

TRABALHO-1

Eduardo Ouriques e Cristiano Marques

Introdução

Com a utilização de um dataset de áudio chamado “*Urban Sound*”, o grupo desenvolveu um modelo para o problema de classificação de áudio. O *dataset* contém dez classes, são elas: *air_conditioner*, *car_horn*, *children_playing*, *dog_bark*, *drilling*, *engine_idling*, *gun_shot*, *jackhammer*, *siren* e *street_music*.

Features

Ao extrairmos o áudio com a biblioteca “*librosa*”, o grupo teve o trabalho de extrair as dez *features* listadas abaixo para cada áudio:

- `librosa.feature.chroma_stft(signal).mean()`,
- `librosa.feature.chroma_cqt(signal).mean()`,
- `librosa.feature.chroma_cens(signal).mean()`,
- `librosa.feature.mfcc(signal).mean()`,
- `librosa.feature.rms(signal).mean()`,
- `librosa.feature.spectral_centroid(signal).mean()`,
- `librosa.feature.spectral_bandwidth(signal).mean()`,
- `librosa.feature.spectral_contrast(signal).mean()`,
- `librosa.feature.spectral_flatness(signal).mean()`,
- `librosa.feature.zero_crossing_rate(signal).mean()`

Métricas

- TREINAMENTO

```
Epoch 5/100
5238/5238 [=====] - 33s 6ms/step - loss: 0.7254 - accuracy: 0.7566 - val_loss: 5.4239e-04
- val_accuracy: 1238.0000
Epoch 6/100
5238/5238 [=====] - 32s 6ms/step - loss: 0.7052 - accuracy: 0.7631 - val_loss: 5.4146e-04
- val_accuracy: 1242.0000
Epoch 7/100
5238/5238 [=====] - 32s 6ms/step - loss: 0.6904 - accuracy: 0.7688 - val_loss: 5.4105e-04
- val_accuracy: 1253.0000
```

Conforme imagem acima, o treinamento com 60% das amostras, obteve uma acurácia de 76.88%. No código foi informado 100 épocas de iteração, mas ao mesmo tempo foi adicionado uma função de callback de parada no treinamento para aqueles caso em que uma nova época não faça diferença. Importante ressaltar que para a etapa de validação foi especificado 20% das amostras.

- TESTE

```
from sklearn.metrics import accuracy_score
target_x_test = [get_result_index(x) for x in result]
accuracy = accuracy_score(y_test, target_x_test)
print("Acurácia: ", accuracy)
```

```
Acurácia: 0.7006296508299943
```

Conforme imagem acima, o teste com 20% do total de amostras, atingiu uma acurácia de 70%.

Conclusão

O grupo conclui uma série de aprendizados, tal como, a quantidade de features escolhidas e a seleção das mesmas tem total relação com o resultado. Além disso, também tivemos um melhor entendimento das camadas do modelo, como ele funciona, e quais as opções que temos para ajustar e melhorar o mesmo. Neste caso, foi identificado que a quantidade de neurônios de saída da primeira camada tem total relação com o resultado apresentado na acurácia, quando diminuimos a quantidade de neurônios de 100 para 20, obtivemos uma acurácia menor de 59% na etapa de teste. Quando utilizamos 100 neurônios da saída da primeira camada, obtivemos o resultado apresentado no capítulo de Métricas deste documento.

Link do Kaggle:

<https://www.kaggle.com/eduardocouriques/eduardo-ouriques-e-cristiano-marques>