

# PIBIC UNICAMP

## O impacto da variação de tempo de análise e da seleção de features na detecção automática de propagandas em rádio

Aluno: Victor Paganotto victor\_9262@hotmail.com

Orientador: Tiago Fernandes Tavares tavares@dca.fee.unicamp.br



### I. INTRODUÇÃO

O rádio é um dos meios de comunicação em massa mais antigos, e continua sendo um importante difusor de informação e entretenimento ao redor do mundo. Por isso é muito utilizado para a circulação de propagandas. Este projeto busca criar um programa de computador capaz de detectar propagandas em programas de rádio. Para isso, usamos técnicas ligadas a Machine Learning, em especial a seleção automática de features.

A base de dados utilizada nesse projeto foi coletada a partir de programas de rádios brasileiras online. As rádios coletadas foram a Transmérica FM <sup>1</sup>, Jovem Pan FM <sup>2</sup> e Radio Bandeirantes <sup>3</sup>. Esses trechos são classificados manualmente em propaganda e não propaganda.

A distribuição da base de dados pode ser observada na Figura 1. Tomamos o cuidado de ter quantidades próximas de trechos de áudio de propagandas e não-propagandas. Dessa forma temos uma base de dados equilibrada.

Duração [s]	Propagandas	Não-Propagandas
10	1007	946
5	2010	2024
2	5020	5060
1	9840	9842

Fig. 1. Número de trechos em cada categoria de áudio para cada diferente duração de trechos

### II. MÉTODO

A Figura 2 mostra uma visão geral do procedimento de teste realizado. Em nossas avaliações, cada trecho de áudio

analisado foi transformado em um vetor de features usando estatísticas relacionadas a descritores espectrais de baixo nível [1]. O conjunto de vetores de features foi então fornecido como entrada a um classificador SVM [2]. Também, avaliamos as possibilidades de preceder o classificador por um processo de redução de dimensionalidade. Para este caso, usamos duas possibilidades: o PCA e a redução de dimensionalidades por ANOVA.

O classificador SVM consiste em uma técnica de aprendizado para problemas de reconhecimento de padrões. Ele sempre é executado em duas etapas: treino, na qual otimiza seus parâmetros para modelar pares entrada-saída relacionados a uma partição do conjunto de dados, e teste, na qual gera saídas relacionadas à outra partição.

Em nosso experimento, as entradas são os vetores de features relacionados a cada trecho de áudio, e a saída correspondente é a sua classe (propaganda ou não-propaganda).

A redução de dimensionalidade busca reduzir o impacto da maldição da dimensionalidade [3] [4]. Essa redução pode ser feita usando PCA, que consiste em uma projeção linear dos dados para um espaço de menor dimensão de tal forma que o erro ao se reconstituir os dados originais seja minimizado.

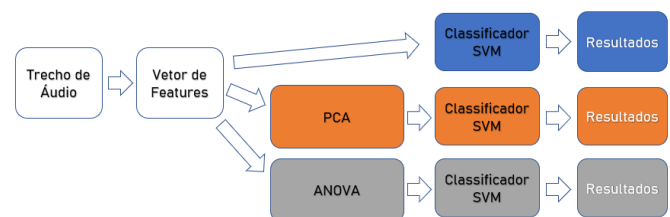


Fig. 2. Diagrama de blocos mostrando os diferentes experimentos realizados

<sup>1</sup><http://www.radiotransamerica.com.br/play/sao-paulo>

<sup>2</sup><https://jovempan.com.br/>

<sup>3</sup><https://radiobandeirantes.band.uol.com.br/>

A técnica ANOVA, por sua vez, consiste em escolher as dimensões do vetor de entrada que levam a diferenças mais estatisticamente significativas entre as classes.

### III. RESULTADOS

Para a avaliação dos resultados utilizamos as métricas *Precision* (mostrada na Equação 1), *Recall* (mostrada na Equação 2) e F1-score. O F1-score (mostrada na Equação 3) foi o utilizado para comparação, uma vez que ele pondera as duas outras métricas.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (1)$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (2)$$

$$\text{F1-Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

A partir da análise da Figura 3, podemos notar que o programa obteve um melhor resultado sem a implementação de nenhum método de redução de dimensionalidade. A duração dos trechos de áudio teve um impacto relativamente desprezível na detecção. A maior diferença observada foi de 3%. No entanto de acordo com a tabela o resultado melhora com a diminuição da duração do trecho de áudio.

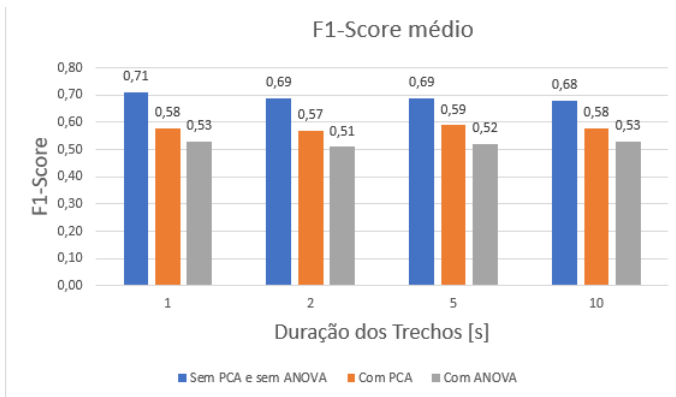


Fig. 3. F1-Score médio para experimentos de classificação com diferentes métodos de redução de dimensionalidade e diferentes durações do trecho de áudio analisado.

### IV. CONCLUSÃO

Neste trabalho, testamos três variações de classificadores para discriminar trechos de áudio contendo propagandas daqueles contendo outros conteúdos. Em nossos resultados, verificamos que é possível fazer essa discriminação apenas com 1s de áudio, e que o aumento do tamanho desse trecho não impacta significativamente os resultados. Essa característica indica que essa ideia pode ser incorporada, futuramente, em sistemas comerciais.

A redução de dimensionalidade com PCA e ANOVA levou à piora do F1-Score das predições. Isso significa que esse tipo de técnica não é interessante para esta aplicação.

Embora os resultados alcançados sejam promissores, não temos indicações de que esses mesmos resultados se generalizem em outras bases de dados, contendo diferentes rádios ou outras propagandas. Esta análise é um caminho interessante para trabalhos futuros.

### REFERENCES

- [1] P. Cook G. Tzanetakis. Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, 10, 2002.
- [2] PUC-Rio. Classificador svm.
- [3] SciKit-Learn Community. Feature selection pipeline.
- [4] Dayong Mei Fengxi Song, Zhongwei Guo. Feature selection using principal component analysis. *International Conference on System Science, Engineering Design and Manufacturing Informatization.*, 2010.