

# XVII JORNADA DE INICIAÇÃO CIENTÍFICA E INICIAÇÃO EM DESENVOLVIMENTO TECNOLÓGICO E INOVAÇÃO

Bicentenário da Independência: 200 anos de Ciência, Tecnologia e Inovação no Brasil

ESTUDO E  
APLICAÇÕES DE  
MÁQUINAS DE  
SUPORTE VETORIAL.



UNIVERSIDADE  
CATÓLICA  
DE SANTOS



# IDENTIFICAÇÃO DO PROJETO

- Estudo e Aplicações de Máquinas de Suporte Vetorial
- LEVY SOUTO SOUSA (VOLUNTÁRIO)
- JOSÉ FONTEBASSO NETO (ORIENTADOR)

# INTRODUÇÃO

- Neste Trabalho, será exposto o Algoritmo de Aprendizado de Máquina “Support Vector Machines” (SVM), seus fundamentos teóricos, aplicações, base de dados utilizada e as ferramentas tecnológicas empregadas.

# OBJETIVOS

- Esta pesquisa tem como objetivo expor as funcionalidades de Máquinas de Suporte Vetorial para problemas de classificação.
- Utilizaremos o repositório MRI , o qual contém diversas neuroimagens, algumas delas com tumores de diferentes tipos
- Ao final, o algoritmo de classificação deverá ser capaz de prever quais das imagens tem tumores e qual seu tipo.

# METODOLOGIA

- Os arquivos com imagens de cérebros no repositório MRI contém 484 para treino, cada uma com suas *labels*, e mais 250 para teste.
- Iremos utilizar a documentação do Scikit-Learn e a Linguagem de Programação Python com suas bibliotecas: nibabel, nilearn, sklearn, numpy, matplotlib e pickle.
- Analisaremos uma imagem individualmente para compreensão dos dados, faremos o pré-processamento, visualização gráfica e ao fim a classificação do dataset com SVM.

# RESULTADOS

- Como dito anteriormente, primeiro carregamos os arquivos .nii.gz do Drive com nibabel, reduzimos sua dimensionalidade e transformamos em um vetor numpy.
- A estrutura das neuroimagens é formada por *voxels* (grades tridimensionais, equivalente a pixels), e na figura abaixo podemos inferir a quantidade de *voxels* em cada uma das 3 dimensões. Esta informação é relevante levando em conta o processamento de diversos arquivos.

# RESULTADOS

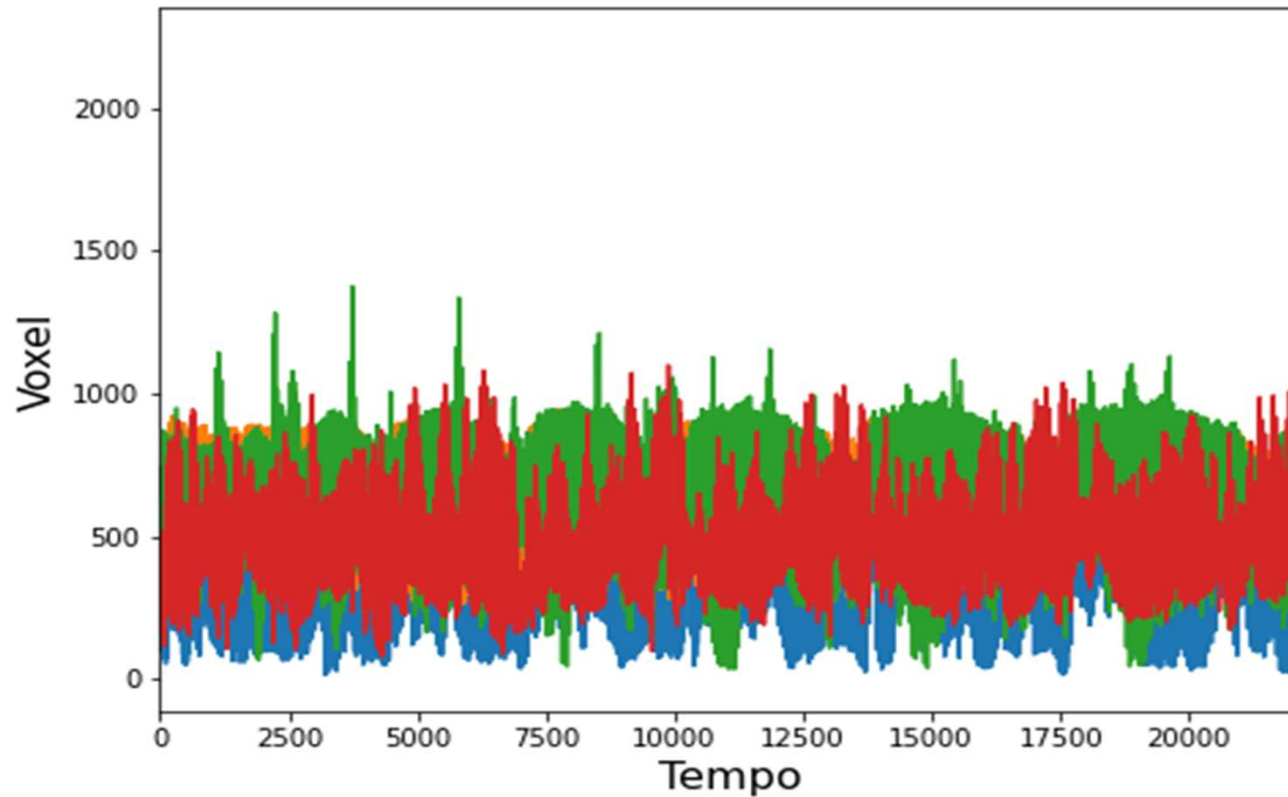


Figura 1.

Fonte: Autor



UNIVERSIDADE  
CATÓLICA  
DE SANTOS



XVII JORNADA DE INICIAÇÃO CIENTÍFICA  
E INICIAÇÃO EM DESENVOLVIMENTO TECNOLÓGICO E  
INOVACÃO

# RESULTADOS

- Com o vetor numpy 3D, podemos agora criar uma visualização gráfica com matplotlib.
- Também podemos visualizar a cada dez camadas separadamente
- A área com maior saturação é a tumoral.



# RESULTADOS

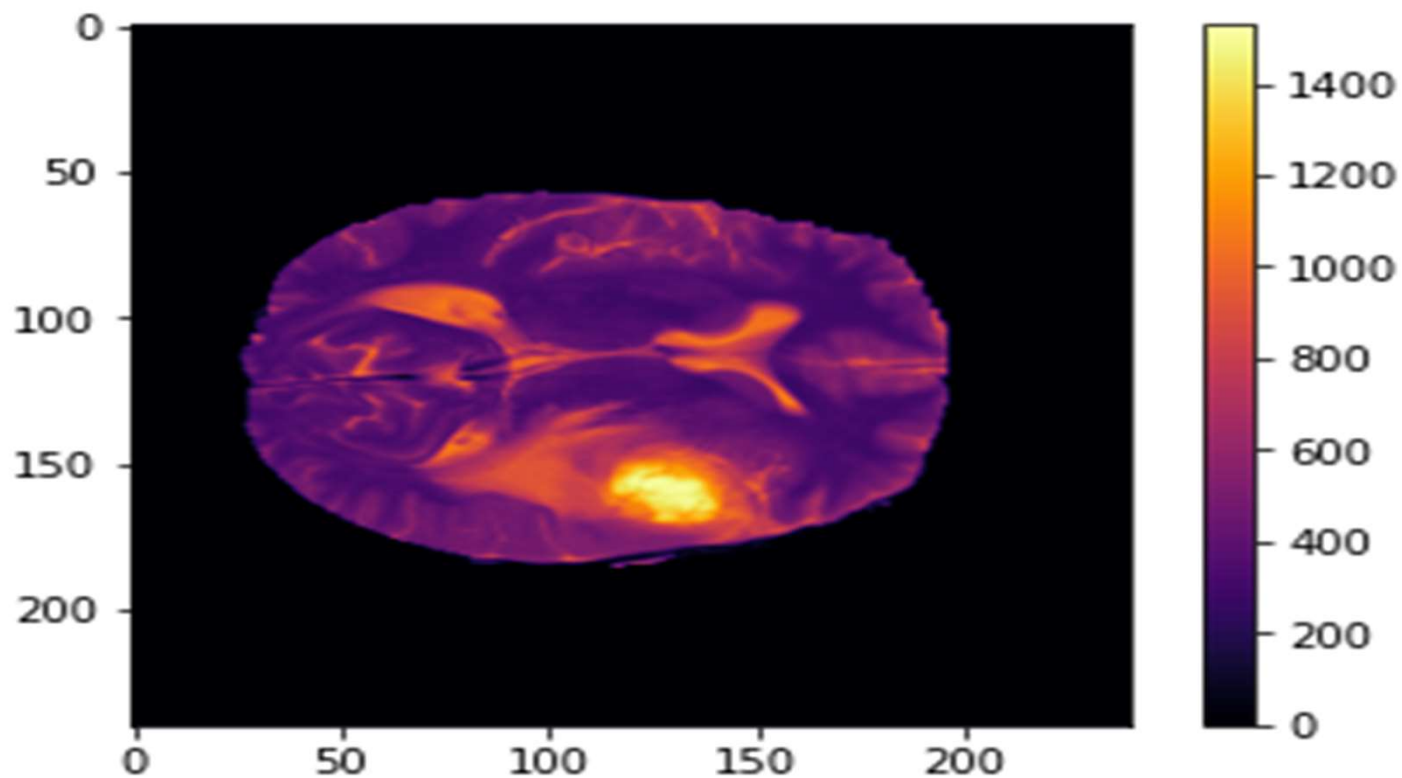


Figura 2.  
Fonte: Autor



UNIVERSIDADE  
CATÓLICA  
DE SANTOS



XVII JORNADA DE INICIAÇÃO CIENTÍFICA  
E INICIAÇÃO EM DESENVOLVIMENTO TECNOLÓGICO E  
INOVACÃO

# RESULTADOS

visualização de camadas no eixo Z

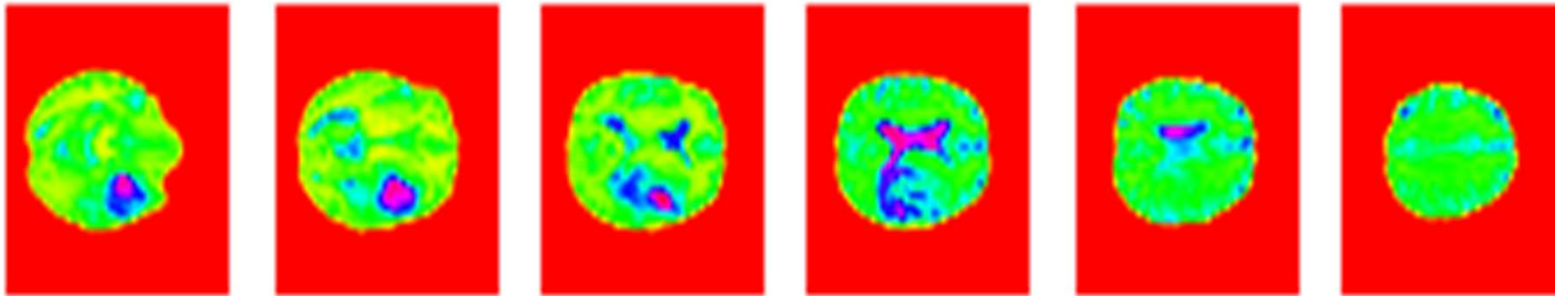


Figura 3.  
Fonte: Autor



UNIVERSIDADE  
CATÓLICA  
DE SANTOS



XVII JORNADA DE INICIAÇÃO CIENTÍFICA  
E INICIAÇÃO EM DESENVOLVIMENTO TECNOLÓGICO E  
INOVACÃO

# RESULTADOS

- Agora, aplicamos a ferramenta de hiperplano do SVM a partir da função plotting do NiLearn.

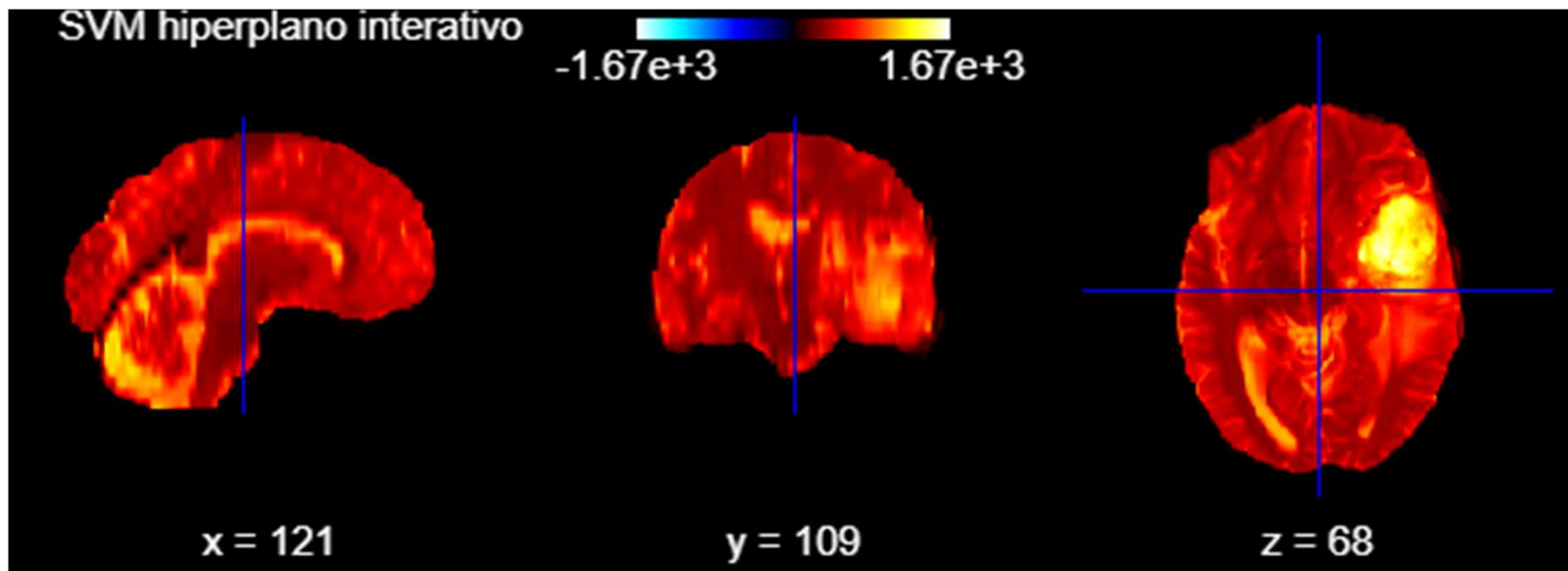


Figura 4

Fonte: Autor

# RESULTADOS

- Agora, para trabalhar com o dataset, carregamos 5 imagens para treino, 5 para labels e 2 para teste, então salvamos as informações no Drive com funções da biblioteca Pickle.
- Transformamos as imagens em vetores numpy e aplicamos a máscara para retirar o fundo, deixando apenas o cérebro.
- Importamos as bibliotecas do Scikit-Learn e então realizamos o treinamento.

# RESULTADOS

- Depois de diversas horas processando, podemos calcular a taxa de acerto do algoritmo com as 2 imagens de teste por meio da função `clf.score()`.
- A taxa de acerto foi de 98%, por mais alta que seja, o resultado não é satisfatório devido ao baixo número de imagens de treino.
- O SVM se utiliza de funções quadráticas para a separação de classes, como os arquivos contém muitas informações, o processamento fica limitado à capacidade do *hardware*.

# CONSIDERAÇÕES FINAIS

- Diante do que foi exposto, podemos concluir que a Máquina de Suporte Vetorial é ideal para problemas de Classificação Binária na separação de diferentes classes.
- Porém, neste exemplo, não foi possível treinar o algoritmo com muitos arquivos devido à grande quantidade de informações em cada um. Para classificar, o SVM se utiliza de equações quadráticas, logo, leva-se muito tempo para processar.
- Ainda assim, podemos obter uma visão positiva da capacidade deste algoritmo, principalmente em ambientes otimizados e com maior potencial de processamento.

# REFERÊNCIAS

- OSUNA E., Freund R., and Girosi F., “**Support Vector Machines: Training and Applications**”, A.I. Memo No.1602, Artificial Intelligence Laboratory, MIT, 1997.
- MORETTIN, Pedro A.; Singer, Julio M. **Introdução à ciência de dados, fundamentos e aplicações**. São Paulo-SP: Departamento de Estatística da Universidade de São Paulo, IME-USP, 2020.
- EVGENIU, Theodorus; PONTIL, Massimiliano. **Support Vector Machines: Theory and Applications** DOI: 10.1007/3-540-44673-7 12. Disponível em: [https://www.researchgate.net/publication/221621494\\_Support\\_Vector\\_Machines\\_Theory\\_and\\_Applications](https://www.researchgate.net/publication/221621494_Support_Vector_Machines_Theory_and_Applications), janeiro de 2001.
- DOBILAS, Saul. **Classificador SVM e kernel RBF**. Disponível em: <https://ichi.pro/pt/classificador-svm-e-kernel-rbf-127248>.

# REFERÊNCIAS

- Michael Joseph, Jerrold Jeyejchandra, and Ein Dickie(eds): **Data Carpentry: Introduction to MRI Data Analysis**. Version 2019.11, November 2019, Disponível em: [github.com/carpentries-incubator/SDC-BIDS-IntroMRI](https://github.com/carpentries-incubator/SDC-BIDS-IntroMRI).
- **SVM RBF parameters**, Scikit Learn Documentation, Disponível em: [scikit-learn.org/stable/auto\\_examples/svm/plot\\_rbf\\_parameters.html](https://scikit-learn.org/stable/auto_examples/svm/plot_rbf_parameters.html). Acesso em Setembro de 2021.
- BAILEY, Stephen. “**Exploring 3D images with matplotlib**”. Dezembro de 2017. Disponível em: [www.youtube.com/watchv=5jQVQE6yfio](https://www.youtube.com/watchv=5jQVQE6yfio).
- PAL, Arghya; RATHI, Yogesh. **A review of deep learning methods for MRI reconstruction**. PDF disponível em: [arxiv.org/abs/2109.08618](https://arxiv.org/abs/2109.08618). Cornwell University. Acesso em setembro de 2021.