

## Curso Ciência de Dados na Prática

### Encontro Modelagem Supervisionada I – 02/07/2022

#### Case VAI Telecom

Este documento contém diversas informações importantes e que vão ajudar na resolução do case do encontro de sábado, 02/07.

## Introdução

A VAI Telecom é uma empresa de telecomunicações com forte atuação nos Estados Unidos. Ela atua na construção de sistemas que permitem a troca de informação por meio de dispositivos eletrônicos em distâncias relativamente grandes.

A necessidade de inovação no setor de telecomunicações é constante, novas tecnologias fazem com que a VAI Telecom precise desenvolver novas estratégias para sustentar a organização, aumentando a eficiência, reformulando novas infraestruturas e buscando novas estratégias para aumentar e fidelizar sua base de clientes.

No cenário da crescente competitividade e maior demanda, a VAI Telecom se viu em um mercado mais exigente em relação à velocidade e qualidade de conexão e ligações. O core de seu negócio se baseia no fornecimento de serviços de telefonia e internet, com o oferecimento de planos customizados para clientes. Nesse cenário, a VAI Telecom vem observando uma piora nos indicadores de retenção de clientes, acompanhada de um aumento considerável nos indicadores de evasão de clientes (**churn rate**).

Com essa evasão de clientes acontecendo a ritmos cada vez mais acelerados, a VAI Telecom montou um comitê para tentar entender as principais causas dessa evasão, bem como para bolar um plano de recuperação. O comitê, formado por uma equipe multidisciplinar com especialistas técnicos e uma equipe estratégica contando com o CEO e CTO da VAI Telecom, disponibilizou uma base de dados histórica do comportamento dos clientes.

A ideia é que, a partir dos dados disponibilizados, uma campanha de retenção de clientes seja colocada em prática de forma a aumentar os índices de retenção, consequentemente impactando outros importantes indicadores de negócio da VAI Telecom, como o LTV (*Lifetime Value*), a ser discutido posteriormente.

## Dados Disponibilizados

### Descrição dos dados:

O comitê da VAI Telecom nos enviou uma base de dados contendo informações sobre seus clientes, configuração de planos e histórico de pagamento em um determinado período. A seguir um descritivo breve das colunas presentes na base de dados.

#### Base:

charge\_records.csv

#### Colunas:

1. **ActiveMonth** – Mês de referência em relação a permanência do cliente na empresa
2. **customerID** – ID do cliente
3. **Charges** – Pagamento por serviços

#### Base:

customers.csv

#### Colunas:

1. **customerID** – ID do usuário
2. **gender** – sexo do usuário
3. **Country** – País do cliente
4. **State** – Estado do cliente
5. **City** – Cidade do cliente
6. **ZipCode** – Código postal do cliente
7. **LatLong** – Latitude e Longitude do cliente
8. **SeniorCitizen** – se o cliente é idoso
9. **Partner** – se cliente tem parceria(a)
10. **Dependents** – se o cliente tem dependentes
11. **PhoneService** – se o cliente possui serviço de telefone
12. **MultipleLines** – se o cliente tem múltiplas linhas
13. **InternetService** – Provedor de serviços de Internet do cliente
14. **OnlineSecurity** – se o cliente possui serviço de segurança online
15. **OnlineBackup** – Se o cliente de backup online
16. **DeviceProtection** – Se cliente tem proteção de dispositivo

- 17. **TechSupport** – Se o cliente tem suporte técnico
- 18. **StreamingTV** – Se o cliente tem streaming de TV
- 19. **StreamingMovies** – Se o cliente tem streaming de filmes ou não
- 20. **Contract** – Termo do contrato do cliente
- 21. **PaperlessBilling** – Se o cliente tem faturamento sem papel
- 22. **PaymentMethod** – O método de pagamento do cliente
- 23. **Churn** – Se o cliente saiu
- 24. **ChurnReason** – Motivos da saída do cliente

## Instruções do Case

### Sobre o encontro

No dia 02/07 (sábado) às 9h, teremos o nosso Encontro de Modelagem Supervisionada com a resolução do case sobre *Churn*.

Esta aula terá um momento de trabalho em grupo entre os alunos, auxiliados pelos facilitadores, como parte da resolução do case.

O encontro será dividido em 3 momentos principais:

- O 1º momento será a apresentação de conceitos importantes sobre o modelo.
- Depois, teremos uma sessão de solução do feature engineering em grupo, com facilitadores apoiando os grupos na solução.
- Ao final, apresentaremos nossa proposta de solução do case.

### Sobre o case

Para tornar a dinâmica mais produtiva e garantir uma boa experiência para todos, é necessário que o grupo esteja familiarizado com as bases que serão utilizadas no case, reforçando a importância do pre-work.

A partir do entendimento do contexto da empresa e das informações disponibilizadas pelo comitê, explore os dados recebidos de forma a gerar gráficos e tabelas que ajudem a entender as principais relações entre as variáveis bem como quais informações podem ser importantes na avaliação da taxa de *Churn* dos clientes.

Sugestão (NÃO EXAUSTIVA) para seu processo de Análise Exploratória:

- Visualizar a distribuição das características dos clientes através do conjunto de dados.
  - Entender onde os clientes moram;
  - Mapear o perfil dos clientes;
  - Descobrir quais os principais serviços e métodos de pagamento utilizados;
- Buscar correlações entre as variáveis;
- Entender o comportamento dos clientes pelo uso dos serviços;
  - Quantidade de meses que utilizou;
  - Valor pago durante o tempo de uso;

Conforme discutido, a avaliação da taxa de Churn impacta diretamente o cálculo do LTV, sendo esta última uma das mais significativas métricas no negócio (disponibilizaremos um material informativo sobre o tema no encontro). Nosso desafio é construir um modelo de classificação para identificar potenciais clientes que podem deixar a empresa e, assim, determinar melhor a taxa de *Churn* potencial.

No primeiro encontro discutimos juntos a construção de algumas análises, porém, observe que a base disponibilizada compreende o histórico dos clientes de maneira temporal. O aluno tem a liberdade de explorar pontos e realizar análises adicionais como preferir.