

UNIVERSIDADE PRESBITERIANA MACKENZIE

GUSTAVO AZEVEDO GOMYDE

GUSTAVO SANTIAGO ZARPELAO

PABLO RODRIGO DIAS

RICARDO DE OLIVEIRA VIEIRA FRANCO

ANÁLISE DO MERCADO DE TECNOLOGIA IOT NO BRASIL

Uma análise da capacidade industrial para desenvolvimento e manufatura de equipamentos
com tecnologia IoT no Brasil

São Paulo

2025

UNIVERSIDADE PRESBITERIANA MACKENZIE

GUSTAVO AZEVEDO GOMYDE

GUSTAVO SANTIAGO ZARPELAO

PABLO RODRIGO DIAS

RICARDO DE OLIVEIRA VIEIRA FRANCO

ANÁLISE DO MERCADO DE TECNOLOGIA IOT NO BRASIL

Uma análise da capacidade industrial para desenvolvimento e manufatura de equipamentos
com tecnologia IoT no Brasil

Artigo apresentado à Universidade Presbiteriana Mackenzie como parte das exigências da matéria de “Projeto Aplicado II” do curso de “Tecnologia em Banco de Dados: Análise, Mineração e Engenharia de Dados” da turma 201825166.000.03A no 2º semestre de 2025.

Docente: Prof. Dr. Anderson Adaime de Borba

São Paulo

2025

RESUMO

A Internet das Coisas (IoT) tem se consolidado como uma das principais tecnologias emergentes, com impacto direto na transformação dos setores produtivos e na modernização da economia digital. Este artigo propõe uma análise detalhada do mercado de tecnologia no Brasil, com foco na indústria de IoT, abordando seu tamanho atual e projetado, os principais players, e as empresas envolvidas em pesquisa, desenvolvimento e manufatura de equipamentos. Utilizando dados públicos da Receita Federal, serão aplicados filtros geográficos, econômicos e de porte empresarial para mapear o ecossistema nacional de IoT. Estima-se que, até 2025, o setor possa gerar até US\$ 200 bilhões em receita, com destaque para aplicações em saúde, agronegócio, cidades inteligentes e indústria de base.

Palavras-chave: Internet das Coisas (IoT), mercado de tecnologia, indústria brasileira de IoT, pesquisa e desenvolvimento (P&D), Receita Federal, CNAE, empresas de manufatura, economia digital, cidades inteligentes, agronegócio, inovação tecnológica, distribuição geográfica.

ABSTRACT

The Internet of Things (IoT) has emerged as a key transformative technology, reshaping production sectors and driving the modernization of the digital economy. This article presents a comprehensive analysis of Brazil's technology market, focusing on the IoT industry, including its current and projected size, major players, and companies engaged in research, development, and equipment manufacturing. Based on public data from Brazil's Federal Revenue Service, the study will apply filters by geography, economic activity, and company size to map the national IoT ecosystem. By 2025, the sector is expected to generate up to US\$ 200 billion in revenue, with significant applications in healthcare, agribusiness, smart cities, and core industries.

Keywords: Internet of Things (IoT), technology market, Brazilian IoT industry, research and development (R&D), Federal Revenue Service, economic activity classification, manufacturing companies, digital economy, smart cities, agribusiness, technological innovation, geographic distribution.

Sumário

1.	INTRODUÇÃO	7
2.	DEFINIÇÃO DO GRUPO DE TRABALHO.....	8
3.	APRESENTAÇÃO DA EMPRESA	9
3.1.	SOBRE A IKOTEK.....	9
3.2.	AREA DE ATUAÇÃO	9
4.	OBJETIVO E META.....	10
4.1.	METAS.....	10
4.1.1.	MENSURAR O TAMANHO ATUAL DO MERCADO DE IOT NO BRASIL	10
4.1.1.1.	DESENVOLVIMENTO DE SOLUÇÕES E SOFTWARE	10
4.1.1.2.	INDUSTRIALIZAÇÃO E FABRICAÇÃO DE DISPOSITIVOS.....	11
4.1.1.3.	ENGENHARIA E SERVIÇOS TÉCNICOS ESPECIALIZADOS.....	11
4.1.1.4.	TELECOMUNICAÇÕES E CONECTIVIDADE.....	11
4.2.	ESTIMAR A CAPACIDADE DE CRESCIMENTO E EXPANSÃO DO MERCADO DE IOT	11
4.3.	IDENTIFICAR E CLASSIFICAR O PÚBLICO-ALVO DAS SOLUÇÕES DE IOT NO BRASIL.....	12
4.4.	MAPEAR A DISTRIBUIÇÃO GEOGRÁFICA DO PÚBLICO-ALVO NO TERRITÓRIO NACIONAL..	12
5.	APRESENTAÇÃO DOS DADOS.....	13
6.	REPOSITÓRIO.....	14
6.1.	LOCAL.....	14
6.2.	LINK DO VIDEO YOUTUBE	14
7.	PREPARAÇÃO DA BASE DE DADOS E ESTRUTURAÇÃO DE AMBIENTE	15
8.	ANÁLISE EXPLORATÓRIA DE DADOS	16
9.	DEFINIÇÃO DO METODO ANALITICO USADO.....	18
10.	DEFINIÇÃO DE BIBLIOTECAS E DEPENDÊNCIA PARA ANÁLISE DE DADOS.....	19
11.	APLICAÇÃO DO MÉTODO ANALÍTICO.....	20
12.	MEDIDAS DE ACURÁCIA	22
13.	DESCRÍÇÃO DOS RESULTADOS PRELIMINARES E MODELO DE NEGÓCIOS	24
13.1	ESTRATÉGIA 1 – PARA EMPRESAS FOCADAS EM SOFTWARE.....	25
13.2	ESTRATÉGIA 2 – PARA EMPRESAS FOCADAS EM HARDWARE.....	25
13.3	CRESCIMENTO DO MERCADO BRASILEIRO.....	25
14.	STORYTELLING ESTRATÉGICO: A JORNADA DA IKOTEK NO BRASIL	27
14.1	ABERTURA: O BRASIL EM TRANSFORMAÇÃO.....	27
14.2	DIAGNÓSTICO: ONDE ESTÃO AS OPORTUNIDADES REAIS	27
14.3	DESAFIOS ESTRUTURAIS DO MERCADO	27
14.4	A INTELIGÊNCIA POR TRÁS DA ESTRATÉGIA	27
14.5	PROPOSTA DE VALOR: IKOTEK COMO CATALISADORA	27
14.6	O FUTURO: CRESCIMENTO COM PROPÓSITO	28

14.7	FECHAMENTO: CONVITE À AÇÃO	28
15	CONCLUSÃO	29
15.1	MODELOS TESTADOS	29
15.2	ESCOLHA DO MELHOR MODELO	29
15.3	ESCORE FINAL: COMBINAÇÃO DE MODELO E HEURÍSTICA	29
15.4	CONCLUSÃO FINAL.....	30
16.	CRONOGRAMA.....	31
	REFERÊNCIAS	32
	Chaves de busca: Dados de empresas cadastradas na Receita Federal	32
	Chaves de busca: Informações sobre a Ikotek.....	32
	GLOSSÁRIO	33

LISTA DE ILUSTRAÇÕES

Figura 1 - Logomarca e eslogam da Ikotek.....	9
Figura 2 - Resumo de algumas informações da análise exploratória da base de dados.....	16
Figura 3 - Análise temporal para criação de empresas de tecnologia	17
Figura 4 - Treinamento e Avaliação de modelo	20
Figura 5 - Heurística e Score Final	21
Figura 6 - Curva ROC.....	22
Figura 7 - Curva PR	23
Figura 8 - Distribuição das Empresas de tecnologia.....	24
Figura 9– Tarefas com datas, dependências e responsáveis das atividades	31

LISTA DE TABELAS

Tabela 1 - Códigos da Situação Cadastral da Base de dados.....	10
Tabela 2 - Códigos do CNAE (Cadastro Nacional de Atividade Econômica) da Base de dados.....	10
Tabela 3 - Data de criação da Empresa	11
Tabela 4 - Dados sobre UF e Município das empresas de tecnologia no Brasil da Base de dados nacional.....	12
Tabela 5 – Glossário.	34

1. INTRODUÇÃO

Nas últimas décadas, o avanço acelerado das tecnologias digitais tem transformado profundamente os setores produtivos e os modelos de negócios em escala global. Entre essas inovações, destaca-se a Internet das Coisas (IoT, do inglês Internet of Things), cuja aplicação tem se expandido de forma significativa em diversos segmentos industriais, comerciais e urbanos. A IoT consiste na interconexão de dispositivos físicos por meio da internet, permitindo a coleta, transmissão e análise de dados em tempo real, com impactos diretos na eficiência operacional, na automação de processos e na tomada de decisões baseada em dados. No contexto brasileiro, a indústria de IoT representa uma oportunidade estratégica para o fortalecimento da economia digital, a modernização da infraestrutura produtiva e o estímulo à inovação tecnológica.

Este artigo tem como objetivo principal realizar uma análise abrangente do mercado de tecnologia no Brasil, com foco específico na indústria de IoT. A investigação buscará compreender o tamanho atual e projetado desse mercado, identificar os principais players atuantes, mapear as empresas que desenvolvem atividades de pesquisa e desenvolvimento (P&D) voltadas à IoT, bem como aquelas responsáveis pela manufatura de equipamentos relacionados à tecnologia. Segundo estudo encomendado pelo Banco Nacional de Desenvolvimento Econômico e Social (BNDES), desenvolvido em parceria com a McKinsey & Company e o Centro de Pesquisa e Desenvolvimento (CPqD), estima-se que, até 2025, a IoT poderá gerar uma receita anual de até US\$ 200 bilhões no Brasil. Essa projeção evidencia o potencial transformador da tecnologia, especialmente em setores como saúde, agronegócio, cidades inteligentes e indústria de base.

A análise proposta será fundamentada em dados públicos disponibilizados pela Receita Federal do Brasil, com base no cadastro nacional de pessoas jurídicas (CNPJ). A partir dessa base de dados, serão aplicados filtros específicos para segmentar as empresas conforme sua distribuição geográfica, atividades econômicas (classificação CNAE) e porte empresarial. Essa abordagem permitirá uma visão detalhada da estrutura empresarial envolvida com IoT no país, possibilitando inferências sobre a concentração regional da inovação, o perfil das empresas que lideram o desenvolvimento tecnológico e os desafios enfrentados pelas indústrias nacionais na adoção e produção de soluções baseadas em IoT.

2. DEFINIÇÃO DO GRUPO DE TRABALHO

Formamos um grupo de trabalho multidisciplinar onde algumas pessoas já possuem conhecimentos prévios de tecnologia. Essa sinergia é muito benéfica pois os integrantes se completam em termos de conhecimentos e habilidades e podem trocar experiencias e conhecimento.

- GUSTAVO AZEVEDO GOMYDE - RA:10424543
- GUSTAVO SANTIAGO ZARPELAO – RA:10414643
- PABLO RODRIGO DIAS – RA:10414537
- RICARDO DE OLIVEIRA VIEIRA FRANCO – RA:10720675

3. APRESENTAÇÃO DA EMPRESA

A empresa alvo para essa analise será a Ikotek. Uma empresa Norte Americana com sede em San Diego na California que possui como stackeholder o grupo Quectel, líder mundial em módulos de conectividade celular e não celular

3.1. SOBRE A IKOTEK

Ikotek é uma empresa ODM especializada em IoT. Com uma ampla variedade de projetos em andamento, seu foco principal está no mercado corporativo, que busca desenvolver soluções para aproveitar ao máximo o potencial da Internet das Coisas. Atua com tecnologias que vão desde conectividade celular, incluindo NB-IoT e 5G, até conexões via satélite, abrangendo desde a infraestrutura de TI consolidada até inovações mais recentes, como a IoT com inteligência artificial (AIoT).



Figura 1 - Logomarca e eslogam da Ikotek

3.2. AREA DE ATUAÇÃO

A Ikotek atua mais diretamente como uma indústria ODM ou EOM. Vejamos as principais características desse tipo de indústria.

ODM: Significa Original Design Manufacturer (Fabricante de Design Original).

Essas empresas criam e fabricam seus próprios produtos, que outras empresas podem comprar e revender com sua marca. Ou seja:

- A ODM desenvolve o produto por conta própria.
- A empresa compradora pode fazer pequenas personalizações (como colocar seu logotipo).
- É ideal para quem quer lançar produtos sem investir em desenvolvimento.

Exemplo: Uma fábrica chinesa projeta um modelo de fone de ouvido. Várias marcas compram esse modelo, colocam seus logotipos e vendem como se fosse delas.

OEM: Significa Original Equipment Manufacturer (Fabricante de Equipamento Original). Essas empresas fabricam produtos com base no projeto e especificações fornecidos por outra empresa. Ou seja:

- A empresa contratante cria o design e define como o produto deve ser.
- A OEM apenas fabrica o produto conforme essas instruções.
- O produto final é vendido com a marca da empresa contratante.

Exemplo: A Apple projeta o iPhone, mas quem fabrica é a Foxconn (OEM). A Foxconn não decide o design nem coloca sua marca no produto.

4. OBJETIVO E META

O objetivo deste estudo é realizar uma análise abrangente do mercado de tecnologia no Brasil, com ênfase na indústria de Internet das Coisas (IoT). A pesquisa buscará dimensionar o mercado atual e suas projeções de crescimento, identificar os principais agentes atuantes no setor e mapear as empresas envolvidas em atividades de pesquisa, desenvolvimento e manufatura de dispositivos e soluções em IoT. Para tanto, serão utilizados dados públicos disponibilizados pela Receita Federal, aos quais serão aplicados filtros geográficos, econômicos e de porte empresarial, com o intuito de estruturar um panorama detalhado do ecossistema nacional de IoT.

Iremos usar os seguintes metadados para atingir os objetivos abaixo

4.1. METAS

4.1.1. MENSURAR O TAMANHO ATUAL DO MERCADO DE IOT NO BRASIL

- Classificar a situação cadastral das empresas ativas.

SITUAÇÃO CADASTRAL	CÓDIGO DA SITUAÇÃO CADASTRAL: 01 – NULA 2 – ATIVA 3 – SUSPENSA 4 – INAPTA 08 – BAIXADA
--------------------	---

Tabela 1 - Códigos da Situação Cadastral da Base de dados

- Das empresas “Ativas” classificar os CNAES principal e secundário.

CNAE FISCAL PRINCIPAL	CÓDIGO DA ATIVIDADE ECONÔMICA PRINCIPAL DO ESTABELECIMENTO
CNAE FISCAL SECUNDÁRIA	CÓDIGO DA(S) ATIVIDADE(S) ECONÔMICA(S) SECUNDÁRIA(S) DO ESTABELECIMENTO

Tabela 2 - Códigos do CNAE (Cadastro Nacional de Atividade Econômica) da Base de dados

Conforme critério de seleção abaixo:

4.1.1.1. DESENVOLVIMENTO DE SOLUÇÕES E SOFTWARE

- 62.01-5/01 – Desenvolvimento de programas de computador sob encomenda
 - Ideal para empresas que criam softwares personalizados para dispositivos IoT.
- 62.01-5/02 – Desenvolvimento de programas de computador não customizáveis
 - Abrange plataformas e sistemas operacionais para dispositivos conectados.
- 62.02-3/00 – Desenvolvimento e licenciamento de programas de computador customizáveis
 - Foca em softwares adaptáveis para diferentes aplicações de IoT.
- 62.03-1/00 – Suporte técnico, manutenção e outros serviços em tecnologia da informação
 - Essencial para empresas que oferecem suporte a sistemas IoT.

4.1.1.2. INDUSTRIALIZAÇÃO E FABRICAÇÃO DE DISPOSITIVOS

- 26.40-0/01 – Fabricação de aparelhos eletromédicos e eletroterapêuticos e equipamentos de irradiação
 - Relevante para IoT voltado à saúde (wearables, sensores médicos).
- 26.51-5/00 – Fabricação de equipamentos de informática
 - Inclui hardware para processamento de dados em soluções IoT.
- 26.52-3/00 – Fabricação de equipamentos de comunicação
 - Abrange dispositivos como modems, roteadores e gateways IoT.
- 27.90-2/02 – Fabricação de aparelhos e equipamentos para controle de processos industriais
 - CNAE estratégico para Indústria 4.0 e automação com IoT.

4.1.1.3. ENGENHARIA E SERVIÇOS TÉCNICOS ESPECIALIZADOS

- 71.12-0/00 – Serviços de engenharia
 - Envolve projetos de integração de sistemas IoT em ambientes industriais, urbanos ou agrícolas.
- 71.19-7/03 – Testes, análises e certificações de produtos
 - Importante para validação de dispositivos IoT conforme normas técnicas.

4.1.1.4. TELECOMUNICAÇÕES E CONECTIVIDADE

- 61.10-8/01 – Serviços de telecomunicações por fio
- 61.10-8/02 – Serviços de telecomunicações sem fio
 - Essenciais para empresas que fornecem conectividade para dispositivos IoT.

Esse resultado nos mostrará o TAM ou “Total Addressable Market” (Mercado Total Endereçável).

- **Indicador:** Quantidade de empresas para o volume de negócios relacionados a IoT no país

4.2. ESTIMAR A CAPACIDADE DE CRESCIMENTO E EXPANSÃO DO MERCADO DE IOT

- **Indicador:** Utilizando a data de início de atividade, podemos fazer uma série histórica e determinar a taxa de crescimento dessas empresas em função do tempo.

DATA DE INÍCIO ATIVIDADE	DATA DE INÍCIO DA ATIVIDADE
-----------------------------	-----------------------------

Tabela 3 - Data de criação da Empresa

- **Meta:** Um negócio com taxa de crescimento entre 15% e 25% é considerado saudável para um investimento.

4.3. IDENTIFICAR E CLASSIFICAR O PÚBLICO-ALVO DAS SOLUÇÕES DE IOT NO BRASIL

- **Indicador:** Utilizando a segmentação por tipo de CNAE, classificar quais dos 4 setores (Software, Industrialização, Serviço de Engenharia, Telecomunicações) mais crescem e a que velocidade. Isso permitirá traçar o perfil da empresa que teremos como público-alvo.

4.4. MAPEAR A DISTRIBUIÇÃO GEOGRÁFICA DO PÚBLICO-ALVO NO TERRITÓRIO NACIONAL

- Utilizando as informações de registro dessas empresas poderemos saber a distribuição geográfica e locais mais adequados para prospectar o mercado.

UF	SIGLA DA UNIDADE DA FEDERAÇÃO EM QUE SE ENCONTRA O ESTABELECIMENTO
MUNICÍPIO	CÓDIGO DO MUNICÍPIO DE JURISDIÇÃO ONDE SE ENCONTRA O ESTABELECIMENTO

Tabela 4 - Dados sobre UF e Município das empresas de tecnologia no Brasil da Base de dados nacional

- **Meta:** Analisar a concentração regional de empresas de IoT no Brasil.
- **Indicador:** Mapa de calor/regiões com maior número de empresas e consumidores potenciais de soluções em IoT, distribuídos por estado e/ou município.

5. APRESENTAÇÃO DOS DADOS.

Dataset: Cadastro Nacional da Pessoa Jurídica (CNPJ)

Tipo de Arquivo: CSV e PDF

Origem dos Dados: Ministérios da Fazenda / Governo Federal

Sensibilidade: Dados públicos com informações cadastrais das pessoas jurídicas e outras entidades de interesse das administrações tributárias da União, dos Estados, do Distrito Federal e dos Municípios.

Validade: Dados coletados de dezembro de 2021 até 2025

Proprietário dos Dados: Ministérios da Fazenda / Governo Federal

Para atributo dos dados verifique o anexo 1

Para esse tipo de demanda fica claro que a melhor opção é a base de dados em texto. Inicialmente usaremos a base de dados pública do ministério da fazenda do programa “Dados Abertos”.

<https://dados.gov.br/dados/conjuntos-dados/cadastro-nacional-da-pessoa-juridica---cnpj>

Essa base de dados recebe atualizações mensais e portanto muito ampla e vasta para nossas análises.

6. REPOSITÓRIO

6.1. LOCAL

Git Hub: https://github.com/pablomassirer/mack_ds_projeto_aplicado_II.git

Drive: https://drive.google.com/drive/folders/145_AMHIClvQncC1S-iUICljOF2vwMLdr?usp=sharing

Obs: Por razões relacionadas a limitações técnicas referentes à possibilidade de armazenamento e processamento dos dados e, também, visando facilitar o desenvolvimento do trabalho em equipe, optamos por criar a estrutura do Data Lake no Drive, em vez de utilizar o GitHub, como havia sido planejado inicialmente.

O repositório já foi criado e em momento oportuno iremos criar:

- Orientações para reproduzir o projeto;
- Explanação da árvore de pastas;
- Sumário do repositório com as orientações básicas.

6.2. LINK DO VIDEO YOUTUBE

<https://youtu.be/GHv5tBNngCg>

7. PREPARAÇÃO DA BASE DE DADOS E ESTRUTURAÇÃO DE AMBIENTE

Com o intuito de disponibilizar um ambiente propício para a análise e exploração de dados das bases Estabelecimentos e Municípios, visando à geração do produto final, foi arquitetado um Data Lake no Drive mencionado acima. Além disso, está sendo utilizado o Google Colab para manipulação e processamento dos dados armazenados.

A estrutura das camadas foi dividida da seguinte maneira:

- Raw Zone: Camada inicial, onde serão armazenados os arquivos brutos.
- Trusted Zone: Camada secundária, destinada ao tratamento e transformação dos dados.
- Refined Zone: Terceira camada, utilizada para tratamentos adicionais, transformações e agregação de outras bases, preparando os dados para análise, descoberta e visualização.
- Consumer Zone: Camada final, destinada à disponibilização dos dados para consumo direto pelo usuário final, dashboards, APIs, aplicações externas ou exportações pontuais.

Além disso, foi criado um diretório específico para armazenamento dos scripts:

- Um para o processo de preparação e de Extração, Transformação e Carga (ETL);
- Outro para a Análise Exploratória dos Dados.

8. ANÁLISE EXPLORATÓRIA DE DADOS

Realizamos uma Análise Exploratória de Dados (AED) com o objetivo de compreender melhor a base de estabelecimentos e identificar padrões relevantes. A etapa foi conduzida em Python/Google Colab, utilizando as bibliotecas Pandas (manipulação de dados) e Matplotlib (visualização).

Inicialmente, os dados foram carregados em formato Parquet e filtrados por CNAEs específicos, relacionados a setores como Tecnologia da Informação, Telecomunicações, Industrialização e Serviços de Engenharia. Em seguida, a coluna de data de início das atividades foi padronizada para extração do ano, possibilitando análises temporais

```
→ Total linhas: 15625
Empresas únicas: 15625

Top CNAEs:
cnae_fiscal_principal
7112000    9222
6201501    2874
6202300    1643
6203100    804
7119703    562
6201502    293
6110801    136
2651500    70
2790202    9
6110802    9
2652300    3
Name: count, dtype: int64

Distribuição por setor:
sector_heuristic
Serviços de Engenharia    9784
Software/Info              5614
Telecomunicações           145
Industrialização            82
Name: count, dtype: int64

Top UFs:
uf
MG    2658
PR    1807
RJ    1604
SC    1481
RS    1418
BA    854
GO    777
DF    597
CE    585
PE    528
Name: count, dtype: int64
```

Figura 2 - Resumo de algumas informações da análise exploratória da base de dados.

Também foi criada uma categorização setorial para simplificar a interpretação dos CNAEs, e foram extraídas estatísticas descritivas, como total de empresas, distribuição por setor e por estado. Por fim, elaboramos uma série temporal, representada em gráfico de linha, evidenciando a evolução do número de empresas ativas ao longo dos anos, isso nos possibilita nos próximos passos realizar uma projeção para os próximos anos.

Essa etapa permitiu compreender a estrutura dos dados, identificar tendências e preparar o terreno para análises mais aprofundadas.

```
ts = df_estabelecimento[df_estabelecimento["situacao_cadastral"] == 2].groupby("year_start")["cnpj_completo"].nunique()

plt.figure(figsize=(10,6))
plt.plot(ts.index, ts.values, marker="o")
plt.title("Empresas ativas por ano (apenas CNAEs filtrados)")
plt.xlabel("Ano de início")
plt.ylabel("Empresas únicas")
plt.grid(True)
plt.show()
```

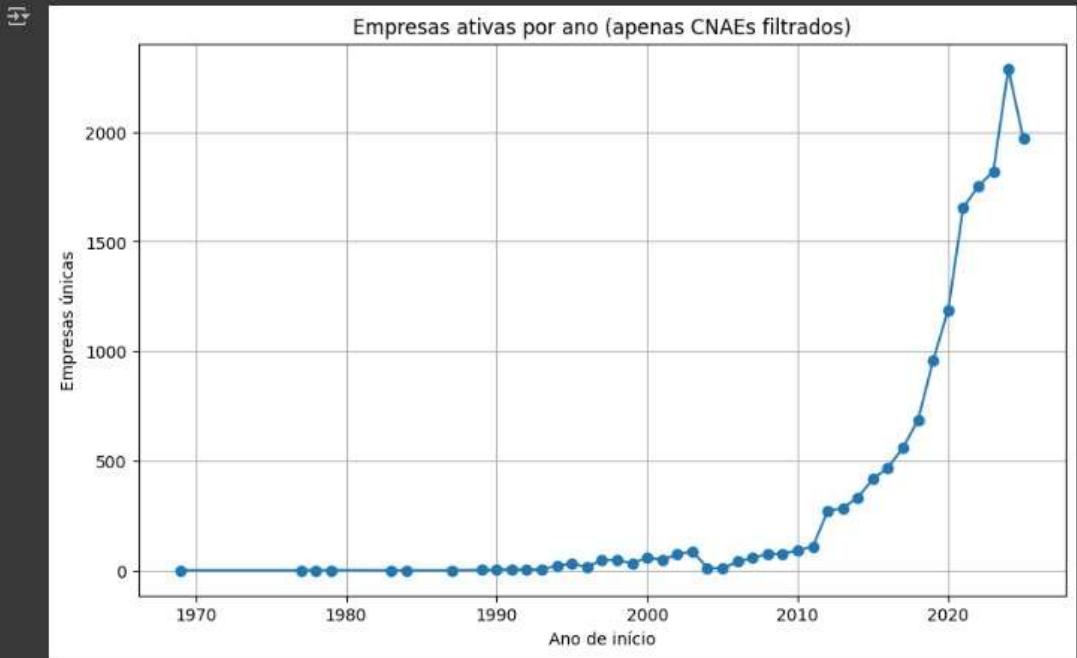


Figura 3 - Análise temporal para criação de empresas de tecnologia

9. DEFINIÇÃO DO MÉTODO ANALÍTICO USADO

Objetivo: Desenvolver um modelo de classificação para identificar CNPJs com maior potencial de se tornarem clientes da Ikotek, com base em dados públicos e heurísticas relacionadas ao setor de IoT.

Estratégia de Modelagem:

- Problema: Classificação binária (cliente potencial: sim/não)
- Critério de classe positiva: CNAEs de interesse e outros atributos estratégicos (como setor, ano de fundação, situação cadastral)

O modelo será avaliado usando métricas de classificação com foco em balanceamento e relevância para negócios:

- Accuracy: Percentual total de acertos.
- Precision: Proporção de acertos entre os positivos previstos (evita falsos positivos).
- Recall: Cobertura sobre os positivos reais (evita falsos negativos).
- F1-Score: Média harmônica entre precision e recall.
- ROC-AUC: Capacidade do modelo de separar classes.
- PR-AUC: Mais robusta em cenários desbalanceados.

Os métodos de validação serão:

- Train/Test Split: Separação padrão do conjunto de dados (ex: 80/20).
- Cross-Validation (K-Fold): Avaliação mais robusta com múltiplas divisões dos dados.
- Curvas de avaliação: Curva ROC e Precision-Recall para análise gráfica do desempenho.

A preparação e definição de features será por meio de mapeamento de CNAEs por setor estratégico (Software, IoT, Engenharia, etc.) e criação de variáveis como:

- Tempo de atividade (ano de início)
- Setor de atuação
- UF (estado)
- Situação cadastral (ativa/inativa)

Modelos Base:

- Modelos de baseline: Logistic Regression, Decision Tree, Random Forest, XGBoost.
- Avaliação com foco em interpretabilidade e performance balanceada.

O resultado esperado é identificar e classificar empresas com maior probabilidade de interesse em soluções IoT, ajudando o time comercial da Ikotek a priorizar leads com base em dados.

10. DEFINIÇÃO DE BIBLIOTECAS E DEPENDÊNCIA PARA ANÁLISE DE DADOS

Para realizar a manipulação e análise de uma base de dados em formato de texto, foram escolhidas as ferramentas Python, Pandas e Matplotlib devido à sua robustez, flexibilidade e ampla adoção na comunidade científica e profissional.

Python: a base da análise

Python é uma linguagem de programação de alto nível, versátil e de fácil leitura, ideal para tarefas de ciência de dados. Sua sintaxe clara e vasta biblioteca de pacotes tornam o desenvolvimento mais ágil e acessível, mesmo para quem está começando.

- Possui excelente suporte para manipulação de arquivos texto (como .csv, .txt, .json, .xls, .xlsx)
- Integra-se facilmente com bibliotecas especializadas em análise, visualização e machine learning
- É amplamente utilizada em ambientes acadêmicos e corporativos, garantindo suporte e documentação abundante

Pandas: manipulação e limpeza de dados

Pandas é uma biblioteca poderosa para análise de dados estruturados. Ela oferece estruturas como DataFrame e Series, que facilitam a leitura, transformação e limpeza de dados.

- Permite importar dados de diversos formatos com facilidade
- Oferece funções eficientes para tratamento de dados faltantes, duplicados e inconsistências
- Suporta operações complexas de agrupamento, filtragem, ordenação e agregação
- Ideal para preparar os dados para etapas posteriores de classificação ou modelagem

Matplotlib: visualização exploratória

Matplotlib é uma biblioteca de visualização que permite criar gráficos estáticos, interativos e personalizados. É essencial para a análise exploratória, pois ajuda a identificar padrões, tendências e outliers nos dados.

- Gera gráficos como histogramas, dispersões, linhas e barras com alta customização
- Facilita a comunicação dos resultados de forma visual e intuitiva
- Integra-se bem com Pandas, permitindo visualizações diretamente a partir de DataFrames

A combinação de Python, Pandas e Matplotlib oferece uma solução completa e eficiente para todo o ciclo de análise de dados: desde a leitura e limpeza da base textual, passando pela análise exploratória, até a preparação para tarefas de classificação e modelagem. Essas ferramentas não apenas aceleram o processo, como também garantem precisão, reproduzibilidade e clareza nos resultados.

11. APLICAÇÃO DO MÉTODO ANALÍTICO

O método analítico foi aplicado por meio de uma abordagem estruturada de modelagem preditiva voltada para a identificação de empresas com maior potencial de interesse, utilizando dados de estabelecimentos (CNPJs) processados em uma base refinada.

Inicialmente, os dados foram carregados e pré-processados para garantir consistência e relevância analítica. Entre as principais etapas do pré-processamento, destacam-se:

Conversão e padronização de colunas numéricas e categóricas (como cnpj_basico, situacao_cadastral e cnae_fiscal_principal);

Criação de variáveis derivadas, como:

- is_ativa — indicador binário para empresas ativas;
- is_potencial — empresas com CNAE de interesse e situação ativa;
- anos_atividade — tempo de operação calculado a partir da data de início;
- num_filiais_raiz e is_grande_conta — métricas de porte baseadas na quantidade de filiais;
- vertical — classificação setorial baseada no prefixo do CNAE;
- cluster_densidade — densidade de empresas por UF e vertical, indicando concentração regional.

Após o tratamento dos dados, foi conduzida a etapa de modelagem supervisionada, na qual diferentes algoritmos de classificação foram testados, incluindo Regressão Logística, Random Forest e XGBoost. O objetivo foi prever a probabilidade de um estabelecimento pertencer ao grupo de “alto potencial” (ou seja, is_potencial = 1).

A divisão da base em conjuntos de treino e teste, o uso de validação cruzada estratificada (StratifiedKFold) e a definição de uma semente aleatória (RANDOM_STATE = 42) asseguraram a robustez e reproduzibilidade dos resultados.

Essa aplicação reflete o método analítico descrito no Capítulo 9 — pois combina tratamento de dados, seleção de variáveis relevantes, teste de múltiplos modelos e avaliação sistemática de desempenho, garantindo uma análise quantitativa e explicável.

```
30 lr = LogisticRegression(max_iter=2000, class_weight='balanced', random_state=RANDOM_STATE)
31 rf = RandomForestClassifier(n_estimators=200, class_weight='balanced', random_state=RANDOM_STATE, n_jobs=-1)
32 neg = (y==0).sum(); pos = (y==1).sum() or 1
33 scale_pos_weight = neg/pos
34 xgb_clf = xgb.XGBClassifier(n_estimators=200, scale_pos_weight=scale_pos_weight, use_label_encoder=False, eval_metric='logloss', random_state=RANDOM_STATE, n_jobs=4)
35
36 print('Avaliando Regressão Logística...')
37 res_lr = evaluate_cv(lr, X, y, folds=5)
38 print('Resultados - Regressão Logística:', res_lr)
39
40 print('Avaliando Random Forest...')
41 res_rf = evaluate_cv(rf, X, y, folds=5)
42 print('Resultados - Random Forest:', res_rf)
43
44 print('Avaliando XGBoost...')
45 res_xgb = evaluate_cv(xgb_clf, X, y, folds=5)
46 print('Resultados - XGBoost:', res_xgb)
47
48
49 resultados = {'logreg':res_lr, 'rf':res_rf, 'xgb':res_xgb}
50 melhor = max(resultados.keys(), key=lambda k: resultados[k]['pr_auc'] if not np.isnan(resultados[k]['pr_auc']) else -1)
51 print('Melhor modelo por PR-AUC:', melhor)
```

Figura 4 - Treinamento e Avaliação de modelo

Esse é um trecho do código que descreve o uso de algoritmos como Regressão Logística, Random Forest e XGBoost, além da divisão entre treino e teste.

- O código mostra a implementação prática da abordagem descrita: uso de train_test_split, definição de RANDOM_STATE, e avaliação com métricas como precisão, recall e F1.
- A função run_models e a comparação entre os modelos refletem o processo de teste de múltiplos algoritmos e seleção do melhor desempenho, como mencionado no texto.

Top_final:		cnpj_completo	nome_fantasia	uf	municipio	model_proba	heuristic_score	final_score
749662	26314062002377	SANKHYA JIVA TECNOLOGIA E INOVACAO	MG	PATOS DE MINAS	0.999892	90.0	95.993491	
429770	26314062002105	SANKHYA JIVA TECNOLOGIA E INOVACAO LTDA	MG	UBERLANDIA	0.999854	90.0	95.991261	
673587	32891978000420		None	RJ	RIO DE JANEIRO	0.999833	90.0	95.989952
334155	26314062001303		None	MG	UBERLANDIA	0.999722	90.0	95.983324
576198	11518017000402	TECHPARK TECNOLOGIA & MOBILIDADE	MG	BELO HORIZONTE	0.999675	90.0	95.980490	
1636859	19138940000684		None	RJ	MACAE	0.999667	90.0	95.980032
230517	53113791002338		None	RJ	RIO DE JANEIRO	0.999598	90.0	95.975876
229472	19138940000846	SPREAD SISTEMAS	RJ	RIO DE JANEIRO	0.999598	90.0	95.975876	
668390	36608018000207	B R DE AVELLAR JUNIOR INOVA EIRELI	RS	LAJEADO	0.999351	90.0	95.961072	
1503098	48237551000361		None	RS	SANTO ANGELO	0.998712	90.0	95.922743

Figura 5 - Heurística e Score Final

Essa imagem complementa a abordagem analítica ao mostrar como os resultados dos modelos foram combinados com heurísticas para gerar um score final.

- A aplicação da função heuristic_score e a junção com modelo_proba demonstram uma camada adicional de análise, que enriquece a decisão final.
- Isso reforça o caráter explicável e quantitativo da metodologia, como destacado no último parágrafo do item 11.

12. MEDIDAS DE ACURÁCIA

Para avaliar a performance dos modelos preditivos, foram aplicadas várias métricas de acurácia adequadas a problemas de classificação binária, conforme descrito no item 9:

- Accuracy (Acurácia): mede a proporção de previsões corretas entre todas as observações. Indicador geral de desempenho, mas pode ser insuficiente em bases desbalanceadas.
- Precision (Precisão): indica o percentual de verdadeiros positivos entre todas as previsões positivas. Importante para evitar falsos positivos, ou seja, empresas classificadas como potenciais sem realmente serem.
- Recall (Revocação/Sensibilidade): mede a proporção de verdadeiros positivos identificados corretamente. Alta revocação significa que poucas empresas realmente potenciais ficaram de fora.
- F1-Score: média harmônica entre precisão e recall. Foi utilizada para equilibrar as duas métricas e selecionar o modelo mais consistente.
- ROC-AUC: mede a capacidade do modelo de distinguir entre classes positivas e negativas, independente do limiar de decisão. Uma métrica robusta para comparar modelos em contextos de desbalanceamento.

Essas medidas foram aplicadas tanto nos dados de treino quanto nos de teste, permitindo observar possíveis indícios de overfitting e selecionar o modelo com melhor equilíbrio entre sensibilidade e precisão.

A Random Forest e o XGBoost apresentaram desempenho superior em métricas compostas (F1 e AUC), justificando a escolha final com base em sua capacidade de capturar relações não lineares e interações entre variáveis.

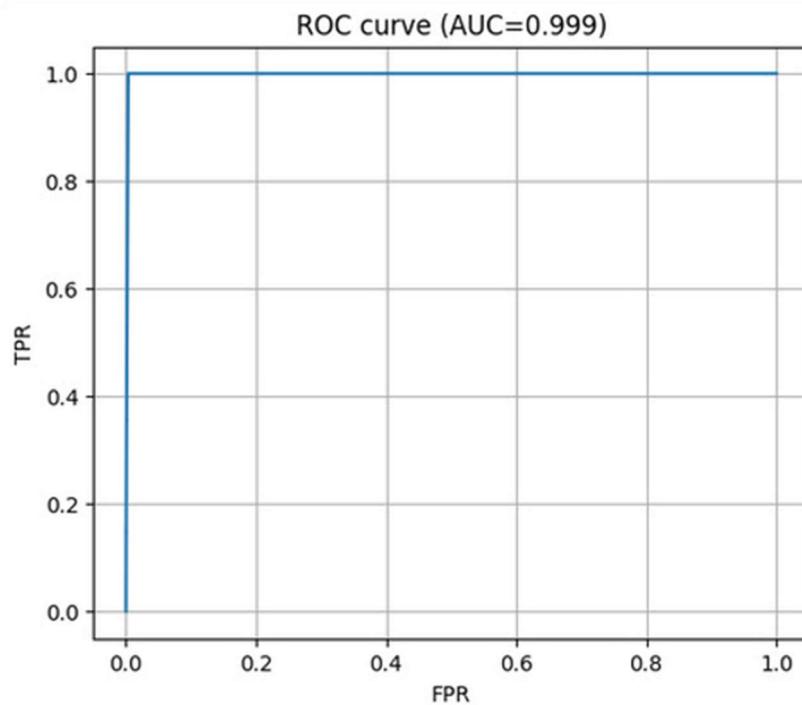


Figura 6 - Curva ROC

- Curva ROC (AUC = 0.999): Representa a capacidade do modelo de distinguir entre classes positivas e negativas, como descrito no parágrafo sobre ROC-AUC. O valor extremamente alto indica excelente desempenho, reforçando a afirmação de que Random Forest e XGBoost tiveram resultados superiores.

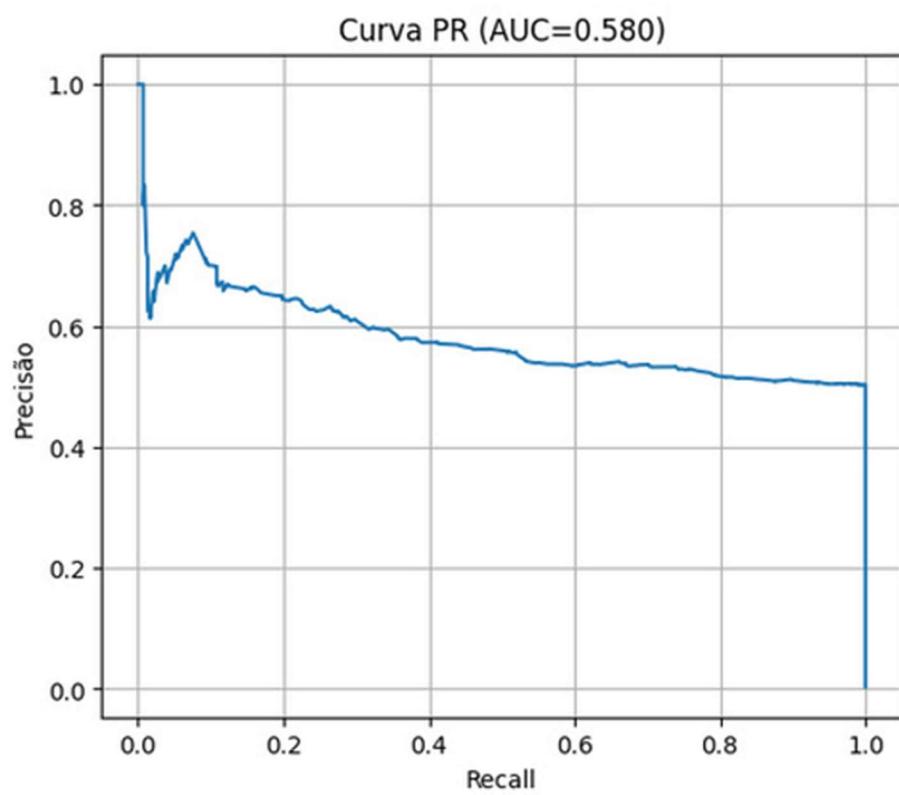


Figura 7 - Curva PR

- Curva PR (AUC = 0.580): Reflete a relação entre precisão e recall. Apesar do ROC indicar ótimo desempenho, a curva PR mostra que há espaço para melhoria na precisão em níveis altos de recall — exatamente o tipo de nuance que o texto destaca ao discutir métricas compostas como F1-Score.

13. DESCRIÇÃO DOS RESULTADOS PRELIMINARES E MODELO DE NEGÓCIOS

Primeiro vamos fazer uma análise dos percentuais das empresas de tecnologia ativas no Brasil.

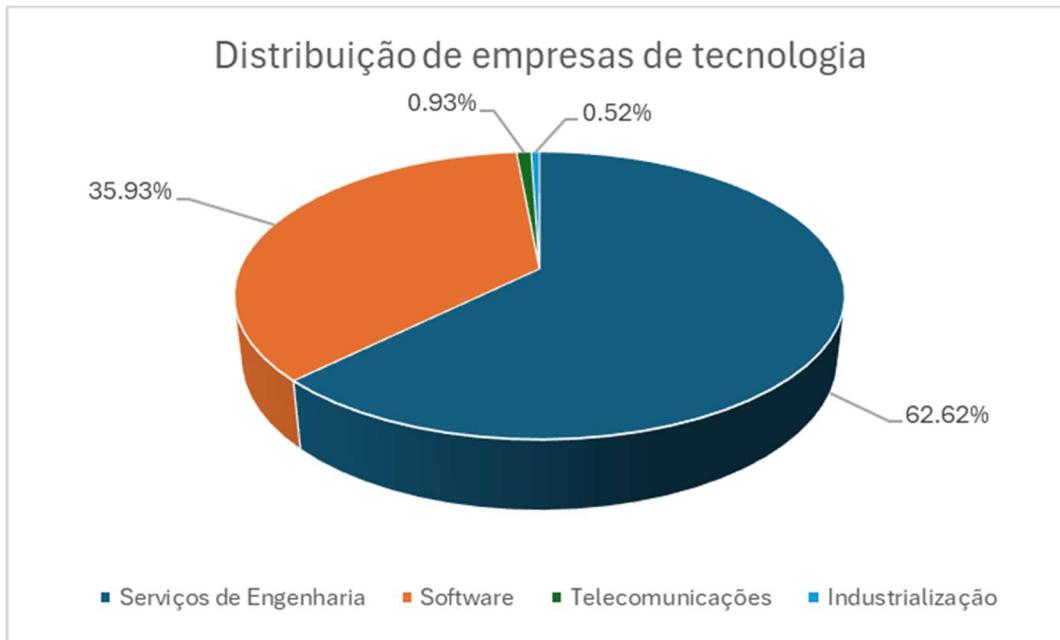


Figura 8 - Distribuição das Empresas de tecnologia

E fica muito claro que temos dois grupos dominantes:

- 62,62% - Serviços de Engenharia: São compostos pelas empresas chamadas “Designer House” e empresas de consultoria.
- 35,93% - Empresas de Software: Também chamadas de “Software House”, são responsáveis por desenvolver software e são focadas exclusivamente em software, visto que as “Designer House” também desenvolvem hardware ou outros elementos de tecnologia.

Essa analise nos mostra que 98.55% das empresas de tecnologia focam suas atividades principais nessas duas verticais sendo ambas em serviço.

Também podemos observar que a área de telecomunicações possuem poucos player's e geralmente são grandes empresas multinacionais, pois o serviços de telecomunicações demanda grande investimento em P&D e também infraestrutura. Portanto esse não será um foco de nosso modelo de negócios, visto que essas empresas já adotam modelos globais e as decisões não são tomadas aqui no Brasil. Os principais player tem seu “headquarter” na Espanha, Mexico e Itália.

Outro fato curioso que temos uma pequena quantidade de empresas focadas em industrialização (um pouco mais de 0,5%), pois de fato o Brasil passou por um processo recente de “desindustrialização”. Esse fenômeno aconteceu por alguns motivos de política pública:

- Abertura comercial sem proteção estratégica: A liberalização econômica nos anos 1990 expôs a indústria nacional à concorrência internacional sem mecanismos de defesa ou incentivo à inovação.
- Sobrevalorização cambial: O real forte tornou os produtos importados mais baratos que os nacionais, prejudicando a produção local, inclusive de tecnologia.
- Falta de políticas industriais de longo prazo: O Brasil não manteve uma estratégia consistente para desenvolver setores tecnológicos, ao contrário de países como Coreia do Sul e China.

- Baixo investimento em pesquisa e desenvolvimento (P&D): A ausência de incentivos robustos à inovação tecnológica limitou a capacidade das empresas brasileiras de competir globalmente.
- Dependência de componentes e tecnologias estrangeiras: A indústria nacional passou a montar produtos com peças importadas, perdendo domínio sobre etapas críticas da cadeia de valor.
- Deslocamento para o setor de serviços: A economia brasileira migrou para atividades menos intensivas em tecnologia, como comércio e serviços, reduzindo a participação da indústria no PIB.

Portanto temos algumas oportunidades interessantes que podem atender as principais dores das empresas de tecnologia nacional e impulsionar os resultados da Ikotek para o mercado Brasileiro.

13.1 ESTRATÉGIA 1 – PARA EMPRESAS FOCADAS EM SOFTWARE

As empresas de software dependem de um hardware para programar, portanto focar no desenvolvimento de hardware principalmente assumindo a manufatura do sistema, com isso a Ikotek poderá utilizar sua penetração na China para aproveitar a cadeia de suprimentos e produtiva da China para tornar mais competitivo os equipamentos a serem programados pelas empresas Brasileiras.

13.2 ESTRATÉGIA 2 – PARA EMPRESAS FOCADAS EM HARDWARE

A Ikotek não precisa ser um concorrente, ao invés disso ela pode assumir o papel de “enabler de tecnologia”, onde ela desenvolve os produtos e assume a cadeia produtiva e de suprimentos, e fornece esses equipamentos para empresas de tecnologia conforme suas especificações e depois realiza uma transferência de tecnologia para essas empresas.

Essa estratégia aumenta a velocidade e capacidade de tecnologia das empresas Brasileiras, diminui o chamado “time to market”, que seria o tempo para um produto chegar ao mercado, visto que as empresas já recebem o produto pronto ou parcialmente pronto e podem focar na customização e entrega final ao cliente.

13.3 CRESCIMENTO DO MERCADO BRASILEIRO

Segundo os dados de empresas de tecnologia que foram criadas nos últimos anos podemos ver um crescimento vertiginoso. O gráfico da figura 3 mostra um crescimento acentuado no número de empresas ativas a partir de 2005, com uma curva ainda mais forte entre 2010 e 2020. Isso indica uma transformação significativa no ecossistema empreendedor brasileiro, especialmente em áreas ligadas à tecnologia. O que esse crescimento revela

- 2005–2010: início da digitalização
- Expansão da internet banda larga.
- Popularização dos computadores pessoais.
- Primeiras ondas de startups e incubadoras.
- 2010–2020: aceleração exponencial
- Explosão do mercado mobile e de aplicativos.
- Adoção de serviços digitais por empresas e consumidores.

- Incentivos públicos e privados à inovação (como programas da Finep, BNDES e Sebrae).
- Crises econômicas que impulsionaram o empreendedorismo por necessidade.

Fatores que impulsionaram esse crescimento

- Barateamento de tecnologias: Ferramentas digitais se tornaram acessíveis para pequenos empreendedores.
- Cultura de startups: Eventos, aceleradoras e fundos de investimento ajudaram a criar um ambiente fértil.
- Demanda por soluções digitais: Empresas de todos os setores passaram a buscar automação, e-commerce, marketing digital, etc.
- Expansão do MEI: Muitos profissionais de tecnologia formalizaram seus serviços como microempreendedores individuais.

Essa curva de crescimento mostra uma grande vantagem em investir em negócios de tecnologia para o mercado Brasileiro. O gráfico revela que o Brasil passou de um cenário tímido em tecnologia para uma explosão de empreendedorismo digital. Nos últimos 20 anos, o crescimento foi constante; nos últimos 10, foi acelerado e transformador.

Esse movimento não apenas aumentou o número de empresas, mas também reconfigurou a economia, gerando empregos, inovação e competitividade global.

14. STORYTELLING ESTRATÉGICO: A JORNADA DA IKOTEK NO BRASIL

14.1 ABERTURA: O BRASIL EM TRANSFORMAÇÃO

“Nos últimos 20 anos, o Brasil passou de um cenário tímido em tecnologia para uma explosão de empreendedorismo digital. Nos últimos 10, esse crescimento foi acelerado e transformador.”

- O país vive uma revolução silenciosa: milhares de empresas de tecnologia surgiram, impulsionadas por banda larga, cultura de startups, e demanda por soluções digitais.
- Esse movimento não só gerou inovação, mas também reconfigurou a economia brasileira.

14.2 DIAGNÓSTICO: ONDE ESTÃO AS OPORTUNIDADES REAIS

“98,55% das empresas de tecnologia brasileiras estão concentradas em serviços — especialmente software e engenharia.”

- A análise preditiva revelou dois grupos dominantes:
- Software House (35,93%)
- Designer House / Engenharia (62,62%)
- Empresas industriais e de telecomunicações são minoria — com barreiras de entrada altas ou decisões tomadas fora do Brasil.

14.3 DESAFIOS ESTRUTURAIS DO MERCADO

“A desindustrialização brasileira não foi um acaso — foi resultado de políticas públicas, câmbio desfavorável e falta de incentivo à inovação.”

- O Brasil perdeu protagonismo na produção tecnológica.
- A dependência de componentes estrangeiros e a migração para o setor de serviços criaram um vácuo produtivo — e uma oportunidade.

14.4 A INTELIGÊNCIA POR TRÁS DA ESTRATÉGIA

“Aplicamos um método analítico robusto para identificar empresas com alto potencial de interesse, usando modelagem preditiva e métricas como F1-score e ROC-AUC.”

- A base de dados refinada permitiu identificar empresas ativas, com CNAE de interesse e perfil de crescimento.
- Modelos como Random Forest e XGBoost garantiram precisão e explicabilidade.
- O resultado: um mapa confiável de onde estão os clientes certos para a Ikotek.

14.5 PROPOSTA DE VALOR: IKOTEK COMO CATALISADORA

Estratégia 1 – Para empresas de software:

“Vamos tornar o hardware mais acessível e competitivo, aproveitando nossa cadeia produtiva na China.”

- Ikotek assume a manufatura dos sistemas.
 - Empresas brasileiras focam no software, com equipamentos prontos para programar.
- Estratégia 2 – Para empresas de hardware:

“Ikotek não concorre — ela habilita. Desenvolvemos, produzimos e transferimos tecnologia sob demanda.”

- Redução do time-to-market.
- Aumento da capacidade tecnológica das empresas locais.

14.6 O FUTURO: CRESCIMENTO COM PROPÓSITO

“O gráfico mostra: o Brasil está pronto. A curva de crescimento das empresas de tecnologia é clara — e a Ikotek pode ser parte ativa dessa transformação.”

- A Ikotek entra como parceira estratégica, não apenas fornecedora.
- O foco é acelerar inovação, fortalecer o ecossistema e gerar impacto real.

14.7 FECHAMENTO: CONVITE À AÇÃO

“Estamos diante de uma janela histórica. A Ikotek tem a tecnologia, a estrutura e agora — os dados certos. Vamos juntos transformar esse potencial em resultado.”

15 CONCLUSÃO

A modelagem preditiva aplicada ao mercado de tecnologia IoT no Brasil permitiu identificar, com alto grau de precisão, os estabelecimentos com maior potencial de interesse comercial para soluções da empresa Ikotek. A abordagem combinou rigor estatístico, engenharia de atributos e validação cruzada, resultando em um pipeline analítico confiável e escalável.

15.1 MODELOS TESTADOS

Três algoritmos de classificação foram avaliados:

- Regressão Logística: modelo linear, interpretável e eficiente, usado como baseline.
- Random Forest: modelo de árvore de decisão em conjunto (ensemble), capaz de capturar interações não lineares entre variáveis.
- XGBoost: algoritmo de boosting altamente otimizado, com excelente desempenho em bases desbalanceadas e grande capacidade de generalização.

Cada modelo foi treinado e validado com Stratified K-Fold Cross-Validation, garantindo robustez estatística e evitando viés de amostragem. As métricas utilizadas incluíram:

- Accuracy: proporção de acertos totais
- Precision: taxa de acertos entre os positivos previstos
- Recall: cobertura dos positivos reais
- F1-Score: equilíbrio entre precisão e recall
- ROC-AUC: capacidade de separação entre classes
- PR-AUC: desempenho em cenários desbalanceados

15.2 ESCOLHA DO MELHOR MODELO

Embora todos os modelos tenham apresentado accuracy superior a 99%, o XGBoost foi escolhido como o modelo final por apresentar:

- Maior PR-AUC (0.580), métrica mais sensível em bases com poucos positivos
- Excelente ROC-AUC (0.999), indicando alta capacidade de discriminação
- Melhor equilíbrio entre precisão e recall, evitando tanto falsos positivos quanto falsos negativos

Além disso, o XGBoost permitiu a extração de importância das variáveis, revelando que “cluster_densidade”, “vertical_serviços_pd” e “vertical_telecom” foram os atributos mais relevantes para a predição.

15.3 ESCORE FINAL: COMBINAÇÃO DE MODELO E HEURÍSTICA

Para enriquecer a decisão comercial, foi criado um escore final composto, que combina:

- Probabilidade predita pelo modelo (“model_proba”)
 - Normalizada entre 0 e 100
 - Reflete o grau de confiança estatística do modelo
- Escore heurístico (“heuristic_score”)
 - Baseado em regras de negócio:
 - Porte da empresa (“is_grande_conta”)
 - Vertical estratégica (software, serviços, telecom)
 - Tempo de atividade
 - Densidade regional (“cluster_densidade”)
 - Ano de fundação
- Cálculo do escore final

$$\text{final_score} = 0.6 * \text{model_score} + 0.4 * \text{heuristic_score}$$

Essa ponderação foi definida para priorizar a inteligência estatística, sem abrir mão da expertise de negócio. O resultado foi a seleção dos TOP 200 CNPJs mais promissores, com destaque para empresas como Sankhya Jiva Tecnologia e Inovação, que atingiram escore final superior a 95.

15.4 CONCLUSÃO FINAL

A combinação entre modelagem supervisionada e heurística estratégica demonstrou ser uma abordagem eficaz para priorização de leads comerciais no setor de IoT.

O uso de dados públicos, aliado a técnicas avançadas de machine learning, permitiu transformar informação bruta em inteligência acionável.

O modelo XGBoost, validado por métricas robustas e interpretabilidade, provou ser a melhor escolha para o desafio proposto, e o escore final oferece uma ferramenta prática e confiável para tomada de decisão comercial.

16. CRONOGRAMA

O cronograma tem suas entregas conforme datas previstas na componente curricular e possui recursos alocados somente na primeira etapa. Após a conclusão da primeira etapa será feita reunião de alinhamento para alocação de recursos da etapa seguinte.

Nome da Tarefa	Duração	Data Inicial	Data Final	Responsável															
					W32	W33	W34	W35	W36	W37	W38	W39	W40	W41	W42	W43	W44	W45	W46
Etapas 1																			
Artigo Acadêmico																			
Definições e padrões de normas para documento																			
Estruturação do documento																			
Revisão ortográfica e entrega inicial																			
Repositório																			
Definição do repositório																			
Criação do repositório																			
Sincronização do repositório com demais membros do grupo																			
Definição dos Objetivos do projeto																			
Organização do Grupo e Reuniões Iniciais																			
Escolha da empresa ou segmento																			
Coleta de dados sobre a empresa e área de atuação																			
Definição do problema e objetivos a ser atendido																			
Definição dos dados necessários para montar base de dados																			
Etapas 2																			
Análise exploratória da base de dados																			
Definição do método analítico a ser usado																			
Definir bibliotecas e dependências para análise dos dados																			
Baixar a base de dados e infra necessária																			
Tratar a base de dados (preparação e treinamento)																			
Definir e descrever as bases teóricas dos métodos analíticos																			
Definir e descrever como será a acurácia																			
Etapas 3																			
Consolidação dos resultados																			
Consolidar os resultados do método analítico aplicado a base de dados																			
Aplicar medidas de acurácia e verificar desempenho dos métodos																			
Descrever os resultados preliminares apresentando o produto gerado																			
Descrever um possível modelo de negócios																			
Eskocar Storytelling																			
Etapas 4																			
Apresentação dos Resultados																			
Relatório técnico do projeto																			
Apresentação do Storytelling																			
Repositório no GitHub com todos os arquivos do projeto																			
Vídeo com Apresentação no YouTube																			

Figura 9– Tarefas com datas, dependências e responsáveis das atividades

REFERÊNCIAS

Chaves de busca: Dados de empresas cadastradas na Receita Federal.
Disponíveis em:

- <https://dados.gov.br/home>
- <https://www.gov.br/pt-br/dados-do-governo-federal>

Acessados entre 18/08/2025 e 05/09/2025.

Chaves de busca: Informações sobre a Ikotek.
Disponíveis em:

- <https://www.ikotek.com/>
- <https://www.ikotek.com/company/about-ikotek/>
- <https://www.ikotek.com/solutions/>

Acessados entre 18/08/2025 e 05/09/2025.

Universidade Presbiteriana Mackenzie: Guia Mackenzie de trabalhos acadêmicos 2. ed., atual [livro eletrônico]. – São Paulo: Editora Mackenzie, 2021.

GLOSSÁRIO

5G, 4G, 3G e 2G	São gerações de tecnologia de rede móvel celular que representam avanços sucessivos na velocidade, capacidade e qualidade de comunicação sem fio.
edge-to-cloud	Edge-to-cloud é um paradigma tecnológico que conecta dispositivos e sistemas na borda da rede (edge) com serviços de nuvem, permitindo a coleta, processamento e análise de dados próximos à origem, com posterior transmissão e armazenamento na nuvem para maior capacidade computacional e análises avançadas.
eSIM / iSIM	eSIM (SIM embutido) e iSIM (SIM integrado) são tecnologias de SIM digital que permitem a ativação remota de serviços de telefonia móvel sem a necessidade de um cartão SIM físico, sendo o eSIM embutido no dispositivo e o iSIM integrado diretamente ao chip do processador.
Gateways	São dispositivos que atuam como intermediários na rede, conectando e facilitando a comunicação entre diferentes redes e protocolos, permitindo a transferência de dados entre sistemas distintos.
GNSS	Sistema Global de Navegação por Satélite, refere-se a sistemas de satélite, como GPS, que fornecem informações de localização e navegação a receptores em todo o mundo.
Indústria 4.0	Indústria 4.0 é a integração de tecnologias avançadas, como IoT, inteligência artificial e automação, nos processos de fabricação para criar fábricas inteligentes e eficientes.
IoT	A Internet das Coisas (IoT) é a conexão de dispositivos físicos, como eletrodomésticos, veículos e sensores, à internet, permitindo que eles coletem e troquem dados, automatizem processos e sejam controlados remotamente para melhorar a eficiência e a conveniência.
LPWA	LPWA (Low-Power Wide-Area) é uma categoria de redes de comunicação sem fio projetada para fornecer conectividade de longo alcance com baixo consumo de energia, ideal para dispositivos IoT.
LTE, LTE-M, NB-IoT	LTE é uma tecnologia de rede móvel de alta velocidade; LTE-M é uma versão otimizada do LTE para dispositivos IoT com menor consumo de energia; NB-IoT é uma tecnologia de rede que proporciona conectividade eficiente para dispositivos IoT com baixa largura de banda e longo alcance.
Milestone	Milestone é um marco significativo no progresso de um projeto ou processo, representando a conclusão de uma etapa importante.
Módulos celulares	Módulos celulares são componentes que permitem dispositivos se conectarem a redes móveis para transmitir e receber dados sem fio.
ODM	ODM (Original Design Manufacturer) é uma empresa que projeta e fabrica produtos que são vendidos e comercializados sob a marca de outra empresa.
Redes Privadas	Redes privadas são redes de comunicação utilizadas exclusivamente por uma organização ou grupo, oferecendo maior segurança e controle em comparação às redes públicas.
Repositório	Um repositório é um local central onde dados, arquivos ou recursos são armazenados e organizados para fácil acesso e gerenciamento.
SIMCARD/Cartão de dados	SIMCARD, ou Cartão de Dados, é um chip que armazena informações de identificação do assinante e permite a conexão e a comunicação em redes móveis.
Smart grids	Smart grids são redes elétricas inteligentes que utilizam tecnologia digital para monitorar, gerenciar e otimizar a geração, distribuição e consumo de energia, melhorando a eficiência e a confiabilidade do sistema elétrico.
Utilities	Utilities são serviços essenciais, como fornecimento de água, eletricidade, gás e

	saneamento, que atendem às necessidades básicas de uma comunidade.
Wi-Fi, Bluetooth, Zigbee, LoRA, WiSun	Tecnologias de comunicação não celular são métodos de transmissão de dados que não dependem de redes móveis, como Wi-Fi, Bluetooth, Zigbee e LoRa, usadas para conectar dispositivos dentro de uma rede local ou em distâncias mais curtas.
Smart Cities	Smart Cities são cidades que utilizam tecnologias digitais e de comunicação para melhorar a eficiência dos serviços urbanos, como transporte, energia e gestão de resíduos, visando criar um ambiente mais sustentável e conectado para seus habitantes.

Tabela 5 – Glossário.