

# Sprint Review 2: Construção da rede neural com wireframe e user stories

O1. Entregas anteriores
Recapitulação das entregas
passadas

O5. Rede neural
Construção da primeira versão da rede neural

Agenda

Sprint 2
Visualização das entregas e sua timeline

O6. Próximos passos
Visão geral do planejamento

User stories
Compreensão dos requisitos para a solução na visão do usuário

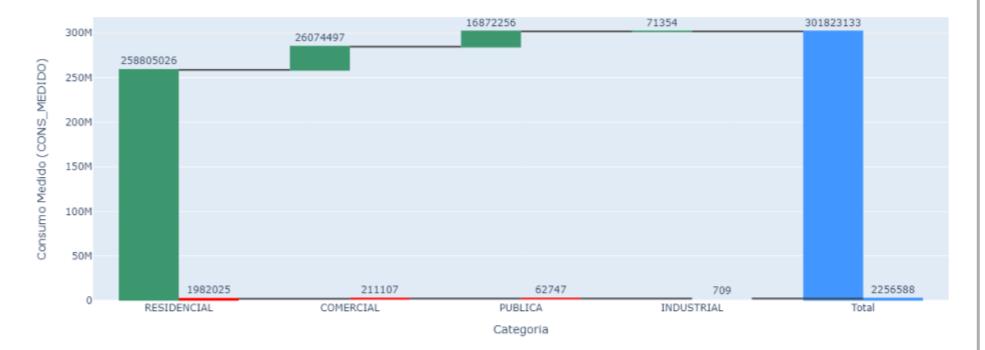
**Wireframe**Esboço inicial do dashboard



# Recapitulando: Sprint 1

## Análise exploratória

Consumo Medido por Categoria e Fraude - Gráfico de Cascata





Compreensão da distribuição dos dados e das variáveis relevantes. Focamos em entender correlações entre consumo e fraudes

#### Personas



#### Augusto Silva

Maximizar a recuperação de receitas e a acurácia na detecção de fraudes para proteger o faturamento da Aegea



Maximizar a recuperação de receitas

Melhorar a acurácia da detecção de fraudes Detectar fraudes de maneira escalável



Sistema gráfico interativo, que simplifique a visualização de padrões de fraude e os correlacione com variáveis como inadimplência, índice de desemprego, entre outros.



Direcionamento das interfaces gráficas e desenvolvimento da arquitetura da rede neural



# Construção da rede neural junto às user stories com wireframde











#### SPRINT 1

Análise exploratória dos dados, Canvas Proposta de Valor, Matriz de Risco, Personas e Anti personas e Jornada do Usuário

#### **SPRINT 2**

User Stories, Wireframes e Modelo de Rede Neural

#### SPRINT 3

Otimização do modelo, protótipo de alta fidelidade, Análise PESTEL e Business Model Canvas

#### SPRINT 4

Avaliação das Métricas e Front End inicial

#### **SPRINT 5**

Disponibilização do modelo, front end final e Análise Financeira do projeto



Persona	Descrição	Prioridade	Esforço	Critérios de Aceitação
Augusto	Eu, como gestor comercial responsável por evitar casos de fraudes, quero obter uma lista com os clientes com maior potencial de fraude, para melhorar a efetividade do processo de investigação de fraudes da empresa.	Alta	Médio	5. Exibição de Localidade e Consumo:  Teste Aceito: A lista apresenta localidade e dados de consumo inicialmente escondidos, assegurando a Privacidade como Configuração Padrão das informações sensíveis dos clientes.  Teste Rejeitado: As informações de localidade e consumo são apresentadas de forma que direta, desrespeitando a Privacidade como Configuração Padrão.

# Privacy by Design

Uso extensivo de dados pessoais dos consumidores na modelagem e detecção de fraudes (matrícula, sequência residencial, localização)

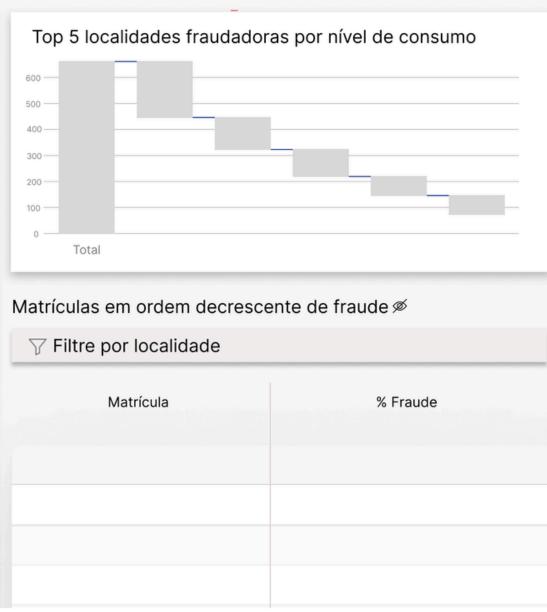


Monitoramento de Fraudes no Consumo de Água

Geral

Raw data





#### Adicione o CSV





CO Monitoramento de Fraudes no Consumo de Água							
Geral Raw data							
Adicione o CSV							
Matrícula	Fraudador	Localidade	Consumo				



# Pipeline de tratamento de dados

categorias



Parquet fará com que haja menor ocupação de espaço e memória



Remoção de colunas que não serão usadas

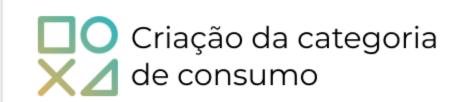
Ex: EMP\_CODIGO / ECO\_OUTRAS / FATURADO\_MEDIA, etc.



Conversão dos tipos de dados

Para inteiros (exceto lat e long) e datetime



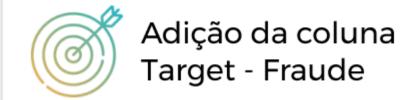


Consumo alto > 0.7 0.2 ≤ Consumo médio ≤ 0.7 Consumo baixo < 0.2



1 matrícula por linha e colunas para cada mês de consumo, de cada ano





1 matrícula por linha e colunas para cada mês de consumo, de cada ano



# Rede neural + Feature Engineering



Identificar a probabilidade de um consumidor ser fraudador

CONSUMO\_MEDIDO

**CONSUMO ALTO** 

DSC\_SIMULTANEA

CONSUMO MÉDIO

**MATRICULA** 

**CONSUMO BAIXO** 

TIPO\_LIGACAO

**CONSUMO TEMPORAL** 

Sub-Categoria

**CATEGORIA** 

fraude



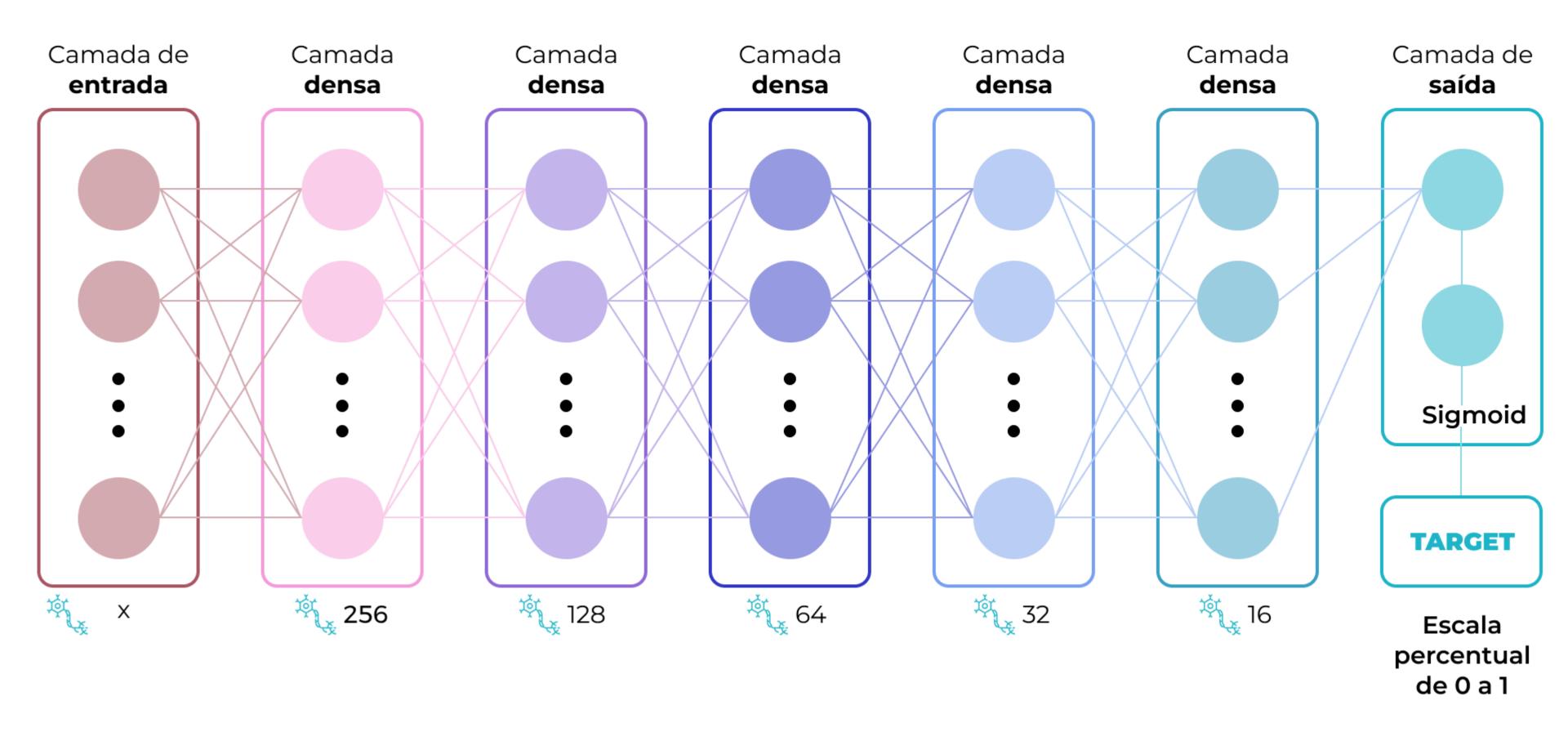
Dados em nível de consumidor e com dados sensíveis, pode inferir viés no modelo com, por exemplo, localidade (lat e long)



os dados foram removidos no treinamento do modelo e reinseridos após o output

# Arquitetura da rede

# Feedfoward (FNN)





# Resultados preliminares

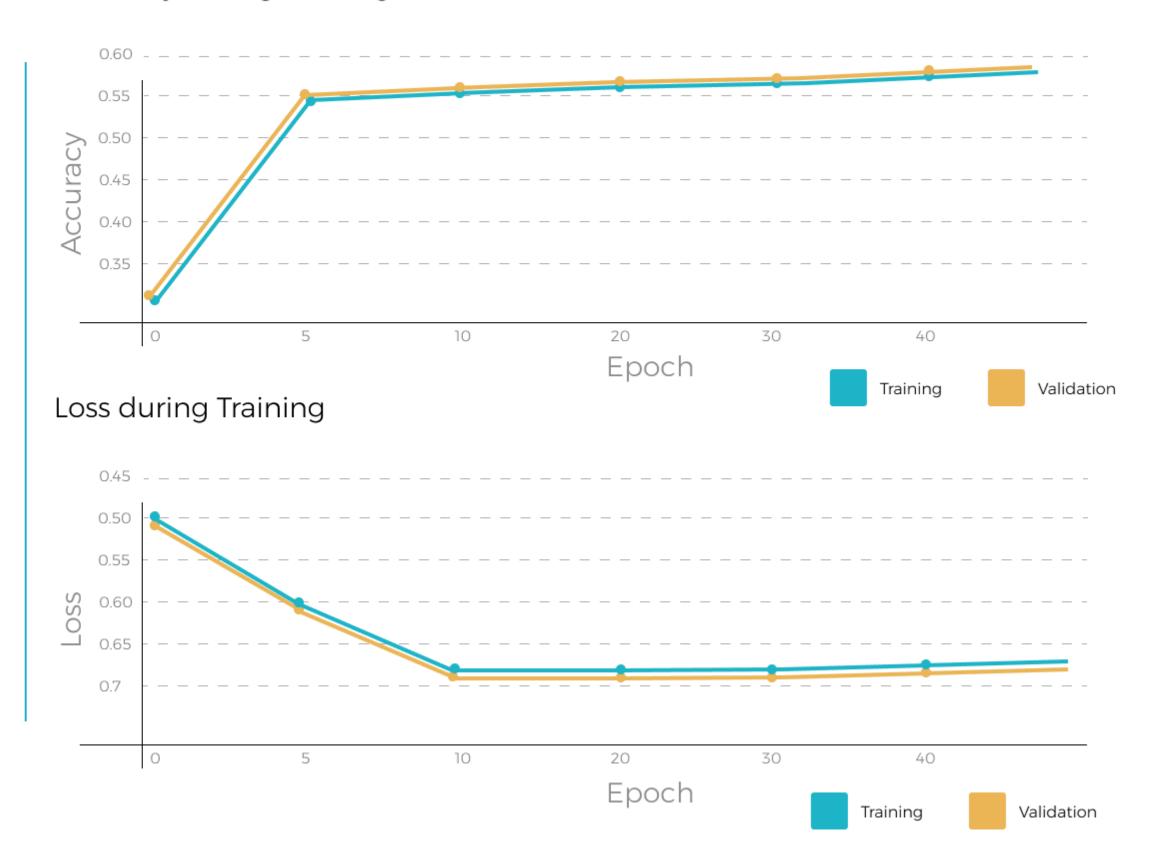
0.55 acurácia de treino

0.56 acurácia de teste

0.68 loss de treino

0.69 loss de teste

#### Accuracy during Training





# Próximos passos



01

## OTIMIZAÇÃO DO MODELO

otimização do modelo de rede neural + relatório técnico



02

## ADIÇÃO DE VARIÁVEIS EXÓGENAS

uso de variáveis externas para avaliar seu impacto no resultado do modelo



03

### PROTÓTIPO DE ALTA FIDELIDADE

representação da interface com foco em design system e acessibilidade



04

## ANÁLISE PESTEL E BUSINESS MODEL CANVAS

avaliar o ambiente macroambiental do projeto e compreender os aspectos fundamentais da solução



# Próximos passos



01

## OTIMIZAÇÃO DO MODELO

otimização do modelo de rede neural + relatório técnico



02

## ADIÇÃO DE VARIÁVEIS EXÓGENAS

uso de variáveis externas para avaliar seu impacto no resultado do modelo



03

## PROTÓTIPO DE ALTA FIDELIDADE

representação da interface com foco em design system e acessibilidade



04

## **ANÁLISE PESTEL E BUSINESS MODEL CANVAS**

avaliar o ambiente macroambiental do projeto e compreender os aspectos fundamentais da solução