

ĐỀ THI VÒNG 3

Business Challenge



Cuộc thi Phân tích Dữ liệu DAZONE 2024

Nội dung

- 1 Giới thiệu Bảo trợ Chuyên môn Vòng 3
- 2 Tổng quan bài toán doanh nghiệp
- 3 Thể lệ Vòng 3: Business Challenge
- 4 Tiêu chí chấm điểm
- 5 Các mốc thời gian quan trọng



1

Giới thiệu

Bảo trợ Chuyên môn Vòng 3: Business Challenge

Tập đoàn Công nghiệp - Viễn thông Quân đội **Viettel**



Tầm nhìn: Sáng tạo vì con người



Sứ mệnh: Tiên phong chủ lực
kiến tạo xã hội số

Tập đoàn Viettel được thành lập vào ngày 1 tháng 6 năm 1989. Các ngành nghề chính của tập đoàn bao gồm: ngành dịch vụ viễn thông & công nghệ thông tin; ngành nghiên cứu sản xuất thiết bị điện tử viễn thông, ngành công nghiệp quốc phòng, ngành công nghiệp an ninh mạng và ngành cung cấp dịch vụ số.

Tập đoàn Công nghiệp - Viễn thông Quân đội **Viettel**

DẤU ẤN THƯƠNG HIỆU

8,902

TỶ USD
GIÁ TRỊ THƯƠNG HIỆU

234

THƯƠNG HIỆU
GIÁ TRỊ NHẤT THẾ GIỚI

16

THƯƠNG HIỆU
VIỄN THÔNG
GIÁ TRỊ NHẤT THẾ GIỚI

9

THƯƠNG HIỆU
VIỄN THÔNG
GIÁ TRỊ NHẤT CHÂU Á

1

THƯƠNG HIỆU
VIỄN THÔNG
GIÁ TRỊ NHẤT ĐNA

Hiện tại, **Viettel** là một trong những nhà khai thác viễn thông phát triển nhanh nhất thế giới với hơn **55.000 nhân viên**. Viettel hiện diện tại **11 quốc gia và vùng lãnh thổ**, phục vụ hơn **270 triệu người dùng** đến từ Châu Á, Châu Phi và Châu Mỹ.



2

Tổng quan

Bài toán doanh nghiệp Vòng 3: Business Challenge

Tổng quan bài toán doanh nghiệp

Bối cảnh thị trường

Trong thời đại số, khả năng kết nối internet mọi lúc mọi nơi đã trở thành nhu cầu thiết yếu của du khách. Internet giúp họ làm việc, giải trí và giữ kết nối một cách liên mạch trong suốt hành trình. Nhằm đáp ứng nhu cầu này, ngành du lịch Việt Nam cần ưu tiên cung cấp các gói 4G được cá nhân hóa phù hợp với nhu cầu sử dụng và lịch trình cụ thể của mỗi du khách.

Tổng quan bài toán doanh nghiệp

Tại Việt Nam, Viettel là nhà cung cấp dịch vụ viễn thông hàng đầu với mạng lưới phủ sóng rộng khắp cả nước, từ các đảo xa xôi đến các vùng núi cao, đảm bảo tín hiệu luôn sẵn có khắp chiều dài đất nước. Với phương châm "Không ai bị bỏ lại phía sau", Viettel mang đến nhiều gói cước đa dạng, đáp ứng nhu cầu sử dụng internet khác nhau của khách hàng.

Trước thực tế lượng du khách đến Việt Nam ngày càng tăng, Viettel đang nỗ lực khai thác phân khúc thị trường tiềm năng này. Cụ thể, nhà mạng cá nhân hóa các gói cước 4G, đảm bảo chất lượng dịch vụ để du khách có thể truy cập internet liên mạch, dễ dàng tìm kiếm thông tin địa điểm du lịch và kết nối với người thân.

Tổng quan bài toán doanh nghiệp

Thách thức

- Giả sử bạn là một thành viên của dự án "**Cá nhân hoá gói cước 4G Viettel**", bạn thấu hiểu thói quen sử dụng internet của khách hàng, đặc điểm riêng biệt của các điểm du lịch và khả năng cung cấp dịch vụ của Viettel trong các khu vực đó.
- Thực hiện **bài toán phân loại** (Classification) để dự đoán những gói cước 4G phù hợp với du khách dựa vào thông tin về chuyến đi của họ. Từ đó, đưa ra những campaign phù hợp để quảng bá các gói cước.

Mô tả bộ dữ liệu

- Bộ dữ liệu được thu thập bằng cách khảo sát các du khách nước ngoài khi họ mua gói cước của Viettel.
- Bộ dữ liệu bao gồm thông tin về chuyến đi của du khách (mục đích, điểm đến, số người đi cùng, ...) và xem rằng họ sẽ chọn gói cước nào phù hợp với bản thân.

Bộ dữ liệu chi tiết tại **đây**.

Lưu ý: Bộ dữ liệu này chỉ được sử dụng để train. Đội thi sẽ nhận bộ dữ liệu để test sau buổi Mentoring Session, dự kiến vào Thứ Hai, 17/06/2024.

Bảng 1: THUỘC TÍNH NGƯỜI DÙNG

No.	Variable	Description
1	id	Mã định danh của bài khảo sát
2	name	Tên khách hàng
3	gender	Giới tính khách hàng
4	education	Trình độ học vấn
5	profession	Nghề nghiệp

Bảng 1: THUỘC TÍNH NGƯỜI DÙNG

No.	Variable	Description
6	income	Thu nhập
7	living_with	Tình trạng hôn nhân
8	nation	Quốc gia
9	phone	Số điện thoại
10	job	Nghề nghiệp

Bảng 1: THUỘC TÍNH NGƯỜI DÙNG

No.	Variable	Description
11	fb_freq	Tần suất sử dụng Facebook mỗi ngày (giờ)
12	yt_freq	Tần suất sử dụng Youtube mỗi ngày (giờ)
13	tik_freq	Tần suất sử dụng TikTok mỗi ngày (giờ)
14	use_less_then_2GB	Số lần sử dụng 4G ít hơn 2GB trong một tháng
15	use_2GB_to_4GB	Số lần sử dụng 4G trong khoảng 2GB - 4GB trong một tháng

Bảng 2: THUỘC TÍNH VỀ CHUYẾN ĐI

No.	Variable	Description
1	id	Mã định danh của bài khảo sát
2	purpose	Mục đích chuyến đi
3	go_with	Người đi cùng
4	weather	Thời tiết
5	time	Thời điểm đáp xuống sân bay

Bảng 2: THUỘC TÍNH VỀ CHUYẾN ĐI

No.	Variable	Description
6	viettel_no_0	Được giới thiệu về SIM Viettel trong 15' đầu tiên sau khi vào tới sân bay
7	viettel_no_1	Được giới thiệu về SIM Viettel trong 30' đầu tiên sau khi vào tới sân bay
8	viettel_no_3	Được giới thiệu về SIM Viettel trong 45' đầu tiên sau khi vào tới sân bay

Bảng 2: THUỘC TÍNH VỀ CHUYỂN ĐI

No.	Variable	Description
9	to_hanoi	Du khách có tới Hà Nội sau khi rời sân bay không
10	to_other	Du khách có tới các tỉnh thành khác sau khi rời sân bay không
11	score	Điểm Viettel++
12	direction	Du khách có tới các tỉnh thành khác sau khi rời sân bay không

Bảng 3: ĐỀ XUẤT GÓI DATA CHO NGƯỜI DÙNG

No.	Variable	Description
1	id	Mã định danh của bài khảo sát
2	mobile_plan	Gói Data đề xuất cho du khách
3	accept	Khách đăng ký gói được đề xuất hay không

Bảng 4: THÔNG TIN VỀ GÓI DATA

No.	Variable	Description
1	mobile_plan	Tên gói Data
2	price	Giá
3	description	Mô tả gói
4	duration	Thời gian hiệu lực



3

Thể lệ

Nộp bài và trình bày bài làm Vòng 3: Business Challenge

Quy định nộp bài

- Thời gian nộp bài:
 - File Notebook và bản Báo cáo phân tích: **trước 23h59 Thứ Tư, 19/06/2024;**
 - Slide thuyết trình: **trước 12h00 Thứ Sáu, 21/06/2024.**
- Các sản phẩm yêu cầu:
 - 01 file Notebook;
 - 01 bản Báo cáo phân tích (tối đa 05 trang, không bao gồm phụ lục và trích dẫn);
 - 01 bản slide thuyết trình (tối đa 20 trang, không bao gồm phụ lục, trang giới thiệu và cảm ơn).
- Đội thi nộp bài qua folder Ban Tổ chức cung cấp và chỉnh chế độ “**Công khai**”, người truy cập có quyền “**Viewer**”;
- Cách thức đặt tên sản phẩm:
[TÊN ĐỘI THI]_[TÊN SẢN PHẨM]_VÒNG 3: BUSINESS CHALLENGE

Quy định trình bày file notebook

- Định dạng: .ipynb;
- Code bằng ngôn ngữ Python;
- Có giải thích code bằng tiếng Anh (giải thích dưới dạng comment);
- File chạy lại được, không lỗi;
- Toàn bộ các biểu đồ phân tích insights và các thử nghiệm model phải xuất hiện ở trong Notebook.

Quy định trình bày báo cáo phân tích

HÌNH THỨC TRÌNH BÀY

- Định dạng: .pdf;
- Không quá 05 trang A4;
- Trình bày dưới dạng Báo cáo, khổ A4 đứng;
- Font chữ: Time News Roman;
- Size: 12;
- Các bảng và biểu đồ sẽ được trình bày trong phạm vi $\frac{1}{3}$ trang, bên trái báo cáo. Đội thi cần có tên và chú thích bảng và biểu đồ;
- Các luận điểm báo cáo sẽ được trình bày trong $\frac{2}{3}$ phạm vi còn lại của trang;
- Đối với các trang không có bảng và biểu đồ, đội thi có thể sử dụng $\frac{1}{3}$ trang còn lại để trình bày luận điểm;
- Mẫu Báo cáo Phân tích được đính kèm trong file sau:

[DAZONE2024] Mẫu Báo cáo Phân tích.pdf

Quy định trình bày báo cáo phân tích

NỘI DUNG

- Data acquisition and understanding;
- Data preprocessing and cleaning;
- EDA;
- Feature selection and engineering;
- Models building and evaluation;
- Deployment and communication of results.

Quy định trình bày báo cáo phân tích

TRÍCH DẪN

Đối với các thông tin bên ngoài đề thi, các đội thi phải trích nguồn của các dữ liệu bên ngoài.

- Trích dẫn tài liệu tham khảo theo hệ thống **Harvard (Harvard Referencing System)**;
- Các trích dẫn được đính kèm trong phần **Appendix**.

Quy định trình bày slide thuyết trình

HÌNH THỨC TRÌNH BÀY

- Định dạng: .pdf;
- Ngôn ngữ trình bày: Tiếng Việt;
- Trình bày dưới dạng slide thuyết trình, không có yêu cầu về mẫu thiết kế;
- Dung lượng: Tối đa 20 slides (chưa bao gồm phụ lục, trang giới thiệu và lời cảm ơn);
- Các biểu đồ phải được đánh số, chú thích và trích dẫn trong các phân tích.

Quy định trình bày slide thuyết trình

NỘI DUNG

- Data acquisition and understanding;
- Data preprocessing and cleaning;
- EDA;
- Feature selection and engineering;
- Models building and evaluation;
- Deployment and communication of results.

Quy định trình bày slide thuyết trình

TRÍCH DẪN

Đối với các thông tin bên ngoài đề thi, các đội thi phải trích nguồn của các dữ liệu bên ngoài.

- Trích dẫn tài liệu tham khảo theo hệ thống **Harvard (Harvard Referencing System)**;
- Các trích dẫn được đính kèm trong phần **Appendix**.

Quy định thuyết trình và phản biện sản phẩm

- Thời gian: **Đêm Chung kết Grand Finale**, Thứ Bảy, 22/06/2024;
- Hình thức: Trực tiếp, 1 vs 1;
- Sau khi nhận được bài làm của các đội thi, các đội thi tiến hành bốc thăm để quyết định cặp phản biện và thứ tự trình bày;
- Phần thi trực tiếp bao gồm thuyết trình sản phẩm, nhận và trả lời câu hỏi phản biện từ đội đối thủ, phản biện sản phẩm của đội đối thủ và vấn đáp với Hội đồng Ban Giám khảo;
- Kết quả: Chọn ra 01 Quán quân, 01 Á quân, 02 Quý quân.

***Hoạt động và thời lượng tương tự với cặp 2.**

THỜI LƯỢNG PHẦN THI TRỰC TIẾP: CẶP 1*	
Hoạt động	Thời lượng
<ul style="list-style-type: none">- Đội thi A trình bày bài làm của mình;- Đội thi B chuẩn bị câu hỏi phản biện dành cho đội thi A.	15 phút
Đội thi B tiến hành đặt câu hỏi phản biện dành cho đội thi A, tối thiểu 1 câu và tối đa 5 câu.	02 phút
Đội thi A trả lời các câu hỏi phản biện của đội thi B.	07 phút

THỜI LƯỢNG PHẦN THI TRỰC TIẾP: CẶP 1

Hoạt động	Thời lượng
<ul style="list-style-type: none">- Đội thi B trình bày bài làm của mình;- Đội thi A chuẩn bị câu hỏi phản biện dành cho đội thi B.	15 phút
Đội thi A tiến hành đặt câu hỏi phản biện dành cho đội thi B, tối thiểu 1 câu và tối đa 5 câu.	02 phút
Đội thi B trả lời các câu hỏi phản biện của đội thi A.	07 phút
Ban Giám khảo đặt câu hỏi dành cho đội thi A và B.	20 phút



4

Tiêu chí

Chấm điểm Vòng 3: Business Challenge

TIÊU CHÍ ĐÁNH GIÁ BÀI LÀM CỦA THÍ SINH

Phần	Tỷ trọng
Bài làm nộp về Ban Tổ chức: <ul style="list-style-type: none">• File Notebook;• Bản Báo cáo phân tích;• Slide thuyết trình.	60%
Phần thi trực tiếp: <ul style="list-style-type: none">• Thuyết trình bài làm;• Phản biện với đội đối thủ;• Vấn đáp với Ban Giám khảo.	40%



5

Mốc thời gian quan trọng

CÁC MỐC THỜI GIAN QUAN TRỌNG

Thời gian	Hoạt động
20h00 ngày 12/6/2024	Ban Tổ chức công bố đề thi Vòng 3
23h59 ngày 19/6/2024	Kết thúc nhận file Notebook và bản Báo cáo phân tích
12h00 ngày 21/6/2024	Kết thúc nhận Slide thuyết trình
14/6/2024 - 17/6/2024	Mentoring Session
22/6/2024	Grand Finale

THÔNG TIN LIÊN HỆ

BAN TỔ CHỨC

Trần Phan Thảo Vy

Trưởng Ban Tổ chức

Tel: (+84) 819 222 509

Mail: tranphanthaovy.rcs@gmail.com

Lưu Thị Khánh Linh

Trưởng Ban Tổ chức

Tel: (+84) 376 916 034

Mail: luuthikhanhlinh.rcs@gmail.com

Mai Thị Quỳnh Anh

Trưởng Ban Chuyên Môn

Tel: (+84)896 951 610

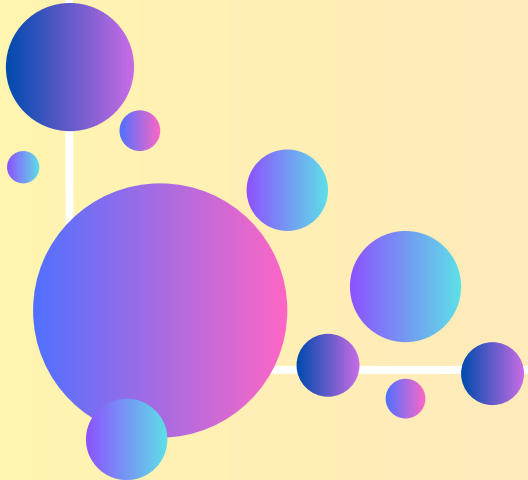
Mail: maithiquynhanh.rcs@gmail.com



FANPAGE: DAZONE -
Cuộc thi Phân tích dữ liệu



GROUP: [DAZONE] DATA
ANALYSIS FOR STUDENTS



**Chúc các đội thi hoàn thành
Vòng 3: Business Challenge
thật tốt!**