

Projeto em Business Intelligence e Analytics
Fase 1

Nome do estudante: Rafael de Menezes Ehlers

Predição de Avaliações Negativas no Ecossistema Olist: Uma Abordagem Integrada de Business Intelligence e Machine Learning

Resumo

Em um cenário em que o comércio eletrônico brasileiro ultrapassou R\$ 204 bilhões em faturamento em 2024, a insatisfação dos clientes representa um risco direto ao Lifetime Value (LTV) e à sustentabilidade financeira das plataformas de varejo digital. O presente projeto propõe uma solução integrada de Business Intelligence e Analytics capaz de diagnosticar os fatores determinantes da insatisfação de clientes no ecossistema da Olist e prever o risco de avaliações negativas em pedidos futuros. A pesquisa adota uma abordagem quantitativa e aplicada, baseada na metodologia CRISP-DM, e utiliza o dataset público da Olist (aproximadamente 100 mil pedidos realizados em marketplaces brasileiros entre 2016 e 2018) como base empírica. A arquitetura proposta combina uma camada descritiva, com modelagem dimensional em Star Schema e dashboards interativos contendo KPIs estratégicos, com uma camada preditiva, baseada em algoritmos de classificação supervisionada (Regressão Logística e Random Forest). Três hipóteses orientam a investigação: (i) o atraso na entrega reduz a probabilidade de avaliação positiva; (ii) em produtos de baixo ticket, o valor do frete impacta mais negativamente a satisfação do que o preço; e (iii) pedidos multi-vendedor apresentam maiores taxas de avaliação negativa devido ao tempo adicional de consolidação logística. Esta Fase 1 estabelece as bases conceituais e o plano de execução, prevendo cronograma trimestral para 2026 e contemplando as etapas de ETL, modelagem dimensional, construção de dashboards, treinamento e validação do modelo preditivo, com avaliação por meio das métricas de acurácia, precisão, recall e F1-score.

Palavras-chave: Business Intelligence; Machine Learning; E-commerce; Olist; Predição de Satisfação; Churn.

1 Introdução

O comércio eletrônico brasileiro consolidou-se como principal motor do varejo digital nacional, alcançando faturamento de R\$ 204,3 bilhões em 2024, com crescimento de 10,5% em relação ao ano anterior, 414,9 milhões de pedidos e 91,3 milhões de compradores únicos [Ass25]. Esse cenário de expansão acelerada eleva a competição entre marketplaces e pressiona os Custos de Aquisição de Clientes (CAC), tornando estratégias focadas em retenção de consumidores e maximização do Lifetime Value (LTV) críticas para a sustentabilidade financeira do setor [PF16]. O ecossistema de marketplaces, exemplificado pela plataforma Olist, enfrenta o desafio adicional de coordenar centenas de vendedores independentes, parceiros logísticos e as expectativas crescentes dos consumidores finais.

Nesse contexto, o presente projeto aborda o problema da insatisfação dos clientes e o consequente risco de churn (abandono da plataforma), gerados por fricções na jornada de compra, com ênfase nos impactos dos prazos logísticos e custos de frete sobre a percepção de valor. Pesquisas recentes evidenciam que o tempo de entrega e a precisão do pedido figuram entre os principais determinantes da satisfação do consumidor em e-commerce, com modelos preditivos atingindo até 92% de acurácia ao incorporar essas variáveis [ZBR24]. A meta central deste trabalho é propor uma solução integrada de Business Intelligence (BI) e Analytics capaz de identificar padrões históricos de insatisfação e prever o comportamento de avaliações futuras. Para tanto, será utilizado o dataset público da Olist, que contém aproximadamente 100 mil pedidos realizados em marketplaces brasileiros entre 2016 e 2018 [OS18].

Este documento está estruturado da seguinte forma: a Seção 2 apresenta o Referencial Teórico com os conceitos fundamentais do projeto; a Seção 3 realiza a Revisão da Literatura mapeando trabalhos correlatos; a Seção 4 detalha o Plano de Trabalho, englobando a descrição do problema, hipóteses, objetivos, cronograma, metodologia e desenho da solução; e a Seção 5 expõe as considerações finais e conclusões desta fase de planejamento.

2 Referencial Teórico

Para fundamentar a compreensão deste projeto, estabelecem-se os pilares teóricos divididos em três domínios principais: Ciência de Dados e Aprendizado de Máquina, Business Intelligence e Visualização de Dados, e Comércio Eletrônico e Comportamento do Consumidor.

- **Ciência de Dados e Aprendizado de Máquina:** De acordo com Provost e Fawcett [PF16], a Ciência de Dados compreende o ciclo de extração de conhecimento a partir de grandes volumes de dados, abrangendo etapas de coleta, limpeza, exploração, modelagem e avaliação. No contexto preditivo deste projeto, será empregado o Aprendizado de Máquina Supervisionado, paradigma em que o algoritmo aprende a mapear variáveis de entrada (recursos logísticos e financeiros) para um desfecho conhecido (avaliação positiva ou negativa). Entre os algoritmos de classificação consolidados na literatura, destacam-se a Regressão Logística, que estima a probabilidade de um evento binário, e o Random Forest, método de ensemble baseado em múltiplas árvores de decisão que reduz overfitting e captura interações não-lineares entre variáveis [PF16]. Antes da modelagem, o pré-processamento dos dados (incluindo tratamento de valores ausentes, codificação de variáveis categóricas e engenharia de features) é etapa crítica para o desempenho do modelo. A avaliação utilizará métricas como acurácia, precisão, recall e F1-score, especialmente relevantes em problemas com classes desbalanceadas, como é o caso de avaliações negativas que representam parcela minoritária dos pedidos.
- **Business Intelligence e Visualização de Dados:** O Business Intelligence (BI) refere-se ao conjunto de técnicas, processos e ferramentas voltados à transformação de dados operacionais em informação útil para a tomada de decisão estratégica. Um dos pilares dessa disciplina é o processo de Extração, Transformação e Carga (ETL), responsável por integrar dados de múltiplas fontes em um repositório analítico unificado. Esse repositório, denominado Data Warehouse, é estruturado segundo princípios de modelagem dimensional. O modelo mais difundido é o Star Schema, composto por uma tabela fato central, que armazena medidas numéricas, conectada a tabelas de dimensão, que contextualizam essas medidas com atributos descritivos [KR13]. A camada de visualização, materializada em ferramentas como Power BI e Tableau, traduz os dados estruturados em dashboards interativos, permitindo a análise de KPIs estratégicos e a identificação de padrões e anomalias por usuários não-técnicos.
- **Comércio Eletrônico e Comportamento do Consumidor:** O comércio eletrônico evoluiu de simples vitrines online para ecossistemas complexos de marketplaces, em que uma plataforma central intermedeia transações entre múltiplos vendedores independentes e consumidores finais. Nesse contexto, a Análise RFM (Recência, Frequência e Valor Monetário) é amplamente utilizada para segmentar a base de clientes a partir do comportamento histórico de compra, identificando perfis de maior e menor valor para o negócio [PF16]. O conceito de churn, originalmente cunhado em setores de assinatura, ganha contornos particulares no varejo digital: caracteriza-se pela interrupção tácita das compras após uma experiência de atrito, sem que haja um evento explícito de cancelamento [Lui19]. A satisfação percebida pelo cliente é frequentemente mensurada por

indicadores como o Net Promoter Score (NPS), métrica que estima a probabilidade de recomendação da plataforma e está diretamente correlacionada à fidelização e ao Lifetime Value (LTV) do consumidor.

3 Revisão da Literatura

A aplicação de Business Intelligence e Analytics para mitigar problemas no varejo eletrônico tem sido amplamente discutida na literatura recente, com trabalhos buscando correlacionar dados operacionais e logísticos com o comportamento do consumidor. Luizon [Lui19] realizou uma revisão bibliométrica no Scopus seguida da aplicação prática de três algoritmos (Árvore de Decisão J48, Multilayer Perceptron e Support Vector Machine) em uma base de e-commerce de varejo, evidenciando que árvores de decisão entregam o melhor custo-benefício entre acurácia e tempo de execução. Wangkiat e Polprasert [WP23] trabalharam diretamente sobre o dataset público da Olist, propondo um método de engenharia de features centrado na duração da entrega e na avaliação média do produto para prever a satisfação do cliente, comparando Random Forest, Regressão Logística e K-Nearest Neighbors. Mais recentemente, Zaghloul, Barakat e Rezk [ZBR24] confrontaram modelos tradicionais de Machine Learning com abordagens de Deep Learning em uma base com mais de 100 mil pedidos, demonstrando que Random Forest atingiu 92% de acurácia e que tempo de entrega e precisão do pedido são os principais determinantes da satisfação. A Tabela 3.1 sintetiza as características desses trabalhos correlatos e as lacunas que justificam a presente pesquisa.

Tabela 3.1: Comparativo de trabalhos correlatos

Autor (Ano)	Metodologia Utilizada	Foco do Problema	Variáveis Analisadas	Lacuna Identificada
Luizon (2019)	Revisão bibliométrica (Scopus) + Árvore de Decisão (J48), MLP e SVM (SMO) aplicados em base de e-commerce de varejo.	Predição de churn em e-commerce sem assinatura.	Comportamento de compra (estático, sem detalhamento logístico).	Análise estática; não integra variáveis logísticas (frete, prazo) nem modelagem dimensional para BI.
Wangkiat & Polprasert (2023)	Random Forest, Regressão Logística e K-NN (Scikit-learn) sobre o dataset Olist.	Predição de satisfação em 4 classes agrupadas (Low, Average, Good, Excellent) a partir das 5 estrelas originais do dataset Olist.	Duração da entrega, rating médio do produto.	Foco apenas no modelo preditivo; sem camada descritiva (BI/Star Schema) nem dashboards integrados.
Zaghloul, Barakat & Rezk (2024)	Comparativo entre Random Forest, SVM e MLP (Deep Learning) em 100k pedidos.	Predição de satisfação futura em e-commerce de varejo.	Tempo de entrega, valor do pedido, localização.	Dataset internacional genérico; não contempla particularidades de marketplace multi-vendedor brasileiro.

Resumo do conhecimento disponível

- A predição de avaliações negativas e de churn no varejo digital é tema ativo na literatura recente, com publicações relevantes entre 2019 e 2024.
- O dataset público da Olist é amplamente utilizado como benchmark para problemas de classificação de satisfação em e-commerce brasileiro.
- Random Forest figura consistentemente entre os algoritmos de melhor desempenho, atingindo até 92% de acurácia em estudos recentes [ZBR24], seguido por Regressão Logística e Árvores de Decisão como baselines competitivos.
- O tempo de entrega e a precisão do pedido são apontados como os principais determinantes da satisfação do consumidor no varejo eletrônico.
- A maioria dos trabalhos foca exclusivamente na camada preditiva (Machine Learning), sem integrar uma camada descritiva de Business Intelligence (Star Schema, dashboards e KPIs) com a camada analítica.

- O contexto específico de marketplaces brasileiros multi-vendedor é pouco explorado na literatura internacional, especialmente quanto à influência da consolidação logística entre múltiplos sellers sobre a satisfação do cliente.

4 Plano de Trabalho

4.1 Descrição do Problema

O ecossistema Olist lida com uma vasta rede de vendedores independentes, o que gera heterogeneidade no tempo de postagem, qualidade de embalagem e acurácia logística. Quando um consumidor final experimenta atrasos na entrega ou custos de frete desproporcionais, a sua insatisfação é manifestada através de notas baixas nas avaliações do pedido (1 a 3 estrelas). O problema reside em mapear quais variáveis operacionais são os reais detratores da experiência do usuário e como agir preventivamente antes que o cliente abandone o canal de compras.

Pergunta de Pesquisa: Quais fatores operacionais, logísticos e financeiros possuem maior peso estatístico na determinação de uma avaliação negativa por parte do cliente no e-commerce brasileiro?

4.2 Hipóteses

Hipótese 1 (H1): Pedidos cujo tempo real de entrega ultrapassa o prazo estimado comunicado no ato da compra apresentam probabilidade significativamente maior de receberem avaliação negativa.

- **Variáveis Independentes (preditoras):**

- **Tempo real de entrega:** número de dias entre a postagem do pedido e a entrega efetiva ao cliente.
- **Prazo estimado de entrega:** número de dias entre a confirmação do pedido e a data prometida ao cliente no ato da compra.
- **Diferença de prazo (delta):** variável calculada como tempo real menos prazo estimado, em dias.
- **Status de atraso:** variável binária (sim/não), derivada de delta maior que zero.

- **Variável Dependente (desfecho):**

- **Avaliação negativa:** variável binária derivada da nota da review (1 se review_score menor ou igual a 3, 0 caso contrário).

Hipótese 2 (H2): Em categorias de produtos de baixo ticket médio (inferior a R\$ 50,00), o valor cobrado pelo frete exerce maior impacto negativo na satisfação do consumidor do que o preço nominal do produto.

- **Variáveis Independentes (preditoras):**

- **Valor do frete:** valor em R\$ cobrado pelo envio do pedido.
- **Preço do produto:** valor unitário do produto, em R\$.

- **Razão frete/preço:** variável calculada como o valor do frete dividido pelo preço do produto.
- **Categoria do produto:** classificação categórica do item (variável categórica).
- **Faixa de ticket:** variável categórica derivada (baixo se inferior a R\$ 50, médio se entre R\$ 50 e R\$ 200, alto se superior).

- **Variável Dependente (desfecho):**

- **Avaliação negativa:** variável binária derivada da nota da review (1 se review_score menor ou igual a 3, 0 caso contrário).

Hipótese 3 (H3): Pedidos que contêm itens de múltiplos vendedores geram maior tempo de consolidação logística e, conseqüentemente, apresentam taxas mais elevadas de avaliações negativas.

- **Variáveis Independentes (preditoras):**

- **Número de vendedores únicos por pedido:** contagem de sellers distintos associados ao mesmo pedido.
- **Pedido multi-vendedor:** variável binária (1 se a contagem de vendedores únicos for maior que 1, 0 caso contrário).
- **Tempo de consolidação logística:** número de dias entre a confirmação do pedido e a alteração do status para “shipped”.
- **Número de itens por pedido:** contagem total de itens no pedido.

- **Variável Dependente (desfecho):**

- **Avaliação negativa:** variável binária derivada da nota da review (1 se review_score menor ou igual a 3, 0 caso contrário).

4.3 Objetivos

Objetivo Geral: Desenvolver uma solução integrada de Business Intelligence e Analytics para diagnosticar os fatores determinantes da insatisfação de clientes no e-commerce e prever o risco de avaliações negativas em pedidos futuros.

Objetivos Específicos:

1. **Implementar** um pipeline de Extração, Transformação e Carga (ETL) para unificar os dados relacionais do dataset público da Olist em uma base analítica única, tratando registros nulos, duplicidades e inconsistências de tipo.
2. **Modelar** a base analítica em formato dimensional (Star Schema), com tabela fato de pedidos e dimensões de cliente, produto, vendedor, tempo e geolocalização, para suportar consultas analíticas eficientes.

3. **Desenvolver** dashboards interativos contendo KPIs estratégicos (faturamento, prazo médio de entrega, NPS estimado e taxa de atrasos), com filtros por dimensão temporal e geográfica.
4. **Treinar e validar** um modelo de Machine Learning supervisionado (Random Forest e Regressão Logística como baselines) para classificar pedidos com alto risco de receberem avaliações negativas, avaliando seu desempenho por meio das métricas de acurácia, precisão, recall e F1-score.
5. **Identificar** os fatores logísticos, financeiros e operacionais com maior peso estatístico na predição de avaliações negativas, por meio da análise de importância de features do modelo treinado, respondendo diretamente à pergunta de pesquisa.

4.4 Atividades e Cronograma

O projeto será executado em regime trimestral ao longo de 2026, com as atividades estruturadas em correspondência direta aos procedimentos detalhados na Seção 4.5. A distribuição contempla os três primeiros trimestres do ano, reservando o quarto trimestre como buffer para revisões finais e eventual extensão de escopo na Fase 2. A Tabela 4.1 apresenta a distribuição temporal das atividades.

Tabela 4.1: Cronograma de Atividades para 2026

Atividades	1º Trim.	2º Trim.	3º Trim.
1. Entendimento do negócio e definição da variável-alvo	X		
2. Coleta dos dados (download do dataset Olist)	X		
3. Preparação dos dados (ETL e variáveis derivadas)	X		
4. Modelagem dimensional e carga em Star Schema	X	X	
5. Construção dos dashboards de BI		X	
6. Treinamento do modelo preditivo (RF e Regressão Logística)		X	X
7. Avaliação e interpretação dos modelos			X
8. Documentação e publicação dos resultados			X

4.5 Metodologia

A pesquisa adota uma abordagem quantitativa e aplicada, estruturada metodologicamente nas dimensões de Materiais, Métodos e Procedimentos.

Materiais

- Dataset público “Brazilian E-Commerce Public Dataset by Olist” disponível na plataforma Kaggle [OS18], composto por nove arquivos CSV relacionais que cobrem aproximadamente 100 mil pedidos realizados entre 2016 e 2018, contemplando informações de clientes, pedidos, pagamentos, avaliações, produtos, vendedores e geolocalização (CEP e coordenadas).

- Ambiente de execução Google Colab, com suporte nativo a notebooks Jupyter.
- Linguagens e bibliotecas: Python 3 com Pandas e NumPy para manipulação dos dados, Scikit-learn para algoritmos de classificação e métricas, Matplotlib e Seaborn para análise exploratória, e SQL para consultas analíticas sobre o Data Warehouse.
- Ferramentas de visualização: Power BI ou Tableau para construção dos dashboards interativos.

Métodos

- **Pesquisa aplicada e quantitativa:** o trabalho produz conhecimento prático aplicável à decisão de negócio no varejo digital, com tratamento numérico das variáveis.
- **Framework CRISP-DM (Cross-Industry Standard Process for Data Mining):** as fases de Entendimento do Negócio, Entendimento dos Dados, Preparação dos Dados, Modelagem, Avaliação e Implantação organizam o ciclo de execução.
- **Modelagem dimensional:** a base preparada é estruturada em Star Schema [KR13], com tabela fato de pedidos e dimensões de cliente, produto, vendedor, tempo e geolocalização.
- **Aprendizado supervisionado:** emprego de algoritmos de classificação binária para a variável-alvo “avaliação negativa” (sim ou não), utilizando Regressão Logística como baseline e Random Forest como modelo principal [PF16].
- **Avaliação de desempenho:** acurácia, precisão, recall, F1-score e matriz de confusão, com atenção especial ao desbalanceamento de classes, dado que avaliações negativas correspondem a parcela minoritária dos pedidos.

Procedimentos

- **Entendimento do negócio:** mapear os KPIs de e-commerce pertinentes à satisfação do cliente (faturamento, NPS estimado, prazo médio de entrega, taxa de atrasos) e definir a variável-alvo do modelo preditivo.
- **Coleta dos dados:** realizar o download dos nove arquivos CSV do dataset Olist diretamente da plataforma Kaggle.
- **Preparação dos dados (ETL):** integrar as tabelas relacionais via Pandas, tratar registros nulos, eliminar duplicidades e calcular variáveis derivadas (diferença em dias entre entrega estimada e real, distância geográfica vendedor-cliente, razão frete sobre preço, contagem de vendedores únicos por pedido).
- **Modelagem dimensional:** estruturar a base preparada em Star Schema e carregá-la em banco SQL para suporte às consultas analíticas e aos dashboards.
- **Construção dos dashboards:** desenvolver visualizações interativas em Power BI ou Tableau, conectadas ao Data Warehouse, contemplando os KPIs estratégicos com filtros por dimensão temporal e geográfica.

- **Treinamento do modelo preditivo:** dividir a base em conjuntos de treino e teste, aplicar Regressão Logística e Random Forest no Scikit-learn, e otimizar hiperparâmetros via validação cruzada.
- **Avaliação e interpretação:** calcular as métricas de desempenho, analisar a matriz de confusão e extrair a importância de features do Random Forest para responder à pergunta de pesquisa.
- **Documentação e publicação:** consolidar os resultados em relatório final e publicar os dashboards conectados à base tratada.

4.6 Desenho da Solução

A arquitetura de dados proposta para este projeto foi concebida para garantir a separação clara de responsabilidades entre as fases descritivas, voltadas ao Business Intelligence, e as fases preditivas, baseadas em Analytics. Conforme ilustrado na Figura 4.1, o pipeline organiza-se em seis camadas sequenciais: (1) coleta dos dados brutos a partir do dataset público da Olist; (2) processamento e ETL em Python e Pandas, produzindo dados limpos e integrados; (3) carga em SQL com modelagem dimensional em Star Schema, viabilizando consultas analíticas eficientes; (4) modelamento de classificação supervisionada, contemplando preparação de features, treinamento e avaliação dos modelos; (5) visualização e BI por meio de dashboards de KPIs, monitoramento de performance e geração de insights; e (6) tomada de decisão de negócio, voltada à predição de risco de churn, ações de retenção e otimização do Lifetime Value (LTV). Essa estrutura espelha as fases do CRISP-DM descritas na Seção 4.5, com a vantagem adicional de explicitar a interface entre a camada descritiva (BI) e a camada preditiva (Machine Learning), que é o principal diferencial deste projeto frente aos trabalhos correlatos.

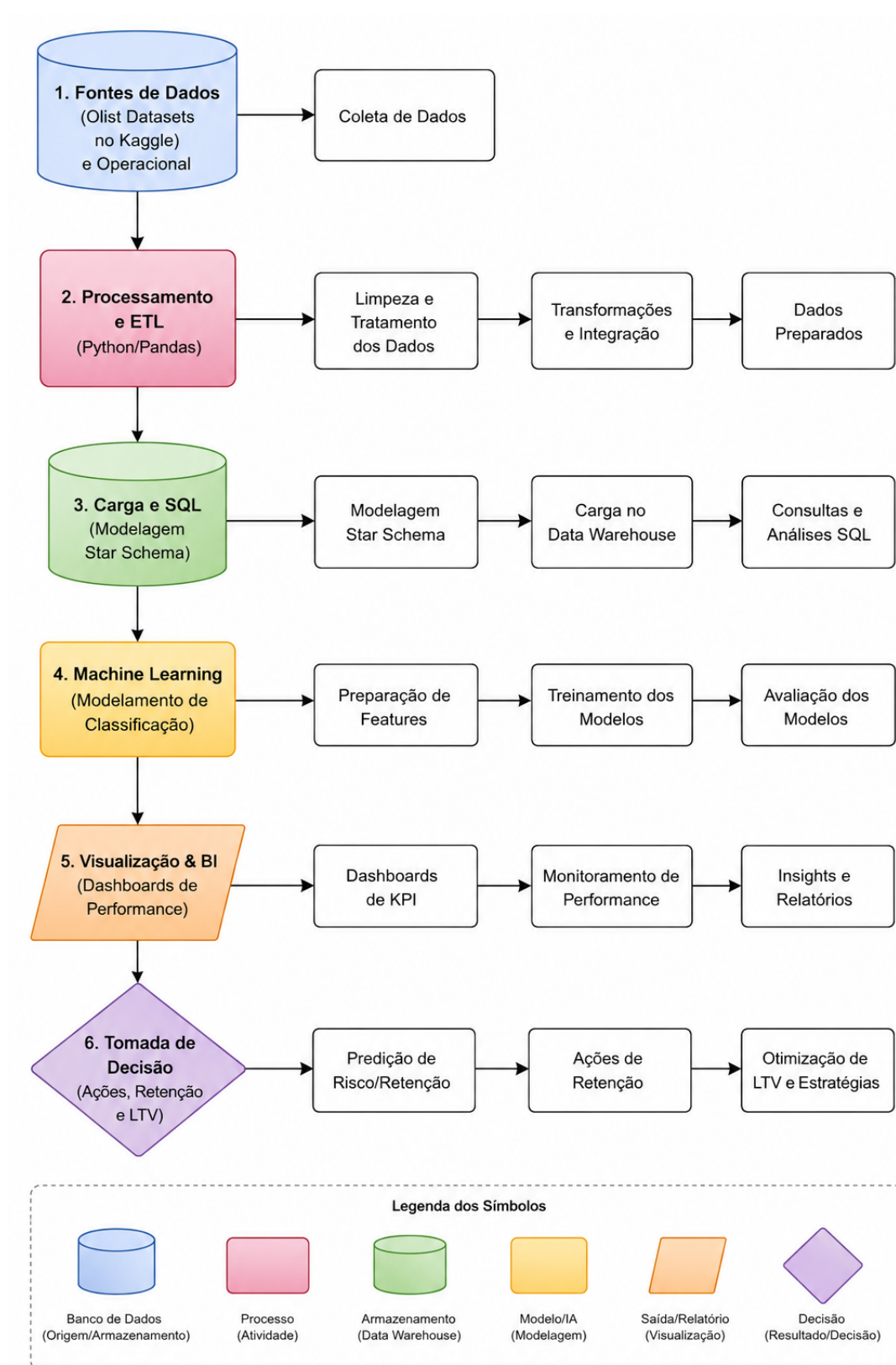


Figura 4.1: Desenho da Solução

5 Conclusão

A presente Fase 1 consolidou as bases conceituais e metodológicas para um projeto integrado de Business Intelligence e Analytics aplicado ao desafio da insatisfação de clientes no e-commerce brasileiro. Foram apresentados os fundamentos teóricos das áreas de Ciência de Dados, Modelagem Dimensional e Comportamento do Consumidor (Seção 2), mapeados três trabalhos correlatos representativos do estado da arte (Seção 3) e detalhado um plano de trabalho contendo pergunta de pesquisa, três hipóteses verificáveis com suas respectivas variáveis preditoras, objetivos gerais e específicos, cronograma trimestral, metodologia ancorada em CRISP-DM e desenho arquitetural da solução (Seção 4).

A relevância do tema justifica-se pela escala do comércio eletrônico brasileiro, que ultrapassou R\$ 204 bilhões em faturamento em 2024, e pelo impacto direto da eficiência logística sobre a satisfação do consumidor, conforme evidenciado tanto pela literatura recente quanto pelos dados operacionais do ecossistema Olist. O diferencial deste projeto frente aos trabalhos correlatos reside na proposta de uma solução integrada, combinando uma camada descritiva de Business Intelligence (Star Schema, ETL e dashboards) com uma camada preditiva de Machine Learning (classificação binária de avaliações negativas), aplicada ao contexto específico de marketplaces multi-vendedor brasileiros, lacuna ainda pouco explorada na literatura.

Para a Fase 2, propõe-se a execução do pipeline planejado, com atenção especial à engenharia de features para captura das interações logísticas multi-vendedor (Hipótese 3) e à mitigação do desbalanceamento de classes característico do problema. Como trabalhos futuros, vislumbra-se (i) a incorporação de análise de sentimentos sobre os comentários textuais das reviews via técnicas de Processamento de Linguagem Natural, ampliando o conjunto de features preditivas; (ii) a exploração da dimensão geoespacial do dataset, correlacionando distância entre vendedor e cliente com taxas de avaliação negativa; (iii) a aplicação do modelo em janelas temporais móveis, permitindo o monitoramento dinâmico do risco de churn ao longo do tempo; e (iv) a generalização da abordagem para outros marketplaces brasileiros, validando a transferibilidade dos resultados.

Referências Bibliográficas

- [Ass25] Associação Brasileira de Comércio Eletrônico. “E-commerce fatura R\$ 204,3 bilhões no Brasil em 2024”. Disponível em: <https://www.ecommercebrasil.com.br/noticias/e-commerce-resultados-2024-brasil-abcomm>, Maio 2026.
- [KR13] Kimball, R.; Ross, M. “The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling”. Indianapolis: John Wiley & Sons, 2013, 3 ed., 600.
- [Lui19] Luizon, B. F. G. “Análise preditiva de churn em um e-commerce”, Trabalho de conclusão de curso (graduação em gestão da informação), Universidade Federal do Paraná, Curitiba, Brasil, 2019, disponível em: <https://hdl.handle.net/1884/64983>. Acesso em: maio 2026.
- [OS18] Olist; Sionek, A. “Brazilian E-Commerce Public Dataset by Olist”. Disponível em: <https://www.kaggle.com/datasets/olistbr/brazilian-ecommerce>, Maio 2026.
- [PF16] Provost, F.; Fawcett, T. “Data Science para Negócios: o que você precisa saber sobre mineração de dados e pensamento analítico de dados”. Rio de Janeiro: Alta Books, 2016, 408.
- [WP23] Wangkiat, P.; Polprasert, C. “Machine Learning Approach to Predict E-commerce Customer Satisfaction Score”. In.: 2023 8th International Conference on Business and Industrial Research (ICBIR), 2023, p. 1176–1181.
- [ZBR24] Zaghloul, M.; Barakat, S.; Rezk, A. “Predicting e-commerce customer satisfaction: Traditional machine learning vs. deep learning approaches”, *Journal of Retailing and Consumer Services*, v. 79, 2024, p. 103865.