**TRƯỜNG ĐẠI HỌC KINH TẾ ĐÀ NẴNG**

**KHOA THƯƠNG MẠI ĐIỆN TỬ**

**🙡🕮🙣**



**HỌC PHẦN: ĐỀ ÁN THỰC HÀNH 1**

**CHỦ ĐỀ: ỨNG DỤNG KỸ THUẬT PHÂN TÍCH KHÁCH HÀNG RỜI BỎ DOANH NGHIỆP CHO NGÀNH VIỄN THÔNG VÀ ĐỀ XUẤT GIẢI PHÁP**

**Giảng viên hướng dẫn:** Nguyễn Văn Chức

**Sinh viên thực hiện:** Trần Thị Như Quỳnh

**Lớp:** 47K29.1

**Đà Nẵng, ngày 5 tháng 5 năm 2024**

**MỤC LỤC**

[LỜI MỞ ĐẦU 1](#_Toc2239)

[A. GIỚI THIỆU TỔNG QUAN: 1](#_Toc28237)

[I. Giới thiệu bài báo cáo: 1](#_Toc27122)

[II. Mục tiêu của bài báo cáo: 2](#_Toc24196)

[III. Đối tượng và phạm vi bài báo cáo: 2](#_Toc17664)

[IV. Phương pháp thực hiện bài báo cáo: 2](#_Toc5348)

[B. CƠ SỞ LÝ THUYẾT: 2](#_Toc27493)

[I. Tỷ lệ khách hàng rời bỏ là gì? 2](#_Toc25803)

[II. Tại sao khách hàng lại rời bỏ doanh nghiệp? 2](#_Toc21391)

[III. Tại sao việc giữ chân khách hàng lại quan trọng? 3](#_Toc16720)

[IV. Công cụ hỗ trợ 3](#_Toc10514)

[1. Python: 3](#_Toc299)

[2. Tableau: 4](#_Toc9075)

[V. Thuật toán Logistic Regression 4](#_Toc29390)

[C. BÀI TOÁN. 5](#_Toc86)

[I. Mục tiêu bài toán: 5](#_Toc19039)

[II. Giới thiệu tổng quan về dữ liệu: 6](#_Toc9237)

[III. Mô tả dữ liệu: 6](#_Toc115)

[D. TRỰC QUAN HÓA DỮ LIỆU 10](#_Toc15364)

[I. Tổng quan về bộ dữ liệu: 10](#_Toc12223)

[II. Tổng quan về khách hàng: 11](#_Toc22517)

[1. Thông tin về nhân khẩu học: 11](#_Toc12561)

[2. Thông tin về gói dịch vụ: 12](#_Toc23734)

[2.1. Dịch vụ điện thoại (PhoneService): 12](#_Toc28751)

[2.2. Đường dây điện thoại (Multiple Line): 13](#_Toc32668)

[2.3. Dịch vụ Internet (Internet Service): 14](#_Toc2650)

[2.4. Dịch vụ bảo mật trực tuyến (Online Security): 15](#_Toc15200)

[2.5. Dịch vụ sao lưu trực tuyến (Online Backup): 15](#_Toc763)

[2.6. Dịch vụ bảo vệ thiết bị (Device Protection): 16](#_Toc28882)

[2.7. Dịch vụ hỗ trợ kỹ thuật (Tech Support): 17](#_Toc6984)

[2.8. Streaming TV: 17](#_Toc14891)

[2.9. Streaming Movies: 18](#_Toc21537)

[2.10. Tổng kết: 19](#_Toc28176)

[3. Thông tin về hợp đồng của khách hàng: 19](#_Toc17852)

[3.1. Thời hạn hợp đồng (Contract): 19](#_Toc30592)

[3.2. Hóa đơn điện tử (Paperless Billing): 20](#_Toc31824)

[3.3. Phương thức thanh toán (Payment Method): 21](#_Toc4664)

[III. Tỷ lệ khách hàng rời bỏ công ty Telco: 21](#_Toc19301)

[1. Churn Rate: 21](#_Toc14178)

[2. Churn Rate Gender: 22](#_Toc21584)

[3. Churn Rate Dependents: 22](#_Toc9348)

[4. Churn Rate SeniorCitizen: 23](#_Toc8675)

[5. Churn Rate Partner: 25](#_Toc30466)

[6. Churn Rate Phone Service/Multiple Lines: 26](#_Toc2658)

[7. Tỷ lệ khách hàng rời bỏ (churn rate) theo dịch vụ Internet (Internet Services) và các dịch vụ đi kèm (Sub Internet Services): 27](#_Toc26812)

[7.1. Internet Service: 27](#_Toc21036)

[7.2. Online Backup: 28](#_Toc24205)

[7.3. Online Security: 29](#_Toc7368)

[7.4. Streaming Movies: 29](#_Toc16301)

[7.5. Streaming TV: 30](#_Toc26415)

[7.6. Tech Support: 31](#_Toc1353)

[7.7. Device Protection: 32](#_Toc23944)

[8. Churn Rate Tenure: 33](#_Toc11222)

[9. Churn Rate Paperless Billing & Payment Method: 34](#_Toc12216)

[10. Churn Rate Contract: 36](#_Toc15328)

[IV. Xây dựng mô hình: 37](#_Toc25454)

[1. Tiền xử lý dữ liệu: 37](#_Toc3684)

[1.1. Đọc và hiểu dữ liệu: 37](#_Toc22420)

[1.2. Kiểm tra dữ liệu null: 38](#_Toc20732)

[1.3. Mã hóa các giá trị nhãn thành các giá trị số nguyên: 39](#_Toc3152)

[2. Xây dựng mô hình Logistic Regression: 39](#_Toc12385)

[3. Đánh giá mô hình: 40](#_Toc7157)

[3.1. Độ chính xác: 40](#_Toc1043)

[3.2. Ma trận nhầm lẫn: 41](#_Toc29845)

[3.3. Đường cong ROC: 41](#_Toc27688)

[4. Dự đoán dựa trên dữ liệu mới: 42](#_Toc7960)

[V. Giải pháp 43](#_Toc6789)

[E. KẾT LUẬN 44](#_Toc2040)

**DANH MỤC HÌNH ẢNH**

[Hình 1 :Dashboard Tổng quan về bộ dữ liệu (Overview) 10](#_Toc8224)

[Hình 2 : Dashboard về Tỷ lệ khách hàng rời bỏ công ty Telco 10](#_Toc21720)

[Hình 3 : Tổng quan về bộ dữ liệu. 11](#_Toc29532)

[Hình 4 : Thông tin về nhân khẩu học 11](#_Toc26775)

[Hình 5 : Số khách hàng sử dụng dịch vụ điện thoại 13](#_Toc29092)

[Hình 6 : Số lượng khách hàng sử dụng dịch vụ đường dây điện thoại. 13](#_Toc848)

[Hình 7 : Phần trăm khách hàng sử dụng dịch vụ internet. 14](#_Toc7895)

[Hình 8 : Số lượng khách hàng sử dụng dịch vụ Bảo mật trực tuyến. 15](#_Toc1406)

[Hình 9 : Số lượng khách hàng sử dụng dịch vụ Sao lưu trực tuyến. 16](#_Toc2542)

[Hình 10 : Số lượng khách hàng sử dụng dịch vụ Bảo vệ thiết bị. 16](#_Toc30080)

[Hình 11 : Số lượng khách hàng sử dụng dịch vụ Hỗ trợ kỹ thuật. 17](#_Toc26847)

[Hình 12 : Số lượng khách hàng sử dụng dịch vụ Streaming TV. 18](#_Toc6814)

[Hình 13 : Số lượng khách hàng sử dụng dịch vụ Streaming Movies. 18](#_Toc11200)

[Hình 14 : Tổng kết tỷ lệ khách hàng sử dụng các dịch vụ Internet. 19](#_Toc18759)

[Hình 15 : Tỷ lệ khách hàng sử dụng các loại hợp đồng. 20](#_Toc19248)

[Hình 16 : Tỷ lệ khách hàng sử dụng hóa đơn điện tử. 20](#_Toc17564)

[Hình 17 : Tỷ lệ khách hàng sử dụng các phương thức thanh toán. 21](#_Toc22003)

[Hình 18 : Số lượng khách hàng rời bỏ công ty. 21](#_Toc22835)

[Hình 19 : Tỷ lệ khách hàng rời bỏ công ty. 22](#_Toc476)

[Hình 20 : Tỷ lệ khách hàng rời bỏ doanh nghiệp theo giới tính. 22](#_Toc5396)

[Hình 21 : Tỷ lệ khách hàng có người phụ thuộc rời bỏ công ty. 23](#_Toc17392)

[Hình 22 : Tỷ lệ rời bỏ công ty của người lớn tuổi. 24](#_Toc16648)

[Hình 23 : Tỷ lệ rời bỏ công ty của khách hàng có vợ/chồng. 25](#_Toc13256)

[Hình 24 : Tỷ lệ rời bỏ công ty của khách hàng sử dụng PhoneService và MultipleLines 26](#_Toc32532)

[Hình 25 : Tỷ lệ rời bỏ công ty của khách hàng sử dụng dịch vụ OnlineBackup. 28](#_Toc15927)

[Hình 26 : Tỷ lệ rời bỏ công ty của khách hàng sử dụng dịch vụ OnlineSecurity. 29](#_Toc12864)

[Hình 27 : Tỷ lệ rời bỏ công ty của khách hàng sử dụng dịch vụ StreamingMovies. 30](#_Toc27437)

[Hình 28 : Tỷ lệ rời bỏ công ty của khách hàng sử dụng dịch vụ StreamingTV. 31](#_Toc5020)

[Hình 29 : Tỷ lệ rời bỏ công ty của khách hàng sử dụng dịch vụ TechSupport. 32](#_Toc15067)

[Hình 30 : Tỷ lệ rời bỏ công ty của khách hàng sử dụng dịch vụ DeviceProtection 33](#_Toc6277)

[Hình 31 : Tỷ lệ khách hàng rời bỏ công ty theo số tháng gắn bó. 34](#_Toc24088)

[Hình 32 : Tỷ lệ khách hàng rời bỏ công ty theo Paperless Billing & Payment Method. 35](#_Toc3621)

[Hình 33 : Tỷ lệ khách hàng rời bỏ công ty theo loại hợp đồng. 36](#_Toc19972)

[Hình 34 : Dữ liệu gốc 37](#_Toc7791)

[Hình 35 : Số lượng khách hàng rời bỏ và không rời bỏ của tập dữ liệu. 37](#_Toc29261)

[Hình 36 : Kiểm tra dữ liệu null. 38](#_Toc25496)

[Hình 37 : Dữ liệu sau khi xóa dữ liệu null. 38](#_Toc31261)

[Hình 38 : Mã hóa các giá trị nhãn thành các giá trị số nguyên 39](#_Toc9078)

[Hình 39 : Chia tập dữ liệu thành tập Train và tập Test. 39](#_Toc15346)

[Hình 40 : Huấn luyện mô hình LogisticRegression. 40](#_Toc13293)

[Hình 41 : Độ chính xác của mô hình. 40](#_Toc23766)

[Hình 42 : Các chỉ số đánh giá mô hình 40](#_Toc882)

[Hình 43 :Ma trận nhầm lẫn. 41](#_Toc26907)

[Hình 44 : Đường cong ROC. 42](#_Toc31753)

[Hình 45 : Nhập dữ liệu dự đoán từ bàn phím 43](#_Toc28399)

[Hình 46 : Dự đoán nhãn cho dữ liệu mới. 43](#_Toc4719)

# **LỜI MỞ ĐẦU**

Trong thế giới kinh doanh ngày nay, việc duy trì mối quan hệ với khách hàng đã trở nên nền tảng quan trọng nhất đối với sự thành công của mọi doanh nghiệp. Tuy nhiên, một thách thức lớn mà nhiều tổ chức đang phải đối mặt là việc giảm thiểu tỷ lệ mất khách hàng. Khách hàng rời bỏ không chỉ ảnh hưởng tiêu cực đến doanh số bán hàng mà còn đặt ra câu hỏi lớn về hiệu suất hoạt động cũng như sức khỏe dài hạn của doanh nghiệp. Trên thực tế khách hàng hiếm khi rời bỏ doanh nghiệp ngay lập tức trừ khi có những lí do đặc biệt nghiêm trọng. Quyết định dừng mua sản phẩm hay dịch vụ của một doanh nghiệp nào đó thường được hình thành qua thời gian, khi những bất mãn với sản phẩm của doanh nghiệp ngày càng nhiều hoặc đơn giản khi những sản phẩm của doanh nghiệp không còn đáp ứng được nhu cầu của họ so với những đối thủ khác có nhiều lựa chọn hơn. Việc dự đoán tỷ lệ rời bỏ ngay từ sớm sẽ giúp doanh nghiệp nhận ra những “dấu hiệu” không hài lòng của khách hàng và tiến hành các biện pháp khảo sát cụ thể hơn để nhanh chóng tìm ra nguyên nhân chính xác và xử lý kịp thời nhằm giảm tỷ lệ này tới mức tối thiểu. Bên cạnh đó việc dự đoán tỷ lệ rời bỏ cũng giúp doanh nghiệp có những chiến lược tốt hơn trong việc xây dựng và bảo vệ hình ảnh thương hiệu để không làm ảnh hưởng đến quá trình thu hút những khách hàng mới tiềm năng. Nếu không sớm phát hiện ra vấn đề thông qua việc dự đoán, những ý kiến bất mãn của khách hàng tại các điểm chạm sẽ tích tụ và có thể khiến hình ảnh doanh nghiệp trở nên xấu đi, thậm chí dẫn đến những nhận định sai lệch về doanh nghiệp.

Trong bài báo cáo này, chúng ta sẽ tiến hành phân tích dữ liệu khách hàng của công ty viễn thông Telco để xác định chân dung của khách hàng, dựa vào đó có thể đưa ra những biện pháp để giảm khả năng rời bỏ doanh nghiệp của khách hàng.

1. **GIỚI THIỆU TỔNG QUAN:**
2. **Giới thiệu bài báo cáo:**

Trong bối cảnh thị trường cạnh tranh khốc liệt hiện nay thì sự lựa chọn của khách hàng ngày càng đa dạng. Cùng với đó họ cũng có xu hướng ít trung thành hơn với các thương hiệu. Các báo cáo hành vi người tiêu dùng cho thấy khách hàng ngày càng thông minh và hiểu nhiều hơn về sản phẩm do sự phát triển của Internet. Mối quan tâm duy nhất của khách hàng không chỉ dừng lại ở việc chất lượng sản phẩm tốt. 60% khách hàng cho biết họ sẽ ưu tiên lựa chọn những thương hiệu đem lại trải nghiệm tuyệt vời trong quá trình trải nghiệm mua sắm của bản thân ở doanh nghiệp.

1. **Mục tiêu của bài báo cáo:**

- Phân tích tập dữ liệu để thấy được chân dung khách hàng đã đăng ký sử dụng dịch vụ của công ty.

- Tập trung vào nhóm khách hàng đã rời bỏ, xác định yếu tố quan trọng ảnh hưởng tới quyết định rời bỏ của khách hàng.

- Sử dụng các thuật toán học máy để dự báo khách hàng rời bỏ doanh nghiệp

Đưa ra các đề xuất để công ty xem xét, cải thiện các yếu tố giúp giữ chân khách hàng.

1. **Đối tượng và phạm vi bài báo cáo:**

- Nhà quản trị cấp tác nghiệp, cấp trung của doanh nghiệp viễn thông này.

- Bộ phận trong công ty: Chăm sóc khách hàng, sales, marketing, R&D.

1. **Phương pháp thực hiện bài báo cáo:**

- Sử dụng công cụ trực quan hóa tableau để phân tích dữ liệu khách hàng bằng các biểu đồ hay Dashboard, Storytelling.

- Sử dụng ngôn ngữ Python để triển khai mô hình dự đoán khách hàng rời bỏ doanh nghiệp bằng thuật toán Logistics Regression.

1. **CƠ SỞ LÝ THUYẾT:**
2. **Tỷ lệ khách hàng rời bỏ là gì?**

Tỷ lệ rời bỏ hay còn gọi là customer churn rate là tỷ lệ khách hàng ngừng mua bán những sản phẩm, dịch vụ của doanh nghiệp và do đó không còn là khách hàng của doanh nghiệp nữa. Tỷ lệ này thường được đo lường trong một khoảng thời gian nhất định, từ một tháng, một quý hoặc một năm tùy theo mục đích đo lường của doanh nghiệp.

1. **Tại sao khách hàng lại rời bỏ doanh nghiệp?**

Việc mang đến trải nghiệm khách hàng tốt nhất luôn là cách để xây dựng thương hiệu hiệu quả cho các doanh nghiệp. Do đó, nếu các sản phẩm, dịch vụ không thể đáp ứng được sự hài lòng của khách hàng thì chắc chắn họ sẽ rời đi và để lại những đánh giá tiêu cực. Điều này gây ra những ảnh hưởng không nhỏ tới uy tín và sự phát triển của doanh nghiệp.

Phản hồi khách hàng chậm chạp: Đối với hầu hết người tiêu dùng, nguyên nhân gây ra sự thất vọng và khiến khách hàng rời bỏ doanh nghiệp chính là không được cập nhật đủ thông tin từ nhân viên bán hàng hoặc phải chờ đợi được phản hồi từ doanh nghiệp quá lâu.

Không lắng nghe khách hàng: Hầu hết các doanh nghiệp để mất đi khách hàng là do doanh nghiệp đó không lắng nghe khách hàng. Bởi điều đó khiến doanh nghiệp không thể thấu hiểu khách hàng đồng thời không đáp ứng được nguyện vọng, mong muốn của họ.

Bỏ qua khách hàng cũ: Nhiều doanh nghiệp đôi khi quá tập trung vào việc thu hút các đối tượng khách hàng tiềm năng mới mà bỏ qua những tập khách hàng cũ đã từng mua sản phẩm dịch vụ của doanh nghiệp. Điều này sẽ khiến các khách hàng cũ chuyển sang một cửa hàng khác có thể thể hiện được sự quan tâm của họ dành cho khách hàng.

1. **Tại sao việc giữ chân khách hàng lại quan trọng?**

Giữ chân khách hàng rất quan trọng vì nó có thể chịu ảnh hưởng rất nhiều từ sự hài lòng của khách hàng. Ngoài ra, nó làm giảm nỗ lực cần thiết để tiếp cận khách hàng mới liên tục, từ đó có thể cải thiện hiệu quả quảng cáo của bạn và là một phần quan trọng trong chiến lược kinh tế của bạn.

Một chiến lược giữ chân khách hàng hiệu quả đảm bảo rằng khách hàng của bạn được cung cấp thông tin chính xác về thương hiệu và sản phẩm của bạn, hài lòng với sản phẩm đã mua và trải nghiệm khách hàng của họ, đồng thời cảm thấy thôi thúc muốn quay lại với thương hiệu của bạn để xem thêm sản phẩm hoặc thông tin cập nhật.

1. **Công cụ hỗ trợ**
2. **Python:**

- Python là một ngôn ngữ lập trình mã nguồn mở đa mục đích, có hỗ trợ các phong cách lập trình cốt lõi khác nhau như các mô hình hướng đối tượng và thủ tục mà triết lý thiết kế của nó rất thuận  lợi cho việc đọc hiểu code, đơn giản và rõ ràng.

- Đặc điểm nổi bật của python:

+ Dễ học và đọc: Python có cú pháp đơn giản và gần gũi với ngôn ngữ tự nhiên, giúp người mới học dễ dàng tiếp cận và hiểu code.

+ Đa năng: Python có thể được sử dụng cho nhiều mục đích khác nhau, từ phát triển web và phần mềm, đến xử lý dữ liệu và trí tuệ nhân tạo.

+ Thư viện phong phú: Python có một số thư viện phong phú cho hầu hết các loại công việc, giúp giảm thiểu thời gian phát triển và tăng hiệu suất.

+ Hỗ trợ đa nền tảng: Python có thể chạy trên nhiều hệ điều hành khác nhau như Windows, macOS và Linux.

+ Code ngắn gọn: Python thường cần ít dòng code hơn so với các ngôn ngữ lập trình khác để thực hiện cùng một nhiệm vụ.

- Trong bài toán này, tôi sẽ sử dụng các thư viện như pandas, numpy, sklearn và nhiều thư viện khác để xây dựng 1 mô hình dự đoán khả năng rời bỏ công ty của khách hàng.

1. **Tableau:**

- Tableau là phần mềm hỗ trợ phân tích và trực quan hóa dữ liệu (Data Visualization), được dùng nhiều trong ngành Business Intelligence. Phần mềm này tổng hợp các dữ liệu giống Excel nhưng ở một cấp độ cao hơn, thể hiện kết quả dưới dạng hình ảnh và biểu đồ trực quan một cách nhanh chóng và hiệu quả.

- Đặc điểm nổi bật của Tableau:

+ Quản lý siêu dữ liệu

+ Hỗ trợ phân tích dữ liệu với Big Data.

+ Dễ dàng xây dựng các Business Dashboard, khám phá dữ liệu hoạt động của doanh nghiệp.

+ Tính năng mô phỏng dữ liệu, xây dựng các tài liệu về dữ liệu hoạt động (Data Storytelling).

+ Phân tích dữ liệu của mạng xã hội.

1. **Thuật toán Logistic Regression**

Logistic Regression là 1 thuật toán phân loại được dùng để gán các đối tượng cho 1 tập hợp giá trị rời rạc (như 0, 1, 2, ...). Một ví dụ điển hình là phân loại Email, gồm có email công việc, email gia đình, email spam, ... Giao dịch trực tuyến có là an toàn hay không an toàn, khối u lành tính hay ác tình. Thuật toán trên dùng hàm sigmoid logistic để đưa ra đánh giá theo xác suất. Ví dụ: Khối u này 80% là lành tính, giao dịch này 90% là gian lận, …

Trong thống kê, mô hình logistic (hoặc mô hình logit) là một mô hình thống kê mô hình hóa xác suất của một sự kiện đang diễn ra bằng cách xác định tỷ lệ log cho sự kiện là sự kết hợp tuyến tính của một hoặc nhiều biến độc lập. Trong phân tích hồi quy, hồi quy logistic (hay hồi quy logit) là ước lượng các tham số của một mô hình logistic (các hệ số trong tổ hợp tuyến tính). Chính thức, trong hồi quy logistic nhị phân có một biến phụ thuộc nhị phân duy nhất, được mã hóa bởi một biến chỉ báo, trong đó hai giá trị được gắn nhãn là "0" và "1", trong khi mỗi biến độc lập có thể là một biến nhị phân (hai lớp, được mã hóa bởi một biến chỉ báo) hoặc một biến liên tục (bất kỳ giá trị thực nào). Xác suất tương ứng của giá trị được gắn nhãn "1" có thể thay đổi giữa 0 (chắc chắn là giá trị "0") và 1 (chắc chắn là giá trị "1"), do đó có việc dán nhãn; hàm chuyển đổi tỷ lệ log thành xác suất là hàm logistic, do đó có tên này Các biến nhị phân được sử dụng rộng rãi trong thống kê để lập mô hình xác suất của một lớp hoặc sự kiện nhất định đang diễn ra, chẳng hạn như xác suất chiến thắng của một đội, của một bệnh nhân khỏe mạnh, v.v.

1. **BÀI TOÁN.**
2. **Mục tiêu bài toán:**

- Xác định chân dung khách hàng, từ đó có thể giúp các bộ phận sales, marketing, R&D xác định đặc điểm và xu hướng hành vi của khách hàng, đưa ra chiến lược tiếp thị, bán hàng, khuyến mãi, phát triển dịch vụ phù hợp để bắt kịp xu hướng, thị hiếu của khách hàng.

- Theo dõi và phân tích tỷ lệ khách hàng rời bỏ hiện tại của công ty, để đưa ra các biện pháp giúp công ty giữ chân khách hàng và hạn chế những rủi ro không đáng có.

Xây dựng một mô hình dự báo để đưa ra những đặc điểm mà khách hàng có thể rời bỏ công ty hay không.

1. **Giới thiệu tổng quan về dữ liệu:**

- Nguồn dữ liệu: <https://www.kaggle.com/datasets/blastchar/telco-customer-churn>

- Dữ liệu TELCO CUSTOMER CHURN được lấy từ công ty viễn thông Telco. Bộ dữ liệu gồm có 21 cột và 7043 dòng dữ liệu, bao gồm các nội dung sau:

+ Khách hàng đã rời đi trong tháng trước – cột có tên là Churn

+ Các dịch vụ mà mỗi khách hàng đã đăng ký – điện thoại, nhiều đường dây, internet, bảo mật trực tuyến, sao lưu trực tuyến, bảo vệ thiết bị, hỗ trợ kỹ thuật và truyền phát trực tuyến TV và phim.

+ Thông tin tài khoản khách hàng – họ đã là khách hàng được bao lâu, hợp đồng, phương thức thanh toán, thanh toán không cần giấy tờ, phí hàng tháng và tổng phí.

+ Thông tin nhân khẩu học về khách hàng – giới tính, độ tuổi và liệu họ có đối tác và người phụ thuộc hay không

1. **Mô tả dữ liệu:**

|  |  |  |
| --- | --- | --- |
| **Mục** | **Tên cột** | **Ý nghĩa** |
| **Nhân khẩu học** | CustomerID | ID khách hàng |
|  | Gender | Giới tính (nam - 1, nữ - 0) |
|  | SeniorCitizen | Có phải là người cao tuổi không (>65t) (1 có - 0 không) |
|  | Partner | Có sống chung với vợ chồng hoặc người yêu không? |
|  | Dependents | Người phụ thuộc tài chính: con cái, ông bà,... (1 có - 0 không) |
| **Thông tin gói dịch vụ** | Tenure | Tổng số tháng khách hàng gắn bó với công ty |
|  | PhoneService | Khách hàng có đăng ký sử dụng dịch vụ điện thoại với công ty không  (1 có - 0 không) |
|  | MultipleLines | Khách hàng có đăng ký nhiều đường dây điện thoại với công ty không (1 có - 0 không hoặc không có dịch vụ điện thoại)  - Nếu phoneservice là không thì MultipleLines không có dịch vụ điện thoại  - Nếu phoneservice là có thì MultipleLines là có hoặc không |
|  | InternetService | Khách hàng có đăng ký dịch vụ internet của công ty không ( 0 - No, 1 - DSL, 2 - Fiber Optic) |
|  | OnlineSecurity | Có đăng ký dịch vụ bảo mật trực tuyến của công ty không (1 có - 0 không hoặc không có dịch vụ bảo mật trực tuyến)  - Nếu InternetService là không thì OnlineSecurity không có dịch vụ internet  - Nếu InternetService là có thì OnlineSecurity là có hoặc không |
|  | OnlineBackup | Có đăng ký dịch vụ sao lưu trực tuyến của công ty không (1 có - 0 không hoặc không có dịch vụ sao lưu trực tuyến)  - Nếu InternetService là không thì OnlineBackup không có dịch vụ internet  - Nếu InternetService là có thì OnlineBackup là có hoặc không |
|  | DeviceProtection | Có đăng ký dịch vụ bảo vệ thiết bị của công ty không (1 có - 0 không hoặc không có dịch vụ bảo vệ thiết bị)  - Nếu InternetService là không thì DeviceProtection là không có dịch vụ internet  - Nếu InternetService là có thì DeviceProtection là có hoặc không |
|  | TechSupport | Có đăng ký dịch vụ hỗ trợ kỹ thuật của công ty không (1 có - 0 không hoặc không có dịch vụ hỗ trợ kỹ thuật)  - Nếu InternetService là không thì TechSupport không có dịch vụ internet  - Nếu InternetService là có thì TechSupport là có hoặc không |
|  | StreamingTV | Khách hàng có sử dụng các chương trình tivi cung cấp bởi bên thứ 3 hay không(1 có - 0 không hoặc không dùng  dịch vụ internet để stream chương trình TV từ bên thứ ba)  - Nếu InternetService là không thì StreamingTV không có dịch vụ internet  -Nếu InternetService là có thì StreamingTV là có hoặc không |
|  | StreamingMovies | Khách hàng có sử dụng dịch vụ xem phim cung cấp bởi bên thứ 3 hay không (1 có - 0 không hoặc không dùng  dịch vụ internet để stream phim từ bên thứ ba)  - Nếu InternetService là không thì StreamingMovies không có dịch vụ internet  - Nếu InternetService là có thì StreamingMovies là có hoặc không |
| **Thông tin hợp đồng** | Contract | Thời hạn hợp đồng (0 month-to-month; 1 one Year; 2 two year) |
|  | PaperlessBilling | Có dùng hóa đơn điện tử không (1 có - 0 không) |
|  | PaymentMethod | Phương thức thanh toán (2 Electronic check; 3 Mailed check; 0 Bank transfer (automatic); 1 credit card (automatic) |
|  | MonthlyCharges | Tổng tiền hằng tháng khách hàng phải trả cho toàn bộ dịch vụ của công ty |
|  | TotalCharges | Tổng tiền khách hàng phải trả cho toàn bộ dịch vụ của công ty vào cuối quý được điều tra |
| **Churn** | Churn | Có churn hay không ( 1-có, 0-không) |

1. **TRỰC QUAN HÓA DỮ LIỆU**

A screenshot of a data report

Description automatically generated

Hình 1:Dashboard Tổng quan về bộ dữ liệu (Overview)

A screenshot of a computer

Description automatically generated

Hình 2: Dashboard về Tỷ lệ khách hàng rời bỏ công ty Telco

1. **Tổng quan về bộ dữ liệu:**

A blue text on a white background

Description automatically generated

Hình 3: Tổng quan về bộ dữ liệu.

Trong bộ dữ liệu Telco Churn, về tổng quan thì bộ dữ liệu này sẽ có 7043 khách hàng đang sử dụng dịch vụ của công ty và số tháng mà khách hàng gắn bó với công ty rời và khoảng 32 (~2.6 năm). Khoảng chi phí trung bình trong một tháng mà khách hàng phải trả khi sử dụng tất cả các dịch vụ ở công ty là 456.12 nghìn, sau khi tổng kết tất cả các dịch vụ mà khách hàng đang sử dụng của công ty cuối quý thì số tiền mà khách hàng phải trả cho công ty là 16.06 triệu.

1. **Tổng quan về khách hàng:**
2. **Thông tin về nhân khẩu học:**

A screenshot of a graph

Description automatically generated

Hình 4: Thông tin về nhân khẩu học

Đối với những thông tin về nhân khẩu học của công ty Telco cho thấy, tỷ lệ giới tính giữa Nam và Nữ trong bộ dữ liệu này ở mức cân bằng tốt xấp xỉ 50-50. Bên cạnh đó, 3 yếu tố về nhân khẩu học là Tỷ lệ người cao tuổi, Tỷ lệ khách hàng đang sống cùng vợ chồng và Tỷ lệ khách hàng phụ thuộc tài chính được thể hiện rõ thông qua các con số. Cụ thể như sau:

* Đối với Tỷ lệ khách hàng là người cao tuổi chỉ chiếm 16,21% so với tổng số khách hàng hiện tại đang sử dụng dịch vụ của công ty, đây là một con số khá nhỏ cho thấy những người cao tuổi (>65 tuổi) thường ít hoặc không sử dụng dịch về công nghệ viễn thông, điều này có thể là do họ là người lớn tuổi không thích hoặc không rành về công nghệ nên tỷ lệ người cao tuổi sử dụng dịch vụ viễn thông ít hơn nhóm người trẻ.
* Đối với Tỷ lệ khách hàng đang sống cùng với vợ/chồng có tỷ lệ sử dụng dịch vụ viễn thông của công ty ở mức trung bình là 48.30% xấp xỉ 50%. Điều này cho thấy những khách hàng có sống chung với vợ/chồng hay là sống một mình đều có nhu cầu sử dụng dịch vụ công nghệ viễn thông ở mức gần như nhau vì sự tiện ích của nó.
* Cuối cùng, Tỷ lệ khách hàng có người phụ thuộc tài chính chỉ chiếm khoảng ⅓ trên tổng số khách hàng sử dụng dịch vụ của công ty với con số là 29.96%. Có thể thấy rằng trên tổng số khách hàng đang sử dụng dịch vụ của công ty thì những khách hàng không bị ràng buộc bởi gia đình, con cái,.. thì họ tự do tài chính hơn nên họ sử dụng các dịch vụ công nghệ viễn thông nhiều hơn so với những khách hàng có người phụ thuộc về tài chính.

1. **Thông tin về gói dịch vụ:**
   1. **Dịch vụ điện thoại (PhoneService):**

A graph with a number of blue squares

Description automatically generated with medium confidence

Hình 5: Số khách hàng sử dụng dịch vụ điện thoại

Trong tổng số khách hàng sử dụng dịch vụ tại công ty, có tới 6361 khách hàng sử dụng dịch vụ điện thoại Phoneservice chiếm khoảng 90% trên tổng số khách hàng. Dựa vào đây có thể thấy dịch vụ điện thoại hiện tại đang khá phổ biến nên đa số khách hàng đều sử dụng dịch vụ điện thoại để có thể nghe gọi với mọi người, điều này rất tiện cho việc liên hệ xa hay phục vụ cho công việc.

* 1. **Đường dây điện thoại (Multiple Line):**

**A graph with blue squares

Description automatically generated with medium confidence**

Hình 6: Số lượng khách hàng sử dụng dịch vụ đường dây điện thoại.

Có 2971 khách hàng sử dụng dịch vụ nhiều đường dây điện thoại và chiếm khoản 42.18% so với tổng số khách hàng. Khi khách hàng sử dụng nhiều đường dây điện thoại thì khách hàng có thể thực hiện hoặc nhận nhiều cuộc gọi cùng một lúc, điều này khá tiện lợi cho việc nghe gọi cho nhiều người nên có khá nhiều khách hàng chọn dịch vụ này. Ngược lại thì có 3390 khách hàng không sử dụng dịch vụ này và chiếm tỷ lệ 48.13% trên tổng số khách hàng. Có thể nhiều khách hàng cho rằng, nghe gọi nhiều cuộc cùng 1 lúc không cần thiết với họ nên họ sẽ không sử dụng dịch vụ này. Ngoài ra, có 682 không sử dụng dịch vụ điện thoại nên sẽ không sử dụng dịch vụ đường dây điện thoại.

* 1. **Dịch vụ Internet (Internet Service):**

A blue circle with numbers and text

Description automatically generated

Hình 7: Phần trăm khách hàng sử dụng dịch vụ internet.

Đối với dịch vụ Internet, tỷ lệ khách hàng sử dụng internet bằng sợi quang học (Fiber optic) nhiều hơn internet bằng đường dây thuê bao kỹ thuật số (DSL) là 9,59%. Lý do vì sao có sự chênh lệch này? Có thể là do khi sử dụng dịch vụ Internet bằng sợi quang học thì tốc độ truyền dẫn dữ liệu sẽ cao hơn đáng kể so với khi sử dụng DSL. Hoặc là dịch vụ bằng sợi cáp quang sẽ có dung lượng truy cập lớn hơn nhiều so với dịch vụ DSL. Ngoài ra còn có 21.67% số khách hàng không sử dụng dịch vụ internet của công ty có thể là do giá cả cao hoặc khách hàng không thích các dịch vụ internet mà công ty đang sở hữu.

* 1. **Dịch vụ bảo mật trực tuyến (Online Security):**

A graph of a security system

Description automatically generated with medium confidence

Hình 8: Số lượng khách hàng sử dụng dịch vụ Bảo mật trực tuyến.

Số lượng khách hàng có tham gia dịch vụ bảo mật trực tuyến của công ty là 2019 khách hàng thấp hơn nhiều so với số khách hàng không tham gia dịch vụ bảo mật 1479 khách hàng. Lý do dẫn đến sự chênh lệch này có thể là do dịch vụ này của công ty có giá khá cao và dịch vụ này của công ty chưa đảm bảo được mức độ bảo mật mà khách hàng mong muốn. Và cuối cùng là có 1526 khách hàng không sử dụng dịch vụ internet nên họ cũng sẽ không dùng dịch vụ bảo mật trực tuyến. Nếu công ty muốn thu hút nhiều khách hàng sử dụng dịch vụ bảo mật thì công ty cần nâng cấp hệ thống bảo mật của công ty được mạnh mẽ và đáng tin cậy hơn hoặc đưa ra các chính sách ưu đãi và khuyến mãi, tăng cường dịch vụ hỗ trợ khách hàng.

* 1. **Dịch vụ sao lưu trực tuyến (Online Backup):**

A screenshot of a data analysis

Description automatically generated with medium confidence

Hình 9: Số lượng khách hàng sử dụng dịch vụ Sao lưu trực tuyến.

Có 2429 khách hàng tham gia dịch vụ sao lưu tài chính ít hơn so với số khách hàng không tham gia dịch vụ này (3088 khách hàng) là 659 khách hàng. Và đối với dịch vụ này cũng có 1526 khách hàng không tham gia dịch vụ internet nên cũng không tham gia dịch vụ sao lưu.

* 1. **Dịch vụ bảo vệ thiết bị (Device Protection):**

A screenshot of a graph

Description automatically generated**.**

Hình 10: Số lượng khách hàng sử dụng dịch vụ Bảo vệ thiết bị.

Có 2422 khách hàng sử dụng dịch vụ bảo vệ thiết bị và 3095 khách hàng không sử dụng, với dịch vụ này cũng có sự chênh lệch giữa khách hàng dùng dịch vụ và không dùng cũng khá là cao. Có thể là do khách hàng chưa rõ về dịch vụ này hoặc giá cả cao, vấn đề bảo mật chưa tốt hoặc quy trình phức tạp khiến cho khách hàng chưa vừa ý với dịch vụ và quyết định không sử dụng.

* 1. **Dịch vụ hỗ trợ kỹ thuật (Tech Support):**

A blue squares with white text

Description automatically generated

Hình 11: Số lượng khách hàng sử dụng dịch vụ Hỗ trợ kỹ thuật.

Với dịch vụ này cũng giống với các dịch vụ trên, số lượng khách hàng dùng dịch vụ có sự chênh lệch khá cao so với khách hàng không tham gia dịch vụ. Số lượng khách hàng sử dụng dịch vụ này chỉ có 2044 khách hàng, trong khi đó khách hàng không sử dụng dịch vụ có tới 3473 khách hàng. Có thể thấy công ty chưa thật sự cung cấp các dịch vụ kỹ thuật tốt để người dùng cảm thấy yên tâm khi sử dụng dịch vụ của công ty. Công ty nên có những thay đổi tích cực hơn về hệ thống, hệ thống bảo mật hay các kỹ thuật công nghệ để thu hút khách hàng nhiều hơn.

* 1. **Streaming TV:**

A graph of blue and green squares

Description automatically generated

Hình 12: Số lượng khách hàng sử dụng dịch vụ Streaming TV.

Dịch vụ xem chương trình TV được khách hàng sử dụng khá ổn định, tỷ lệ khách hàng sử dụng và khách hàng không sử dụng là xấp xỉ nhau 50-50. Có thể khách hàng ưa chuộng các dịch vụ giải trí hơn nên đăng ký dịch vụ xem chương trình TV cũng nhiều hơn các dịch vụ khác.

* 1. **Streaming Movies:**

A graph of a number of blue squares

Description automatically generated with medium confidence

Hình 13: Số lượng khách hàng sử dụng dịch vụ Streaming Movies.

Dịch vụ xem phim từ bên thứ 3 cũng giống với dịch vụ xem chương trình TV, dịch vụ này có số khách hàng sử dụng gần bằng với số khách hàng không sử dụng. Có vẻ khách hàng ưu tiên chọn các dịch vụ giải trí hơn là các dịch vụ về bảo mật hay về kỹ thuật.

* 1. **Tổng kết:**

**A group of blue text

Description automatically generated**

Hình 14: Tổng kết tỷ lệ khách hàng sử dụng các dịch vụ Internet.

Số khách hàng sử dụng Internet Service chiếm 78.33% trên tổng số khách hàng, trong đó khách hàng sử dụng các dịch vụ:

+ Online Backup chiếm 34.49%

+ Online Security chiếm 28.67%

+ Streaming Movies chiếm 38.79%

+ Streaming TV chiếm 38.44%

+ Tech Support chiếm 29.02%

+ Device Protection chiếm 34.39%

⇒Tuy số lượng khách hàng sử dụng dịch vụ Internet Service khá cao chiếm tới 78.33% nhưng mà các dịch vụ khác liên quan đến dịch vụ này lại chiếm tỷ lệ khác ít chỉ nằm trong khoảng 20-40%. Có thể thấy công ty chưa thực sự thành công trong việc thu hút khách hàng tham gia các dịch vụ của công ty.

1. **Thông tin về hợp đồng của khách hàng:**
   1. **Thời hạn hợp đồng (Contract):**

**A blue rectangles with white text

Description automatically generated**

Hình 15: Tỷ lệ khách hàng sử dụng các loại hợp đồng.

Phần lớn khách hàng sử dụng dịch vụ tại công ty ưa chuộng việc cam kết hợp đồng theo từng tháng, với tỉ lệ khá cao khoảng 55,02% trên tổng số khách hàng. Đối với loại hợp đồng kéo dài 1 năm và 2 năm, tỷ lệ khách hàng lựa chọn để ký kết lần lượt là 20,91% và 24,07% và dễ dàng nhận thấy tỷ lệ khách hàng đăng ký hợp đồng 1 năm có tỉ lệ thấp nhất.

* 1. **Hóa đơn điện tử (Paperless Billing):**

**A blue circle with white text

Description automatically generated**

Hình 16: Tỷ lệ khách hàng sử dụng hóa đơn điện tử.

Trong tổng số khách hàng, phần trăm khách hàng có sử dụng hóa đơn điện tử cao nhất là 59,22%. Có thể thấy khách hàng ưa chuộng hóa đơn điện tử hơn là hóa đơn giấy hay các loại khác. Công ty nên phát triển loại hóa đơn này hơn.

* 1. **Phương thức thanh toán (Payment Method):**

**A blue rectangles with numbers and a few black arrows

Description automatically generated**

Hình 17: Tỷ lệ khách hàng sử dụng các phương thức thanh toán.

Tỷ lệ khách hàng sử dụng phương thức thanh toán bằng séc điện tử (electronic check) là cao nhất (chiếm 33,58%), tiếp đến là khách hàng sử dụng séc qua mail (chiếm 22,89%), kế tiếp là khách hàng sử dụng phương thức thanh toán tự động qua chuyển khoản ngân hàng (chiếm 21,92%) và thấp nhất là khách hàng sử dụng thanh toán qua thẻ tín dụng (chiếm 21,61%).

1. **Tỷ lệ khách hàng rời bỏ công ty Telco:**
2. **Churn Rate:**

A white background with blue numbers

Description automatically generated

Hình 18: Số lượng khách hàng rời bỏ công ty.

A number and text on a white background

Description automatically generated

Hình 19: Tỷ lệ khách hàng rời bỏ công ty.

Trong 7034 khách hàng sử dụng dịch vụ của công ty, có 1869 khách hàng đã rời bỏ công ty và chiếm 26,54% trên tổng số khách hàng. Có thể thấy, đối với một doanh nghiệp thì tỷ lệ khách hàng rời bỏ công ty với con số 26,54% là rất lớn. Vì vậy, công ty cần phải đưa ra các biện pháp giữ chân khách hàng ở lại với công ty.

1. **Churn Rate Gender:**

**A screenshot of a graph

Description automatically generated**

Hình 20: Tỷ lệ khách hàng rời bỏ doanh nghiệp theo giới tính.

Tỷ lệ rời bỏ của công ty giữa khách hàng nam và khách hàng nữ gần như bằng nhau, không có sự chênh lệ nào giữa 2 nhóm khách hàng này cho thấy rằng giới tính gần như không ảnh hưởng gì nhiều đến sự rời bỏ của khách hàng.

1. **Churn Rate Dependents:**

**A graph with blue squares and numbers

Description automatically generated**

Hình 21: Tỷ lệ khách hàng có người phụ thuộc rời bỏ công ty.

Dựa vào biểu đồ trên ta có thể thấy, những khách hàng có người phụ thuộc thì thường có xu hướng là không rời bỏ công ty nên ta có thể thấy tỷ lệ rời bỏ của nhóm khách hàng này thấp hơn nhóm khách hàng không phụ thuộc với con số là 15,45%. Dựa vào điều này, ta có thể khẳng định được xu hướng tiêu dùng dịch vụ viễn thông của nhóm những người có gia đình, sống chung với vợ chồng, con cái hoặc cha mẹ thì họ thường thích sự cố định, ít thay đổi dịch vụ hoặc có thể dùng giống nhau đối với tất cả các thành viên trong nhà. Công ty cần phải nắm bắt thói quen, xu hướng này để chăm sóc, giữ chân họ cũng như đưa ra những gói khuyến mãi phù hợp. Còn những khách hàng không có người phụ thuộc thì tỷ lệ rời bỏ công ty của nhóm khách hàng này khá là cao chiếm tới 31,28%. Vì vậy, đối với nhóm khách hàng rời bỏ doanh nghiệp, thì ta dễ dàng nhận thấy họ ít có người phụ thuộc, sống tự do, không ràng buộc nên thường có xu hướng dễ thay đổi hơn.

1. **Churn Rate SeniorCitizen:**

**A blue rectangles with black text

Description automatically generated**

Hình 22: Tỷ lệ rời bỏ công ty của người lớn tuổi.

Tỷ lệ rời bỏ công ty của nhóm khách hàng có độ tuổi trên 65 tuổi chiếm 41,68% cao hơn nhiều so với nhóm khách hàng dưới 65 tuổi có tỷ lệ rời bỏ là 23,61%. Có thể hiểu rằng, đối với nhóm khách hàng trên 65 tuổi thì họ có thể ít linh hoạt hơn trong việc thích nghi và sử dụng các công nghệ mới. Họ có thể gặp khó khăn trong việc hiểu và sử dụng các tính năng mới hoặc các ứng dụng công nghệ mới. Hay là những người lớn tuổi họ thường không có sự quan tâm về internet hay dịch vụ viễn thông nhiều như những người trẻ tuổi hoặc là họ sống với mức lương hưu không được cao nên nhu cầu sử dụng dịch vụ viễn thông không cao lắm. Ngược lại với nhóm khách hàng này là nhóm khách hàng trẻ dưới 65 tuổi, vì họ còn trẻ nên họ có thể linh hoạt hơn khi sử dụng các dịch vụ về công nghệ, họ có tiền lương cao hơn để sử dụng các dịch vụ viễn thông cao cấp hơn, hay là công việc của họ rất cần đến dịch vụ viễn thông. Vậy nên dựa vào những điều kiện trên có thể thấy khách hàng có độ tuổi trên 65 sẽ có tỷ lệ rời bỏ cao hơn những khách hàng trẻ dưới 65 tuổi. Công ty nên có các biện pháp cải thiện trải nghiệm người dùng của khách hàng trên 65 tuổi bằng cách đơn giản hóa giao diện, cung cấp hướng dẫn và hỗ trợ dễ hiểu, và tối ưu hóa tính khả dụng của dịch vụ cho người cao tuổi. Tăng cường dịch vụ hỗ trợ khách hàng để giúp người cao tuổi giải quyết mọi vấn đề hoặc thắc mắc một cách nhanh chóng và hiệu quả. Và đưa ra các gói dịch vụ có giá cả phù hợp với người lớn tuổi có lương hưu không được cao để nhóm khách hàng này có thể lựa chọn những dịch vụ phù hợp và hạn chế tỷ lệ rời bỏ cao của nhóm này.

1. **Churn Rate Partner:**

**A blue rectangles with black text

Description automatically generated**

Hình 23: Tỷ lệ rời bỏ công ty của khách hàng có vợ/chồng.

Nhóm khách hàng có tỷ lệ rời bỏ cao nhất đó là nhóm khách hàng không sống chung với vợ/chồng chiếm tới 32,96% cao hơn nhiều so với nhóm người đã kết hôn sống chung với vợ/chồng tới 13,3%. Có thể thấy rằng nhóm người đã kết hôn thường có sự ổn định gia đình hơn, và do đó, họ có thể ít có xu hướng thay đổi địa điểm hoặc dịch vụ viễn thông. Hay công ty thưởng có các gói dịch vụ hoặc ưu đãi đặc biệt dành cho khách hàng đã kết hôn, giúp họ cảm thấy hài lòng hơn và duy trì mối quan hệ dài hạn với công ty. Ngược lại, nhóm khách hàng không sống cùng vợ/chồng thường tìm kiếm nhiều sự mới mẻ hơn, họ thường tìm những dịch vụ mới và cao cấp hơn để đáp ứng nhu cầu của họ nên khi dịch vụ công ty chưa đủ mới mẻ hay chưa đáp ứng nhu cầu của họ thì họ sẽ sẵn sàng rời bỏ công ty để đi tìm những nói đáp ứng đủ nhu cầu đó. Hay là nhóm khách hàng này chưa có nguồn tài chính ổn định mà công ty chưa có những ưu đãi tối ưu cho nhóm này nên họ cũng sẵn sàng rời bỏ công ty vì không lợi ích chưa được tối ưu. Để giảm tỷ lệ rời bỏ trong nhóm khách hàng chưa kết hôn, công ty có thể cần tập trung vào cung cấp các dịch vụ và ưu đãi đặc biệt phù hợp với nhu cầu đặc thù của họ. Ngoài ra, tăng cường cam kết và tạo ra các chương trình khuyến mãi có thể giúp thu hút và giữ chân khách hàng trong thời gian dài hơn.

1. **Churn Rate Phone Service/Multiple Lines:**

**A screenshot of a graph

Description automatically generated**

Hình 24: Tỷ lệ khách hàng rời bỏ theo PhoneService và MultipleLines

**Phone Service:** Tỷ lệ khách hàng rời bỏ công ty khi có sử dụng dịch vụ điện thoại của công ty là 53,65% cao hơn khoảng gấp 2 lần so với khách hàng không sử dụng dịch vụ điện thoại. Dịch vụ phone service thường được xem là một phần thiết yếu của cuộc sống hàng ngày, đặc biệt đối với những người dùng có nhu cầu liên lạc nhiều hoặc cần sự kết nối liên tục. Do đó, nhu cầu sử dụng và yêu cầu về chất lượng dịch vụ phone service có thể cao hơn, và bất kỳ vấn đề hoặc không hài lòng nào về dịch vụ có thể dẫn đến sự rời bỏ nhanh chóng. Khi mà công ty không đáp ứng được yêu cầu về chất lượng dịch vụ hoặc hỗ trợ khách hàng đối với dịch vụ phone service. Nếu khách hàng gặp phải vấn đề về chất lượng cuộc gọi, tính khả dụng của dịch vụ, hoặc không hài lòng với hỗ trợ khách hàng, họ có thể quyết định chuyển đổi sang các nhà cung cấp dịch vụ khác. Khi đó doanh nghiệp cần tập trung vào cải thiện chất lượng dịch vụ phone service, bao gồm cả chất lượng cuộc gọi, tính khả dụng và tốc độ phản hồi của hệ thống. Hoặc tạo ra các chương trình khuyến mãi và ưu đãi đặc biệt cho khách hàng sử dụng dịch vụ phone service, như giảm giá, tặng kèm, hoặc dịch vụ miễn phí trong một khoảng thời gian nhất định để có thể giữ chân khách hàng tránh tỷ lệ khách hàng rời bỏ cao.

**Multiple Lines:** Tỷ lệ khách hàng rời bỏ công ty khi sử dụng dịch vụ Multiple Lines là 28,61% và không sử dụng là 25,04%. Dịch vụ Multiple Lines có thể tăng chi phí hàng tháng cho khách hàng nên một khi khách hàng cảm thấy chi phí này không xứng đáng hoặc không được bù đắp bởi giá trị mà họ nhận được từ dịch vụ, họ có thể quyết định rời bỏ công ty. Nếu dịch vụ Multiple Lines không đáp ứng được mong đợi của khách hàng về chất lượng cuộc gọi, tính khả dụng của dịch vụ, hoặc hỗ trợ khách hàng không đủ tốt, họ có thể cảm thấy không hài lòng và quyết định chuyển đổi sang các nhà cung cấp dịch vụ khác. Trong trường hợp này, công ty nên tập trung vào cải thiện giá trị và tiện ích của dịch vụ Multiple Lines bằng cách cung cấp các tính năng mới, tối ưu hóa chất lượng cuộc gọi, và tăng cường tính khả dụng của dịch vụ. Công ty có thể cung cấp thêm các gói dịch vụ Multiple Lines được tùy chỉnh và phân loại dựa trên nhu cầu cụ thể của từng khách hàng.

1. **Tỷ lệ khách hàng rời bỏ (churn rate) theo dịch vụ Internet (Internet Services) và các dịch vụ đi kèm (Sub Internet Services):**
   1. **Internet Service:**

Tỷ lệ khách hàng rời bỏ công ty của khách hàng có dùng dịch vụ Internet services cao hơn nhiều so với khách hàng không sử dụng dịch vụ này (7.4%). Như vậy, dịch vụ Internet services đang có vấn đề khá lớn khi những người sử dụng dịch vụ này không tiếp tục gắn bó với công ty.

Mặc dù có nhiều khách hàng sử dụng dịch vụ Internet bằng sợi quang học (fiber optic) hơn bằng đường dây thuê bao kỹ thuật số (DSL), nhưng nhóm này cũng có tỷ lệ khách hàng rời bỏ công ty cũng nhiều hơn hẳn. Như vậy, trong 2 lựa chọn của dịch vụ Internet, dịch vụ Internet bằng fiber optic đang có vấn đề lớn hơn khi gây ra tỷ lệ rời bỏ của khách hàng cao.

* 1. **Online Backup:**

**A screenshot of a service

Description automatically generated**

Hình 25: Tỷ lệ rời bỏ công ty của khách hàng sử dụng dịch vụ OnlineBackup.

Đối với dịch vụ Online Backup thì nhóm khách hàng không sử dụng các dịch vụ internet bằng DSL và Fiber optic có tỷ lệ rời bỏ cao hơn nhóm khách hàng sử dụng 2 loại dịch vụ internet này với con số lần lượt là 25,69% và 50,77%, con số này được đánh giá là rất cao, cao hơn khoảng hơn 2 lần so với khách hàng có sử dụng dịch vụ đi kèm này (10.68% và 30,31%).

Ngoài ra, khách hàng sử dụng dịch vụ Internet bằng cáp quang có sử dụng dịch vụ sao lưu trực tuyến cũng có tỷ lệ rời bỏ công ty khá cao (30.31%), thậm chí còn cao hơn cả khách hàng dùng dịch vụ DSL nhưng không dùng dịch vụ sao lưu trực tuyến (25,69%). Điều này cho thấy rằng dịch vụ sao lưu trực tuyến dành cho nhóm khách hàng sử dụng dịch vụ Internet bằng cáp quang có thể đang gặp vấn đề. Nguyên nhân dẫn đến vấn đề này có thể là do dịch vụ của công ty không ổn định, có thể gây ra sự không hài lòng từ phía khách hàng, đặc biệt là trong việc sao lưu dữ liệu quan trọng.

Khi khách hàng bỏ tiền để mua dịch vụ tốt nhất, hài lòng nhất mà công ty không đáp ứng được thì sẽ dẫn đến tình trạng tỷ lệ khách hàng rời bỏ công ty ngày càng cao. Công ty nên cung cấp dịch vụ sao lưu tốt hơn, bảo mật mạnh hơn và nâng cấp hệ thống cao cấp hơn để phục vụ được nhu cầu của khách hàng.

* 1. **Online Security:**

**A screenshot of a computer

Description automatically generated**

Hình 26: Tỷ lệ rời bỏ công ty của khách hàng sử dụng dịch vụ OnlineSecurity.

Tỷ lệ khách hàng rời bỏ công ty khi không sử dụng dịch vụ bảo mật trực tuyến lần lượt là 27,96% và 49,36% (DSL và Fiber optic), tỷ lệ này đang ở mức khá cao, cao gấp 2 lần so với tỷ lệ khách hàng sử dụng dịch vụ này (9,49% và 21,81%). Có thể thấy khi khách hàng sử dụng đồng thời cả dịch vụ Internet và dịch vụ bảo mật trực tuyến thì tỷ lệ khách hàng rời bỏ ở mức khá an toàn. Tuy nhiên, riêng khách hàng sử dụng dịch vụ bảo mật trực tuyến của Internet cáp quang thì tỷ lệ rời bỏ của nhóm này cũng đang ở mức khá là cao tới 21,81% cho thấy dịch vụ bảo mật trực tuyến của nhóm này đang gặp vấn đề. Vấn đề này có thể là do dịch vụ bảo mật trực tuyến của công ty không đáp ứng được nhu cầu của khách hàng, bảo mật dữ liệu kém, chi phí đắt đỏ dẫn đến khách hàng không hài lòng với dịch vụ của công ty.

* 1. **Streaming Movies:**

**A screenshot of a service

Description automatically generated**

Hình 27: Tỷ lệ rời bỏ công ty của khách hàng sử dụng dịch vụ StreamingMovies.

Đối với các khách hàng không sử dụng dịch vụ internet để Stream phim từ bên thứ ba Streaming Movies, tỉ lệ rời bỏ lần lượt là 22.36% và 45.80% (DSL và Fiber Optic) và ngay cả khi có sử dụng dịch vụ, tỉ lệ cũng lần lượt là 13.97% và 38.89%. Với tỷ lệ cao như vậy có thể thấy dù khách hàng có sử dụng dịch vụ này hay không thì tỷ lệ khách hàng rời bỏ vẫn rất cao. Có thể thấy dịch vụ Streaming Movies khi chọn Fiber optic đang có chất lượng khá thấp khi mà ngay cả khi khách hàng có đăng ký dịch vụ hay là không thì tỉ lệ khách hàng ngừng sử dụng dịch vụ đều đạt tỉ lệ khá đáng kể.

* 1. **Streaming TV:**

**A screenshot of a computer

Description automatically generated**

Hình 28: Tỷ lệ rời bỏ công ty của khách hàng sử dụng dịch vụ StreamingTV.

Tỷ lệ rời bỏ công ty của các khách hàng kể cả khi không sử dụng dịch vụ Streaming TV hay có sử dụng dịch vụ này để có con số cao ngất, cụ thể đối với khách hàng có sử dụng dịch vụ Streaming TV có tỷ lệ rời bỏ lần lượt là 13,27% và 39,26% (DSL và Fiber optic), đối với khách hàng không sử dụng dịch vụ Streaming TV thì tỷ lệ khách hàng rời bỏ lần lượt là 22,68% và 45,32%. Cho nên, khi tỷ lệ rời bỏ của 1 dịch vụ của công ty kể cả khi khách hàng có sử dụng dịch vụ đó hay không cao như vậy có nghĩa là dịch vụ Streaming TV mà công ty đang sở hữu gặp vấn đề nghiêm trọng. Có thể là về chất lượng đường truyền, chất lượng internet, giá cá khá cao dẫn đến khách hàng không hài lòng với dịch vụ và cuối cùng sẽ lựa chọn từ bỏ doanh nghiệp để chọn nơi có dịch vụ tốt hơn. Công ty cần điều chỉnh lại chất lượng của dịch vụ, chất lượng đường truyền internet và có thể đưa ra những chương trình khuyến mãi để giữ chân khách hàng ở lại công ty.

* 1. **Tech Support:**

**A screenshot of a computer

Description automatically generated**

Hình 29: Tỷ lệ rời bỏ công ty của khách hàng sử dụng dịch vụ TechSupport.

Đối với các khách hàng không sử dụng dịch vụ Tech Support, tỉ lệ rời bỏ lần lượt là 27.76% và 49.37% ( DSL và Fiber optic) cao hơn rất nhiều so với khách hàng có sử dụng dịch vụ đi kèm này (9.68% và 22.63%). Theo số liệu cho thấy, dịch vụ Tech Support đang không thu hút được khách hàng, tỉ lệ khách hàng rời bỏ cao, công ty nên có các biện pháp cải thiện dịch vụ để thu hút và giữ chân được khách hàng.

Khách hàng sử dụng dịch vụ tech support bằng sợi quang học (fiber optic) có tỷ lệ rời bỏ công ty nhiều hơn so với khách hàng dùng dịch vụ Tech support bằng đường dây thuê bao kỹ thuật số (DSL). Trong 2 dịch vụ Internet, dịch vụ Internet bằng fiber optic đang có vấn đề lớn hơn dịch vụ Internet bằng DSL khi mà tỉ lệ rời bỏ công ty cao hơn rất nhiều.

* 1. **Device Protection:**

**A screenshot of a computer

Description automatically generated**

Hình 30: Tỷ lệ rời bỏ công ty của khách hàng sử dụng dịch vụ DeviceProtection

Về dịch vụ đăng ký bảo vệ thiết bị, đối với khách hàng có đăng ký sử dụng bởi hình thức DSL có tỷ lệ khá tốt, chỉ khoảng 10% tỷ lệ khách hàng sẽ rời bỏ. Tuy nhiên đối với dịch vụ sử dụng sợi quang học, dù có hay không đăng ký dịch vụ thì tỉ lệ khách hàng rời bỏ rất cao (lần lượt là 31,54% và 49,97%). Thậm chí kể cả khi khách hàng đăng ký dịch vụ Tech Support thì tỷ lệ vẫn cao hơn so với tỉ lệ khách hàng rời bỏ khi không đăng ký dịch vụ sử dụng hình thức DSL là 25,22%. Dù khách hàng có đăng ký dịch vụ bảo vệ thiết bị hay không, tỉ lệ khách hàng rời bỏ đối với hình thức sợi quang học khá cao, nằm ở mức 31.54% đối với khách hàng có sử dụng và xấp xỉ 50% đối với khách hàng không sử dụng.

1. **Churn Rate Tenure:**

**A graph of a number of blue and white bars

Description automatically generated**

Hình 31: Tỷ lệ khách hàng rời bỏ công ty theo số tháng gắn bó.

Dựa vào biểu đồ, ta có thể thấy khi số tháng mà khách hàng gắn bó với công ty là 0 thì cũng đồng nghĩa với việc là tỷ lệ khách hàng rời bỏ công ty cao nhất với con số là 54,93%. Khi mà số tháng khách hàng gắn bó với công ty ngày càng cao thì tỷ lệ khách hàng rời bỏ càng thấp. Trong biểu đồ cho thấy số tháng khách hàng gắn bó với công ty bằng 0 là thấp nhất và số tháng mà khách hàng gắn bó với công ty lớn nhất là 69 tháng, đây cũng là số tháng mà có tỷ lệ rời bỏ thấp nhất chỉ có 3.53%. Như vậy có thể rút ra rằng, khi mà khách hàng càng gắn bó lâu với công ty thì khách hàng khó có thể rời bỏ công ty, còn khi khách hàng gắn bó với công ty càng ít thì khách hàng càng dễ rời bỏ công ty tìm một công ty mới lý tưởng hơn. Cho nên, công ty nên đưa ra các biện pháp giữ chân khách hàng, để khách hàng gắn bó với công ty càng lâu càng tốt.

1. **Churn Rate Paperless Billing & Payment Method:**

**A graph of a bar chart

Description automatically generated with medium confidence**

Hình 32: Tỷ lệ khách hàng rời bỏ công ty theo Paperless Billing & Payment Method.

Về việc phân tích tỷ lệ rời bỏ khách hàng theo hóa đơn điện tử và phương thức thanh toán Paperless Billing & Payment Method:

* Tỷ lệ rời bỏ khách hàng theo hóa đơn điện tử Paperless Billing:từ biểu đồ cho thấy khách hàng có sử dụng hóa đơn điện tử Paperless Billing cao hơn khách hàng không sử dụng hóa đơn điện tử Paperless Billing (Yes > No). Xét theo trung bình tổng tiền hằng tháng khách hàng chi trả cho toàn bộ dịch vụ của công ty Avg. Monthly Charges, khách hàng có sử dụng dịch vụ Paperless Billing đã phải chi trả một khoản phí cao hơn.
* Tỷ lệ rời bỏ khách hàng theo phương thức thanh toán Payment Method: Từ biểu đồ cho thấy phần lớn khách hàng sử dụng phương thức thanh toán séc điện tử Electronic check, tiếp đó là Mailed check, chuyển khoản ngân hàng Bank transfer (automatic) và thẻ tín dụng Credit card (automatic). Khách hàng thanh toán theo séc điện tử Electronic check chiếm tỷ lệ cao nhất đồng thời cũng là nhóm khách hàng có tỷ lệ churn cao nhất lên đến 50%, tiếp đến là Mailed check 27%, Bank transfer 21% và thấp nhất là Credit card 19%, Bank transfer và Credit Card là hai nhóm khách hàng có tỷ lệ churn thấp nhất thì lại có số lượng khách hàng không nhiều. Xét theo trung bình tổng tiền hằng tháng khách hàng chi trả cho toàn bộ dịch vụ của công ty Avg. Monthly Charges, khách hàng thanh toán theo séc điện tử Electronic check cũng chi trả một khoản phí cao hơn (79.52), tiếp đó là Bank transfer (75.56), Credit card 74.75 và cuối cùng là Mailed check 53.37

1. **Churn Rate Contract:**

**A graph of a bar

Description automatically generated with medium confidence**

Hình 33: Tỷ lệ khách hàng rời bỏ công ty theo loại hợp đồng.

Ở phần thông tin về hợp đồng cho thấy số khách hàng sử dụng loại hợp đồng theo tháng là nhiều nhất, thì ở biểu đồ trên cũng cho thấy tỷ lệ khách hàng rời bỏ công ty khi sử dụng hợp đồng theo tháng cũng là cao nhất đạt 42,71%. Và tỷ lệ rời bỏ khi sử dụng hợp đồng 1 năm cao thứ 2 với con số là 11,27%, thấp nhất là hợp đồng 2 năm với 2,83%. Dù 2 loại hợp đồng này cũng có tỷ lệ khách hàng rời bỏ công ty nhưng không đáng kể so với hợp đồng theo tháng. Có thể là do hợp đồng theo tháng thường mang lại tính linh hoạt cao cho khách hàng, cho phép họ chuyển đổi hoặc hủy bỏ dịch vụ một cách dễ dàng khi có một số điều không hài lòng với các dịch vụ mà công ty cung cấp hơn so với hợp đồng dài hạn. Hoặc là khách hàng có thể không muốn cam kết với một nhà cung cấp dịch vụ trong thời gian dài, đặc biệt là khi họ chưa thử nghiệm hoặc chưa có kinh nghiệm với dịch vụ của công ty đó. Cho nên doanh nghiệp nên tập trung hơn vào nhóm khách hàng này, nên có phương thức thu hút khách hàng về các ưu đãi cũng như quyền lợi đặc quyền khi tham gia vào các loại hợp đồng lâu năm để tăng sự cam kết khách hàng đồng thời giảm đi tỉ lệ rời bỏ của khách hàng.

1. **Xây dựng mô hình:**
2. **Tiền xử lý dữ liệu:**
   1. **Đọc và hiểu dữ liệu:**

A screenshot of a computer

Description automatically generated

Hình 34: Dữ liệu gốc

Bộ dữ liệu gồm có 21 cột và 7043 dòng dữ liệu, bao gồm các nội dung sau:

- Khách hàng đã rời đi trong tháng trước – cột có tên là Churn

- Các dịch vụ mà mỗi khách hàng đã đăng ký – điện thoại, nhiều đường dây, internet, bảo mật trực tuyến, sao lưu trực tuyến, bảo vệ thiết bị, hỗ trợ kỹ thuật và truyền phát trực tuyến TV và phim.

- Thông tin tài khoản khách hàng – họ đã là khách hàng được bao lâu, hợp đồng, phương thức thanh toán, thanh toán không cần giấy tờ, phí hàng tháng và tổng phí.

- Thông tin nhân khẩu học về khách hàng – giới tính, độ tuổi và liệu họ có đối tác và người phụ thuộc hay không

- Trong tập dữ liệu Telco Churn có 5174 khách hàng không rời bỏ công ty (No) và có 1869 khách hàng đã rời bỏ công ty (Yes)

A black screen with white text

Description automatically generated

Hình 35: Số lượng khách hàng rời bỏ và không rời bỏ của tập dữ liệu.

* 1. **Kiểm tra dữ liệu null:**

**A screenshot of a computer

Description automatically generated**

Hình 36: Kiểm tra dữ liệu null.

Sau khi kiểm tra thì bộ dữ liệu Telco Churn có tổng là 11 giá trị null ở cột dữ liệu TotalCharges.

A screenshot of a computer

Description automatically generated

Hình 37: Dữ liệu sau khi xóa dữ liệu null.

Sau khi xóa null, kiểm tra và xóa giá trị trùng lặp thì bộ dữ liệu sẽ còn lại là 7032 dòng và 21 cột.

* 1. **Mã hóa các giá trị nhãn thành các giá trị số nguyên:**

A screenshot of a computer

Description automatically generated

Hình 38: Mã hóa các giá trị nhãn thành các giá trị số nguyên

Tiếp theo sẽ chuyển các giá trị yes/No thành các giá trị 0 và 1, trong đó 0 là No và 1 là Yes. Việc chuyển hóa này sẽ giúp cho quá trình huấn luyện mô hình được hiệu quả hơn

1. **Xây dựng mô hình Logistic Regression:**

A computer screen shot of a black screen

Description automatically generated

Hình 39: Chia tập dữ liệu thành tập Train và tập Test.

Đầu tiên chia tập dữ liệu Telco Churn thành tập huấn luyện có 4922 dòng và 20 cột, tập kiểm tra có 2110 dòng và 20 cột

A screenshot of a computer program

Description automatically generated

Hình 40: Huấn luyện mô hình LogisticRegression.

Tiếp theo sẽ sử dụng thư viện LogisticRegression của Sklearn để tiến hành huấn luyện mô hình trên tập Train và dự đoán trên tập Test.

1. **Đánh giá mô hình:**
   1. **Độ chính xác:**

A screen shot of a computer

Description automatically generated

Hình 41: Độ chính xác của mô hình.

Mô hình Logistic Regression cho bộ dữ liệu Telco Churn có độ chính xác là 81,18%, đây là con số khá tốt đối với 1 mô hình dự đoán. Với độ chính xác này, dữ liệu được dự đoán từ mô hình này có thể cho kết quả khá là chính xác.

A screen shot of a computer

Description automatically generated

Hình 42: Các chỉ số đánh giá mô hình

Các chỉ số đánh giá mô hình:

- Đối với lớp 0 (Negative), mô hình có Precision khá cao (85%) và Recall (90%). F1-score cũng khá cao (88%), cho thấy mô hình hoạt động rất tốt trong việc dự đoán các trường hợp không rời bỏ doanh nghiệp.

- Đối với lớp 1 (Positive), Precision chỉ có 67%, Recall là 57%, và F1-score là 62%. Mô hình chưa hoạt động tốt lắm trong việc dự đoán các trường hợp khách hàng rời bỏ doanh nghiệp

* 1. **Ma trận nhầm lẫn:**

**A screenshot of a blue and white color scheme

Description automatically generated**

Hình 43:Ma trận nhầm lẫn.

Trong ma trận nhầm lẫn của mô hình Logistic Regression thì số khách hàng không rời bỏ công ty được dự đoán đúng là 1395 khách hàng và dự đoán sai 154 khách hàng. Số khách hàng rời bỏ công ty được dự đoán đúng là 318 khách hàng và dự đoán sai là 243 khách hàng.

* 1. **Đường cong ROC:**

**A graph showing a logistic regression

Description automatically generated**

Hình 44: Đường cong ROC.

Đường cong ROC được biểu diễn bằng màu xanh lam. Một mô hình dự báo chính xác càng cao thì đường ROC của nó trên đồ thị càng gần góc bên trái trên cùng. Đường cong này cho thấy rằng mô hình có độ nhạy là 0.856, có nghĩa là mô hình có thể phân loại chính xác 85.6% các lớp Yes và No. Nhìn chung đường cong ROC của mô hình trên cho thấy rằng mô hình có hiệu suất tốt, và có thể phân loại khá chính xác các lớp Yes và No trong bộ dữ liệu Telco Churn.

⇒ Tóm lại, mô hình dường như hoạt động khá tốt trong việc phát hiện các trường hợp khách hàng không rời bỏ công ty với độ chính xác cao. Tuy nhiên, đối với trường hợp khách hàng có rời bỏ công ty mô hình có thể được cải thiện để giảm số lượng dự đoán sai và tăng độ recall.

1. **Dự đoán dựa trên dữ liệu mới:**

Đầu tiên sẽ nhập dữ liệu bất kỳ từ bàn phím để dự đoán:

A screenshot of a computer

Description automatically generated

Hình 45: Nhập dữ liệu dự đoán từ bàn phím

Sau đó sẽ tiến hành dự đoán dựa trên mô hình Logistic Regression đã được huấn luyện:

A screenshot of a computer

Description automatically generated

Hình 46: Dự đoán nhãn cho dữ liệu mới.

Sau khi dự đoán thì mô hình đã dự đoán được trường hợp dữ liệu được nhập vào là trường hợp khách hàng rời bỏ công ty.

1. **Giải pháp**

Doanh nghiệp không thể cố gắng giữ chân khách hàng bằng cách làm hài lòng tất cả các khách hàng của mình, vì nguồn lực của công ty dù lớn bao nhiêu cũng đều có hạn. Thay vào đó, họ cần nhìn nhận được lượng khách hàng mà nếu níu giữ sẽ mang lại lợi nhuận cao nhất cho doanh nghiệp. Trong bức tranh toàn cảnh về số khách hàng rời bỏ, bằng cách sử dụng các phân tích trên, tôi muốn mang đến những giải pháp hiệu quả nhất để giúp doanh nghiệp vừa có những chính sách giữ chân khách hàng hợp lý, vừa tiết kiệm được chi phí để tối đa hóa lợi nhuận.

Đầu tiên phải nói đến nhóm khách hàng là người già (>=65 tuổi) với tỉ lệ rời bỏ rất cao lên đến 41,68%. Đối với nhóm khách hàng này, công ty nên xem xét lại các chính sách hay các dịch vụ để phù hợp với người lớn tuổi hơn vì những người lớn tuổi họ thường không linh hoạt và hiểu biết nhiều về các dịch vụ viễn thông. Công cần hỗ trợ nhóm khách hàng này nhiều hơn để giảm tỷ lệ rời bỏ ở nhóm khách hàng này.

Thứ 2, công ty cần kiểm tra lại dịch vụ Internet sợi cáp quang mà công ty đang cung cấp. Cần kiểm tra lại về chất lượng đường truyền internet, băng thông hay quá trình hỗ trợ khách hàng như thế nào để tỷ lệ khách hàng rời bỏ cao hơn so với bình thường là 41,89%. Nếu có vấn đề doanh nghiệp cần nhanh chóng giải quyết và liên hệ với khách hàng ngay lập tức.

Thứ 3, đối với nhóm khách hàng sử dụng hợp đồng theo tháng (month-to-month) có tỷ lệ khách hàng rời bỏ tận 42,71%, công ty cần có các chính sách, ưu đãi phù hợp với nhóm khách hàng này, hoặc công ty có thể đưa ra các ưu đãi, các chương trình khuyến mãi để khuyến khích khách hàng có thể chuyển sang loại hợp đồng dài hạn 1 năm hoặc 2 năm

Cuối cùng, công ty nên rà soát, cải thiện lại phương thức thanh toán "Electronic Check" xem có đang gây khó khăn, bất tiện gì cho khách hàng không khi mà tỷ lệ khách hàng rời bỏ khi sử dụng phương thức này cao nhất với mức 45,29%.

1. **KẾT LUẬN**

Hiện nay, sau khi đại dịch covid chấm dứt thì nền kinh tế của Việt Nam nói chung và của thế giới nói riêng đang dần phục hồi lại thời kì đỉnh cao. Viễn thông là một thị trường đã bão hòa, để có được lợi thế cạnh tranh khi doanh nghiệp hoạt động trong lĩnh vực này phải hiểu rõ được điểm mạnh – điểm yếu của mình để có thể tận dụng cũng như cải thiện hiệu suất kinh doanh. Và dữ liệu chính là sức mạnh của mọi doanh nghiệp, một khi doanh nghiệp có thể sử dụng tốt nguồn sức mạnh này, họ sẽ có thể chiếm lĩnh được thị trường. Việc hoạt động trong một thị trường cạnh tranh khốc liệt như Viễn thông thì quản trị tốt mối quan hệ với khách hàng là một yêu cầu cấp thiết đối với mọi công ty. Để làm được điều đó, công ty phải giữ chân được khách hàng cũ và tìm kiếm thêm khách hàng mới. Dự án “ỨNG DỤNG KỸ THUẬT PHÂN TÍCH KHÁCH HÀNG RỜI BỎ DOANH NGHIỆP CHO NGÀNH VIỄN THÔNG VÀ ĐỀ XUẤT GIẢI PHÁP” đã phần nào giúp doanh nghiệp làm tốt hơn công tác quản trị quan hệ khách hàng. Trong bài báo cáo, cá nhân tôi đã đưa ra đặc điểm chung của những khách hàng rời bỏ doanh nghiệp, nguyên nhân khiến họ rời bỏ và đề xuất một số giải pháp để khắc phục tình trạng đó. Ngoài ra, tôi đã xây dựng được một mô hình dự đoán khả năng rời bỏ, đưa ra các cảnh báo cho doanh nghiệp để ngăn chặn, cải thiện tình trạng rời bỏ của khách hàng. Tuy nhiên, do vốn kiến thức và kinh nghiệm còn hạn chế nên bài báo cáo sẽ còn tồn tại nhiều thiếu sót. Tôi rất hy vọng nhận được những nhận xét, đóng góp để bài báo cáo được hoàn chỉnh hơn. Xin chân thành cảm ơn.

**TÀI LIỆU THAM KHẢO**

1. Đinh, N. (2024, February 18). Tỷ lệ rời bỏ của khách hàng: dự đoán và khắc phục chỉ số CX này ra sao?*Filum*.

<https://blog.filum.ai/ty-le-roi-bo-khach-hang-cach-du-doan-va-khac-phuc#:~:text=T%E1%BB%B7%20l%E1%BB%87%20r%E1%BB%9Di%20b%E1%BB%8F%20hay,%C4%91o%20l%C6%B0%E1%BB%9Dng%20c%E1%BB%A7a%20doanh%20nghi%E1%BB%87p>.

1. *Hướng dẫn đầy đủ về giữ chân khách hàng*. (2024, April 24). Amazon Ads. [https://advertising.amazon.com/vi-vn/library/guides/customer-retention#:~:text=Gi%E1%BB%AF%20ch%C3%A2n%20kh%C3%A1ch%20h%C3%A0ng%20r%E1%BA%A5t,l%C6%B0%E1%BB%A3c%20kinh%20t%E1%BA%BF%20c%E1%BB%A7a%20b%E1%BA%A1n](https://advertising.amazon.com/vi-vn/library/guides/customer-retention" \l ":~:text=Gi%E1%BB%AF%20ch%C3%A2n%20kh%C3%A1ch%20h%C3%A0ng%20r%E1%BA%A5t,l%C6%B0%E1%BB%A3c%20kinh%20t%E1%BA%BF%20c%E1%BB%A7a%20b%E1%BA%A1n)
2. Nguyễn T. (2023, July 12). *8 lý do khiến cho khách hàng rời bỏ doanh nghiệp*. <https://bizfly.vn/techblog/khach-hang-roi-bo-doanh-nghiep.html>

**PHỤ LỤC**

<https://public.tableau.com/views/Sssss/Story1?:language=fr-FR&publish=yes&:sid=&:display_count=n&:origin=viz_share_link>

***-------------------------------------------The End-------------------------------------------***