O Impacto de Algoritmos Inteligentes em uma Empresa de Tecnologia: Méliuz

Arilo C. Dias-Neto[1,2], Daniel Fernandes[1], Kaio Wagner[1]

[1]Méliuz Veiculação e Divulgação Virtual (Méliuz)
Manaus – AM – Brasil

[2]Instituto de Computação – Universidade Federal do Amazonas
Manaus – AM – Brasil.

{arilo,daniel,kaio}@meliuz.com.br

Abstract. *Méliuz is a technology startup that in 2020 became the first to open shares (IPO) on the Brazilian Stock Exchange. This growth is due to the adoption of products based on intelligent algorithms and the adaptation of the development process of such products to the company's business strategy. This paper describes the history of the emergence of a team responsible for developing such solutions using intelligent algorithms, the process for developing such solutions, exemplified in two products, and some measurable results of the impacts that these products have generated on the company's business metrics.*

Resumo. *Méliuz é uma startup de tecnologia que se tornou em 2020 a primeira a abrir ações (IPO) na Bolsa de Valores Brasileira. Muito do seu crescimento se deve à adoção de produtos baseados em algoritmos inteligentes e a adequação do processo de desenvolvimento de tais produtos à estratégia de negócio da empresa. Este artigo descreve o histórico do surgimento de um time responsável por desenvolver tais soluções usando algoritmos inteligentes, o processo para o desenvolvimento de tais soluções, exemplificado em dois produtos, e alguns resultados mensuráveis dos impactos que tais produtos têm gerado em métricas de negócio da empresa.*

1. Sobre a Méliuz

Méliuz é uma startup de tecnologia que nasceu em 2011 em Belo Horizonte com foco em recompensar usuários que realizam compras em sites de comércio eletrônico devolvendo parte do dinheiro usado para compras, chamado de *cashback*. Desde seu surgimento, diversos novos produtos foram lançados, sempre visando prover benefícios financeiros aos seus usuários. Em 2016, foi lançada a oferta de *cashback* no caixa de lojas físicas, e em 2019 expandiu a oferta de cashback em compras físicas por meio do escaneamento de notas/cupons fiscais em qualquer tipo de estabelecimento. Em 2018 foi feito o lançamento do cartão de crédito Méliuz, em parceria com um banco já existente, no qual oferece aos usuários até 1,8% de cashback em todas as compras feitas a partir de sua utilização. Em 2020, foi criado um produto para aquisição de créditos em serviços oferecidos de forma digital (ex: iFood, Uber, Spotify, dentre outros) e para recargas de créditos para aparelhos celulares com contas pré-pagas, ambos retornando parte do valor usado em forma de *cashback*. Ainda em 2020, a Méliuz se tornou a primeira startup brasileira a abrir capital na Bolsa de Valores do Brasil (B3), com o código *CASH3*, representando um marco para os diversos empreendedores espalhados pelo Brasil, seja na área de tecnologia ou não.

Ao longo dos anos, a Méliuz tem investido bastante não apenas em tecnologia, mas também em seu time de tecnologia. De 12 engenheiros de software que atuavam em 2016

no time de tecnologia, passou a ter, em 2021, 70 engenheiros de software, 12 designers, 8 gerentes de produto, 12 engenheiros/cientistas de dados, e com uma projeção de abrir ao menos 100 vagas apenas para o time de tecnologia nos próximos semestres. Em seu pool de tecnologias, podem ser citadas plataformas como *React Native* para plataforma híbrida móvel, *React* como framework para o *frontend* de suas aplicações baseadas em *web browsers*, e *Phyton* e *Node* como plataformas para o *backend* das aplicações.

Sua base de contas cadastradas hoje é composta por mais de 18 milhões de contas. Como meta de evolução tecnológica, novos produtos estão sendo constantemente estudados e desenvolvidos com foco em **Serviços Financeiros** e em ***MarketPlace***.

2. O Caminho para o Uso de Algoritmos Inteligentes

Em 2016, a Méliuz passou por um processo de expansão de seu time de Engenharia, e junto com esse novo braço de desenvolvimento, havia também o sonho da criação de produtos e funcionalidades que fizessem uso de inteligência de dados. Ainda naquele ano, o primeiro microsserviço desenvolvido e colocado em produção a utilizar inteligência de dados foi o Sistema de Busca com *Autocomplete*, baseado em modelos de Recuperação de Informação, tornando-se um marco dentro da área de Engenharia e dando início a uma série de outros produtos que estariam por vir. Essa funcionalidade foi desenvolvida pelo time na época denominado de **Engenharia de Dados**, cuja formação inicial tinha apenas 3 pessoas.

Para continuar o desenvolvimento de produtos inteligentes, era necessário ter a matéria prima básica: dados dos usuários e de suas interações com a plataforma do Méliuz, de forma que isso pudesse servir de insumo para alimentar análises e algoritmos de recomendação, bem como para a geração de *insights* e metrificação dos produtos. No entanto, naquele momento não havia réplica do banco de produção ou um *Data Warehouse* que pudesse suprir tal necessidade, e a instrumentação das aplicações em produção também era algo que estava passando por um processo de amadurecimento. Para vencer os desafios da falta de visão sobre determinados fluxos e interações dos usuários, o time priorizou iniciativas de coletar, processar e armazenar dados em um ambiente centralizado que pudesse facilitar análises exploratórias em um primeiro momento. Tais dados passaram a ser consolidados em uma plataforma de dados em nuvem, robusta, escalável e que nos permitiu cruzar dados de diferentes fontes. Dessa forma, foi possível ter um arcabouço mais consistente para a realização de análises, experimentos e geração de *insights*.

A partir desse momento, o time de Engenharia de Dados cresceu em número de integrantes e uma série de produtos foram desenvolvidos posteriormente: coletores especializados em cupons e ofertas; mecanismos de expansão de audiência; diversas funcionalidades envolvendo recomendação de produtos e ofertas voltadas para segmentos de varejo e e-commerce. Tais produtos inteligentes são utilizados nas três plataformas de interação com os usuários: app (iOS e Android), site e extensão para navegadores web.

Com foco na crescente demanda por produtos inteligentes, bem como na necessidade de aumentar o potencial analítico das diversas áreas da Méliuz, foi criada em 2021 a **Área de Dados**, formada por três times: ***Analytics*** (cuida da modelagem e preparação dos dados do *Data Warehouse)*; ***Platform*** (cuida da infraestrutura e pipeline de dados analíticos)*;* ***Solutions*** (atua no desenvolvimento de modelos por meio de técnicas de *data science* e de produtos de inteligência de dados, como busca e recomendação.

3. Processo de Desenvolvimento de Produtos Inteligentes na Méliuz

A criação de produtos inteligentes dentro do time de *Data Solutions* começa a partir de uma ideia bruta ou necessidade vinda da área de produto. Em seguida, diversas etapas técnicas são realizadas até o momento de se colocar em produção a solução desenvolvida, para que esta seja avaliada do ponto de negócio. O processo em si é formado por 7 etapas, conforme a Figura 1. Tais etapas serão mais detalhadas nas próximas subseções.

Um dos grandes diferenciais deste processo é que não tentamos resolver todos os problemas na primeira versão do produto. São feitas validações de uma única hipótese por vez e versões prontas para produção são criadas a cada nova hipótese validada, como pode ser visto na Figura 1, pelo modelo cíclico de processo que é adotado. A primeira versão corresponde ao nosso baseline e, geralmente, representa a solução mais simples possível, que ajude a resolver parte do problema e a gerar impacto positivo em nossos usuários e nas métricas de negócio. Na prática, ao longo do projeto são testadas muitas hipóteses e ideias, com base na intuição, experiência e conhecimento prévio dos membros do time.

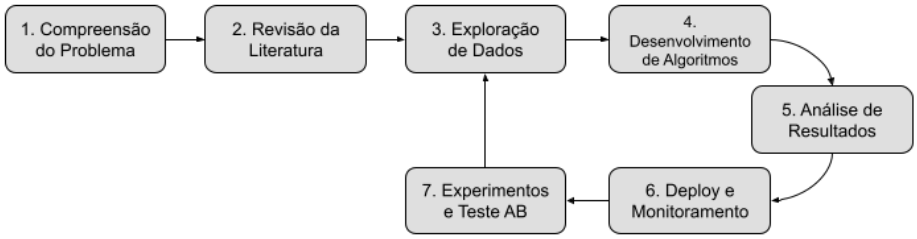

Figura 1. Processo de Desenvolvimento de Produtos Inteligentes na Méliuz.

3.1. Compreensão do Problema e Revisão da Literatura

Na etapa de Compreensão do Problema, o time de *Data Solutions* se reúne com o Gerente de Produto responsável pela demanda para alinhar objetivos e escopo do projeto, KPIs[1], prazos e discutir diferentes hipóteses para resolvê-lo. Assim é possível definir quais serão os entregáveis a cada ciclo e ter uma ideia das próximas hipóteses a serem testadas.

Na etapa de revisão da literatura, o time visa procurar por artigos acadêmicos no Google Scholar[2], pesquisar por projetos semelhantes no Github[3] ou Kaggle[4] e até mesmo buscar por soluções de empresas que resolveram problemas similares ao que está sendo solucionado no projeto em questão. O objetivo ao final desta etapa é obter uma lista de possíveis técnicas ou algoritmos que podem ser experimentadas para prover soluções ao problema que está sendo tratado.

3.2. Exploração de Dados, Desenvolvimento de Algoritmos e Análise de Resultados

Na etapa de exploração de dados, são realizadas análises nos conjuntos de dados obtidos e com isso extraímos hipóteses a partir de dados brutos. As melhores candidatas para resolver o problema são selecionadas e priorizadas.

Na etapa de desenvolvimento de algoritmos, são testados vários métodos clássicos de algoritmos de aprendizado de máquina, redes neurais e até mesmo combinações de algoritmos para gerar modelos ainda mais robustos.

---

[1] Key Process Indicator.

[2] https://scholar.google.com

[3] https://github.com

[4] https://www.kaggle.com

Na fase de análise de resultados, é calculada a qualidade do modelo treinado, explorando métricas técnicas, como precisão, revocação e F1. É de extrema importância que as métricas de negócio (ou KPIs) definidas no início do projeto estejam diretamente ligadas às métricas técnicas, pois qualquer melhoria na qualidade dos algoritmos, vai impactar diretamente a experiência dos usuários que utilizarem tais produtos.

3.3. Deploy, Monitoramento, Experimentos e Teste AB

A etapa seguinte é o deploy e monitoramento de soluções. É nesta etapa que acontece o que chamamos de "produtização", que significa adaptar o código do modelo desenvolvido para que funcione de forma escalável em ambientes reais de produção. Além disso, é de extrema importância que o time faça o monitoramento do desempenho do modelo, pois isso pode tanto ajudar na detecção de problemas não percebidos durante o desenvolvimento quanto na mudança na distribuição dos dados sobre o qual o modelo opera, o que impactaria negativamente no seu desempenho.

Ao término desta etapa, o ciclo de desenvolvimento encerra com a validação da hipótese que tenta resolver o problema com mais eficiência/qualidade do que a versão atual em produção. É importante ressaltar que temos uma cultura de testes muito forte dentro da empresa, e isso chega até o time de *Data Solutions*. Sendo assim, para testar a eficácia de um novo modelo, é aplicado um método chamado Teste A/B. Nele, o novo modelo é implantado para uma parte da base dos usuários, enquanto o restante da base continua exposta à versão anterior de forma simultânea. Essa metodologia é fundamental no nosso processo, pois nos permite uma comparação do efeito de qualquer métrica de negócio entre dois ou mais grupos na base de usuários.

O processo de desenvolvimento de produtos inteligentes na Méliuz é fruto de uma experiência prática construída de forma contínua ao longo do tempo e que está em constante evolução. Com seu amadurecimento, será possível construir uma cultura de resolução de problemas por meio de hipóteses baseadas em dados, o que nos permitirá errar menos e agregar rapidamente valor aos nossos usuários.

4. Resultados de Produtos Inteligentes na Méliuz

A contínua evolução no desenvolvimento de produtos inteligentes fez com que estes passassem a desempenhar um papel importante na Méliuz em relação ao impacto no faturamento e também na retenção de usuários. Atualmente, compras que passam por fluxos impactados por produtos inteligentes representam cerca de 38% do faturamento do E-Commerce na Méliuz e a meta é chegar a 45% até o fim de 2021. Dentre as principais funcionalidades que se destacam nesse contexto, podem ser citadas: o Rastreio Inteligente na Extensão (Figura 2 - esquerda) e a Sidebar de Recomendação (Figura 2 - direita).

Figura 2. Tela das funcionalidades Rastreio Inteligente e *Sidebar* de Recomendação.

4.1. Rastreio Inteligente

Compreensão do Problema: o fluxo de atribuição de uma compra (e consequentemente seu *cashback*) em um parceiro do E-commerce a um usuário Méliuz ocorre a partir do evento de ativação do *cashback* feito pelo usuário dentro da plataforma Méliuz (Site, App ou Extensão para navegadores web), sendo este redirecionado em seguida para o site do parceiro. O problema é que se durante o fluxo de eventos entre a ativação do *cashback* e a finalização da compra o usuário acessar outras plataformas comissionadas no E-commerce (ex: acesso a sites de cupons, de comparação de preços ou de outros benefícios), há um risco de sua compra ser atribuída a esta outra plataforma, e desta forma a Méliuz não é comissionada e o usuário não ganharia o seu *cashback*, visto que a regra comum dos sites de E-commerce é atribuir a compra ao último local visitado pelo usuário (chamado de *last-click*). Desta forma, foi desenvolvida uma funcionalidade que ajuda o usuário da Extensão Méliuz a evitar a potencial perda de um *cashback* devido a um fluxo indevido de navegação do usuário, como citado acima.

Revisão da Literatura, Extração de Dados e Desenvolvimento de Algoritmos: com base em análises de dados históricos de sessões de usuários providos pós ativação da Extensão, optou-se por tratar o problema com *Machine Learning,* fazendo uso do algoritmo de Regressão Logística para indicar se uma determinada ação do usuário poderá impactar na perda de atribuição da comissão para a Méliuz e, consequentemente, impactando também no *cashback*. Uma vez identificado esse potencial risco, o cashback na Extensão fica com status desativado e uma mensagem é exibida ao usuário para que ele possa optar por uma nova ativação, conforme pode ser observado na Figura 2 - esquerda.

Análise de Resultados: apenas no 1° trimestre de 2021, foram recuperadas 2.172 compras que potencialmente teriam sido perdidas sem a intervenção desta funcionalidade junto ao usuário para proceder com a reativação do cashback.

4.2. *Sidebar* de Recomendação

Compreensão do Problema: um grande volume de consultas por produtos a serem comprados por meio da Internet ocorre nos motores de busca, como por exemplo o Google. Muitas vezes, usuários Méliuz possuem dificuldade em descobrir que um determinado produto é vendido por algum parceiro Méliuz e com um valor de *cashback* que possa ser tão atraente, a ponto de se tornar uma excelente oportunidade de economia. Assim, esta funcionalidade tem o comportamento de uma vitrine virtual para apresentar ofertas dos parceiros Méliuz quando os nossos usuários estão procurando por produtos em motores de busca. A recomendação é feita pela Extensão Méliuz, por meio de notificações suspensas que aparecem após o usuário digitar a sua pesquisa, conforme mostrado na Figura 2 - direita.

Revisão da Literatura, Extração de Dados e Desenvolvimento de Algoritmos: o grande desafio dessa funcionalidade é recomendar ofertas relacionadas ao produto desejado a partir de uma base com dezenas de milhões de ofertas dos nossos parceiros. Este é um problema aberto e bastante discutido na literatura técnica, que não apresenta solução trivial, conhecido como *product matching* (Köpcke et al., 2012*)*. A nossa abordagem envolve processamento de linguagem natural (PLN) para representar e extrair as características das ofertas a partir de conteúdo textual não estruturado, como o título e a descrição. Uma vez representadas, as ofertas são então submetidas a um modelo preditivo, capaz de determinar se elas correspondem ou não ao mesmo produto. Este modelo foi treinado usando o algoritmo de *Machine Learning* supervisionado SVM

(*Support Vector Machine*)(Drucker et al., 1997) a partir de uma base de ofertas rotulada. Tal abordagem de *matching* encontra-se em constante evolução, a fim de garantir a qualidade dos resultados apresentados aos usuários.

Análise de Resultados: no primeiro trimestre de 2021, a Sidebar de Recomendação foi responsável por 12.947 compras de usuários.

4.3. Resultados de Outros Produtos Inteligentes

A Tabela 1 descreve resultados de outros produtos da plataforma Méliuz que utilizam algoritmos inteligentes:

Tabela 1. Resultados de Produtos Inteligentes na Méliuz no 1° Trimestre de 2021.

Produto	Métrica de Sucesso (KPI)	Solução Técnica Utilizada	Resultado Obtido (Q12021)
Busca por parceiros e produtos com autocomplete	Compras realizadas	Busca: Modelo Vetorial Autocomplete: Variação da técnica proposta em (Deng et al., 2016)	292 mil compras no 1tri de 2021
Categorizador automático de cupons e ofertas	Cupons categorizados	Modelo utilizando uma combinação de algoritmos de classificação: *Random Forest* e *XGBoost*.	1.7 mil cupons e ofertas categorizados
Score de engajamento de usuários Méliuz	Usuários com score calculado	Modelo estatístico calculado com base em eventos de interação do usuário com a plataforma Méliuz.	100% dos usuários clusterizados em grupos de acordo com score calculado

Como observado ao longo do artigo, a adoção de algoritmos inteligentes tem sido uma constante dentro da área de Produto e Dados da Méliuz, causando não apenas um diferencial tecnológico, que contribui para a captação de novos talentos, desenvolvimento do time de engenheiros e cientistas de dados, mas também provendo um impacto positivo no negócio da empresa. Isso demonstra que o equilíbrio entre o tempo requerido para o desenvolvimento de inovação e pesquisa, e a urgência para que resultados de valor sejam obtidos, é algo que pode ser alcançado dentro de uma empresa, e esses dois universos, academia e indústria, são impactados positivamente um pelo outro.

Referências Bibliográficas

Deng, Dong, et al. "META: an efficient matching-based method for error-tolerant autocompletion." Proceedings of the VLDB Endowment 9.10 (2016): 828-839.

Drucker, Harris; Burges, Christ. C.; Kaufman, Linda; Smola, Alexander J.; and Vapnik, Vladimir N. (1997); "Support Vector Regression Machines", in Advances in Neural Information Processing Systems 9, NIPS 1996, 155–161, MIT Press.

Köpcke, H; Thor, A.; Thomas, S.; Rahm, E. (2012) “Tailoring entity resolution for matching product offers”. In: EDBT '12: Proceedings of the 15th International Conference on Extending Database Technology. March 2012 Pages 545–550.