Các đặc trưng có mối tương quan yếu với biến mục tiêu:**
* Tuổi ("age"): Tương quan rất yếu với "deposit" (~0.13), cho thấy tuổi tác không phải là yếu tố quyết định lớn đối với khả năng gửi tiền.
* Nghề nghiệp ("job") và trình độ học vấn ("education"): Cả hai đều có tương quan thấp với "deposit", ngụ ý rằng các yếu tố này không ảnh hưởng nhiều đến việc khách hàng có gửi tiền hay không.
* Số ngày từ lần liên hệ trước ("previous"): Cũng có mối tương quan rất yếu với "deposit", cho thấy thời gian từ lần tiếp xúc cuối cùng đến hiện tại không quá quan trọng.
* Ngày gọi điện ("day") và phương thức liên hệ ("contact"): Tương quan rất thấp, gần như không ảnh hưởng lẫn nhau.
* Cân bằng tài khoản ("balance") và default (khả năng vỡ nợ): Tương quan thấp, điều này có thể cho thấy số dư tài khoản không liên quan nhiều đến việc có vỡ nợ hay không.

### Các biến quan trọng nhất ảnh hưởng đến quyết định đăng ký tiền gửi có kỳ hạn.

Chúng tôi cho rằng có một vài vấn đề khi sử dụng mô hình tuyến tính như sau:

- Ma trận tương quan chỉ đánh giá mối quan hệ tuyến tính giữa các biến. Do đó, những mối quan hệ phi tuyến hoặc tương tác phức tạp có thể không được phát hiện, làm cho một số yếu tố có vẻ ít quan trọng hơn so với thực tế.

- Ma trận tương quan chỉ đo lường mối quan hệ tuyến tính giữa từng cặp biến, và giữa từng biến với biến mục tiêu, không xem xét sự tương tác giữa nhiều biến cùng lúc. Điều này làm hạn chế khả năng phát hiện những tương tác tiềm ẩn quan trọng.

- Ma trận tương quan: Dễ bị ảnh hưởng bởi các yếu tố nhiễu hoặc biến dư thừa, làm giảm khả năng phát hiện các yếu tố quan trọng thực sự.

Vì vậy chúng tôi đã sử dụng thêm mô hình Random Forest để loại bỏ những vấn đề tiềm ẩn trên, do Random Forest là mô hình phi tuyến tính, Random Forest có khả năng phát hiện và xử lý những mối quan hệ phức tạp và phi tuyến giữa các biến. Nó không chỉ xem xét mối quan hệ trực tiếp mà còn phát hiện các tương tác giữa các biến, giúp xác định tầm quan trọng của từng yếu tố trong việc dự đoán quyết định.

A graph of a model

Description automatically generated with medium confidence

##### Duration (Thời gian liên lạc)

* Tầm quan trọng: Đây là yếu tố quan trọng nhất có tác động đến khả năng gửi tiền của khách hàng. Điều này cho thấy rằng thời gian mà khách hàng đã dành để trao đổi với nhân viên tư vấn là yếu tố then chốt trong việc ảnh hưởng đến quyết định đăng ký.
* Giải thích: Việc khách hàng dành nhiều thời gian để tương tác có thể phản ánh mức độ quan tâm của họ đối với sản phẩm tài chính hoặc cho thấy nhân viên tư vấn có khả năng giải thích rõ ràng và thuyết phục.

##### Month (Tháng liên lạc)

* Tầm quan trọng: Tháng liên hệ đứng thứ hai về tầm quan trọng trong cả hai biểu đồ, mặc dù month có tương quan thấp với deposit (0.04), từ đó cho thấy mối quan hệ phi tuyến tính giữa 2 biến.
* Giải thích: Thời gian trong năm có thể ảnh hưởng đến tâm lý và nhu cầu tài chính của khách hàng. Ví dụ, các tháng cuối năm có thể là thời điểm khách hàng cân nhắc về việc đầu tư hoặc tiết kiệm nhiều hơn, hoặc có thể có những biến động theo chu kỳ nhất định.

##### Balance (Số dư tài khoản)

* Tầm quan trọng: Biến này có ảnh hưởng lớn, mặc dù tương quan giữa balance và deposit là 0.12 (thể hiện mối quan hệ không đáng kể), từ đó thể hiện mối quan hệ phi tuyến tính giữa hai biến.
* Giải thích: Khả năng tài chính của khách hàng, được phản ánh qua số dư tài khoản, rõ ràng có ảnh hưởng trực tiếp đến việc họ có khả năng hoặc có hứng thú tham gia vào các sản phẩm tài chính như tiền gửi có kỳ hạn. Khách hàng có số dư tài khoản cao hơn có xu hướng dễ dàng đưa ra quyết định hơn.

##### Age (Tuổi)

* Từ Random Forest: Biến này có ảnh hưởng tương đối lớn, nằm trong nhóm các biến quan trọng trung bình.
* Từ ma trận tương quan: age có tương quan âm với deposit (-0.11), cho thấy những khách hàng trẻ hơn có xu hướng dễ đăng ký tiền gửi hơn.
* Kết luận: Độ tuổi là một yếu tố ảnh hưởng, với khách hàng trẻ có khả năng đăng ký cao hơn. Do đó, nên nhắm vào phân khúc khách hàng trẻ trong các chiến dịch tiếp cận.

##### Pdays (Số ngày từ chiến dịch trước)

* Từ Random Forest: Biến này có mức độ quan trọng thấp trong mô hình, nhưng không thể bỏ qua hoàn toàn.
* Từ ma trận tương quan: Tương quan giữa pdays và deposit là -0.50, cho thấy mối quan hệ ngược. Những khách hàng được liên hệ sau khoảng thời gian dài từ chiến dịch trước có khả năng đăng ký thấp hơn.
* Kết luận: Khoảng thời gian giữa các chiến dịch cần được chú ý. Nếu thời gian liên hệ quá xa, khả năng đăng ký của khách hàng giảm rõ rệt.

##### Previous (Số lượng tương tác trước đó)

* Từ Random Forest: Biến này có ảnh hưởng trung bình trong mô hình.
* Từ ma trận tương quan: previous có tương quan cao với poutcome (0.51) và cả với deposit (0.12), cho thấy rằng việc có nhiều lần tương tác trước đó có thể làm tăng khả năng đăng ký của khách hàng.
* Kết luận: Các chiến dịch cần chú trọng vào việc theo dõi và duy trì mối quan hệ với khách hàng thông qua các tương tác trước đó.

### Kết luận chung

Tóm lại, tệp khách hàng nên được chú ý sẽ giúp tăng khả năng đăng ký tiền gửi có kỳ hạn, nâng cao hiệu quả chiến dịch marketing. bao gồm:

* Khách hàng độc thân, trình độ đại học, không có nợ quá hạn, và không có khoản vay.
* Những người trong độ tuổi 30-50 với số dư tài khoản cao.
* Phương thức liên hệ qua di động và thời gian cuộc gọi dài hơn có thể tăng tỷ lệ thành công.
* Nên hạn chế liên hệ quá nhiều lần và tập trung vào những khách hàng chưa được liên hệ gần đây.

Đặc biệt tập trung vào những đối tượng sau để tối ưu hóa nhất hiệu quả chiến dịch marketing;

* Những khách hàng có thời gian tư vấn ở các chiến dịch trước lâu.
* Những khách hàng có số dư tài khoản lớn.
* Những khách hàng trẻ tuổi.
* Những khách hàng đã có tương tác trước đó với ngân hàng.
* Những khách hàng được liên hệ vào những tháng chiến dịch có hiệu quả cao.

## 3. Đề xuất chiến lược

#### 3.1 Kế hoạch nhắm đến phân khúc khách hàng cụ thể

* Tập trung vào nhóm tuổi 30-50 với số dư tài khoản cao:

Khách hàng trong độ tuổi 30-50 thường có thu nhập ổn định, có khả năng tiết kiệm và đầu tư lâu dài. Đây là giai đoạn họ đạt đỉnh cao trong sự nghiệp, và cũng là lúc họ có nhu cầu sử dụng các sản phẩm tài chính an toàn, có lãi suất ổn định như tiền gửi có kỳ hạn. Thêm vào đó, nhóm khách hàng có số dư tài khoản cao thường ít bị áp lực tài chính và có khả năng đầu tư vào các sản phẩm tài chính dài hạn.

Chiến lược cụ thể: Phát triển các sản phẩm tiền gửi có kỳ hạn phù hợp với nhu cầu và ưu tiên của nhóm khách hàng này, kèm theo lợi ích về lãi suất và tính bảo mật. Nhắm vào thông điệp về sự an toàn, ổn định và tích lũy tài sản dài hạn.

* Nhắm vào các nhóm khách hàng có trình độ học vấn cao và không có nợ:

Những khách hàng có trình độ học vấn cao thường có kiến thức về tài chính tốt hơn, do đó dễ dàng nhận ra lợi ích của các sản phẩm tiền gửi có kỳ hạn. Đồng thời, việc tập trung vào khách hàng không có nợ xấu hoặc không có các khoản vay giúp giảm rủi ro cho ngân hàng và đảm bảo tính khả thi của việc đăng ký sản phẩm.

Chiến lược cụ thể: Thiết kế các chương trình marketing riêng cho đối tượng khách hàng này, với các thông điệp tập trung vào các giá trị lợi nhuận dài hạn và tính bền vững tài chính. Điều này sẽ làm cho các khách hàng có trình độ cảm thấy sản phẩm phù hợp với mục tiêu tài chính cá nhân của họ.

#### 3.2. Lựa chọn thời gian liên hệ và tối ưu hóa lịch gọi

Chọn đúng thời điểm liên hệ giúp tăng khả năng phản hồi tích cực từ khách hàng. Ví dụ, trong các tháng mà chiến dịch trước đạt hiệu quả cao, hoặc vào các khung giờ mà khách hàng có nhiều thời gian để trao đổi (như giờ nghỉ trưa hoặc sau giờ làm việc). Điều này giúp tối ưu hóa sự tương tác và hiệu quả của cuộc gọi.

Chiến lược cụ thể: Xây dựng lịch trình liên hệ phù hợp dựa trên dữ liệu lịch sử của các chiến dịch trước để nhắm vào những tháng và giờ cụ thể có hiệu quả cao hơn. Tập trung vào các thời điểm khách hàng có khả năng tương tác tốt hơn, giúp tăng tỷ lệ thành công.

#### 3.3. Phương thức liên lạc tối ưu

Thống kê cho thấy kênh liên lạc qua di động có tỷ lệ thành công cao hơn so với điện thoại bàn hoặc các kênh liên lạc khác. Điều này là do tính thuận tiện và trực tiếp của di động, giúp khách hàng dễ dàng tiếp nhận và phản hồi ngay lập tức.

Chiến lược cụ thể: Ưu tiên sử dụng điện thoại di động cho các cuộc gọi tư vấn và tăng thời gian cuộc gọi để có thể tư vấn chi tiết hơn về lợi ích của sản phẩm. Đồng thời, sử dụng tin nhắn SMS hoặc email nhắc nhở trước và sau cuộc gọi để giữ sự liên tục trong quá trình giao tiếp với khách hàng.

#### 3.4. Tối ưu số lần liên hệ và thời gian tư vấn

Liên hệ quá nhiều lần có thể gây phiền toái cho khách hàng và làm giảm uy tín của ngân hàng. Ngược lại, nếu liên hệ vừa phải và tập trung vào những khách hàng chưa được liên hệ gần đây hoặc có khả năng quan tâm cao, ngân hàng sẽ tối ưu hóa được nguồn lực và tăng hiệu quả chiến dịch.

Chiến lược cụ thể:

* Hạn chế số lần liên hệ từ 1-3 lần cho một khách hàng cụ thể. Tập trung vào những khách hàng có dấu hiệu quan tâm hoặc chưa được liên hệ trong các chiến dịch gần đây. Điều này giúp tăng khả năng thành công mà không gây phiền hà cho khách hàng.
* Tối ưu hóa thời gian tư vấn: Nhân viên được đào tạo để nắm bắt thông tin của khách hàng nhanh chóng và tư vấn hiệu quả trong khoảng thời gian tối ưu. Đối với những khách hàng có sự quan tâm đặc biệt, thời gian tư vấn có thể kéo dài hơn.

#### 3.5. Theo dõi và đánh giá hiệu quả chiến dịch

Hoạt động chi tiết:

* Xây dựng hệ thống theo dõi: Thiết lập hệ thống theo dõi các chỉ số hiệu quả chiến dịch như tỷ lệ phản hồi, tỷ lệ chuyển đổi (từ liên hệ sang đăng ký), chi phí cho mỗi cuộc gọi, và thời gian tư vấn trung bình.
* Phân tích dữ liệu sau chiến dịch: Đánh giá dữ liệu từ chiến dịch để phân tích những yếu tố thành công và những yếu tố cần cải thiện, từ đó rút ra bài học và điều chỉnh cho các chiến dịch tiếp theo.

Tối ưu quy trình dựa trên dữ liệu: Áp dụng các bài học từ chiến dịch trước để tiếp tục tối ưu hóa quy trình liên hệ và tư vấn cho khách hàng, từ đó gia tăng hiệu quả trong những lần tiếp theo.

=> Việc theo dõi và phân tích các chỉ số hiệu quả giúp ngân hàng nắm bắt được tình hình thực tế, từ đó điều chỉnh chiến lược kịp thời và phù hợp, tối ưu hóa chi phí và nguồn lực.

### 4. Đánh giá tác động:

* Tăng tỷ lệ đăng ký lên khoảng 15-25%:

Bằng cách tập trung vào các nhóm khách hàng có tiềm năng cao (độ tuổi 30-50, trình độ học vấn cao, không có nợ và có số dư tài khoản lớn), cùng với việc chọn kênh liên lạc và thời điểm phù hợp, khả năng thành công của chiến dịch marketing sẽ được cải thiện. Dữ liệu lịch sử cho thấy rằng những khách hàng này thường có tỷ lệ đăng ký sản phẩm cao hơn khi được tư vấn đúng cách và vào thời điểm thích hợp.

Ước tính tác động: ***Việc áp dụng chiến lược này có thể làm tăng tỷ lệ đăng ký sản phẩm từ 15-25%***, dựa trên kết quả phân tích dữ liệu khách hàng tiềm năng và lịch sử chiến dịch.

* Giảm chi phí liên hệ và tăng hiệu quả:

Việc tối ưu số lần liên hệ và lựa chọn đúng phương thức liên lạc (đặc biệt là qua di động) giúp giảm chi phí cho mỗi cuộc gọi và tối ưu hóa nguồn lực. Khi tập trung vào khách hàng tiềm năng và sử dụng phương pháp tiếp cận hiệu quả, ngân hàng sẽ tiết kiệm được nguồn lực liên lạc mà vẫn đạt được tỷ lệ chuyển đổi cao hơn.

Ước tính tác động: ***Chi phí cho mỗi cuộc gọi hoặc mỗi khách hàng có thể giảm từ 10-20%,*** trong khi tỷ lệ thành công của các cuộc gọi lại tăng, giúp cải thiện đáng kể ROI (Return on Investment) cho chiến dịch.

* Tăng tỷ lệ giữ chân khách hàng:

Nhắm mục tiêu đến những khách hàng đã có tương tác trước đó với ngân hàng giúp tăng cường mối quan hệ và xây dựng lòng trung thành. Khách hàng đã có trải nghiệm tốt với ngân hàng sẽ dễ dàng tin tưởng và tiếp tục sử dụng các sản phẩm tài chính khác. Điều này không chỉ giúp ngân hàng tăng doanh thu ngắn hạn mà còn tạo ra giá trị lâu dài thông qua việc giữ chân khách hàng.

Ước tính tác động: ***Tỷ lệ giữ chân khách hàng có thể tăng khoảng 10-15%***, nhờ vào việc tạo dựng mối quan hệ lâu dài thông qua các sản phẩm tiền gửi có kỳ hạn và các chương trình chăm sóc khách hàng hiệu quả hơn. Điều này giúp tăng độ trung thành của khách hàng, giảm chi phí cho các chiến dịch thu hút khách hàng mới.

## KẾT LUẬN

Kết quả phân tích cho thấy chiến dịch tiếp thị qua điện thoại của ngân hàng đã đạt được những thành công đáng kể, nhưng vẫn còn tiềm năng để tối ưu hóa nhằm đạt hiệu quả cao hơn. Thời gian cuộc gọi là yếu tố quan trọng nhất, với những cuộc gọi dài hơn có xu hướng dẫn đến quyết định đăng ký tiền gửi cao hơn từ khách hàng. Ngoài ra, phương thức liên hệ qua điện thoại di động tỏ ra vượt trội so với điện thoại cố định về hiệu quả tương tác, trong khi thời điểm tiếp cận khách hàng, đặc biệt vào tháng 5, cũng đóng vai trò quan trọng trong việc nâng cao tỷ lệ đăng ký.

Phân tích sâu hơn về nhân khẩu học cho thấy nhóm khách hàng trong độ tuổi 30-50, có trình độ học vấn cao và không có nợ xấu hoặc khoản vay nhà, là những đối tượng tiềm năng nhất cho sản phẩm tiền gửi có kỳ hạn. Đặc biệt, những khách hàng có số dư tài khoản lớn và đã có lịch sử tương tác tích cực với ngân hàng là những ứng viên có khả năng cao sẽ tham gia vào sản phẩm này.

Ngoài ra, chiến dịch cũng gặp phải một số thách thức, bao gồm tỷ lệ không đăng ký tiền gửi cao ở những khách hàng được liên hệ quá nhiều lần hoặc vào những thời điểm không phù hợp. Do đó, việc tối ưu hóa tần suất liên hệ và tập trung vào chất lượng tư vấn trong mỗi cuộc gọi là điều cần thiết để không chỉ giảm chi phí tiếp thị mà còn nâng cao trải nghiệm của khách hàng.

Việc áp dụng các đề xuất trong báo cáo, bao gồm tối ưu hóa phân khúc khách hàng, tăng cường thời gian và chất lượng tương tác, và lựa chọn thời gian tiếp cận thích hợp, có thể giúp tăng tỷ lệ đăng ký lên 15-25%. Hơn nữa, chiến lược này sẽ góp phần giảm chi phí liên hệ từ 10-20%, đồng thời cải thiện tỷ lệ giữ chân khách hàng thông qua việc duy trì mối quan hệ lâu dài và xây dựng lòng trung thành. Điều này không chỉ mang lại giá trị ngắn hạn trong việc nâng cao doanh thu mà còn tạo dựng nền tảng vững chắc cho sự phát triển bền vững của ngân hàng trong tương lai.