

# Projeto Criminalidade Chicago

<sup>1</sup>Alvaro Cristian da Silva Botelho

<sup>1</sup>MBA Data Science e Big Data – Universidade do Vale dos Sinos (UNISINOS)  
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brazil

Cristian.ucpel@gmail.com

## INTRODUÇÃO

O presente documento visa relatar a aplicação do algoritmo Prophet no conjunto de dados sazonais obtidos através da plataforma Kaggle <https://www.kaggle.com/currie32/crimes-in-chicago>. A base contém 7.941.282 Linhas e 23 atributos. Mapas e os scripts foram anexados juntos deste documento. Mais informações estão no <https://github.com/alcrisian/MachineLearningTask10-23-2018>.

## Experimento

Primeiro testaremos como o algoritmo se comporta com os dados sazonais de somente 1 ano.

```
In [25]: #agrupando total de crimes
dfy = df_crime_data_cartheft2016.groupby(['Date']).size()
dfy = dfy.reset_index()
dfy.columns = ['Date', 'Qtd_Crimes']
dfy = dfy.groupby(dfy['Date'].dt.date).size()
dfy = dfy.reset_index()
dfy.columns = ['Date', 'Qtd_Crimes']
dfy['Date'] = pd.to_datetime(dfy['Date'])

dfy['Qtd_Crimes'] = dfy['Qtd_Crimes'].astype(float)
dfy = dfy.reset_index()
dfy.head()
```

```
Out[25]:
```

	index	Date	Qtd_Crimes
0	0	2016-01-01	22.0
1	1	2016-01-02	22.0
2	2	2016-01-03	24.0
3	3	2016-01-04	25.0
4	4	2016-01-05	19.0

```
In [26]: dfy.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 366 entries, 0 to 365
Data columns (total 3 columns):
index      366 non-null int64
Date       366 non-null datetime64[ns]
Qtd_Crimes 366 non-null float64
dtypes: datetime64[ns](1), float64(1), int64(1)
memory usage: 8.7 KB
```

Aplicação do algoritmo fbProphet aos dados criminais do ano de 2016.

```
In [27]: #aplicando FbProphet para prever os dados utilizando somente 2016 como base.
from fbprophet import Prophet
dfy = dfy.rename(columns={'Date': 'ds', 'Qtd_Crimes': 'y'})
CarTheft_model = Prophet(interval_width=0.95)
CarTheft_model.fit(dfy)

INFO:fbprophet.forecaster:Disabling yearly seasonality. Run prophet with yearly_seasonality=True to override this.
INFO:fbprophet.forecaster:Disabling daily seasonality. Run prophet with daily_seasonality=True to override this.
C:\Users\chaos\Anaconda3\lib\site-packages\pystan\misc.py:399: FutureWarning: Conversion of the second argument of issubdtype
from 'float' to 'np.floating' is deprecated. In future, it will be treated as 'np.float64 == np.dtype(float).type'.
    elif np.issubdtype(np.asarray(v).dtype, float):
```

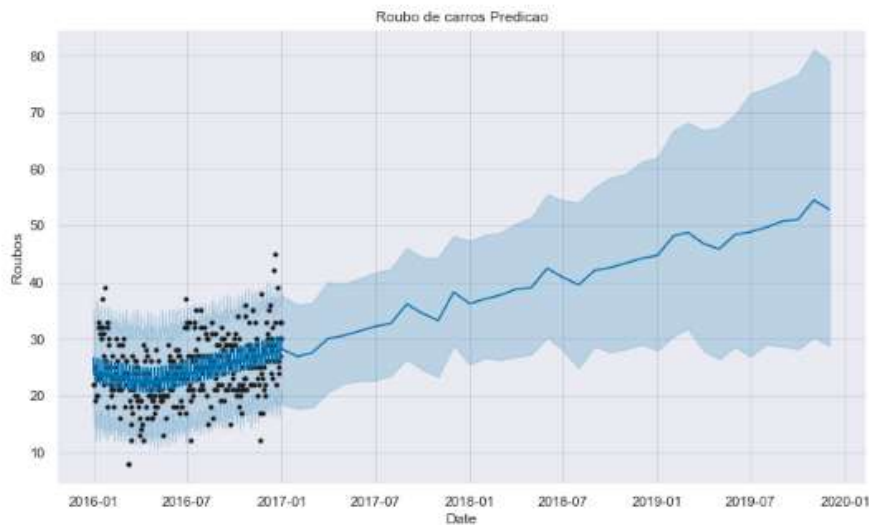
```
Out[27]: <fbprophet.forecaster.Prophet at 0x1cb85d68fd0>
```

```
In [28]: CarTheft_forecast = CarTheft_model.make_future_dataframe(periods=36, freq='MS')
CarTheft_forecast = CarTheft_model.predict(CarTheft_forecast)
```

```
In [28]: CarTheft_forecast = CarTheft_model.make_future_dataframe(periods=36, freq='MS')
CarTheft_forecast = CarTheft_model.predict(CarTheft_forecast)
```

```
In [29]: plt.figure(figsize=(18, 6))
CarTheft_model.plot(CarTheft_forecast, xlabel = 'Date', ylabel = 'Roubos')
plt.title('Roubo de carros Predicao');
```

<Figure size 1296x432 with 0 Axes>



Após a execução do algoritmo fbProphet para criação do forecast, podemos ver que utilizando somente 1 o ano de 2016 e previsto um aumento no roubo de carros com a predição ate 2020.

```
In [31]: #aplicando FbProphet para predizer os dados utilizando toda a base.
```

```
dfall = dfall.rename(columns={'Date': 'ds', 'Qtd_Crimes': 'y'})
CarTheftAll_model = Prophet(interval_width=0.95)
CarTheftAll_model.fit(dfall)
```

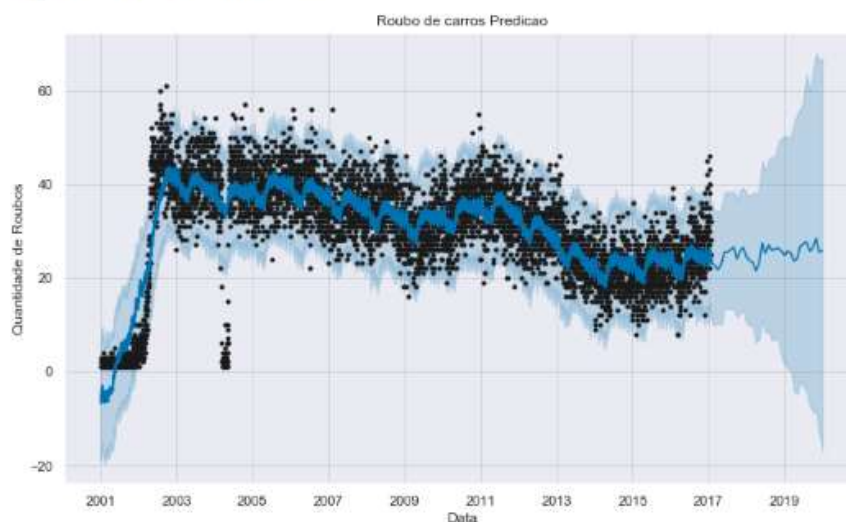
INFO:fbprophet.forecaster:Disabling daily seasonality. Run prophet with daily\_seasonality=True to override this.  
C:\Users\chaos\Anaconda3\lib\site-packages\pystan\misc.py:399: FutureWarning: Conversion of the second argument of issubdtype from 'float' to 'np.floating' is deprecated. In future, it will be treated as 'np.float64 == np.dtype(float).type'.  
elif np.issubdtype(np.asarray(v).dtype, float):

```
Out[31]: <fbprophet.forecaster.Prophet at 0x1c6fb6a5c50>
```

```
In [32]: CarTheftAll_forecast = CarTheftAll_model.make_future_dataframe(periods=36, freq='MS')
CarTheftAll_forecast = CarTheftAll_model.predict(CarTheftAll_forecast)
```

```
In [33]: plt.figure(figsize=(18, 6))
CarTheftAll_model.plot(CarTheftAll_forecast, xlabel = 'Data', ylabel = 'Quantidade de Roubos')
plt.title('Roubo de carros Predicao');
```

<Figure size 1296x432 with 0 Axes>



Utilizando os dados de toda a base de dados podemos ver que a linha de predição se mante estável, então podemos inferir que existe uma tendência na diminuição da criminalidade.



A análise de componentes do algoritmo Prophet mostra todas as tendências que foram utilizadas quando o algoritmo rodou.

Plotando os *trending* captados pelo algoritmo podemos ver que existe uma tendência, se olharmos semanalmente a sexta feira existe uma probabilidade maior de ocorrer roubo de carros. Já no plano mensal, Julho e Agosto, são os meses onde mais ocorrem roubos. E o primeiro gráfico confirma a tendência leve de aumento no roubo de carros como podemos ver no plot da predição inicial.

```
In [43]: #Visualizando a predicao
forecast = CarTheftAll_forecast.rename(columns={'ds': 'Date'})
forecast.head()

plt.figure(figsize=(10, 7))
plt.plot(forecast['Date'], forecast['trend'], 'b-')

Out[43]: [<matplotlib.lines.Line2D at 0x1cb81c570b8>]
```



No último gráfico gerado com dados preditos, podemos ver o aumento na criminalidade.