

# Número de mortes por poluição do ar utilizando RNA MLP

Matheus Fontinele Alves Vieira  
João Victor Fernandes Pinto





# Sumário

- Base de dados
- Problemática
- Materiais e métodos utilizados
- Resultados
- Conclusão



# Base de dados

- A base de dados informa a quantidade de mortes causadas pela poluição do ar de 208 países durante o período de 1990 a 2019.
- Será utilizado dessa base de dados 1 entrada e 1 saída, esta saída consistirá na previsão do aumento ou diminuição de mortes por poluição em um ano posterior.
- Por razões de demonstração foi utilizado o Brasil dos 208, o que resulta em 30 instâncias
- Por razões de demonstração foi utilizado o Brasil dos 208, o que resulta em 30 instâncias.



## Base de dados utilizada

- A entrada terá a coluna referente aos 30 anos, conforme a base de dados.
- A saída será o indicador do número de mortes por poluição de cada ano.

	Entity	Year	Air pollution
780	Brazil	1990	86653
781	Brazil	1991	82606
782	Brazil	1992	80515
783	Brazil	1993	79970
784	Brazil	1994	77697
785	Brazil	1995	75808
786	Brazil	1996	74643
787	Brazil	1997	72801
788	Brazil	1998	72541
789	Brazil	1999	71693
790	Brazil	2000	70631
791	Brazil	2001	70380
792	Brazil	2002	70325
793	Brazil	2003	70275
794	Brazil	2004	69869
795	Brazil	2005	68220
796	Brazil	2006	68260
797	Brazil	2007	68088



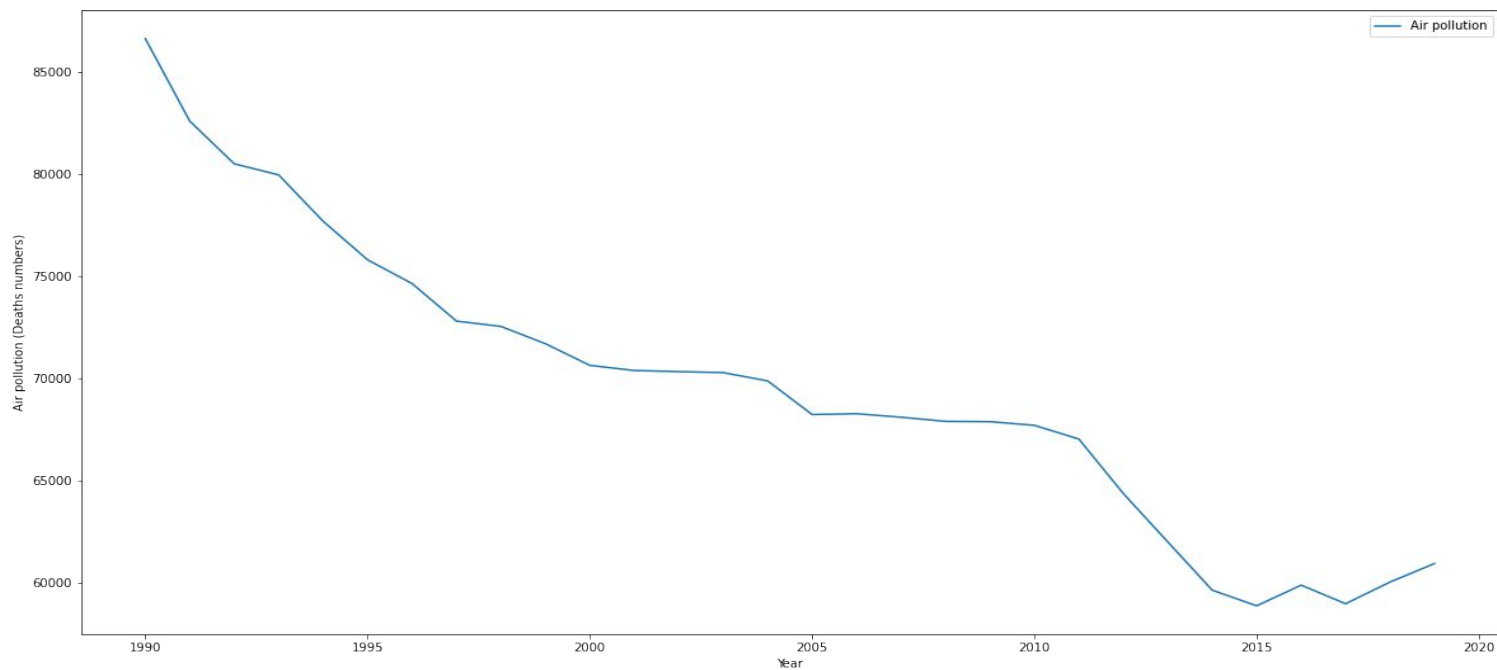
# Problemática

- Uso de um modelo de regressão linear para prever o comportamento do número de mortes causadas pela poluição do ar no Brasil.
- A poluição do ar é um dos principais fatores de risco de mortes no mundo, devido aos gases emitidos na combustão de resíduos químicos como petróleo e derivados serem altamente tóxicos e nocivos à saúde.
- Uma pessoa exposta a esses gases podem causar doenças respiratórias e até mesmo câncer.



# Problemática

Gráfico referente aos registros de mortes por poluição no Brasil





# Materiais e métodos utilizados

- Foi utilizado a biblioteca **pandas** e o **numpy** para o tratamento de dados.
- Foi excluído as colunas desnecessárias para o modelo, com exceção das colunas **Entity**, referente ao nome do país, **Year**, o ano do registro e **Air Pollution**, exibindo a quantidade de mortes por poluição.
- Para a geração dos gráficos foi utilizada a biblioteca **matplotlib.pyplot**.
- Foi utilizado o **SGDRegressor** como modelo para a predição dos valores.



# Processamento

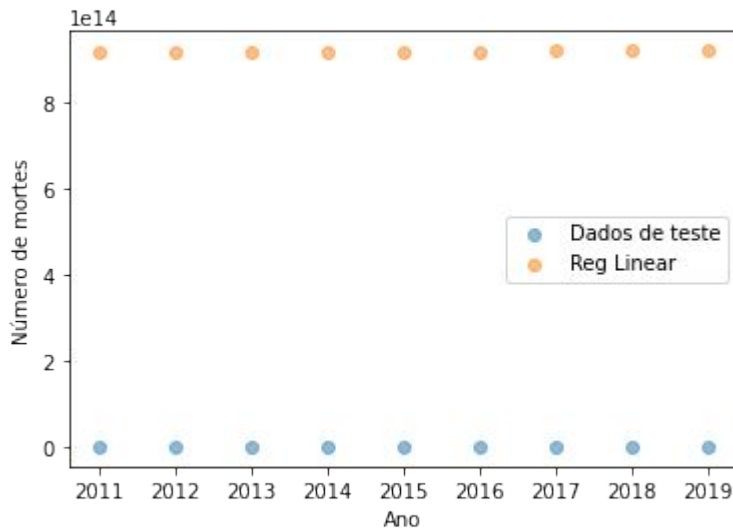
- Em seguida foi feito o processamento dos dados com o SGDRegressor, em que o número máximo de iterações será 2000, uma tolerância de 0.0000001, uma taxa de aprendizado de 0.01 constante.





# Resultados

- Feito o treinamento e testes, o algoritmo gerou o resultado do gráfico ao lado.
- O resultado obtido do coeficiente R2 da regressão linear foi:  
**-1.2494969730007154e+23**





## Conclusão

- Com isso, executado todos os algoritmos, para o ano de 2020 a previsão é de que haja um aumento de **9.20931438e+14** pessoas mortas por poluição do ar no Brasil.

