

DESAFIO CLUSTERING

DESAFIO IFOOD DE CLUSTERING

Por Gustavo Pilatti



05:24



iFood: quem somos?

Somos muito mais que um app. Somos uma empresa de tecnologia aplicada ao universo da alimentação. Se quiser um nome moderninho, também pode nos chamar de foodtech.

Queremos revolucionar o universo da alimentação por uma vida mais prática e prazerosa. O iFood é uma empresa brasileira que aproxima clientes, restaurantes e entregadores de forma simples e prática. E para proporcionar uma experiência incrível a cada um deles, nossa entrega vai muito além do delivery.

Por incrível que pareça, a nossa história começou fora do mundo digital, em 2011, com a Disk Cook. Um guia (impresso!) de cardápios, com uma central telefônica para onde você ligava e fazia o seu pedido. A ideia do iFood surgiu aí, para melhorar essa experiência.

Não demorou muito pra gente sair do papel e ganhar o Brasil. Já no ano seguinte, lançamos o nosso aplicativo e o site. E a paixão por alimentação e

DESAFIO CLUSTERING

As fusões com outras empresas fizeram o iFood crescer rapidamente. Em 2015, alcançamos o nosso primeiro milhão de pedidos. No ano seguinte, quase triplicamos esse número, impulsionados pela fusão com a Spoon Rocket.

Queremos continuar alimentando e desenvolvendo a comunidade que faz o iFood. Por isso, investimos em novas tecnologias e em expansão. Hoje, somos a FoodTech referência na América Latina e estamos presentes em dois países (Brasil e Colômbia), com o maior orgulho de ser uma empresa brasileira.

AVANÇAR

DESAFIO CLUSTERING

CONTEXTO DO DESAFIO

Vejam estas notícias:

Dados públicos de abril/21 - Em 2021, a empresa atingiu uma nova marca de 60 milhões de pedidos mensais. Crescimento de 100% em um ano; mais de 110.000 restaurantes se inscreveram na plataforma nos últimos 12 meses. Só em março, o mês mais letal da pandemia para o Brasil, 270 mil restaurantes operavam com a plataforma iFood.

Fonte: labsnews.com/en/articles/business/ifood-60-million-monthly-orders-amid-the-pandemic/

... o iFood confirma a preferência pelos hambúrgueres no delivery. Foram, em média, quatro unidades por segundo, totalizando mais de 100 milhões de pedidos de opções como x-tudo, x-bacon, hambúrguer, x-burguer e x-salada.

...E, no iFood, conveniência de comprar marmitas prontas também entrou no ranking dos 10 mais pedidos pelos clientes.

Fonte: www.gazetadopovo.com.br/bomgourmet/negocios-e-franquias/mais-pedidos-delivery-2021

... "De zero a dez, quanto você recomendaria esta empresa a um parente ou amigo?" o método de "promoção e/ou programa de fidelidade" é desaprovado por 30,5% dos respondentes, seguido de agilidade e tempo de entrega (17,7%) e facilidade do uso dos aplicativos (14,0%). Em contrapartida, a qualidade de produtos é a característica mais aprovada pelos respondentes, com 93,2% de "likes", seguida dos próprios restaurantes e estabelecimentos (89,0%) e entrega e/ou entregadores (86,9%).

A pergunta que dá vida à medição de NPS, de qualidade no atendimento, mostra o iFood como o aplicativo de delivery com melhor avaliação do mercado, com nota 59,8.

DESAFIO CLUSTERING

Fonte: exame.com/negocios/nps-ifood-e-melhor-e-mais-popular-app-de-delivery-diz-pesquisa/

A diversidade e amplitude de clientes que utilizam a plataforma do iFood é gigante - a cobertura do aplicativo se estende por muitas geografias no Brasil, de múltiplos gostos, regionalismos e bolsos. Alguns buscam uma experiência no momento da refeição, outros só querem receber a marmita na hora certa para não se atrasar para a primeira reunião pós-almoço.

Neste contexto, precisamos entender melhor nossos clientes! Quem é o mais fiel e quem mais compra conosco? Quem mais gosta de cupom? Quem gosta de entrega grátis? Qual o padrão de consumo? Qual o padrão de comportamento? Quanto os clientes vão gastar? Onde?

AVANÇAR

DESAFIO CLUSTERING

BASE DE DADOS

Aqui você pode ter acesso às bases de dados que você poderá usar neste desafio:

DOCUMENTAÇÃO BASE -
CLUSTERING

0 MB

MARKETING

1.6 MB

ORDERS

111.2 MB

SESSIONS

581.1 MB

Total de **20.249 clientes** diferentes, únicos e anonimizados, na base, que realizaram **226.269 pedidos**, em **7 meses** (fev20 a set20), através de mais

DESAFIO CLUSTERING

AVANÇAR

DESAFIO CLUSTERING

ROTEIRO PARA O DESAFIO

ENTENDIMENTO DO PROBLEMA

Está claro quem é o iFood, contexto e problema? O problema está definido de maneira SMART? É hora de usar todas as ferramentas que você tem ao seu dispor - de problem solving a ciclo analítico, tudo! Não menospreze esta etapa.

TRATAMENTO DOS DADOS

Muita calma! É importante conhecer os dados antes de analisá-los. Não há pegadinha na base de dados, mas ela é uma base real, então pode conter algum erro, duplicidade, missing, etc. Faça as validações necessárias (dados faltantes, duplicados, etc). Em alguns casos, devemos aplicar padronização ou normalização dos dados para que a técnica de clusterização tenha bons resultados.

EDA

Hora de olhar a relação entre as variáveis. Lembrem-se que quando queremos clusterizar alguma coisa, queremos a maior variância possível entre os grupos e a menor variância dentro de cada grupo. O EDA é o momento de entender a relação entre as variáveis. Sugestão: gere um entregável dessa etapa, algum dash ou apresentação com a explicação e definição da variável que você está definindo como objetivo.

PRODUTIZAÇÃO

Gerem os clusters em si! Definam variáveis que serão utilizadas, e o porque - evitem cair em alguma relação espúria.

Uma vez com a análise feita, não esqueça do “So what?” - como o iFood pode usar o que vocês estão entregando?

Apontar os pontos falhos e de risco do projeto - quais críticas conseguimos adicionar aos números e aos KPIs?

EDA PÓS CLUSTERING

DESAFIO CLUSTERING

voces proponham para cada cluster :

AVANÇAR

DESAFIO CLUSTERING

TOOLBOX, DICAS DE FERRAMENTAS

1. Problem solving é essencial!

Pense no problema antes de pensar e mexer nos dados, segure a ansiedade!
Passe pelas etapas do Problem Solving

Para te ajudar:

https://www.leadershipnow.com/leadingblog/2019/10/7_steps_to_builetproof_problem.html

<https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/how-to-master-the-seven-step-problem-solving-process>

2. Storytelling

Nem todo mundo sabe o que é clustering, e nem todo mundo tem tempo para aprender. A definição de “como” vocês vão convencer a sua audiência que a análise que foi feita é correta e deve ser implantada e qual o ganho que vai ser gerado, é muito importante!

Para te ajudar:

<https://mitsloan.mit.edu/ideas-made-to-matter/next-chapter-analytics-data-storytelling>

<https://www.forbes.com/sites/brentdykes/2016/03/31/data-storytelling-the-essential-data-science-skill-everyone-needs/?sh=52096b9e52ad>

<https://www.youtube.com/watch?v=Ez5yS4Q5AS>

3. Menos é mais!

Vocês receberam um monte de informação! Se atenha ao que é importante, não desperdice tempo tentando decifrar cada enigma que encontrarem na base de dados!

DESAFIO CLUSTERING

Essa pergunta é retórica, só para deixar mais claro um conceito que é um pouco abstrato: ética. No iFood, a ética é fundamental. Vamos falar um pouco sobre isso?

'Isso é a coisa certa a fazer?', 'Estou bem se alguém descobrir que eu fiz isso?', 'A sociedade aceitará o que estou fazendo se isso aparecer na primeira página de um grande jornal?'

A ética dos dados é a resposta “sim” às três perguntas anteriores.

Formalmente, “a ética representa a conversa contínua da sociedade – o mecanismo para atualizar as normas morais, regulamentos e leis da sociedade para se adaptar a questões emergentes, como avanços tecnológicos”. A ética nos dados abrange as obrigações morais de coletar, proteger e usar informações e como isso afeta os indivíduos e a sociedade.

Os dados e as tecnologias de dados estão em constante evolução. Pessoas e organizações podem não entender completamente que tipo de comportamentos e quais limites devem ser aplicados neste ambiente tão novo, com tantas oportunidades.

Às vezes, as regras e políticas em vigor podem não fornecer a orientação e o controle adequados sobre o comportamento das pessoas, ou por falta de especificidade ou porque estão desatualizados devido aos rápidos desenvolvimentos. Muitas vezes caímos em dilemas como: se eu fizer isso, não estou infringindo nenhuma lei, mas eu deveria fazer isso?

Apesar de aqui ser só um exercício didático, vale praticarmos esse tópico!

5. Por onde começar

A. Algumas sugestões de bibliotecas a serem utilizadas:

i. pandas

ii. matplotlib

iii. numpy

iv. seaborn

DESAFIO CLUSTERING

B. Exemplos de análise passo a passo

- i. <https://realpython.com/k-means-clustering-python/>
- ii. <https://machinelearningmastery.com/clustering-algorithms-with-python/> (exemplos de algoritmos)
- iii. <https://jakevdp.github.io/PythonDataScienceHandbook/05.11-k-means.html> (olhando k-means mais de perto)
- iv. <https://towardsdatascience.com/k-means-clustering-in-python-4061510145cc>
- v. <https://towardsdatascience.com/machine-learning-algorithms-part-9-k-means-example-in-python-f2ad05ed5203>

Mais algumas leituras gerais sobre o tema de clusterização:

<https://towardsdatascience.com/practical-steps-and-considerations-to-successfully-profile-your-customer-base-19f3f0991407>

<https://towardsdatascience.com/clustering-algorithms-for-customer-segmentation-af637c6830ac>



O QUE ACHOU DESTA AULA?

Deixe seu feedback para continuarmos melhorando sua experiência.

 1 MIN

AVALIAR

VOLTAR PARA O CURSO

DESAFIO CLUSTERING