

# Maior Dúvida da Aula

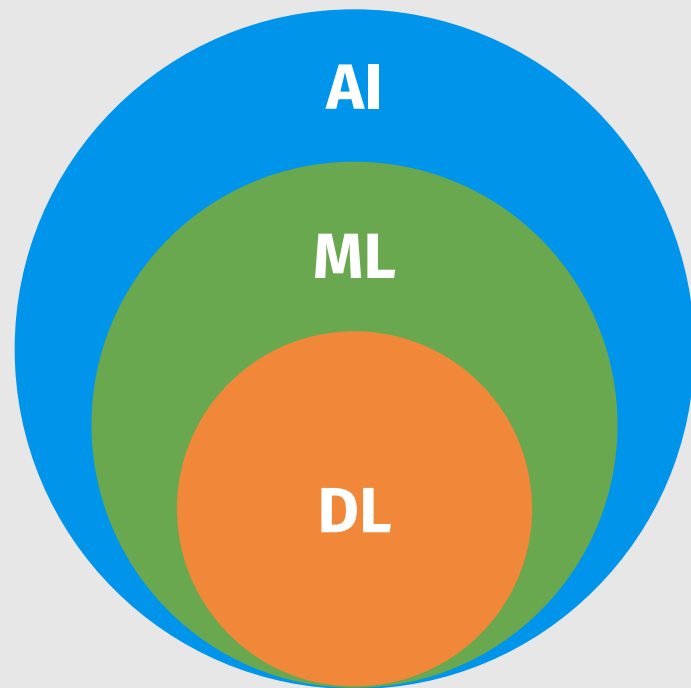
Machine Learning: A High Level Overview

# Supervisionado vs. Não-supervisionado

1. Quando eu tenho um problema real, como eu escolho qual tipo de aprendizado e algoritmo utilizar? A partir de quais características do meu problema que eu analiso e decido o que utilizar e quais testar?
2. Por que seria melhor resolver um problema com um modelo supervisionado ao invés de um modelo não supervisionado?

# AI, ML & DL

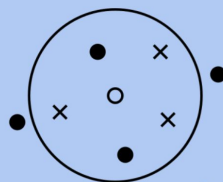
3. Não ficou claro para mim se Inteligência Artificial contém Machine Learning que contém Deep Learning ou não.



# Instance based vs. Model based learning

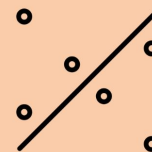
4. Minha maior dúvida foi sobre a diferença entre Instance based vs. Model based learning.

## Instance-based Learning <sup>02</sup>



Instance-based learning algorithms use the entire dataset as the model. For example: k-Nearest Neighbors aka kNN algorithm looks at the close neighborhood of the input example in the space of feature vectors and outputs the label that it saw the most often in this close neighborhood.

## Model-based Learning



Model-based learning algorithms use the training data to create a model that has parameters learned from the training data. For example: In Support Vector Machines aka SVM, we have  $w^*$  (learned weights value) and  $b^*$  (learned bias value). After the model is built, the training data can be discarded.

robotified.com

# Perguntas Gerais

5. Se os datasets são bem comportados e fogem do que é encontrado na vida real, por que eles ainda são utilizados ao invés de focar em outros que refletem melhor as condições reais?
6. Dado o potencial impacto da inteligência artificial no século 21, é possível compará-la com a eletricidade, em relação ao poder de impacto na vida humana, como meio de mudança social?
7. Por que é preferível usar a linguagem Python? (quando se fala em inteligência artificial mesmo a linguagem deixando a desejar quando se fala de performance)

# Câncer de Pele

8. Fiquei curiosíssima sobre qual foi o motivo encontrado pelo modelo de melanoma ser preciso mesmo sem a informação da lesão!!

# CONHEÇA O **ABCDE**

## A

ASSIMETRIA



Um lado não é parecido com o outro

## B

BORDAS



Bordas irregulares

## C

COR



Variedade de cores, geralmente tons de preto e marrom

## D

DIÂMETRO



Tamanhos maiores que 6 mm

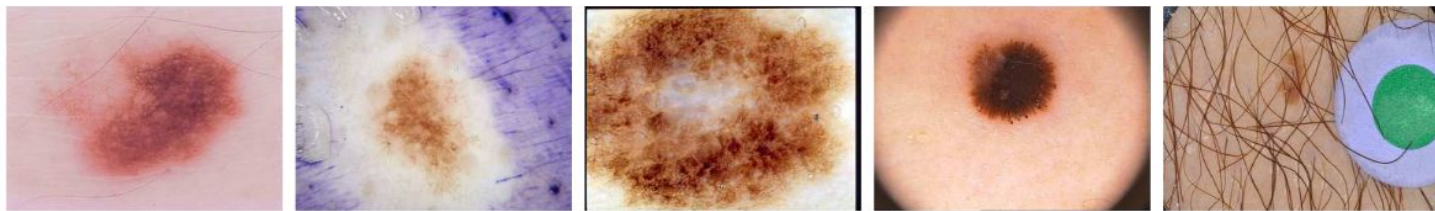
## E

EVOLUÇÃO



Mudanças em tamanho, forma, cor; sangramento ou novo sintoma

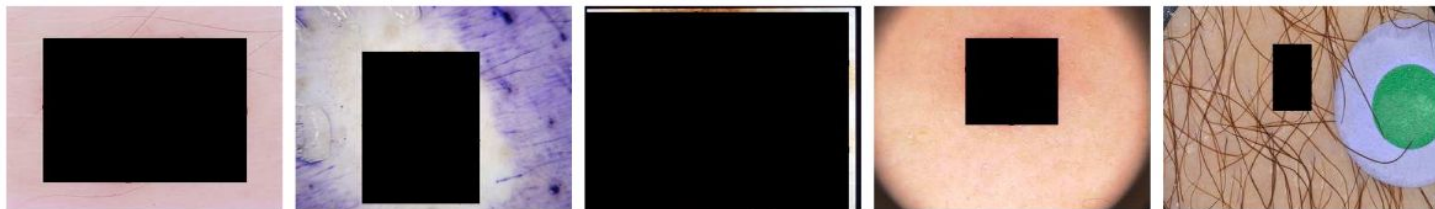
Traditional



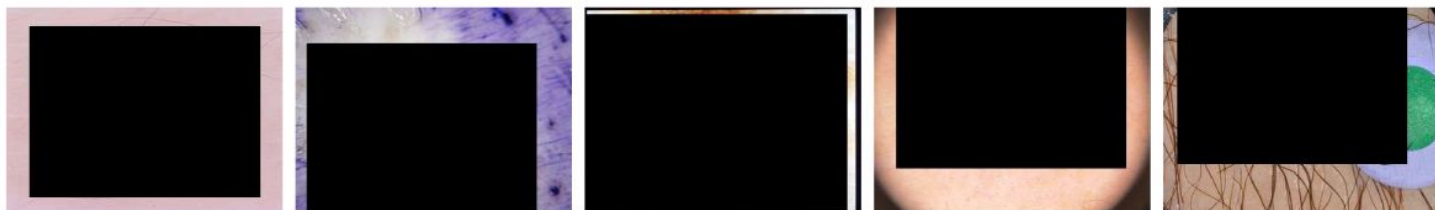
Only skin



Bbox

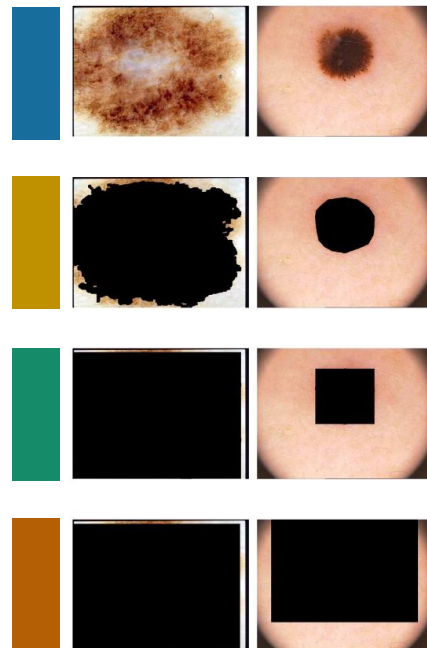
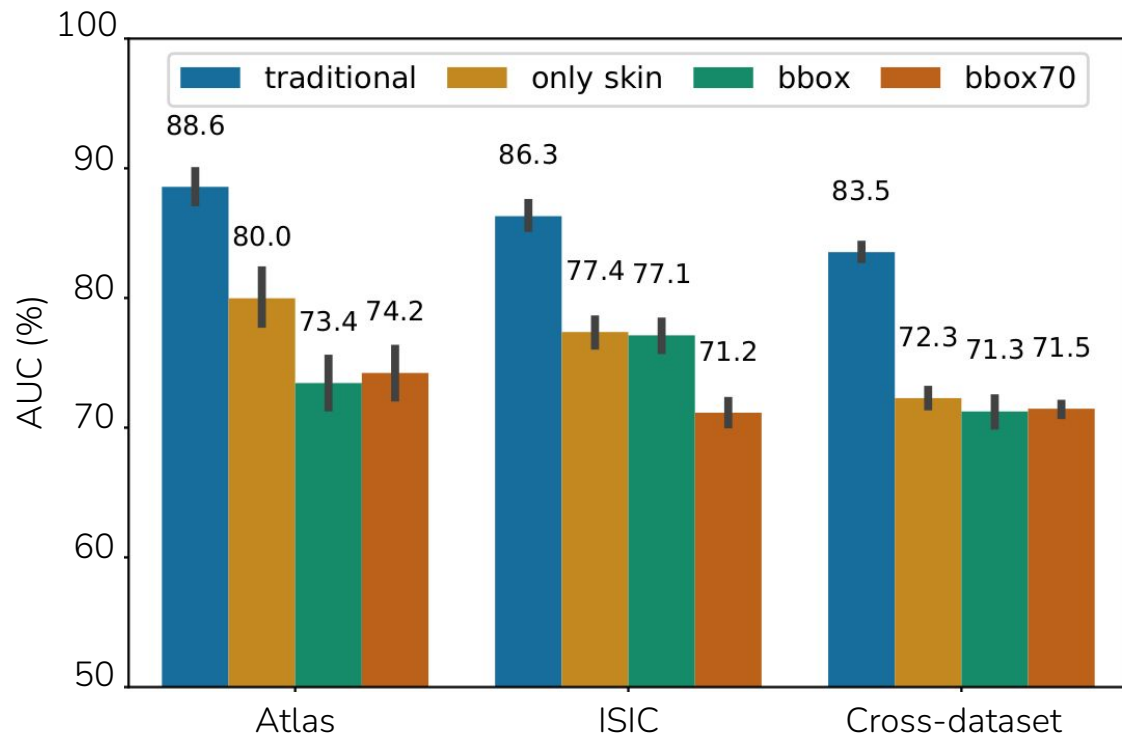


Bbox70





# 157 dermatologists on ISIC dataset: 67% AUC

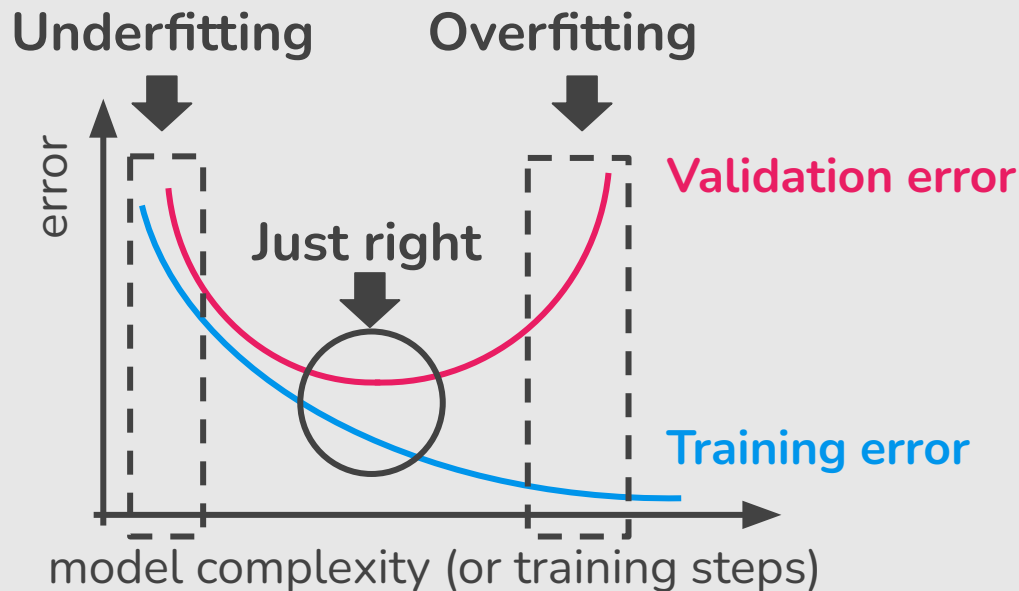


Bissoto et al., “(De)Constructing bias on skin lesion datasets”, CVPR, 2019. 🌟 Best paper award.

# Underfitting vs. Overfitting

9. Como diferenciar um modelo que contém overfitting de um modelo que tem bons resultados devido a uma boa modelagem? E como diferenciar um modelo com Underfitting de apenas falta de treinamento?
10. Quando um modelo obtém 100% de acurácia no conjunto de treino, podemos com certeza afirmar que houve overfitting? (Mesmo sem conhecer o desempenho nos casos de teste)

# Underfitting vs. Overfitting



# Treino, Validação & Teste

11. Não entendi muito bem a diferença do conjunto de validação e de teste. Se ambos são usados na fase de treino, o que difere os dois?
12. Se a escassez de dados já é um problema, quando eu faço divisão de dados de treino/validação, eu não aumento esse problema por estar diminuindo a quantidade de dados no treino?
13. Existe alguma forma estatística de determinar o número de amostras na validação e no teste?

# Maior Dúvida da Aula

## Artificial Unintelligence: ML & Titanic

# Machine Learning

<https://youtu.be/5q87K1WaoFI>



<https://youtu.be/7epCIU2bjY4>



# Dados Faltantes

- 14. Como preencher valores faltantes no conjunto de teste? É recomendável fazê-lo usando como base (de medianas) apenas dados do treino?
- 15. Como proceder a partir de dados faltantes? Especificamente, na dualidade entre preencher os dados faltantes ou eliminar linhas da nossa base de dados.
- 16. Existem várias técnicas para preencher os dados faltantes, mas até que ponto você não está enviesando o seu modelo com esses dados novos?

# Perguntas Gerais

- 17. Uma vez encontrado um modelo com uma boa precisão, esse modelo fica “solidificado” para ser utilizado como um programa ou toda vez que esse modelo for executado ele continua se aprimorando com as entradas recebidas?
- 18. Existem tipos de tarefas que são impossíveis de serem “aprendidas”!?



Oi professora, eu acredito que as ideias do significado de ML estão ficando bastante claras, os seus materiais de introdução foram muito bons.  
... Mas, consegui entender a mensagem sobre ética da aula, que os números precisam de interpretação, pois eles podem acabar confirmando desigualdade tão presentes em nossa realidade, então precisamos atuar sobre esses resultados e não simplesmente acatá-los como verdade absoluta.