

NeuralTraffic

- Disciplina: IN0997 - Redes Neurais
- Data: 27/08/2021
- Docente: Aluizio Fausto Ribeiro Araújo
- Discentes:
 - Adriano Marabuco de A. Lima - amal@cin.ufpe.br
 - Jair Paulino de Sales - jps4@cin.ufpe.br
 - Sara Bandeira Coutinho - sbc2@cin.ufpe.br

Contexto

- Densidade populacional e trânsito;
- Problema de mobilidade;
- O trânsito de Recife é o 15º pior do mundo e o 1º pior do Brasil (Traffic Index Ranking, 2019) ;
- Acidentes de trânsito são um problema de saúde mundial;

Problema

- É possível prever com confiança a ocorrência de acidentes a partir dos dados coletados por medidores de velocidade?

Objetivo

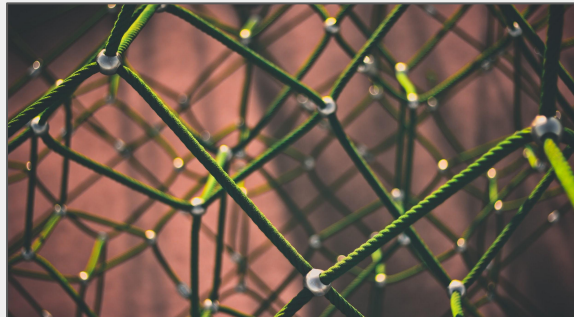
- Fazer previsões com confiança sobre a ocorrência de acidentes em determinada localidade e intervalo de tempo!

Proposta de inovação

Dados reais de trânsito



Aprendizado profundo



Previsão espacial-temporal de
acidentes



Dados

Velocidade
das vias 2015
a 2021

Acidentes
2015 a 2021

The screenshot shows the 'DADOS RECIFE' website interface. At the top, there's a header with the 'PREFEITURA DO RECIFE' logo and navigation links: 'A PREFEITURA', 'SERVIÇOS', 'SECRETARIA E ÓRGÃOS', and 'CONHEÇA O RECIFE'. Below this is a blue navigation bar with 'DADOS RECIFE' logo and links: 'Conjuntos de dados', 'Organizações', 'Grupos', 'Visualizações', 'Consultas Livres', 'Contato', and 'Sobre'. A search bar is on the right. The main content area shows the profile of 'CTTU - AUTARQUIA DE TRÂNSITO E TRANSPORTE URBANO DO RECIFE'. It includes the CTTU logo, a description of the organization, and statistics: '0' seguidores and '21' conjuntos de dados. Below this is a list of organizations with 'Autarquia de Trânsito...' highlighted. There are also sections for 'Grupos' (Mobility, Health, Urbanism) and a list of data sets. The '21 conjuntos de dados encontrados' section lists 'Malha Cicloviária do Recife' and 'Acidentes de Trânsito com e sem vítimas 2015 a 2021'. The 'Velocidade das Vias' section is also visible.

PREFEITURA DO
RECIFE

A PREFEITURA | SERVIÇOS | SECRETARIA E ÓRGÃOS | CONHEÇA O RECIFE

DADOS RECIFE

Conjuntos de dados | Organizações | Grupos | Visualizações | Consultas Livres | Contato | Sobre

Pesquisar

/ Organizações / Autarquia de Trânsito e ...

CTTU
AUTARQUIA DE TRÂNSITO
E TRANSPORTE URBANO DO RECIFE

Autarquia de Trânsito e Transporte Urbano do Recife - CTTU

A Autarquia de Trânsito e Transporte Urbano do Recife (CTTU) <http://ctu.recife.pe.gov.br/> (0800 081 1078) é uma sociedade de economia mista integrante da Administração... Leia mais

Seguidores **0** Conjuntos de dados **21**

Organizações

Autarquia de Trânsi... **21**

Grupos

Mobilidade **18**

Saúde **4**

Urbanismo **3**

Conjuntos de dados | Fluxo de Atividades | Sobre

Pesquisar conjuntos de dados...

21 conjuntos de dados encontrados Ordenar por: Relevância

Malha Cicloviária do Recife

O Conjunto de dados contém os dados georreferenciados da malha cicloviária do Recife. São 186 quilômetros de ciclovias, ciclofaixas e ciclorrotas na cidade, sendo 158 km...

CSV GeoJSON JSON

Acidentes de Trânsito com e sem vítimas 2015 a 2021

Dados dos acidentes com e sem vítimas do trânsito de Recife. Os dados estão disponíveis no formato csv, agrupado por ano. Os dados de 2015 são a partir de junho 2015.

CSV JSON

Velocidade das Vias - Quantitativo por Velocidade Média - 2021

Esse conjunto de dados apresenta os quantitativos de veículos por velocidade média no intervalo de 15 minutos. Os dados apresentados descrevem quantos veículos por equipamento...

CSV JSON

Visão geral dos dados e Análise exploratória

41.8%

Lombada

58.2%

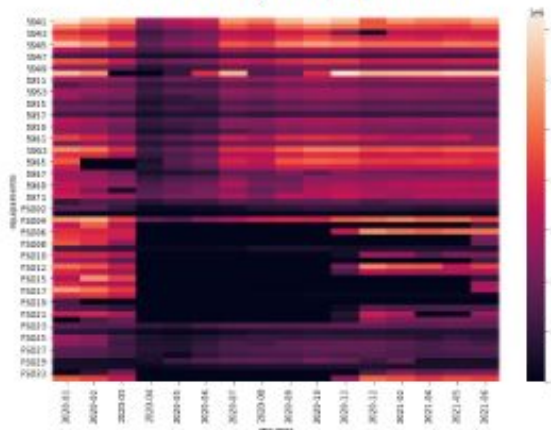
Fotossensor

Quantidade de Veículos por Mês a cada Faixa de Velocidade

2020 , 2021

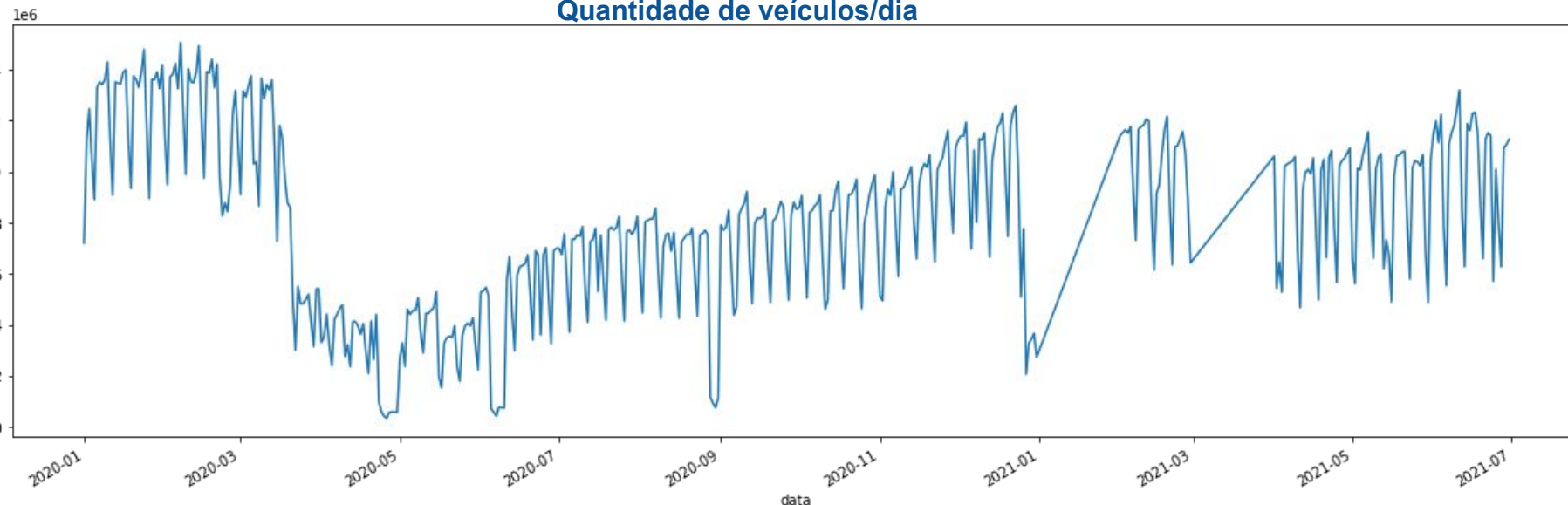


Quantidade por Equipamento



ACIDENTES DE TRÂNSITO NO RECIFE

Quantidade de veículos/dia

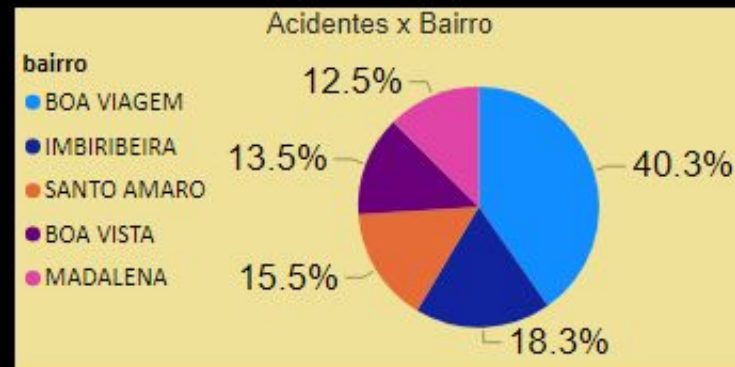
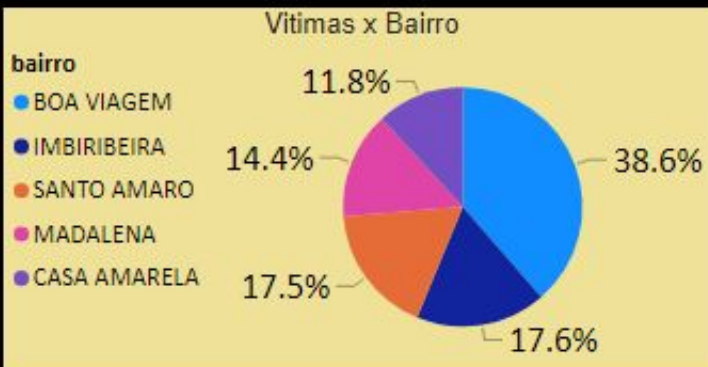


3232

Vítimas

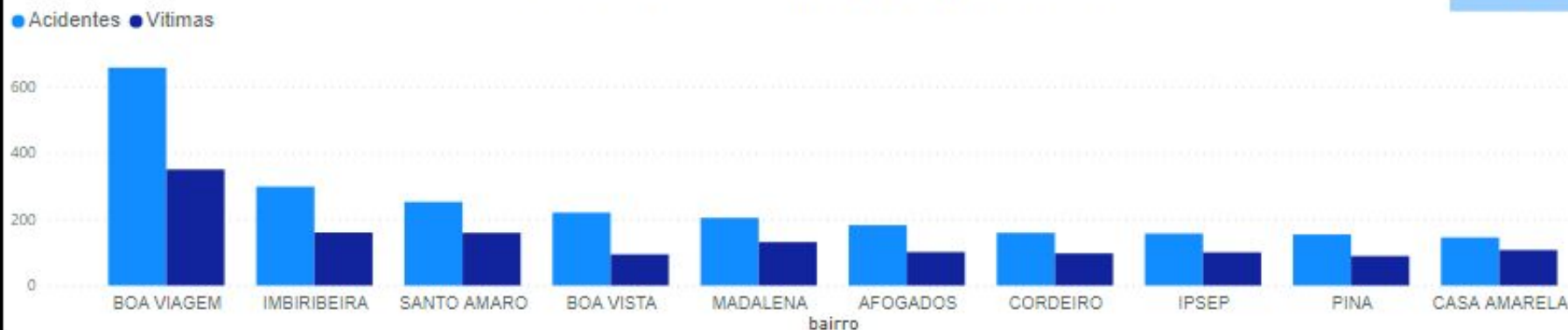
5359

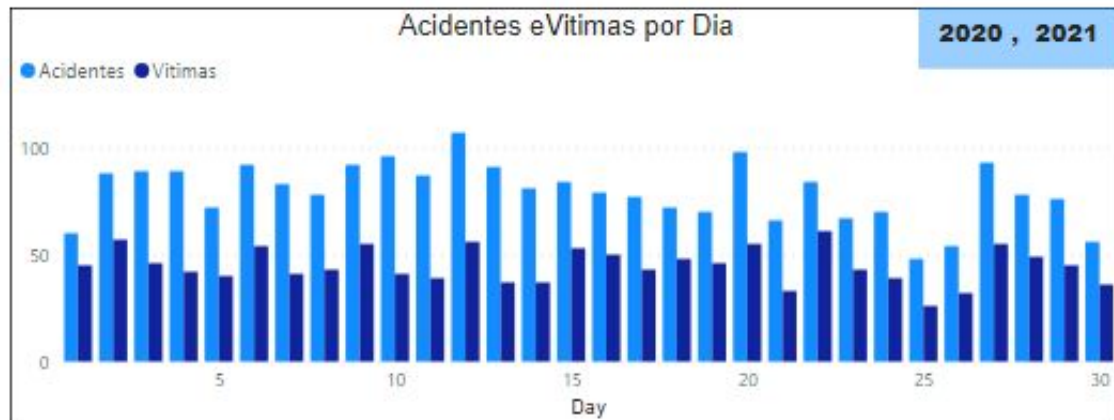
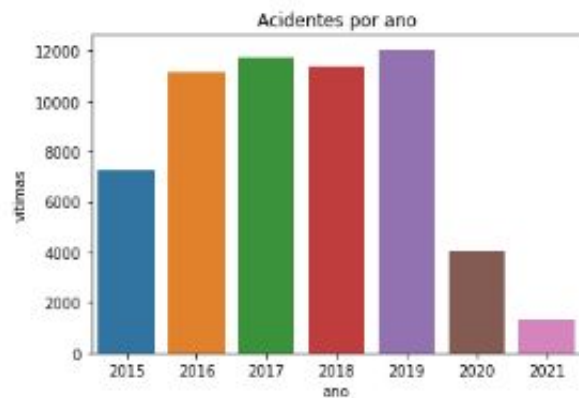
Acidentes



Acidentes e Vítimas pelos principais Bairros

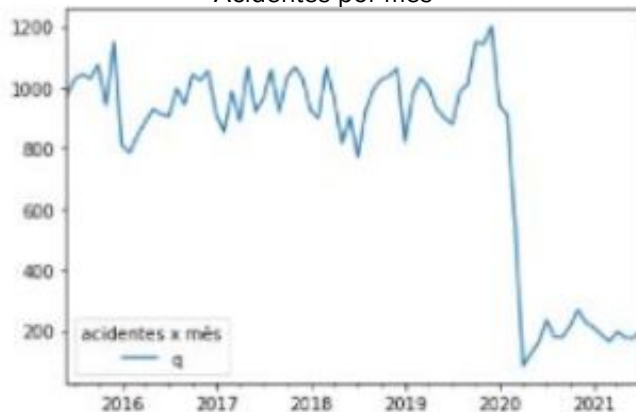
2020 , 2021



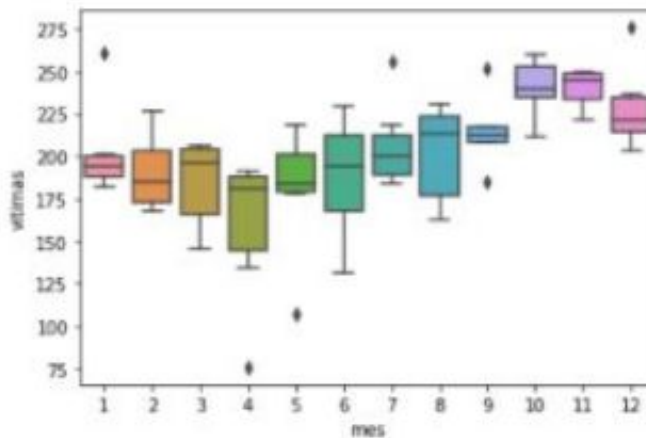
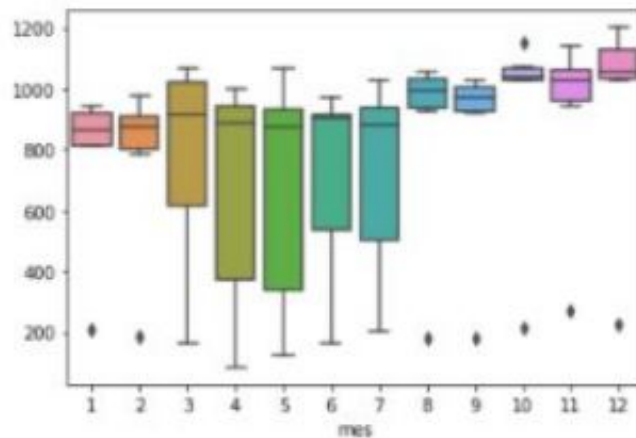
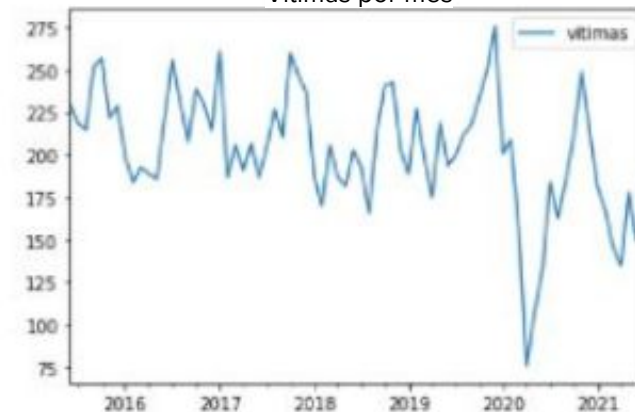


ACIDENTES DE TRÂNSITO NO RECIFE

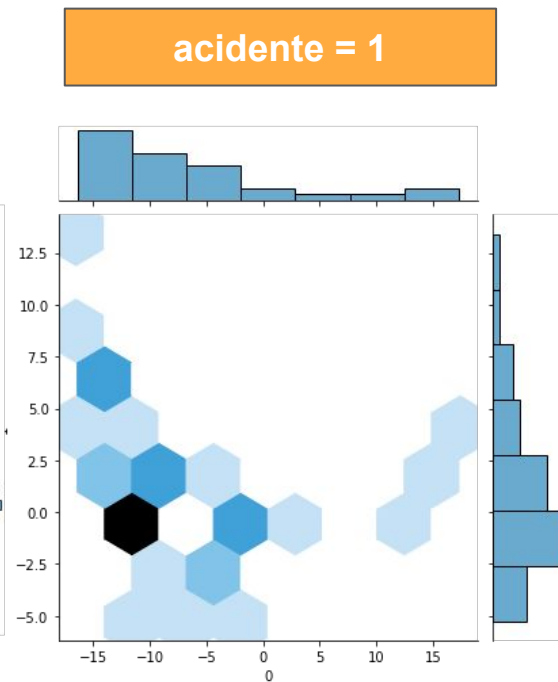
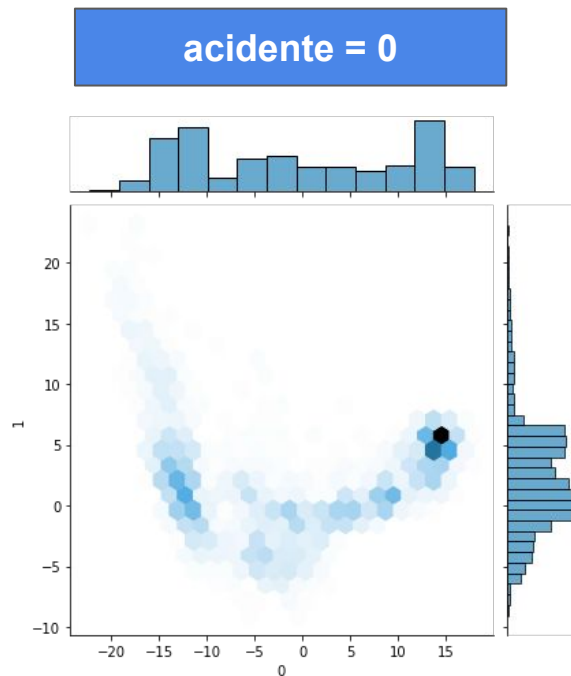
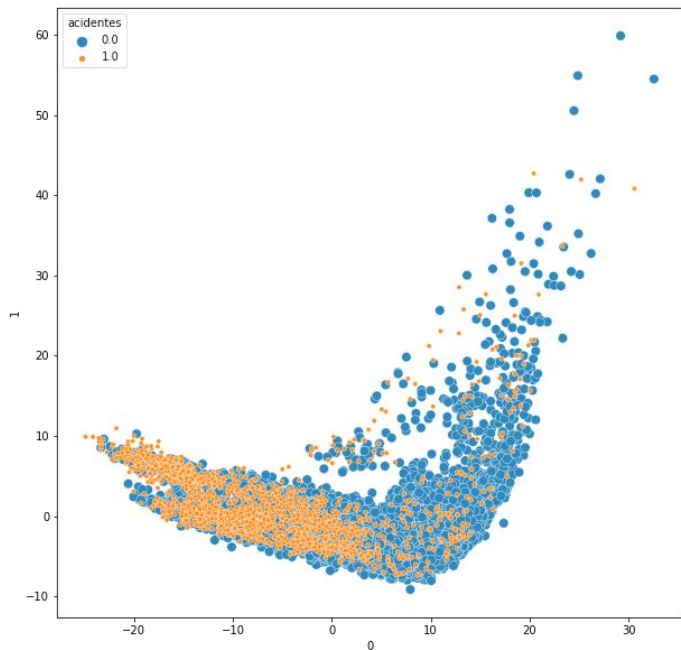
Acidentes por mês



Vítimas por mês



Distribuição das Classes - PCA



Desbalanceamento de classes

Acidente x Hora - Treino

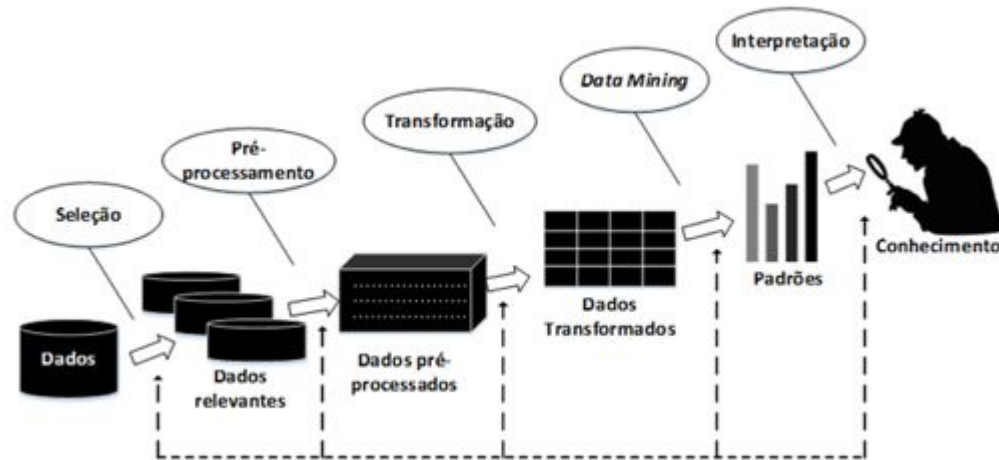
- Classe 0: 243650 (99,08%)
- Classe 1: 2274 (0,92%)

Acidente x Hora - Teste

- Classe 0: 60797 (99,5%)
- Classe 1: 327 (0,5%)

Experimento

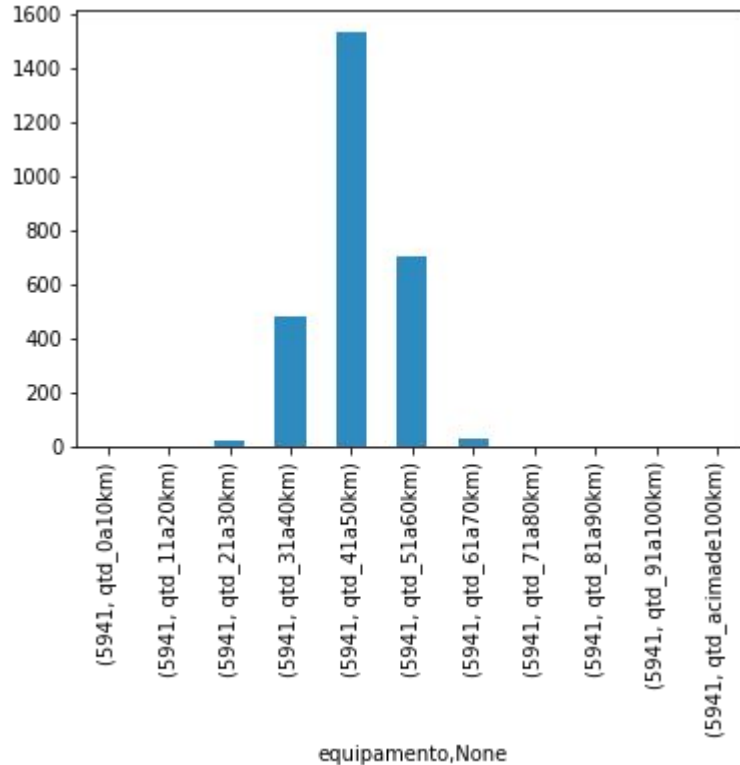
Descoberta de conhecimento



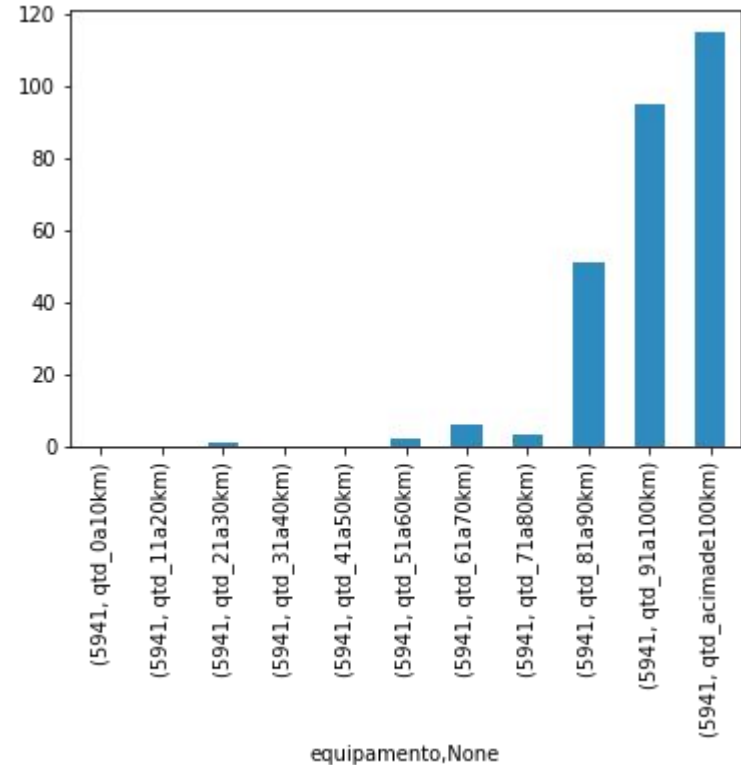
Dados de entrada

AV. MAL. MASCARENHAS DE MORAES

22/02/2020 10:00 am - lomb 5941



14/12/2020 4:00 am - lomb 5941



Transformações

Normalização



$$\mu = 0$$
$$\sigma = 1$$

Imputação



$$\emptyset = 0$$

1/4 do total

Frequência

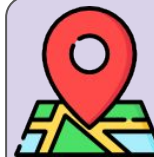


$$\Sigma_{15\text{min}, \text{faixa}} = \text{hora}$$

Treino: 2020

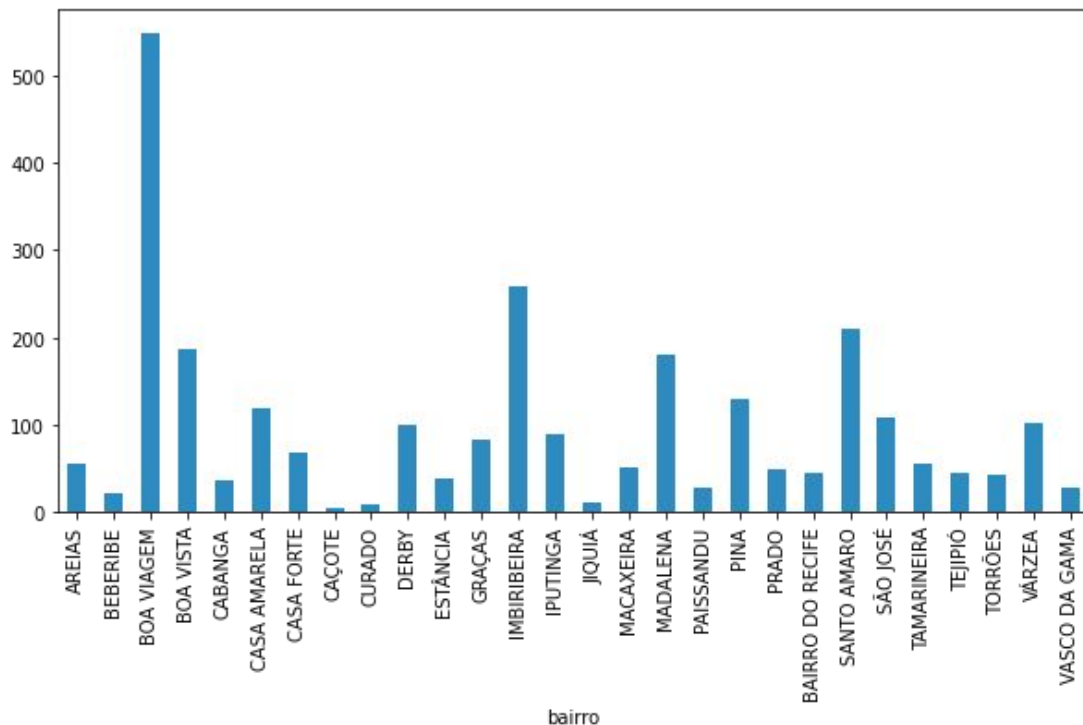
Teste: 04 a 06/2021

Divisão espacial



Filtro por bairro x
localização

Dados de saída



Input: 12 lags

t_{-12}

t_{-n}

t_{-1}

t_0

Forecast: 12 steps ahead, acid = 1

t_{+1}

t_{+2}

t_{+n}

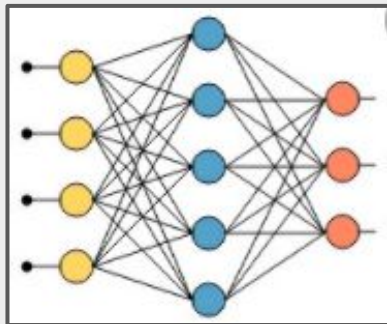
t_{+12}

30 treinamentos e previsões por modelo

Modelos

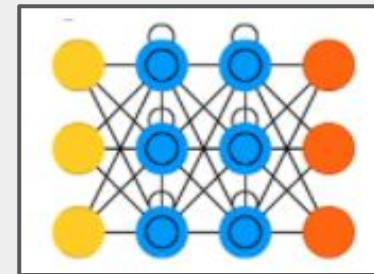
MLP

```
Dense(128, activation='relu'),
Dropout(0,5),
Dense(128, activation='relu'),
Dropout(0,5),
Dense(128, activation='relu'),
Dropout(0,5),
Flatten(),
Dense(28, activation='sigmoid')
```



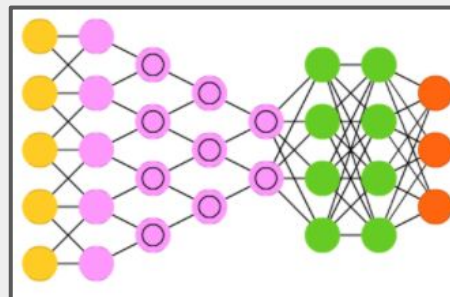
LSTM

```
LSTM(128,
return_sequences=True),
Dropout(0,5),
LSTM(128, return_sequences=False),
Dropout(0,5),
Flatten(),
Dense(28, activation='sigmoid')
```



CNN

```
Conv1D(filters=64,
kernel_size=(CONV_WIDTH,), activation='relu'),
Dropout(0,5),
Dense(128, activation='relu'),
Dropout(0,5),
Flatten(),
Dense(28, activation='sigmoid')
```

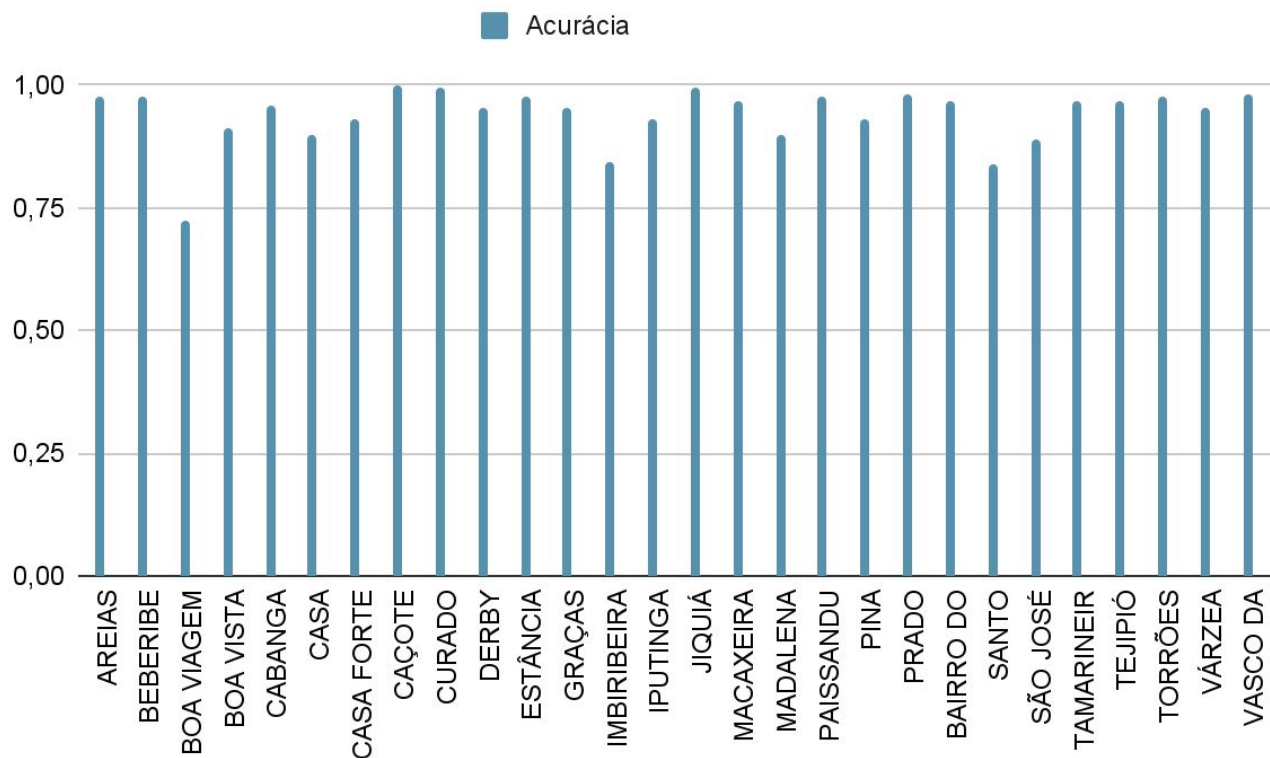


Resultados

Resultados

- Métricas;
- Tabelas para comparação dos modelos;
- Teste de hipótese.

Acurácia *baseline*



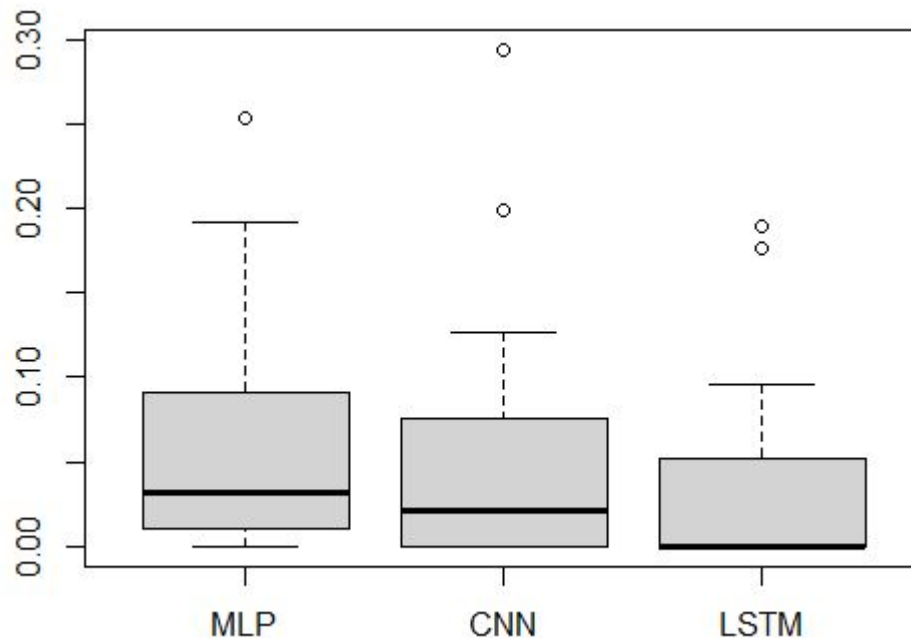
Valor médio por bairro
para trinta iterações.

Bairro	MLP	CNN	LSTM	MÉDIA BAIRRO
AREIAS	2,30%	0,66%	0,00%	0,98%
BEBERIBE	0,00%	0,00%	0,00%	0,00%
BOA VIAGEM	25,37%	29,37%	17,59%	24,11%
BOA VISTA	11,45%	11,36%	9,55%	10,79%
CABANGA	5,81%	2,82%	0,00%	2,88%
CASA AMARELA	6,66%	3,48%	0,77%	3,64%
CASA FORTE	2,83%	0,22%	0,00%	1,02%
CAÇOTE	0,00%	0,00%	0,00%	0,00%
CURADO	0,00%	0,00%	0,00%	0,00%
DERBY	9,37%	12,25%	9,59%	10,40%
ESTÂNCIA	1,37%	0,66%	0,00%	0,68%
GRAÇAS	1,67%	2,01%	0,00%	1,23%
IMBIRIBEIRA	12,14%	12,30%	8,76%	11,07%
IPUTINGA	0,69%	2,33%	0,00%	1,01%
JIQUEIÁ	0,00%	0,00%	0,00%	0,00%
MACAXEIRA	2,71%	2,07%	0,00%	1,59%
MADALENA	19,17%	19,85%	18,91%	19,31%
PAISSANDU	4,53%	0,26%	0,00%	1,59%
PINA	11,17%	10,29%	8,44%	9,96%
PRADO	0,00%	0,00%	0,00%	0,00%
BAIRRO DO RECIFE	3,51%	2,49%	0,00%	2,00%
SANTO AMARO	13,07%	12,66%	8,48%	11,40%
SÃO JOSÉ	8,82%	3,57%	0,00%	4,13%
TAMARINEIRA	2,44%	0,00%	0,00%	0,81%
TEJIPIÓ	2,36%	0,00%	0,00%	0,79%
TORRÕES	5,99%	4,71%	0,00%	3,57%
VÁRZEA	7,92%	0,64%	2,00%	3,52%
VASCO DA GAMA	0,75%	0,00%	0,00%	0,25%
MÉDIA	5,79%	4,79%	3,00%	4,53%

Resultado - F1

Bairro	MLP	CNN	LSTM	Média
BOA VIAGEM	25,37%	29,37%	17,59%	24,11%
BOA VISTA	11,45%	11,36%	9,55%	10,79%
IMBIRIBEIRA	12,14%	12,30%	8,76%	11,07%
MADALENA	19,17%	19,85%	18,91%	19,31%
SANTO AMARO	13,07%	12,66%	8,48%	11,40%
VÁRZEA	7,92%	0,64%	2,00%	3,52%
BEBERIBE	0,00%	0,00%	0,00%	0,00%

Boxplot - F1



Dados ***não seguem distribuição normal.***
(Teste Shapiro-Wilk: $W = 0.74$, $p < 0.001$).

Teste de Kruskal-Wallis

H0: Os modelos apresentaram desempenho igual.

H1: Há pelo menos um modelo melhor.

Resultado: chi-squared = 10.011, df = 2, **p-value = 0.0067.**

Interpretação: a um nível nível de significância de 5%, não se aceita a hipótese nula de que todos os modelos apresentam o mesmo desempenho, ou seja, pelo menos um modelo difere dos demais.

Teste de Mann-Whitney (não paramétrico)

H0: Os modelos possuem o mesmo desempenho.

H1: Os modelos possuem desempenho diferente.

Modelo	MLP	LSTM	CNN
MLP	-	0.003 (564)	0.226 (466)
LSTM	0.003 (564)	-	0.021 (259)
CNN	0.226 (466)	0.021 (259)	-

* adotou-se $\alpha = 0.05$

Teste de Mann-Whitney

Há evidências para:



Não há evidências para:



Limitações

Limitação	Descrição
Dados incompletos	Dados das lombadas eletrônicas de 01 e 03/2021 estão ausentes
Dados ausentes	Disponibilidade de informação dos fotossensores e lombadas varia com tempo
Mudança de padrão	Alteração no padrão de 2015 a 2019 x 2020 e 2021
Viés dos dados	Medidores funcionam como redutores de velocidade, não registram condições reais de velocidade
Restrição espacial	Quantidade e distribuição espacial dos fotossensores reduz amostra dos dados
Falta de material de referência	Não foram encontrados na literatura trabalhos com abordagens semelhantes;

Melhorias futuras

- *Big data;*
- Balanceamento dos dados;
- *Data augmentation;*
- Aprimoramento das redes.

Referências

- Classificação de dados desequilibrados:
https://www.tensorflow.org/tutorials/structured_data/imbalanced_data
- Previsão de série temporal:
https://www.tensorflow.org/tutorials/structured_data/time_series
- Contexto do problema:
<https://www.scielo.org/article/rbepid/2017.v20n4/727-741/pt/> e
<http://revistas.poli.br/index.php/repa/article/view/1707/740>
- Rede CNNLSTM:
<https://medium.com/swlh/reading-3d-cnn-lstm-deep-neural-networks-for-no-reference-video-quality-assessment-e70359cce64c>