# Projeto 2 – Conclusão

O Projeto 2 de Ciência dos Dados teve como objetivo fazer um programa para uma determinada empresa que analisa as mensagens disponíveis, essas mensagens foram obtidas pelo Twitter.

Para isso, primeiramente, foi escolhido a empresa em que iriamos avaliar, TESLA, em seguida pegou-se twetts sobre ela e, manualmente, classificou-se as mensagens em Relevante e Irrelevante.

Feito isso, fomos para a parte de programação em que criamos várias funções a fim de: limpar as mensagens, tirar caracteres que não irão influenciar se as mensagens são ou não relevantes; obter as porcentagens e as proporções das palavras aparecerem no geral e nas classificações; para então, através do Teorema de Bayes conseguir fazer com que a máquina aprenda a analisar as mensagens de forma automática.

Tendo em vista o percentual bom que tivemos ao comparar as classificações mecânicas com as que a máquina nos retornou, nos lapidamos um pouco mais o programa, fazendo com que fossem ignoradas as palavras com menos de duas letras – essas palavram não são relevantes para a empresa, pois a maioria é artigo, conjunção, ... -, além de adicionar mais colunas as nossas classificações, tendo os seguintes termos: "Irrelevante", "Negativo", "Neutro", "Positivo". Assim, obtivemos, novamente, um resultado bom de: (porcentagem do resultado).

Com isso, concluímos que o programa irá ajudar a empresa que nos contratou, além de nos abrir portas para outras empresa, pois como nossa programação está escrito com funções, permite que analisamos não só os dados desse contexto quanto os demais.