

Data Science Workflow Canvas*

Start here. The sections below are ordered intentionally to make you state your goals first, followed by steps to achieve those goals. You're allowed to switch orders of these steps!

| Title: | | |
|---|--|---|
| <div>1 Problem Statement What problem are you trying to solve? What larger issues do the problem address?</div> <div>Melanoma é o câncer de pele mais mortal. Identifica-lo rápida e previamente pode significar a sobrevivência de milhares todos os anos. Entretanto o diagnóstico leva tempo e demanda diversos exames, em alguns casos biopsia é necessário.</div> | <div>2 Outcomes/Predictions What prediction(s) are you trying to make? Identify applicable predictor (X) and/or target (y) variables.</div> <div>Este projeto visa mais facilmente identificar e classificar lesões normais de pele ou lesões melanoma, que são lesões de pele mais severas. Gerando um alerta e recomendando uma visita ao médico.</div> | <div>3 Data Acquisition Where are you sourcing your data from? Is there enough data? Can you work with it?</div> <div>A base de dados envolveu imagens da HAM1000, usando apenas 7818 imagens divididas em lesões normais e melanoma. Além desta, também foi usada a base MED-NODE contendo 170 imagens divididas também em Nevus e melanoma.</div> |
| <div>4 Modeling What models are appropriate to use given your outcomes?</div> <div>Foram usados 3 modelos de redes neurais CNN pré-treinadas, dentre elas: VGG16, Inceptionv3 e ResNet50.</div> | <div>5 Model Evaluation How can you evaluate your model's performance?</div> <div>Usando acurácia, perda e curva ROC (Curva Característica de Operação do Receptor), afim de avaliar o melhor modelo.</div> | <div>6 Data Preparation What do you need to do to your data in order to run your model and achieve your outcomes?</div> <div>Foi necessário unir as duas bases de dados, tratando: exclusões e transformações de maneira que ambas as bases ficassem iguais. Ao usar a base de imagens em treinamentos fez-se necessário passar por um processo de aumento de dados.</div> |

Activation

When you finish filling out the canvas above, now you can begin implementing your data science workflow in roughly this order.

1 Problem Statement → 2 Data Acquisition → 3 Data Prep → 4 Modeling → 5 Outcomes/Preds → 6 Model Eval

* **Note:** This canvas is intended to be used as a starting point for your data science projects. Data science workflows are typically nonlinear.