



Preditiva.ai

Live #09

Diminuição de Churn com Análise de Dados e Machine Learning

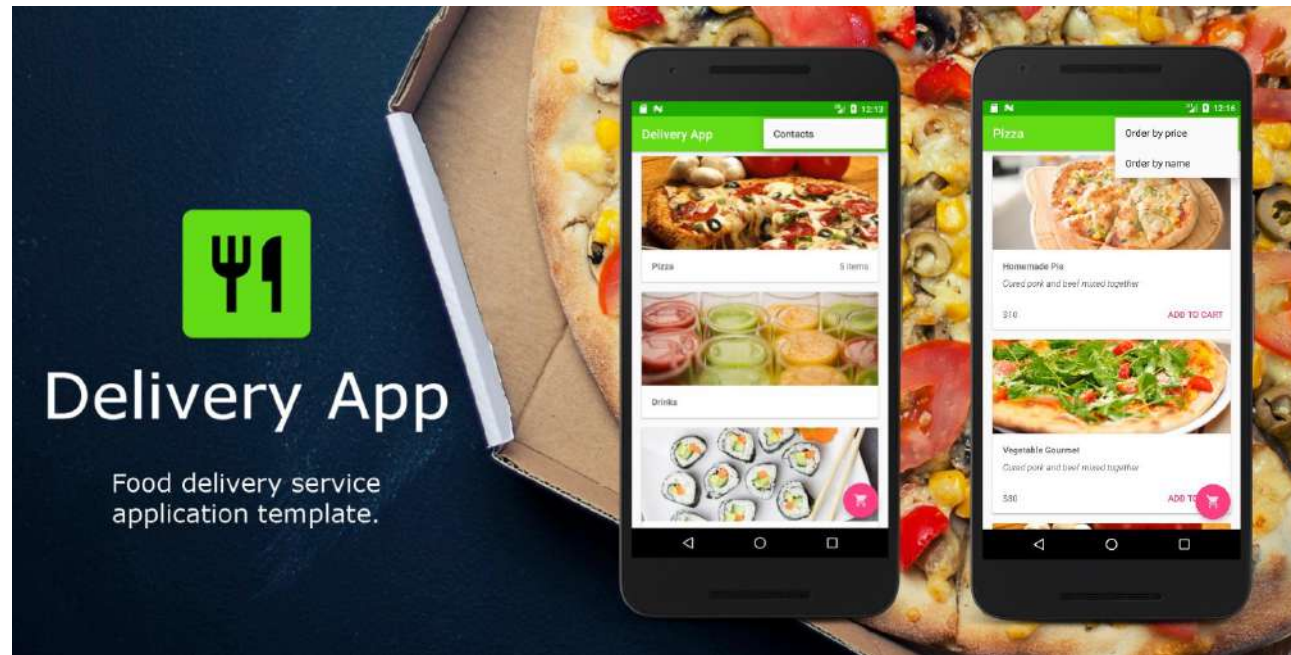
14/02/2023

Live 09 – Diminuição de Churn

Problema de negócio



Nesta live, vamos propor uma solução para um dos problemas comuns de **CRM Analytics: Diminuição de Churn de Clientes.**



Antes de mais nada, já se inscreveu em nosso grupo para ter acesso aos materiais e avisos das lives?



<https://t.me/preditiva>

Ciência de Dados

A área de Ciência de Dados engloba uma infinidade de assuntos e conhecimentos. No entanto, é possível resumir-la como **“a junção das técnicas de matemática, estatística e computação com o conhecimento do negócio”** de forma a otimizar o processo decisório das empresas ou mesmo resolver problemas de qualquer espécie com o apoio dos dados.



Matemática e Estatística

Cálculo, Álgebra, Otimização, Estatística Descritiva, Probabilidades e Inferência.



Computação

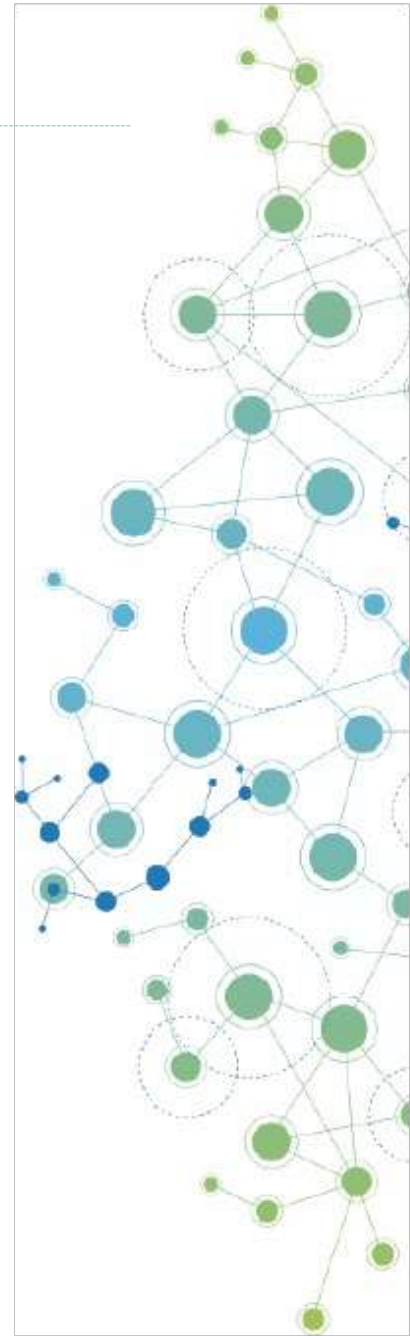
Linguagens de programação, conceitos de banco de dados relacionais e não relacionais, ferramentas de Visualização de Dados.



Conhecimento do Negócio

Sem o conhecimento das particularidades do negócio, a prática de Ciência de Dados fica sem sentido e não agrega o valor necessário.

Um overview do problema de negócio



Contexto de negócio

Lembra do funil de marketing e vendas?



Blog, e-book, Lives, Mídias Sociais, campanhas de tráfego pago, cursos gratuitos.

Formulários, landing pages.

Contato, CRM, follow up e fechamento da venda.

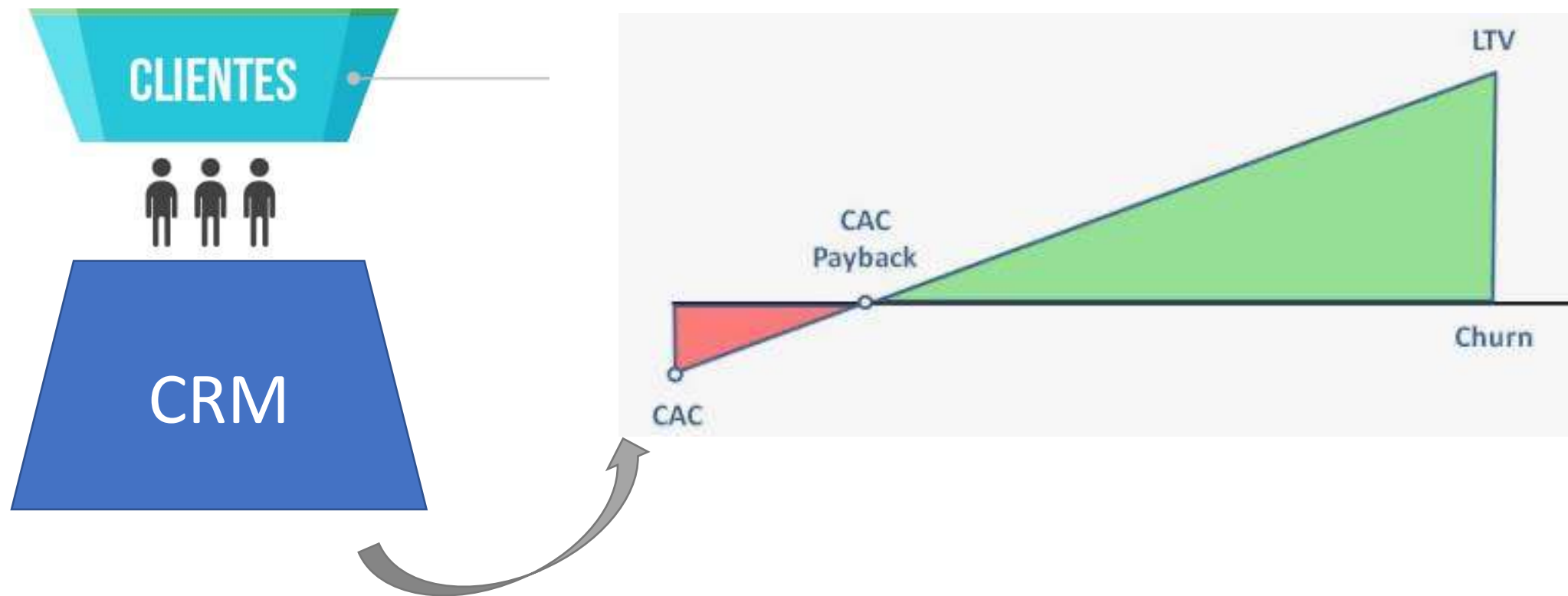
Pós venda, identificação e segmentação do perfil, avaliar novas compras, pesquisa de satisfação, novos conteúdos.

Contexto de negócio

Agora queremos manter o cliente...



Preditiva.ai



LTV (Life Time Value): A grosso modo, quanto o cliente vai trazer de receita ao longo do tempo

CAC (Custo de Aquisição do Cliente): Qual foi o custo para trazer esse cliente (ex: mídia paga, equipe de vendas etc)

Churn (Taxa de Cancelamento): Percentual de clientes que cancelam ou deixam de usar seu produto no tempo

Maximizar

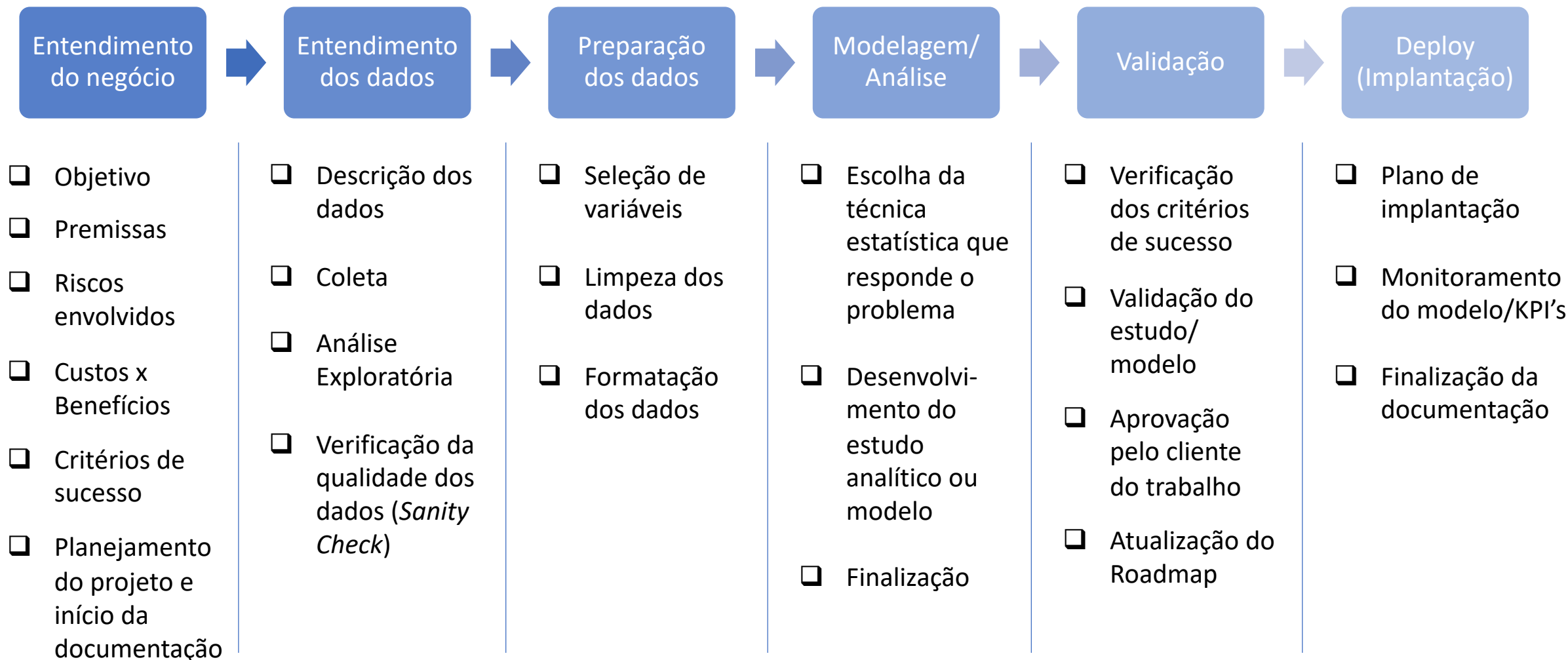
$$\frac{\text{LTV} - \text{Churn}}{\text{CAC}}$$

Framework de Análise de Dados

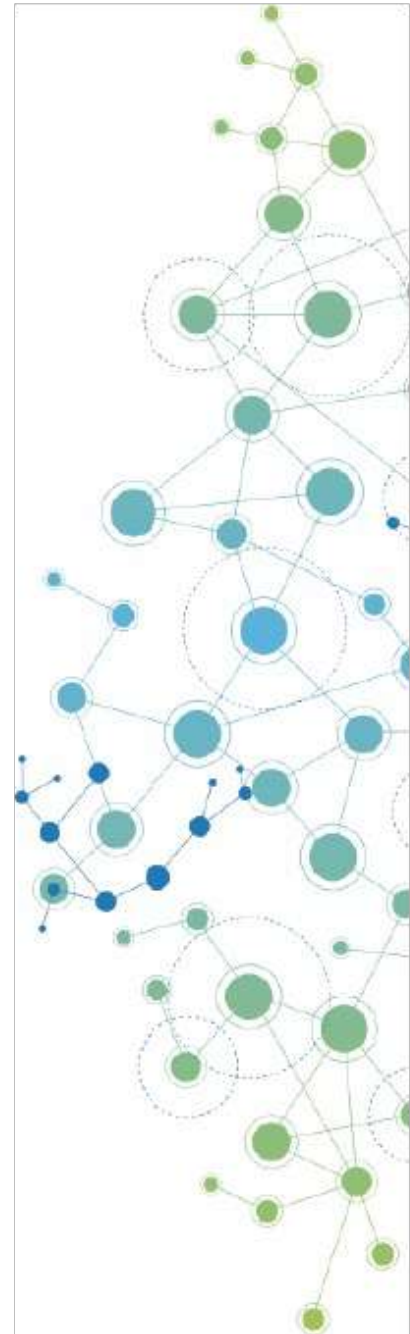
Overview



Utilize essas etapas como um “**check-list**” do seu projeto:



1) Contexto de negócio e objetivo do trabalho



Projeto: Diminuição de Churn

Contexto de negócio e objetivo do trabalho



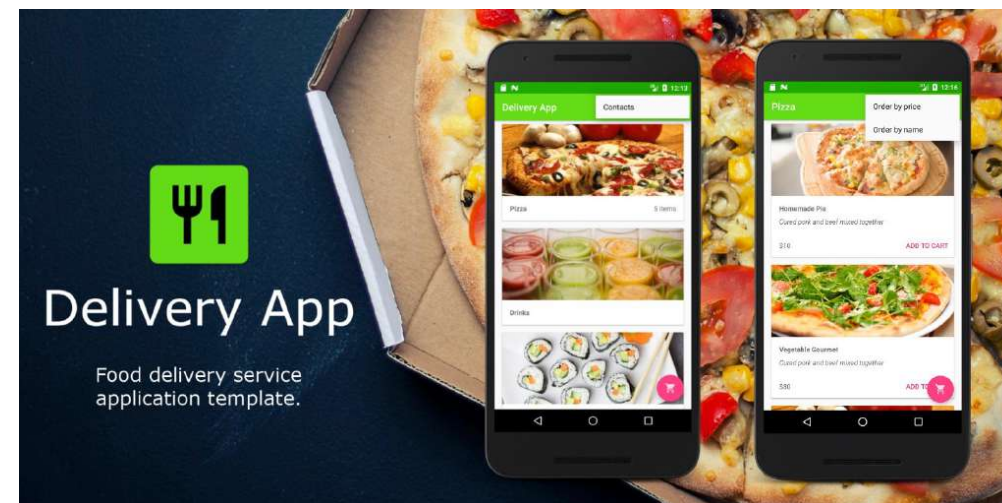
O aplicativo **ToComFome** é um app de entrega de comida por delivery. Como todo aplicativo de varejo, a empresa se preocupa com o Churn de clientes, ou seja, da perda desse cliente para a concorrência. Para entender o perfil de risco do cliente com maior chance de churn, a área de CRM gostaria de entender e classificar o cliente com risco de trocar de aplicativo nos próximos meses.

Para isso, a área de CRM passou uma amostra de cerca de **10 mil** clientes com suas respectivas informações de cadastro e transações nos **próximos 4 meses** a contar da data de extração usada como referência.

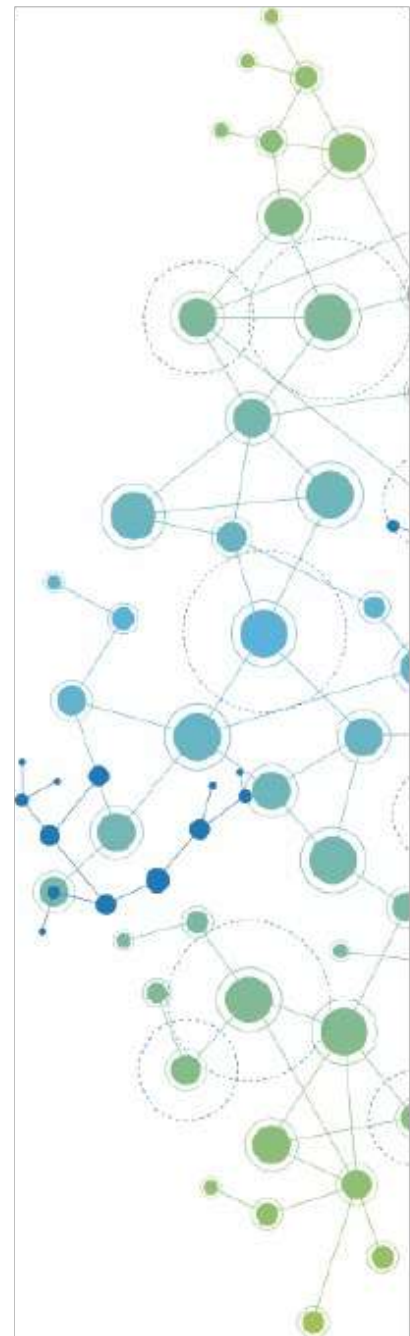
Obs: Um cliente é classificado como Churn quando nos próximos 4 meses em relação à data de referência, ele não transacionou por um período de 30 dias.

Objetivo do trabalho:

- ☐ Descobrir quais fatores de risco estão associados com o Churn de Clientes;
- ☐ Atribuir a probabilidade de darem Churn nos próximos **4 meses** em relação à data de extração de referência e;
- ☐ propor os possíveis planos de ação que a empresa pode tocar para diminuir esse problema.



2) Análise exploratória



Projeto: Diminuição de Churn

Descrição das variáveis



Variável	Descrição
ClientId	Identificador de cliente
DataExtracao	Data da extração dessa visão de clientes (foto)
Score_Credito	Score de Crédito no mercado. A empresa comprou essa variável de um Bureau de Crédito. Quanto mais próximo de 1000, melhor é o cliente em relação a inadimplência.
Estado	Estado do cliente
Gênero	Genero do cliente
Idade	Idade do cliente
Tempo_Cliente	Tempo em meses em que o cliente utiliza o aplicativo
Limite_Credito_Mercado	Limite de Crédito que o cliente tem no mercado. A empresa comprou essa variável de um Bureau de Crédito. Quanto maior é o limite, maior é a chance do cliente ter um bom poder aquisitivo.
Qte_Categorias	Qte de categorias de produtos (ex: lanches, comida japonesa, doces etc) pedidos nos ultimos 3 meses anteriores à data de extração.
Usa_Cartao_Credito	Flag se cliente cadastrou um cartão de crédito no aplicativo. 1 = Cadastrou e 0 = Não Cadastrou
Programa_Fidelidade	Flag se cliente participa de um programa de fidelidade no aplicativo. 1 = Participa e 0 = Não Participa.
Sum_Pedidos_Acumulados	Valores gastos no aplicativo nos ultimos 4 meses anteriores à data de extração.
DataUltimaTransacao	A data da última transação realizada por cada cliente, extraída no dia 30/10/2019 às 23h59min do cliente no aplicativo

Overview sobre técnicas

O que temos disponível no mercado?



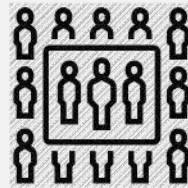
Arsenal analítico



**Analytics e
Dashboards**

- Resumir uma base de dados
- Entender a relação entre os variáveis

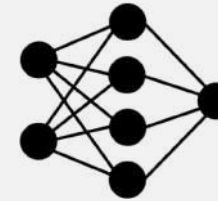
Tomar decisão com os dados completos de forma **manual**



**Inferência
Estatística**

- Estimar valores através de amostras
- Testar hipóteses através de amostras

Tomar decisão com os dados amostrais de forma **manual**



**Machine Learning
e AI**

- Estimar valores de forma automática
- Classificar “coisas” de forma automática

Tomar decisão com dados completos ou amostrais de forma **automática**

Quando usar?

Overview sobre técnicas

O que temos disponível no mercado?



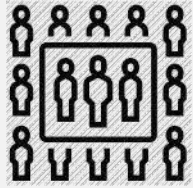
**Analytics e
Dashboards**

Técnicas mais usadas:

- Matemática
- **Estatística Descritiva:**
 - Tabelas de Frequência
 - Medidas Resumo (Média, Mediana, Quartis, Desvio Padrão etc)
 - Gráficos analíticos (Barras, Linhas, Pareto, Histograma, Box-Plots etc)
- **Análise de Associação:**
 - Correlação de Pearson
 - Coeficiente de Determinação (R Quadrado)
 - Information Value (IV)
 - Estudo de Causalidade e Contrafactual
- Dashboards Operacionais, Analíticos e Executivos
- Probabilidade e Forecasting

Overview sobre técnicas

O que temos disponível no mercado?



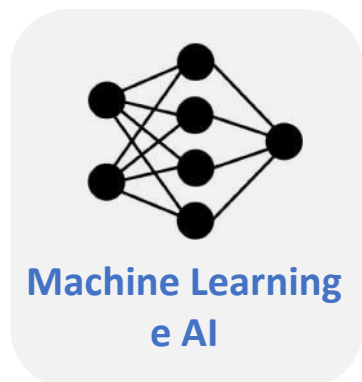
Inferência
Estatística

Técnicas mais usadas:

- Cálculo do tamanho da Amostra
- Extração de amostras randomizadas e com o menor viés possível
- Estimação pontual
- Estimação por Intervalos de Confiança
- Controle da Margem de Erro
- Experimentos (Grupos de Controle e Teste)
- Testes de Hipóteses / Testes A/B (p-valor)
- Regressão Linear Múltipla
- Regressão Logística

Overview sobre técnicas

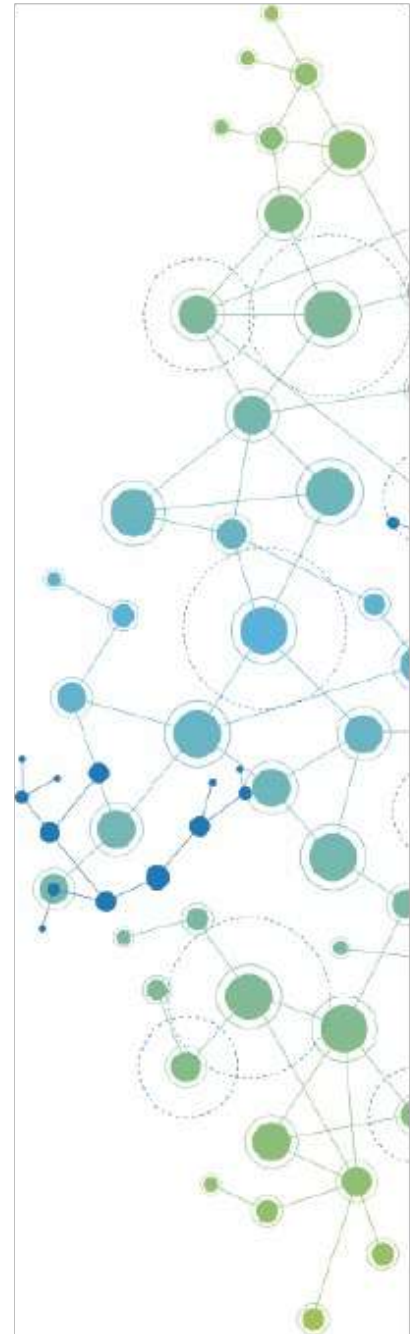
O que temos disponível no mercado?



Técnicas mais usadas:

- **Modelos Supervisionados:**
 - Regressões
 - Árvores de Decisão
 - Redes Neurais Artificiais
 - Diagnóstico de Modelos
- **Modelos Não Supervisionados:**
 - Clusterizações
 - Detecção de Anomalias
- **Deep Learning / AI:**
 - NLP (Machine Learning com dados de texto e áudio)
 - Computer Vision (Machine Learning com dados de imagem e vídeos)
 - Automação e dispositivos inteligentes

Demonstração da Análise Exploratória



Projeto: Diminuição de Churn

Análise exploratória da base



Para verificar quais fatores estão mais relacionados com o Churn, vamos utilizar a técnica do **Information Value (IV)**. Essa medida é responsável por mensurar o “**grau de influência**” que uma variável possui sobre nossa variável target (no caso, o churn).

Se uma variável tem poder forte, isso significa que uma ou mais categorias da variável tem um alto ou baixo nível de churn, sendo útil estudá-la com mais profundidade.

Variável	IV	Grau de Influência no Churn
Qte Categoria	0,93	Muito forte
Idade	0,85	Muito forte
Estado	0,17	Médio
Programa Fidelidade	0,15	Médio
Limite Credito	0,11	Médio
Genero	0,07	Fraca
Score Credito	0,01	Fraca
Tempo Cliente	0,01	Fraca
Soma Pedidos Acum	0,00	Fraca
Usa Cartao Credito	0,00	Fraca

Empresa pode descartar essa variável, não precisando ter o custo de comprá-la para esta necessidade.

Projeto: Diminuição de Churn

Análise exploratória da base



Preditiva.ai

Qte Categoria	Churn			% Sim	% Não	Frequência de Churn	IV
	1	0	Total Geral				
1	1395	3689	5084	69%	46%	27%	0,09
2	343	4247	4590	17%	53%	7%	0,41
3	217	49	266	11%	1%	82%	0,29
4	58	2	60	3%	0%	97%	0,14
Total Geral	2013	7987	10000			20%	
IV Total							0,93

Variável inconclusiva. Vale entender o que está por trás das quantidades. Que produtos são esses?

Idade	Churn			% Sim	% Não	Frequência de Churn	IV
	1	0	Total Geral				
18-22	20	206	226	1%	3%	9%	0,02
23-27	53	741	794	3%	9%	7%	0,08
28-32	138	1632	1770	7%	20%	8%	0,15
33-37	247	2050	2297	12%	26%	11%	0,10
38-42	366	1653	2019	18%	21%	18%	0,00
43-47	411	773	1184	20%	10%	35%	0,08
48-52	346	324	670	17%	4%	52%	0,19
53-57	226	159	385	11%	2%	59%	0,16
>58	206	449	655	10%	6%	31%	0,03
Total Geral	2013	7987	10000			20%	
IV Total							0,85

Clientes acima de 43 anos têm maior taxa de Churn.

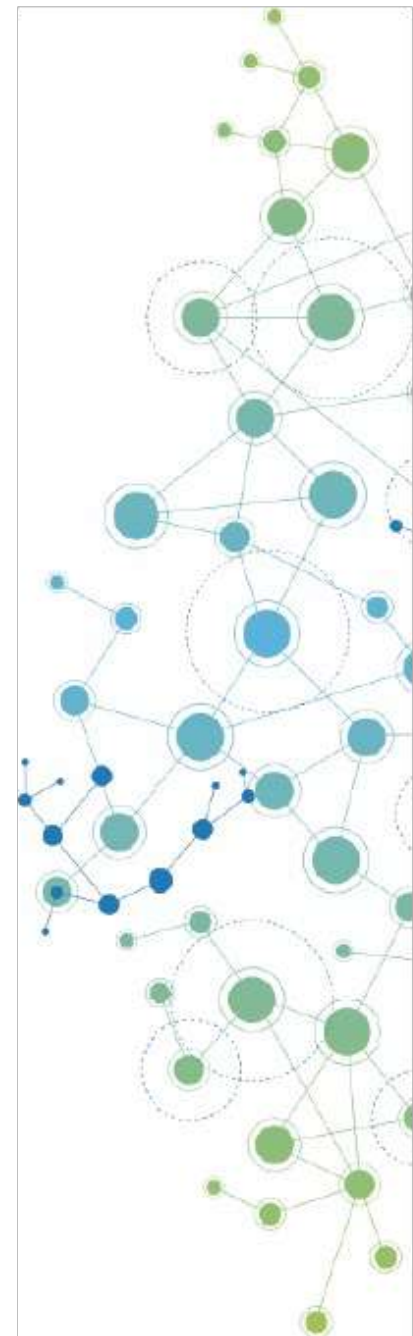
Estado	Churn			% Sim	% Não	Frequência de Churn	IV
	1	0	Total Geral				
Minas Gerais	804	1705	2509	40%	21%	32%	0,12
Rio de Janeiro	407	2070	2477	20%	26%	16%	0,01
São Paulo	802	4212	5014	40%	53%	16%	0,04
Total Geral	2013	7987	10000			20%	
IV Total							0,17

Clientes de Minas Gerais têm maior taxa de churn (32% - ou seja, 12 pp acima da média geral)

Programa Fidelidade	Churn			% Sim	% Não	Frequência de Churn	IV
	1	0	Total Geral				
0	1286	3563	4849	64%	45%	27%	0,07
1	727	4424	5151	36%	55%	14%	0,08
Total Geral	2013	7987	10000			20%	
IV Total							0,15

Clientes fora do programa de fidelidade têm maior taxa de churn (27% - ou seja, 7 pp acima da média geral)

3) Planos de ação sugeridos e próximos passos



Projeto: Diminuição de Churn

Planos de ação



Com base nos *insights* obtidos anteriormente, sugerimos os seguintes planos de ação:

Prioridade	O que fazer?	Por que?	Quem?	Como fazer?	Qual o custo?
1	Pesquisa Qualitativa entre usuários	Clientes acima de 43 anos ou clientes de Minas Gerais estão entre os maiores Churns	Time Discovery ou consultoria	Pesquisa Qualitativa entre usuários com perguntas sobre seu comportamento de uso do App e Concorrência na Região. Além disso, entender o que buscam e por que usam o App.	Salários e Tempo dos times envolvidos ou de consultoria externa
2	Incentivar a adesão ao Programa de Fidelidade	Clientes fora do programa de fidelidade têm maior taxa de churn (27% - ou seja, 7 pp acima da média geral)	Time de CRM	Estudar formas de estender e melhorar os benefícios de Programa de Fidelidade entre os usuários além de maior investimento em Remarketing para esses grupos de clientes.	Custo de Marketing, Ferramentas.
3	Revisar ou implementar uma política de descontos	Pode ser uma estratégia de curto prazo para incentivar o consumo	Time de CRM	Criar um grupo controle e teste com a nova política cupoms de desconto e analisar sua efetividade.	Diminuição da Margem Líquida. Além disso pode “viciar” os clientes em descontos e não ser sustentável a longo prazo.
4	Entrar em contato com clientes que estão com alta chance de Churn	Como alternativa à política de descontos, poderia-se mandar e-mails ou Push no App sobre ofertas.	Time de CRM e TI	Criar um grupo controle e teste com clientes que receberam o contato pró ativo e avaliar a efetividade.	Uso de uma ferramenta de Marketing para disparos

Projeto: Diminuição de Churn

Contexto de negócio e objetivo do trabalho



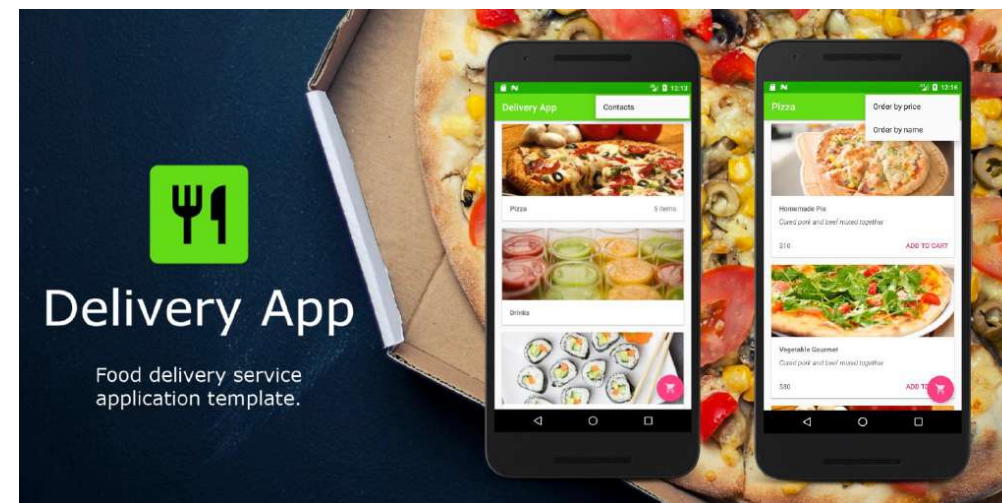
O aplicativo **ToComFome** é um app de entrega de comida por delivery. Como todo aplicativo de varejo, a empresa se preocupa com o Churn de clientes, ou seja, da perda desse cliente para a concorrência. Para entender o perfil de risco do cliente com maior chance de churn, a área de CRM gostaria de entender e classificar o cliente com risco de trocar de aplicativo nos próximos meses.

Para isso, a área de CRM passou uma amostra de cerca de **10 mil** clientes com suas respectivas informações de cadastro e transações nos **próximos 4 meses** a contar da data de extração usada como referência.

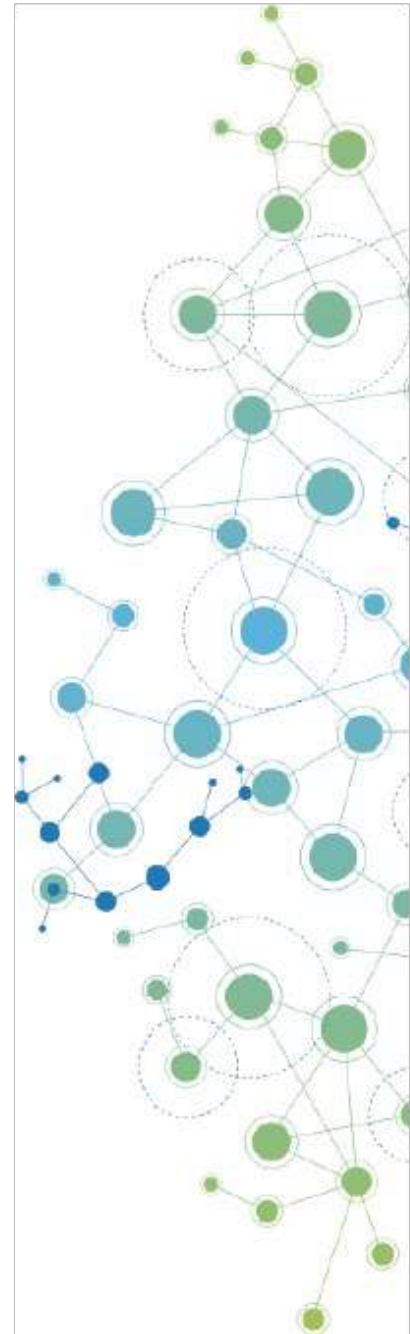
Obs: Um cliente é classificado como Churn quando nos próximos 4 meses em relação à data de referência, ele não transacionou por um período de 30 dias.

Objetivo do trabalho:

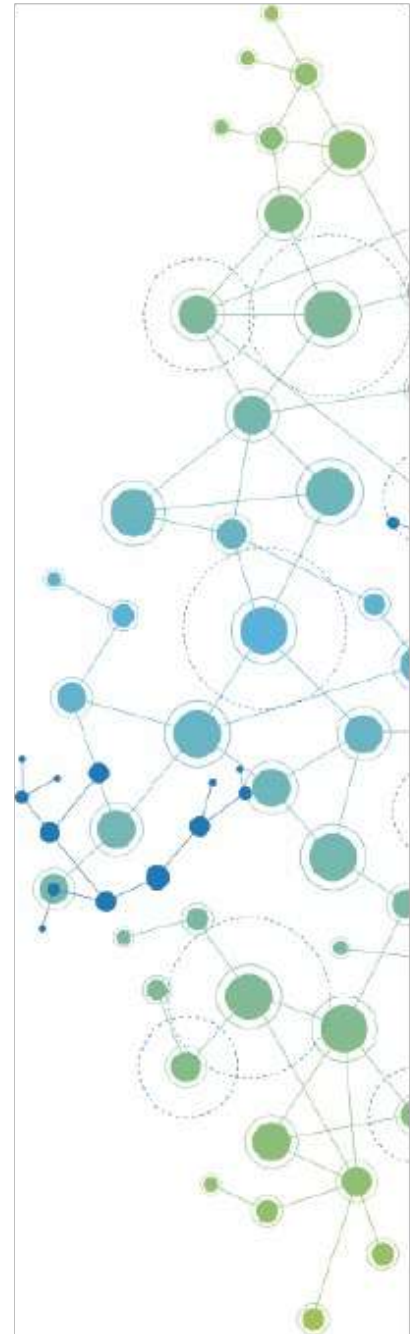
- ☐ Descobrir quais fatores de risco estão associados com o Churn de Clientes;
- ☐ Atribuir a probabilidade de darem Churn nos próximos **4 meses** em relação à data de extração de referência e;
- ☐ propor os possíveis planos de ação que a empresa pode tocar para diminuir esse problema.



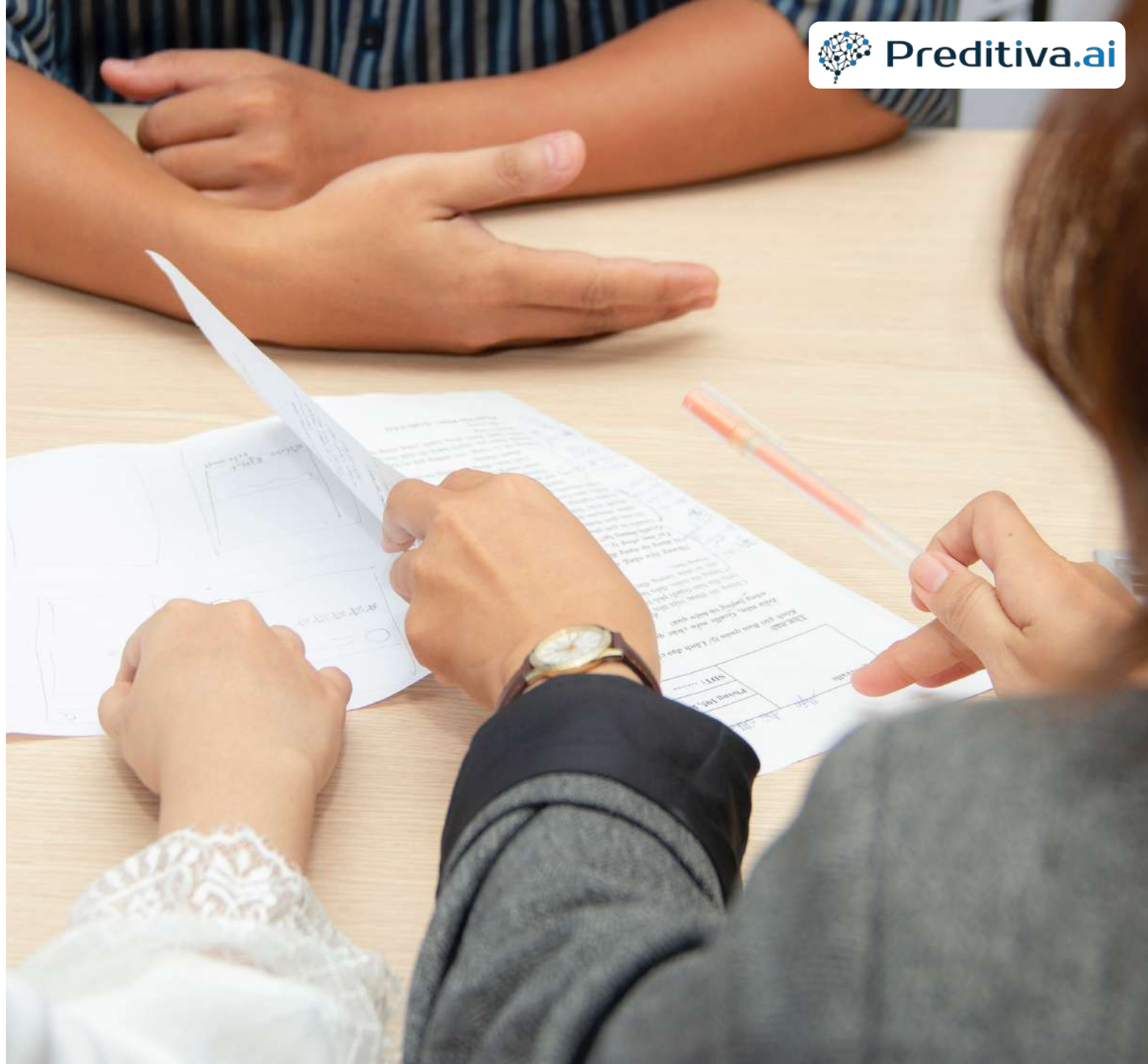
Criando um modelo de Machine Learning para atribuir a probabilidade



Mensurando o valor financeiro potencial



Você quer
acelerar seu
desenvolvimento
e se sentir **ainda
mais confiante**
para resolver
problemas como
mostrei aqui ?



Como **complemento** à nossa formação principal
“**Gerando Valor com Dados**”, criamos o programa:



ACELERA

Evolua 6 meses de experiência em Dados em apenas 1 mês

Como funciona?

ACELERA

Evolua 6 meses de experiência em dados em apenas 1 mês



○ Acelera é um **programa ao vivo** onde você vai:

- 1** **Resolver problemas reais** utilizando as **técnicas, ferramentas e metodologias** aprendidas na Plataforma Preditiva com o **apoio de mentores experientes**.
- 2** Ter a experiência de trabalhar em um **projeto de dados colaborativo**, com colegas que têm o mesmo objetivo que você e o **apoio de mentores experientes**.
- 3** Treinar seu **Data Storytelling** **apresentando suas análises** para os gestores de diversas áreas e obtendo feedback em tempo real.
- 4** Aprender em **apenas 1 mês**, de forma **100% prática** a extrair **insights dos dados**, e estar pronto para buscar sua transição de carreira.

ACELERA

Evolua 6 meses de experiência em dados em apenas 1 mês



Carga Horária: 15h

Serão 5 encontros ao vivo das 19h30 às 22h30 (3h cada encontro)



Certificado

Participando de **pelo menos 4 encontros**, você conquistará **seu certificado em Projetos de Dados**



Calendário

- **Dia 1:** 03-Abr-23: Onboarding, Team Building e Início do projeto
- **Dia 2:** 10-Abr-23: Desenvolvimento das Análises – Parte 1
- **Dia 3:** 17-Abr-23: Desenvolvimento das Análises – Parte 2
- **Dia 4:** 24-Abr-23: Avaliação dos Resultados e Preparação da Apresentação
- **Dia 5:** 08-Mai-23: Apresentação Final

Inscreva-se agora

Inscrições até 28/02



Dúvidas finais



Preditiva.ai