



Human-Centered Data & AI

Vinicius Caridá, Ph.D.



Head of Digital Customer Service Platforms,
PCP, WFM, Data and AI - Itaú Unibanco

MBA Professor - FIAP

Google Developer Expert – Machine Learning

Co-organizer TFUGSP and AWSUGSP



@vinicius caridá



@vfcarida



@vinicius caridá



@vfcarida



@vinicius caridá



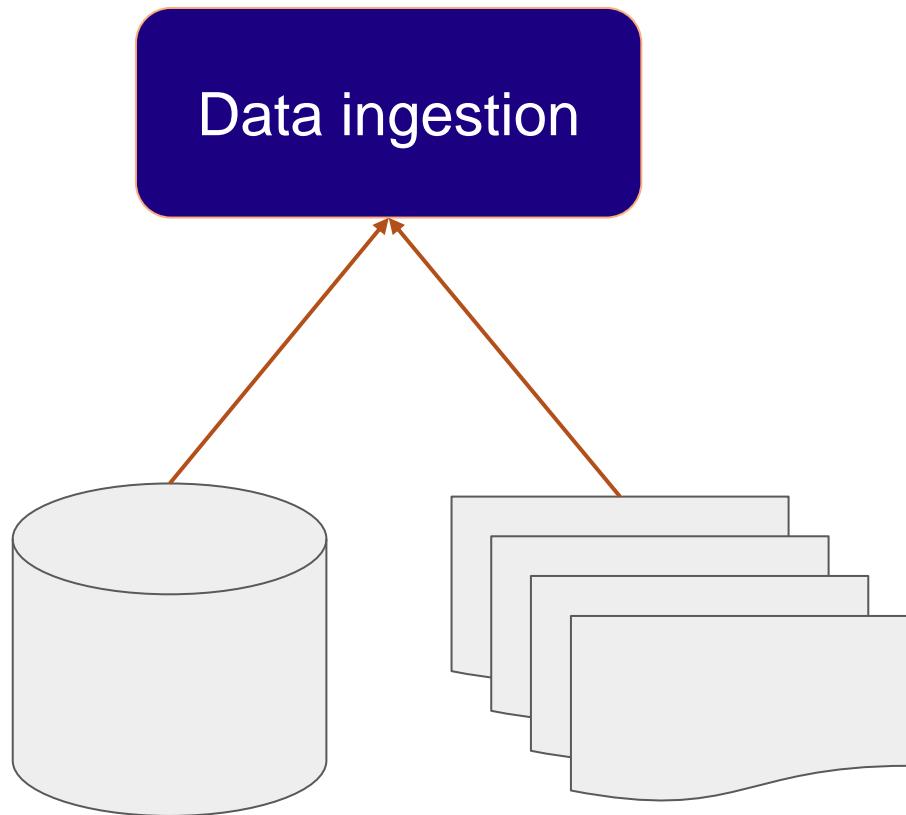
@vfcarida

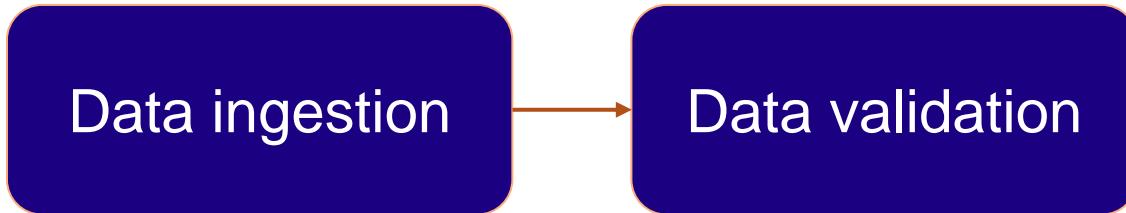
“

Introdução a Machine Learning





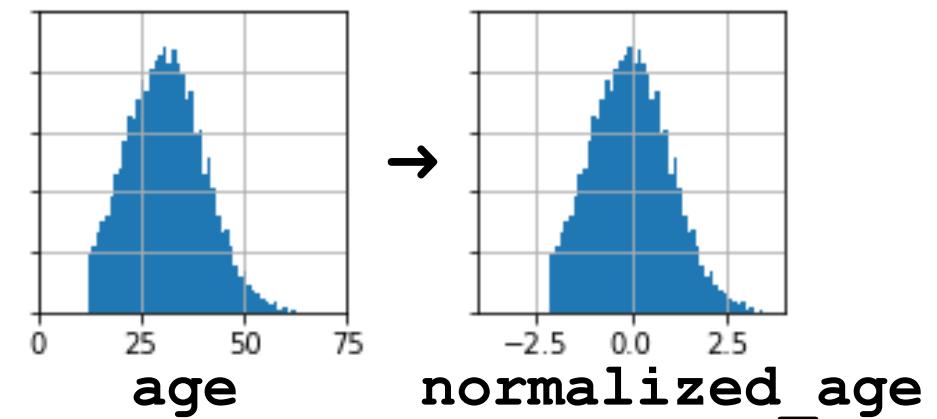




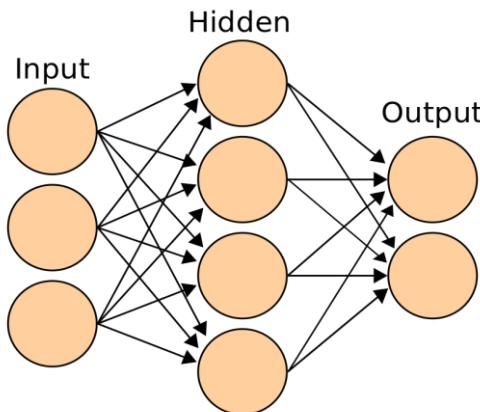
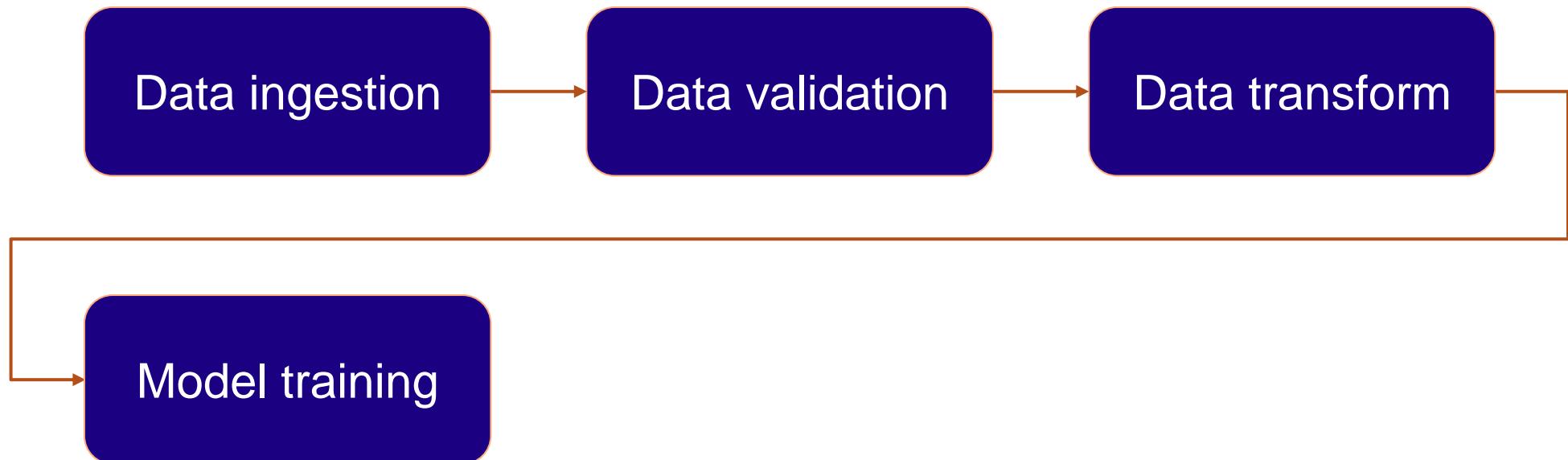
age is missing

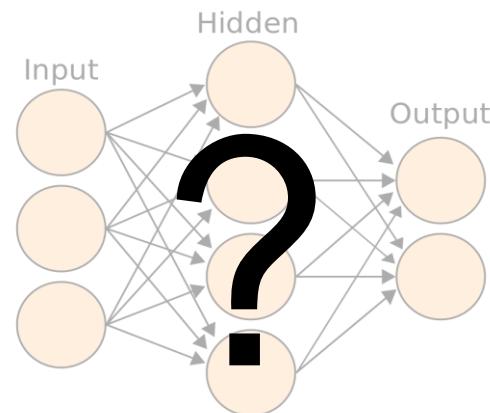
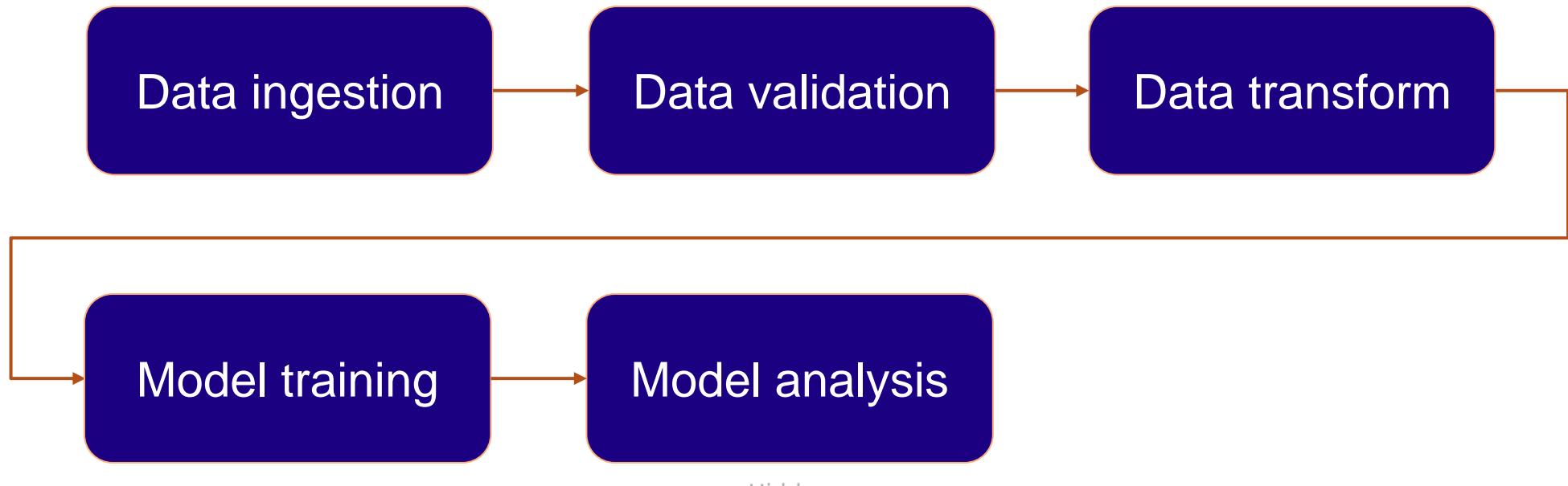


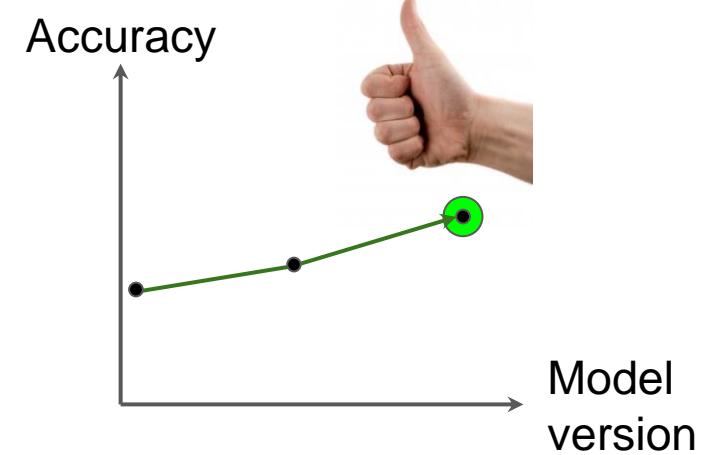
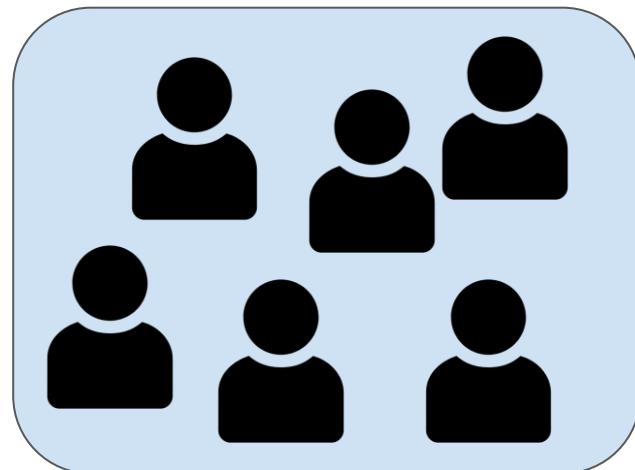
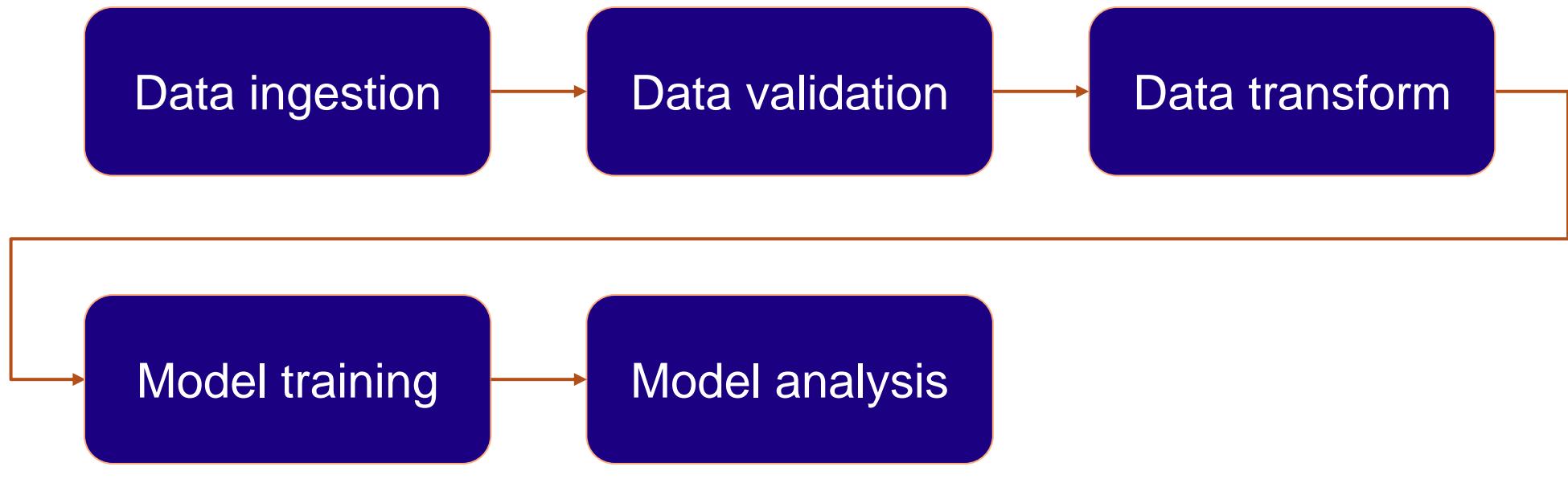
country not in:
• China
• India
• USA

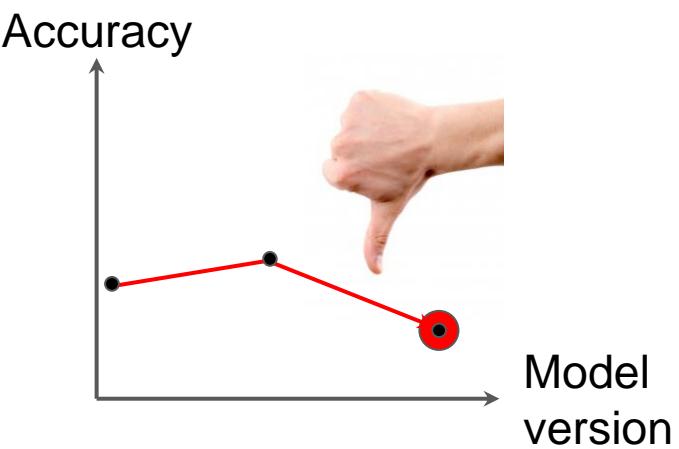
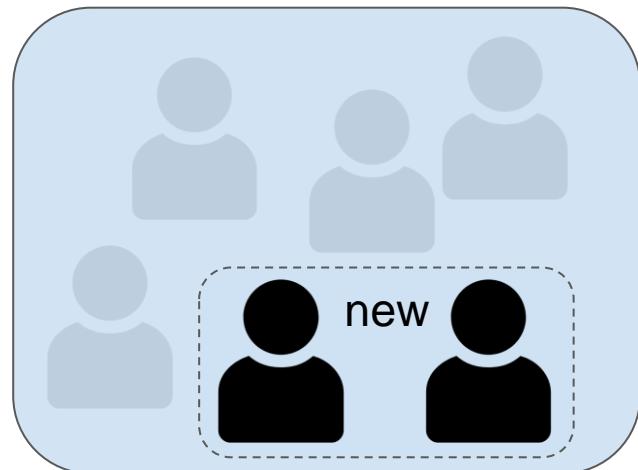
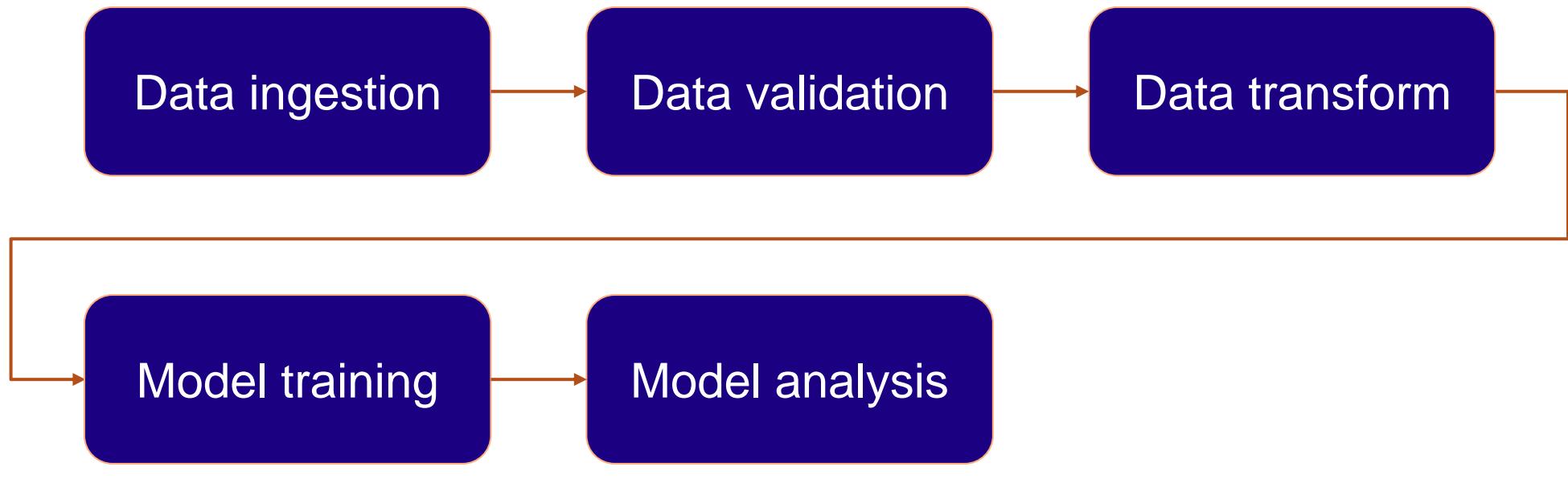


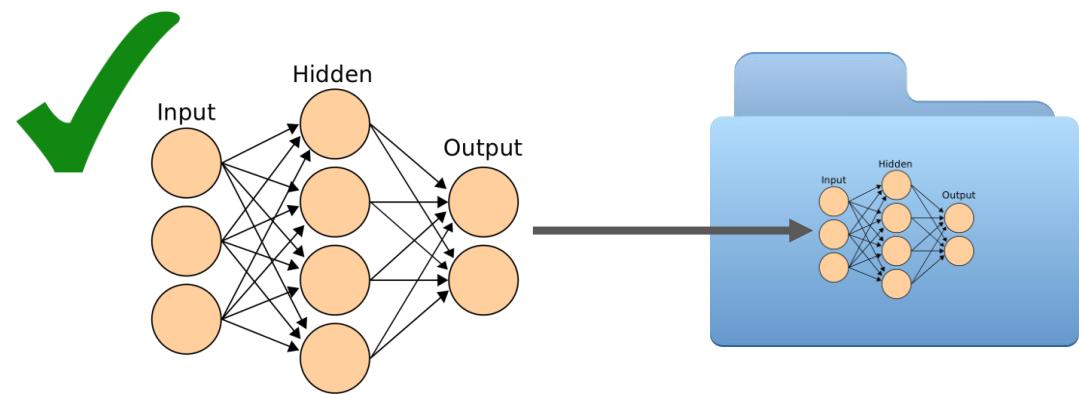
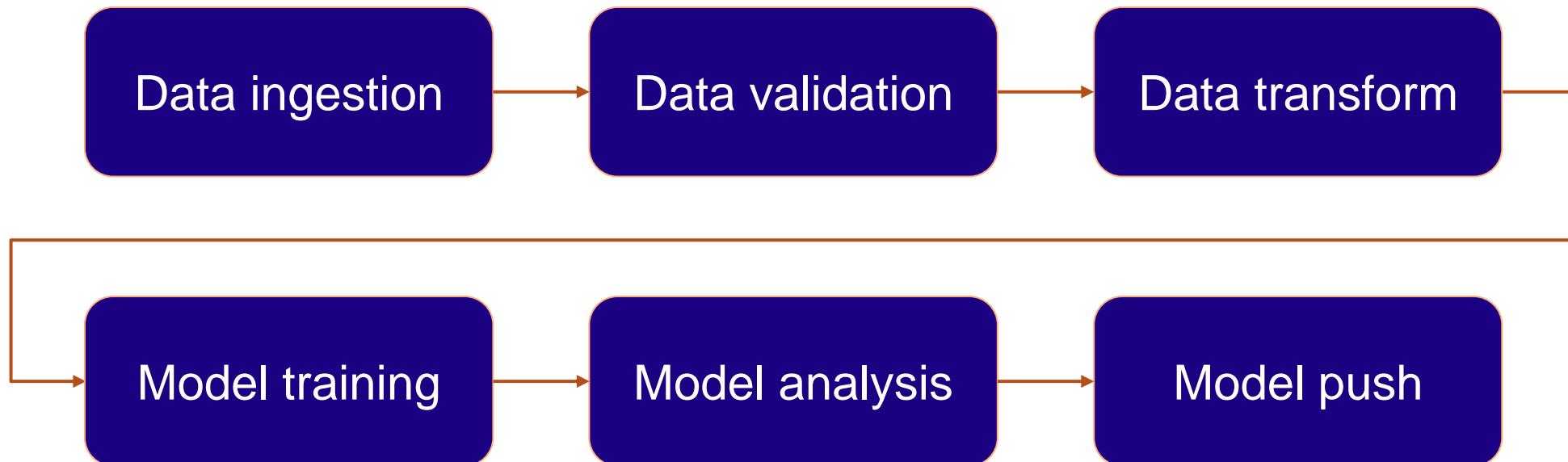
China	→	[1, 0, 0]
India	→	[0, 1, 0]
USA	→	[0, 0, 1]

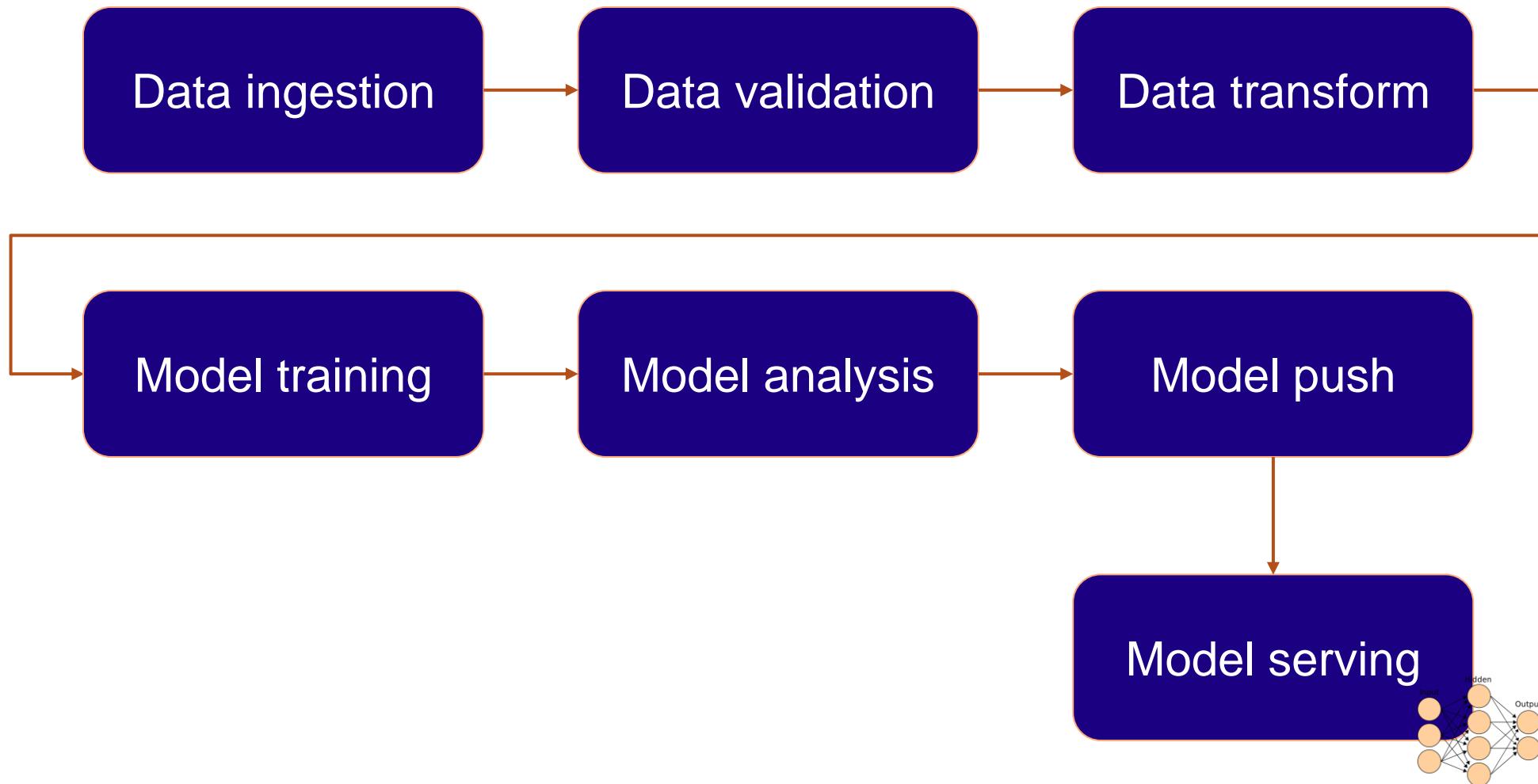


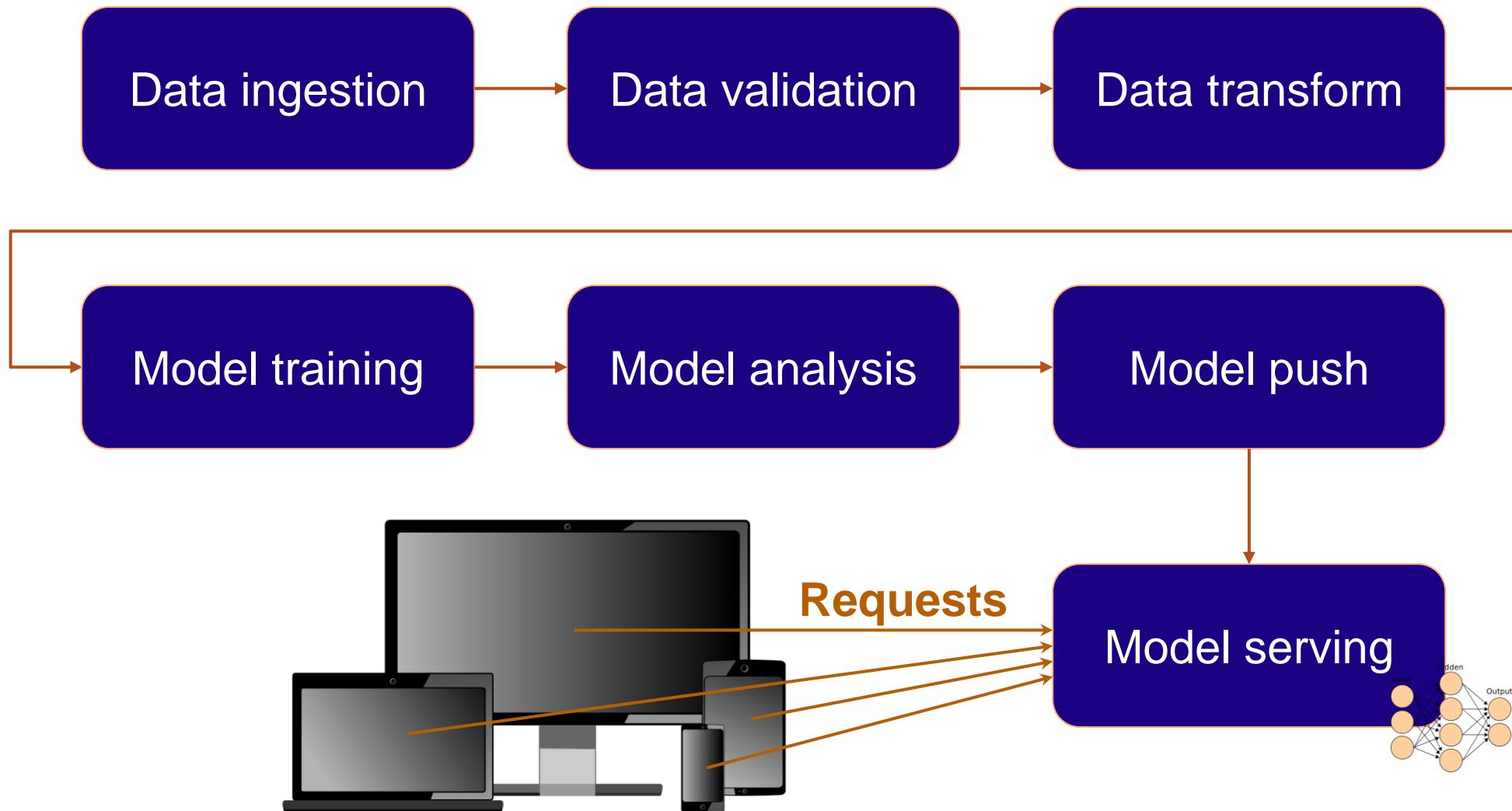


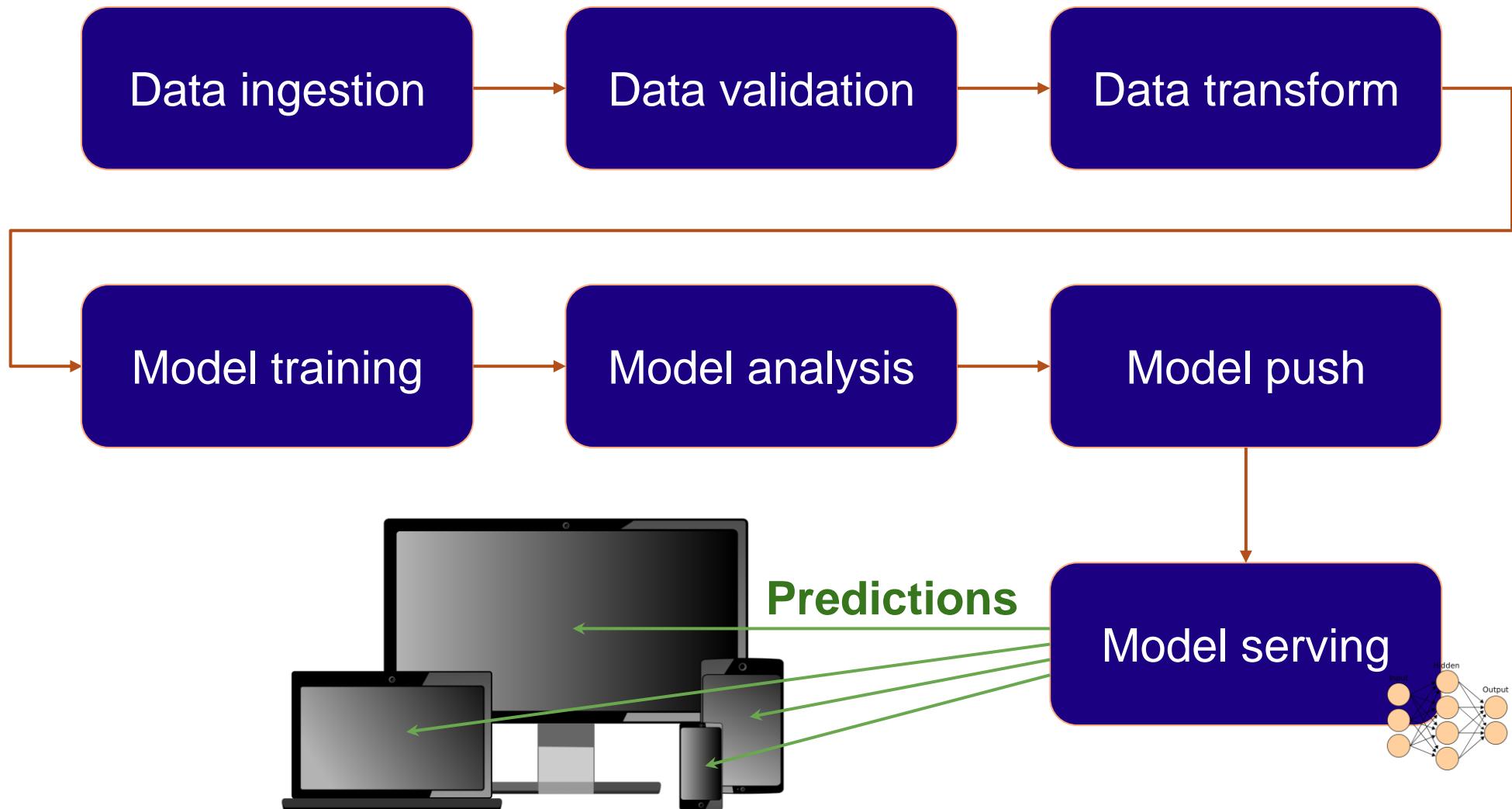


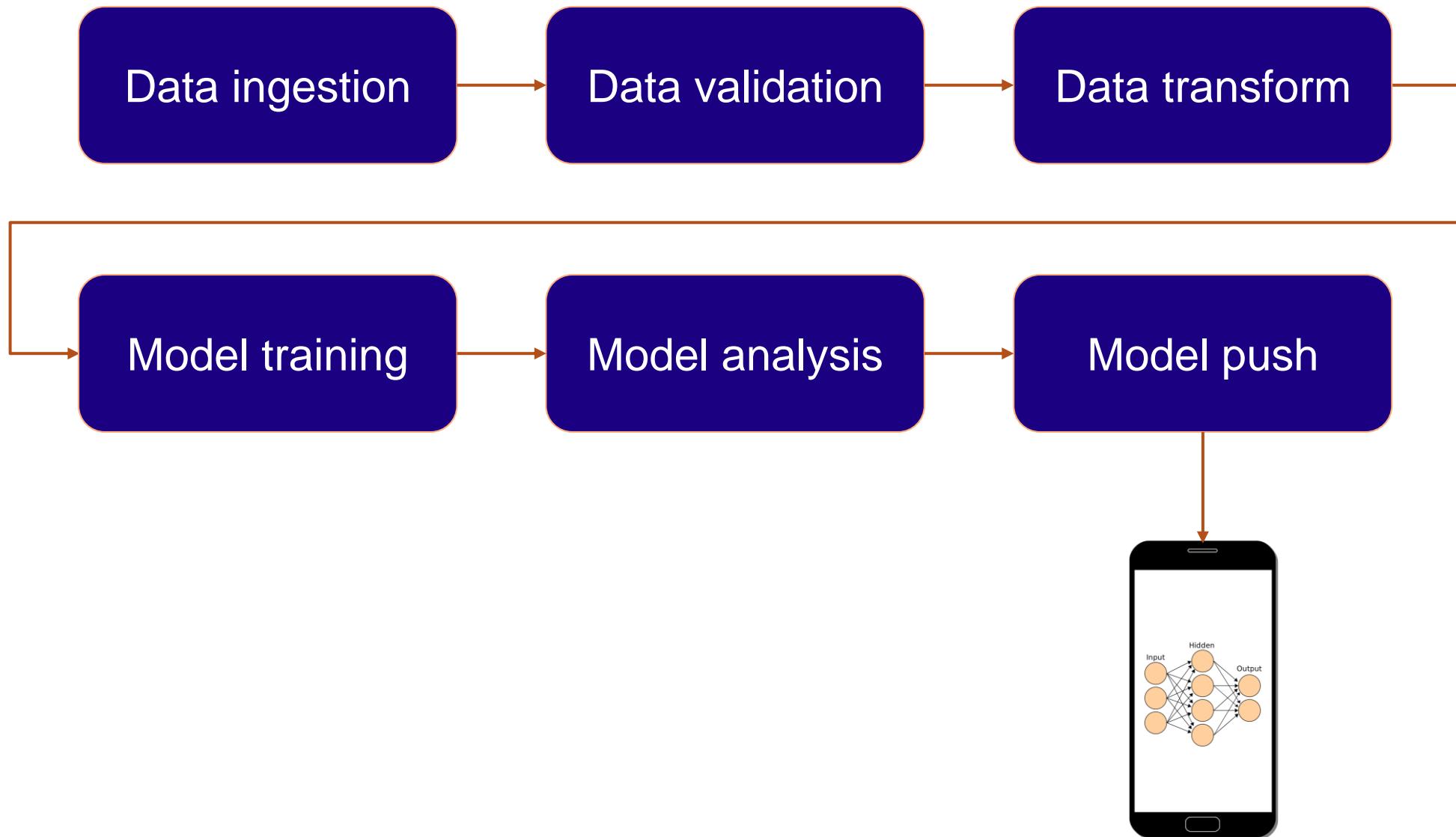


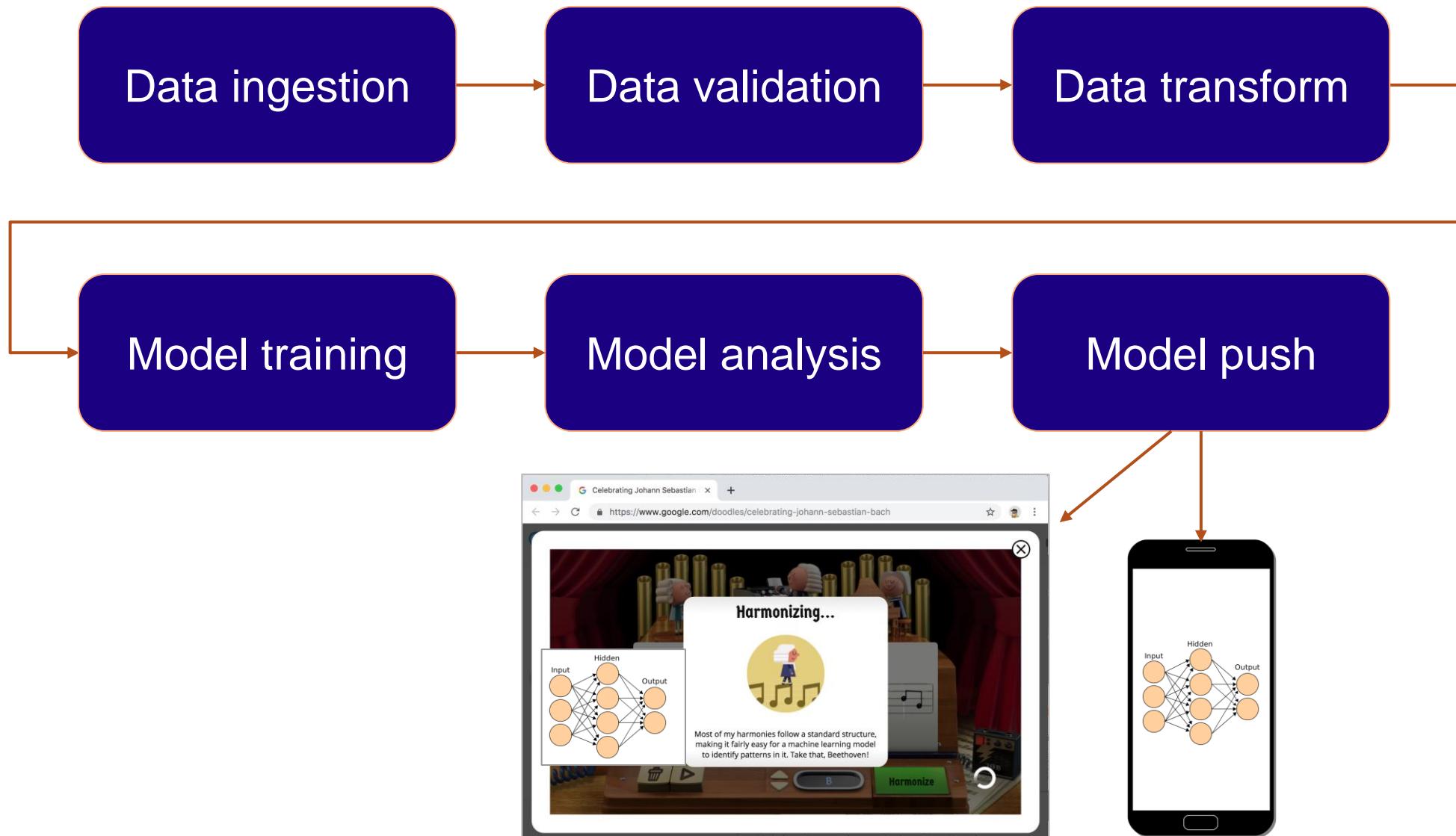


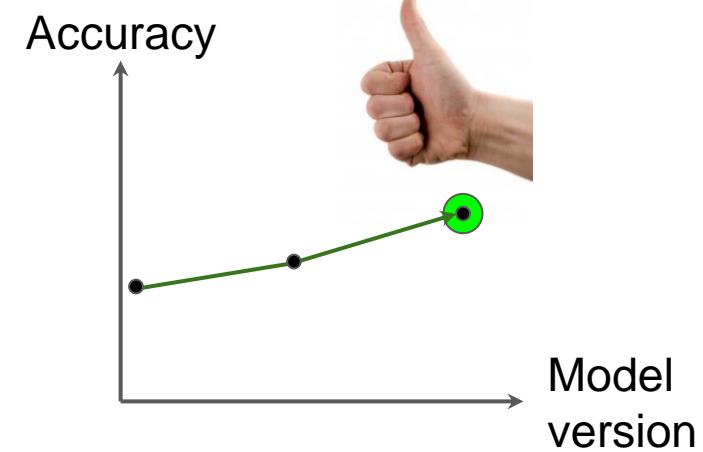
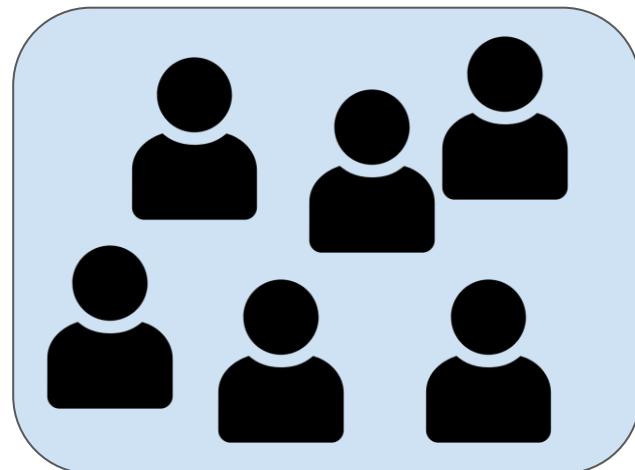
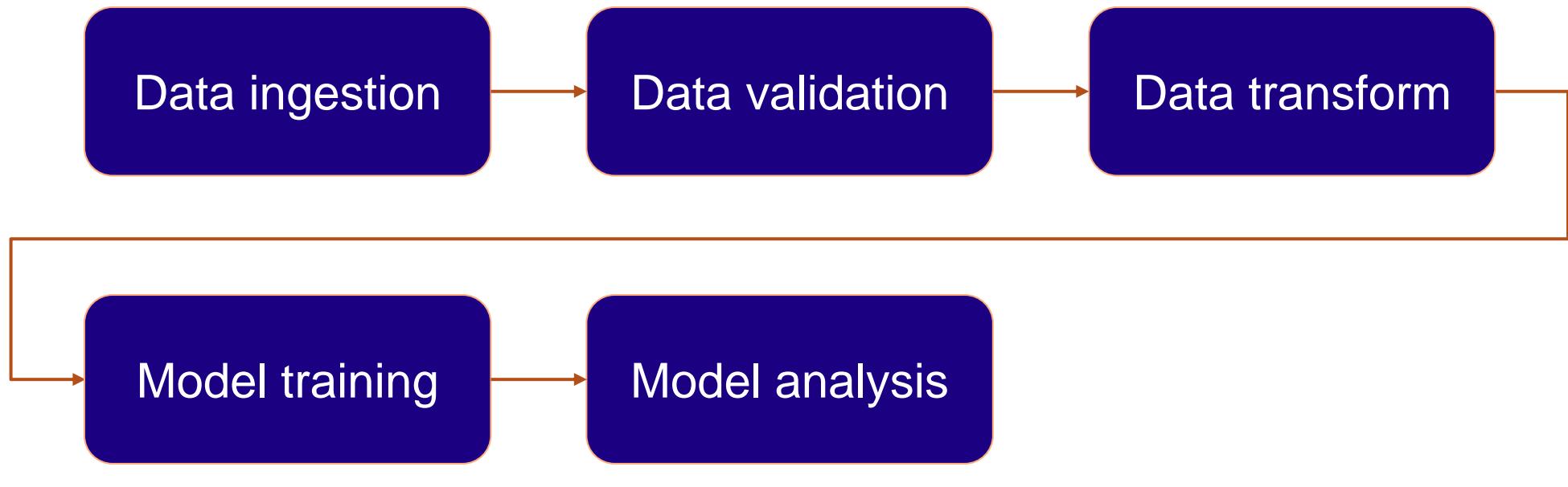










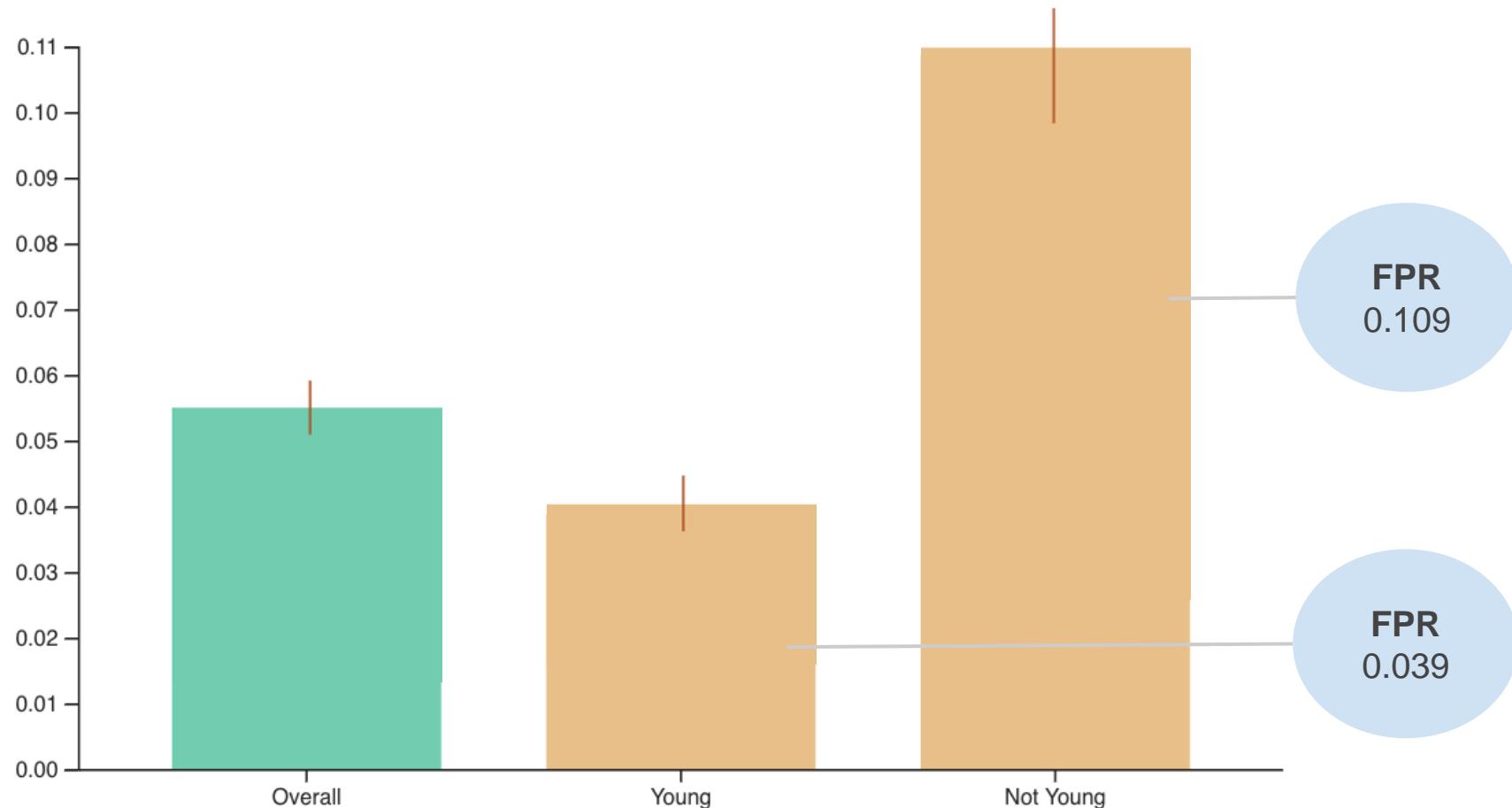




Remediation: Smile Detection on CelebA using TF Constrained Optimization



Results: Unconstrained `tf.keras.Sequential` model





Remediation: CelebA example using TFCO

1. Define subsets of interest

```
context = tfco.rate_context(predictions, labels=lambda:labels_tensor)
context_subset = context.subset(lambda:groups_tensor < 1)
```



Remediation: CelebA example using TFCO

1. Define subsets of interest

```
context = tfco.rate_context(predictions, labels=lambda:labels_tensor)
context_subset = context.subset(lambda:groups_tensor < 1)
```

2. Set constraints on subset using rate helpers

```
constraints = [tfco.false_positive_rate(context_subset) <= 0.05]
```



Remediation: CelebA example using TFCO*

1. Define subsets of interest

```
context = tfco.rate_context(predictions, labels=lambda:labels_tensor)
context_subset = context.subset(lambda:groups_tensor < 1)
```

2. Set constraints on subset using rate helpers

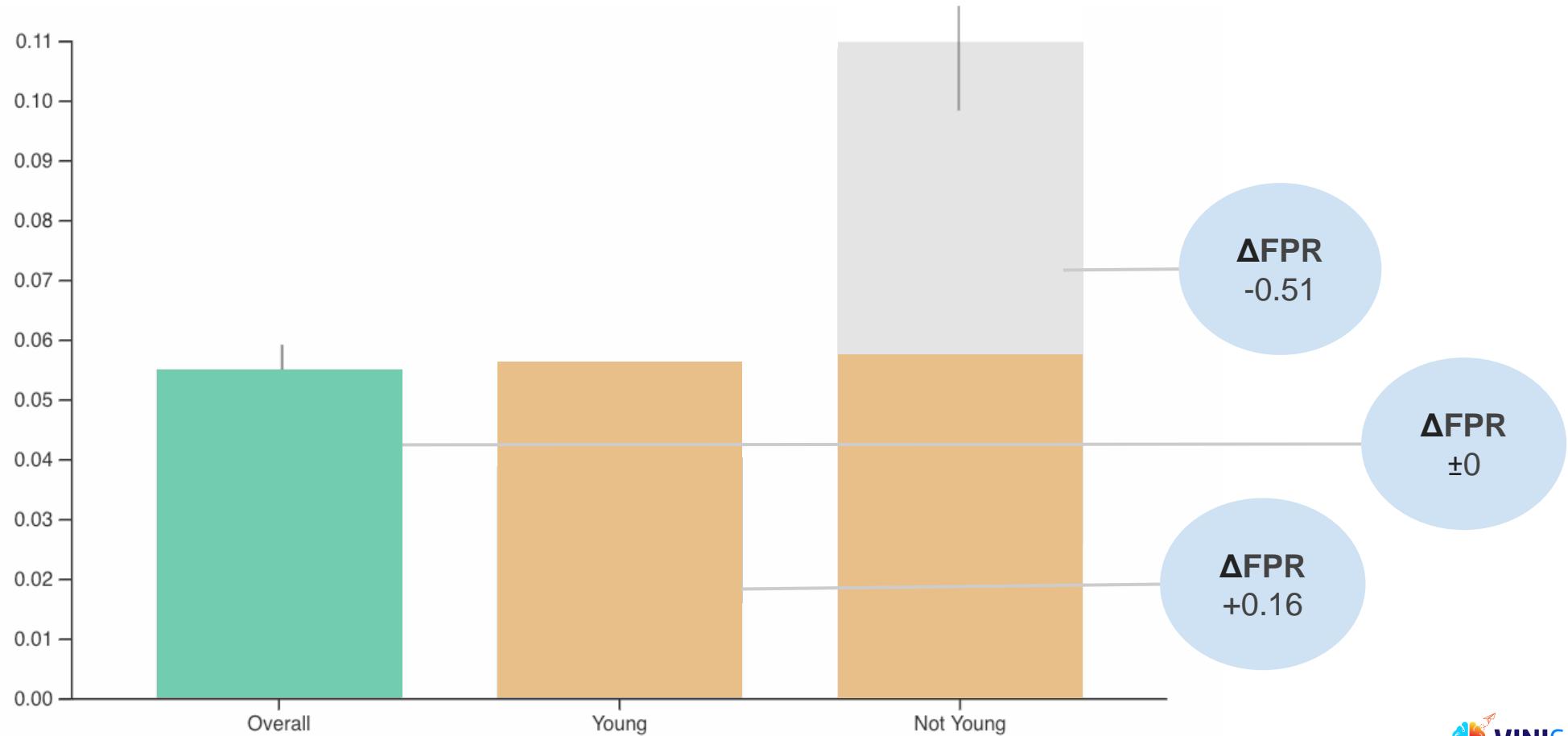
```
constraints = [tfco.false_positive_rate(context_subset) <= 0.05]
```

3. Define optimizer and train

```
problem = tfco.RateMinimizationProblem(tfco.error_rate(context), constraints)
optimizer = tfco.ProxyLagrangianOptimizerV2(
    optimizer=tf.keras.optimizers.Adam(learning_rate=0.001),
    constraint_optimizer=tf.keras.optimizers.Adam(learning_rate=0.001),
    num_constraints=problem.num_constraints)
```

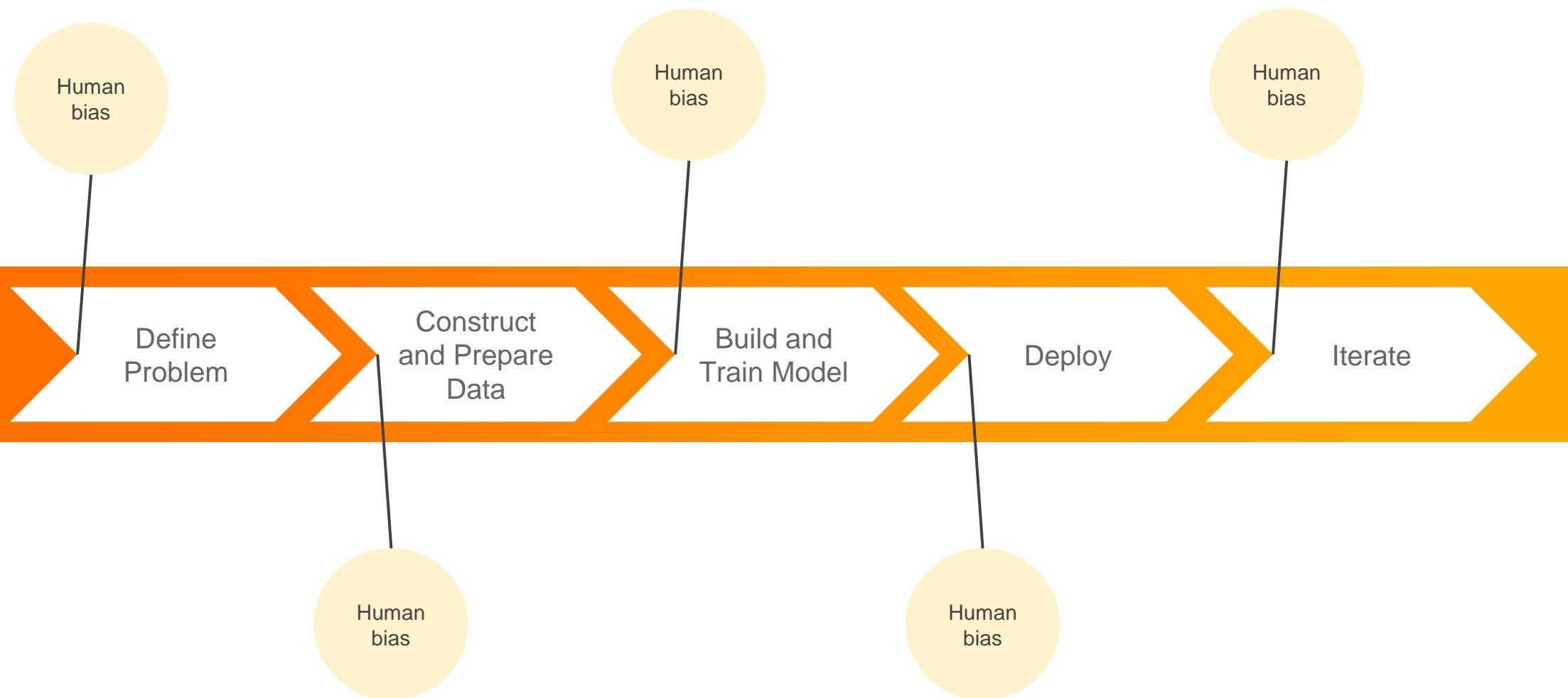
https://github.com/google-research/tensorflow_constrained_optimization

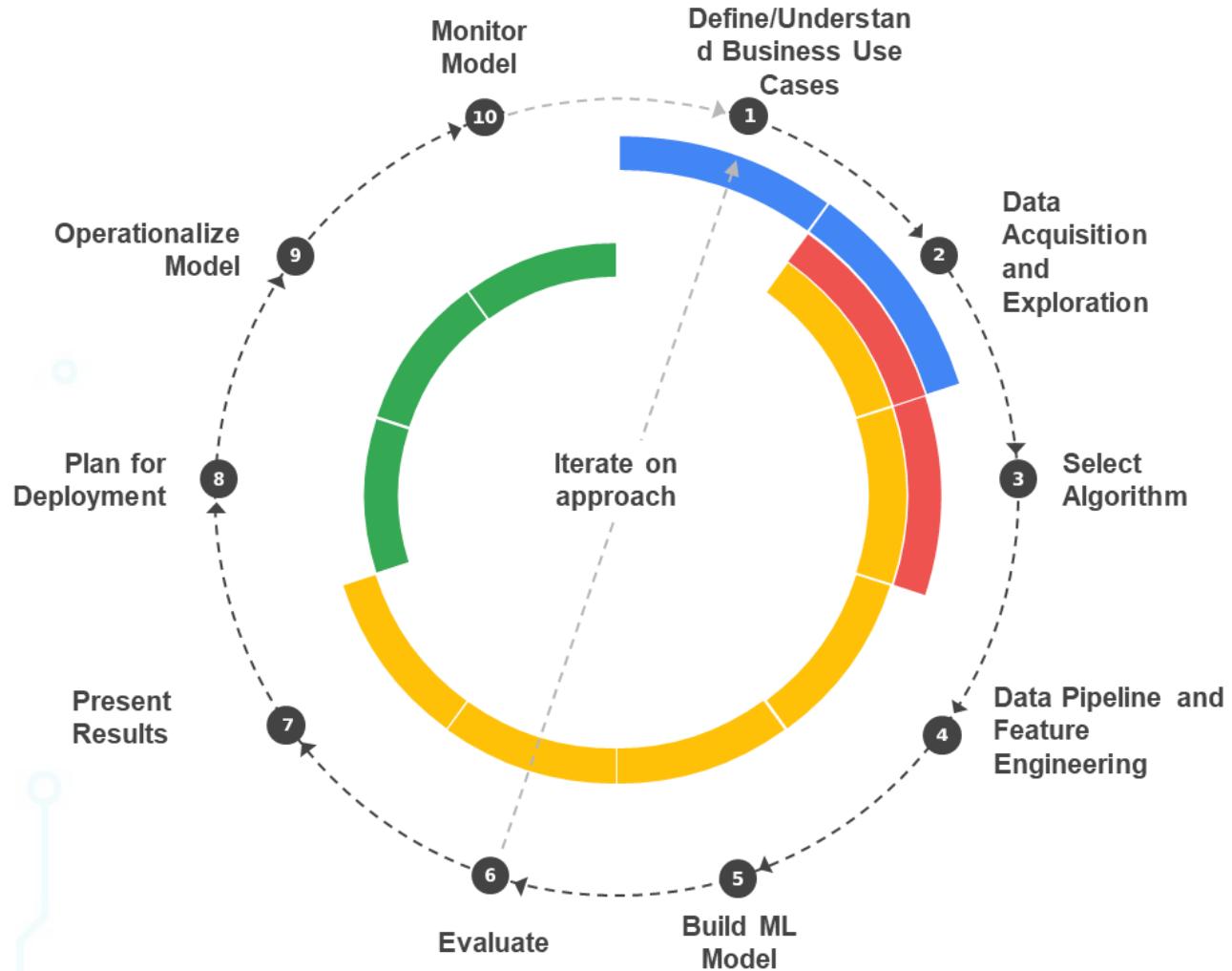
Results: Constrained `tf.keras.Sequential` model



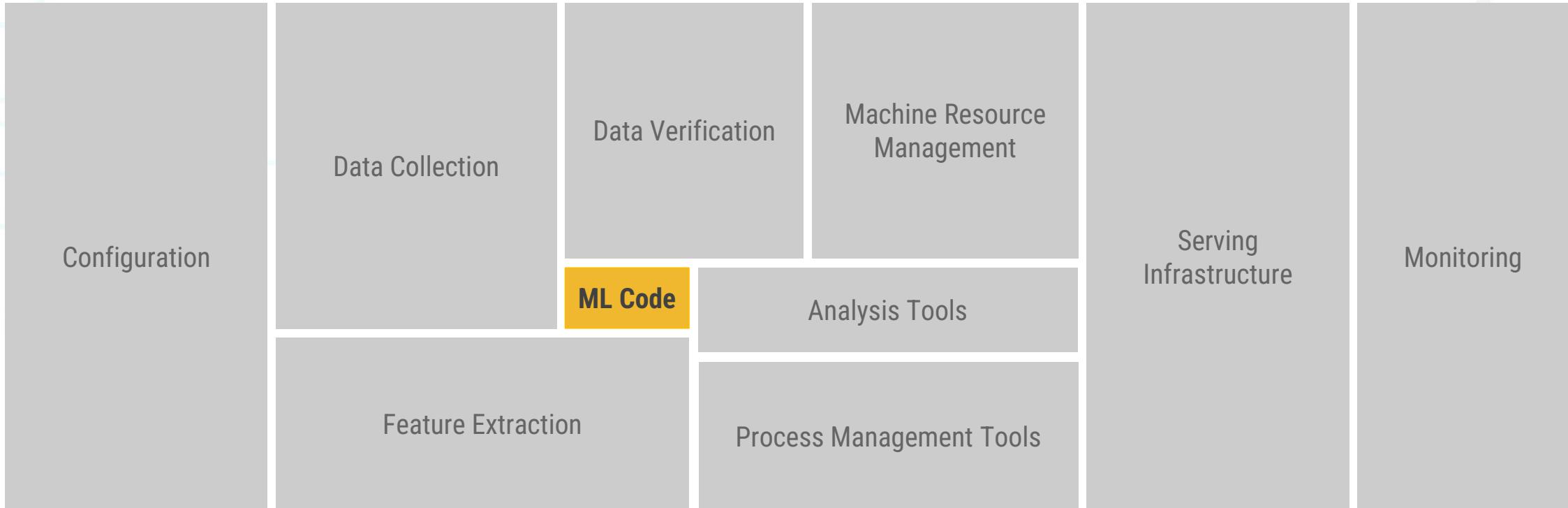


**Humans have a history of making
product design decisions that are *not*
in line with the needs of everyone.**





... a production solution requires so much more



Leading ML | Best Practices

Continuous Training for Production ML in the TFX Platform. OpML (2019).

Slice Finder: Automated Data Slicing for Model Validation. ICDE (2019).

Data Validation for Machine Learning. SysML (2019).

TFX: A TensorFlow-Based Production-Scale Machine Learning Platform. KDD (2017).

Data Management Challenges in Production Machine Learning. SIGMOD (2017).

Rules of Machine Learning: Best Practices for ML Engineering. Google AI Web (2017).

Machine Learning: The High Interest Credit Card of Technical Debt. NeurIPS (2015).

KDD 2017 Applied Data Science Paper

TFX: A TensorFlow-Based Production-Scale Machine Learning Platform

Denis Baylor, Eric Breck, Heng-Tze Chen, Salem Haykal, Mustafa Ispir, Clemens Mewald, Akshay Nareja, Steven Euijong Whang, Martin Zinkevich

D. Sculley¹, Todd Phillips², {dsculley, tdpdphilli}@google.com

ABSTRACT
Creating and maintaining a platform for reliably training and deploying machine learning models requires orchestration of many components – a learner for models based on training data, modules for analyzing both data as well as models, and finally infrastructure for serving models in production. This becomes challenging when data changes over time and needs to be produced continuously. Unfortunately, orchestration is often done ad hoc using glue code scripts developed by individual teams for specific needs leading to duplicated effort and fragile system technical debt.

We present TensorFlow Extended (TFX), a general-purpose machine learning platform at Google. By integrating the learner and infrastructure platforms we were able to standardize, simplify the platform configuration, and time to production from the order of months to providing platform stability that minimizes disruption.

Data Management Challenges in Production Machine Learning

Neoklis Polyzotis, Sudip Roy, {npolyzotis, srujan}@google.com

ABSTRACT
The tutorial discusses data-management issues that arise in machine learning pipelines at production. Informed by our own experience with scale pipelines, we focus on issues related to us validating, cleaning, and enriching training data. The goal of the tutorial is to bring forth these issues, directions to prior work in the database literature, and open research questions that are not addressed.

Continuous Training for Production Machine Learning

Denis Baylor, Eric Breck, Heng-Tze Chen, Salem Haykal, Mustafa Ispir, Clemens Mewald, Akshay Nareja, Steven Euijong Whang, Martin Zinkevich

Abstract
Large organizations rely increasingly on continuous pipelines in order to keep machine-learned models up-to-date with respect to data. In this scenario, disruptions in the pipeline can increase model staleness and thus degrade the quality of downstream services using these models. In this paper we describe the operational mechanisms in TensorFlow Extended (TFX) that we developed and deployed at Google. We present mechanisms in TFX to support this type of production and the lessons learned from the deployment of the platform internally at Google.

Automated Data Slicing for Model Validation

Yeounoh Chung, Tim Kraska, Neoklis Polyzotis, Sudip Roy, {younohc, timkraska, npolyzotis, srujan}@google.com

Abstract
As machine learning systems become more complex, it is increasingly difficult to validate them. This is especially true for problems involving validation across multiple dimensions. Data slicing allows users to analyze the model performance on arbitrary slices, our goal is to find interpretable slices that are both statistically significant and large. We propose Slice Finder, which can find such slices. Applications include diagnosing model failure points and identifying problematic regions in the data. Human intervention is crucial. This research is part of a larger effort to make machine learning systems more transparent and explainable.

Index Terms—data slicing, model validation, model analysis

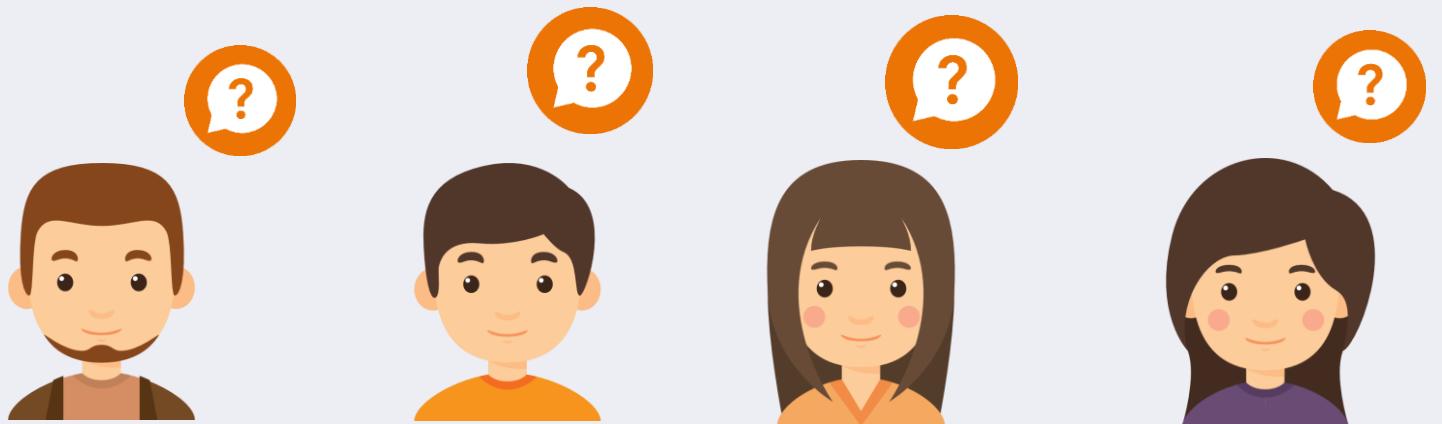
DATA VALIDATION FOR MACHINE LEARNING

Eric Breck¹, Neoklis Polyzotis¹, Sudip Roy¹, Steven Euijong Whang², Martin Zinkevich¹

ABSTRACT
Machine learning is a powerful tool for gleaning knowledge from massive amounts of data. While a great deal of machine learning research has focused on improving the accuracy and efficiency of training and inference algorithms, there is less attention in the equally important problem of monitoring the quality of data fed to machine learning. The importance of this problem is hard to dispute: errors in the input data can nullify any benefits on speed and accuracy for training and inference. This argument points to a data-centric approach to machine learning that treats training and serving data as an important production asset, on par with the algorithm and infrastructure used for learning.

In this paper, we tackle this problem and present a data validation system that is designed to detect anomalies specifically in data fed into machine learning pipelines. This system is deployed in production as an integral part of TFX(Baylor et al., 2017) – an end-to-end machine learning platform at Google. It is used by hundreds of product teams use it to continuously monitor and validate several petabytes of production data per day. We faced several challenges in developing our system, most notably around the ability of ML pipelines to soldier on in the face of unexpected patterns, schema-free data, or training-serving skew. We discuss these challenges, the techniques we used to address them, and the various design choices that we made in implementing the system. Finally, we present evidence from the system’s deployment in production that illustrate the tangible benefits of data validation in the context of ML: early detection of errors, model-quality wins from using better data, savings in engineering hours to debug problems, and a shift towards data-centric workflows in model development.

Afinal,



Como funciona inteligência artificial?

Inteligência Artificial

	Nome	Renda	Investidor?
	João	R\$3.500,00	Sim
	Luiz	R\$1.500,00	Não
	Maria	R\$5.000,00	Sim

Base de dados inicial

Inteligência Artificial

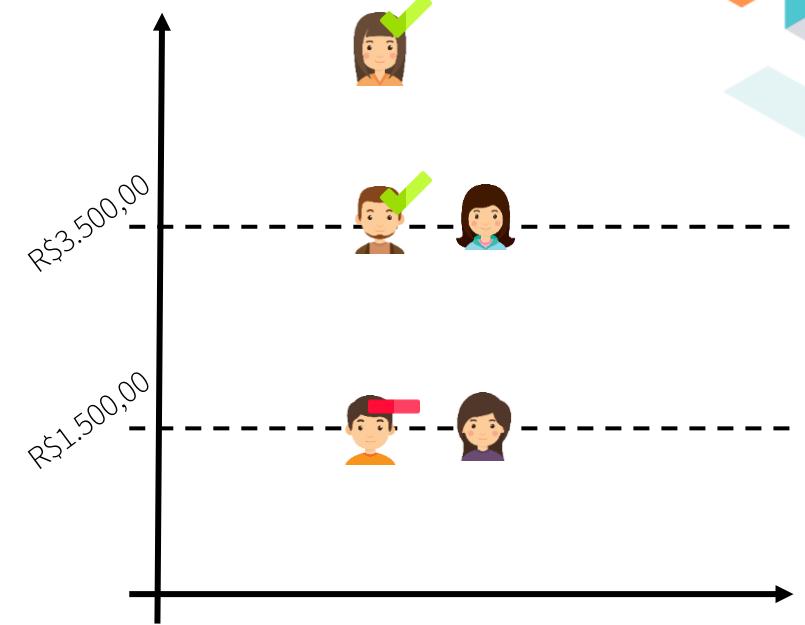
	Nome	Renda	Investidor?	
	João	R\$3.500,00	Sim	
	Luiz	R\$1.500,00	Não	
	Maria	R\$5.000,00	Sim	

Atributos **Classe**

Instâncias

Inteligência Artificial

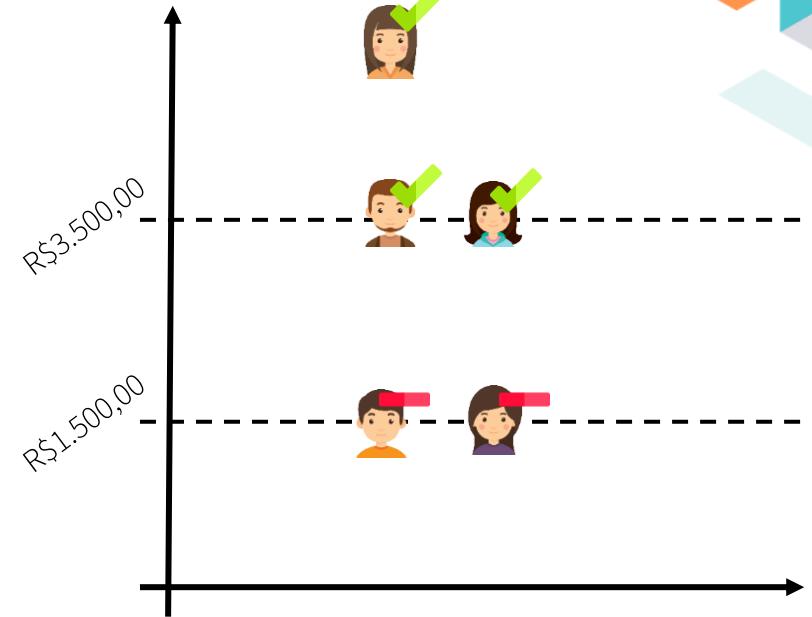
	Nome	Renda	Investidor?
	João	R\$3.500,00	Sim 
	Luiz	R\$1.500,00	Não 
	Maria	R\$5.000,00	Sim 
	Rosa	R\$1.500,00	???
	Ana	R\$3.500,00	???



Novas entradas desconhecidas

Inteligência Artificial

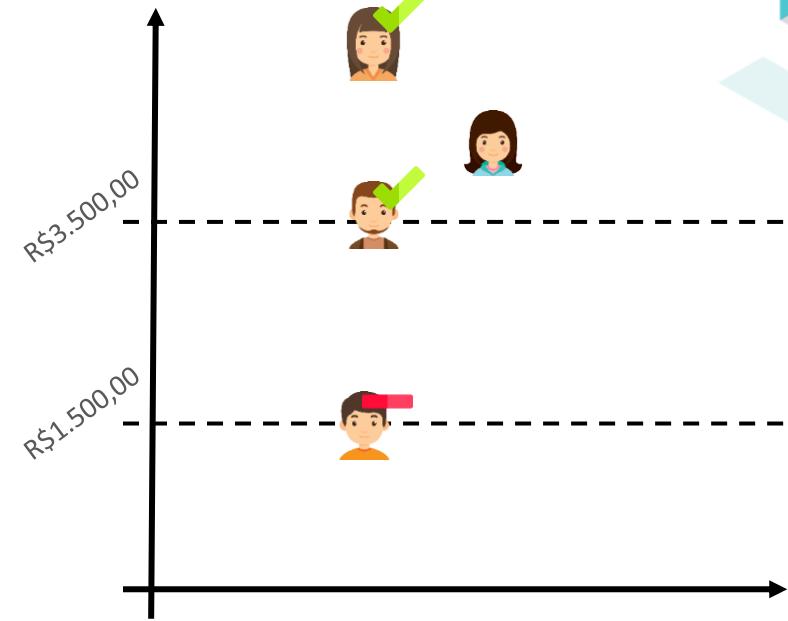
	Nome	Renda	Investidor?
	João	R\$3.500,00	Sim 
	Luiz	R\$1.500,00	Não 
	Maria	R\$5.000,00	Sim 
	Rosa	R\$1.500,00	Não 
	Ana	R\$3.500,00	Sim 



É possível classificar

Inteligência Artificial

	Nome	Renda	Investidor?
	João	R\$3.500,00	Sim 
	Luiz	R\$1.500,00	Não 
	Maria	R\$5.000,00	Sim 
	Ana	R\$3.659,00	???



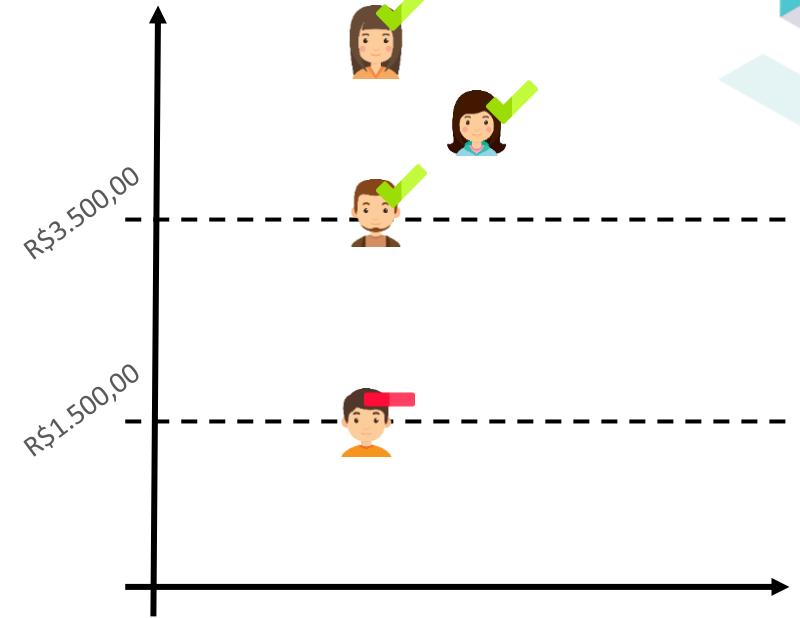
Inteligência Artificial

	Nome	Renda	Investidor?
	João	R\$3.500,00	Sim 

	Luiz	R\$1.500,00	Não 
---	------	-------------	---

	Maria	R\$5.000,00	Sim 
---	-------	-------------	---

	Ana	R\$3.659,00	Sim 
---	-----	-------------	---



É possível classificar

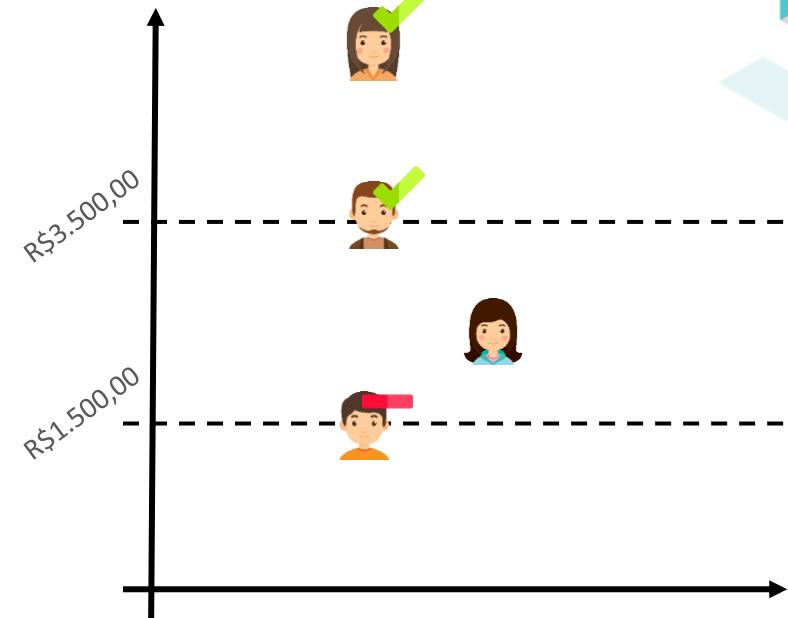
Inteligência Artificial

	Nome	Renda	Investidor?
	João	R\$3.500,00	Sim 

	Luiz	R\$1.500,00	Não 
---	------	-------------	---

	Maria	R\$5.000,00	Sim 
---	-------	-------------	---

	Ana	R\$2.000,00	???
---	-----	-------------	-----



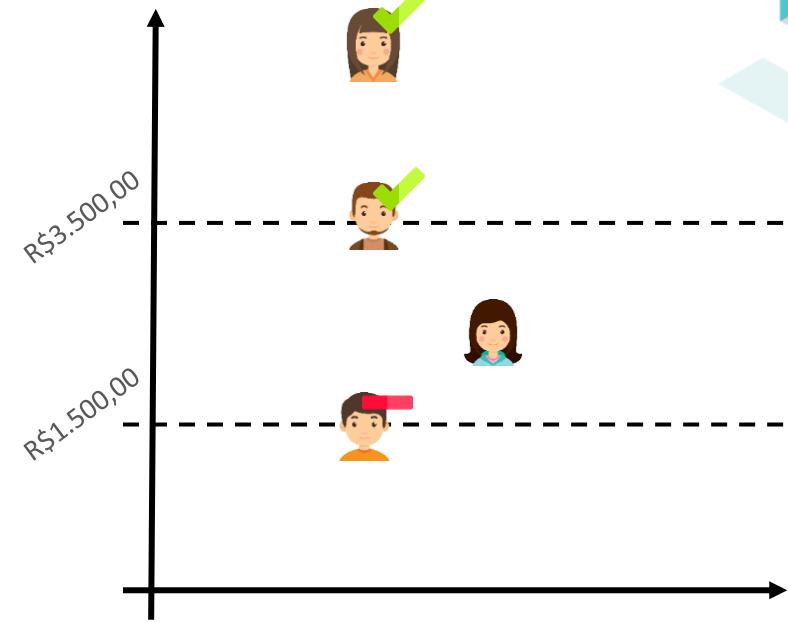
Inteligência Artificial

	Nome	Renda	Investidor?
	João	R\$3.500,00	Sim 

	Luiz	R\$1.500,00	Não 
---	------	-------------	---

	Maria	R\$5.000,00	Sim 
---	-------	-------------	---

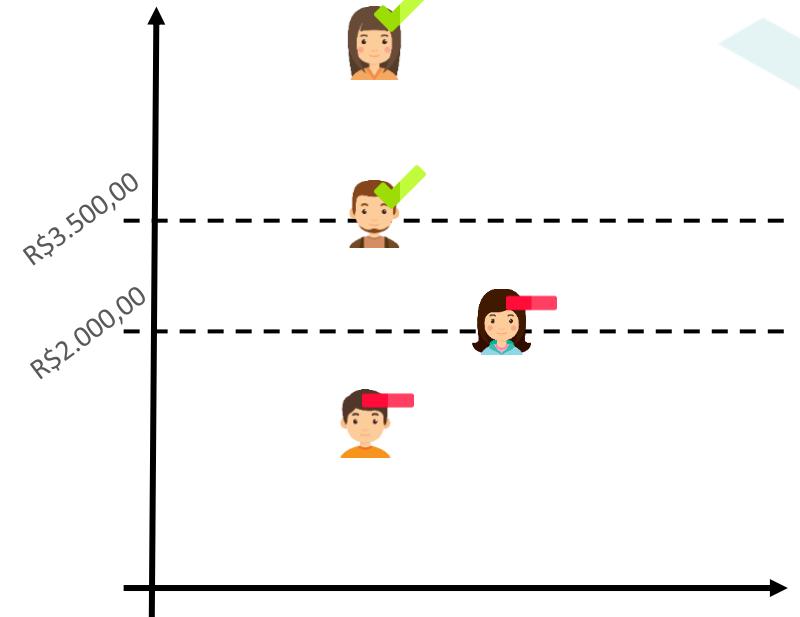
	Ana	R\$2.000,00	???
---	-----	-------------	-----



Não é possível
classificar

Inteligência Artificial

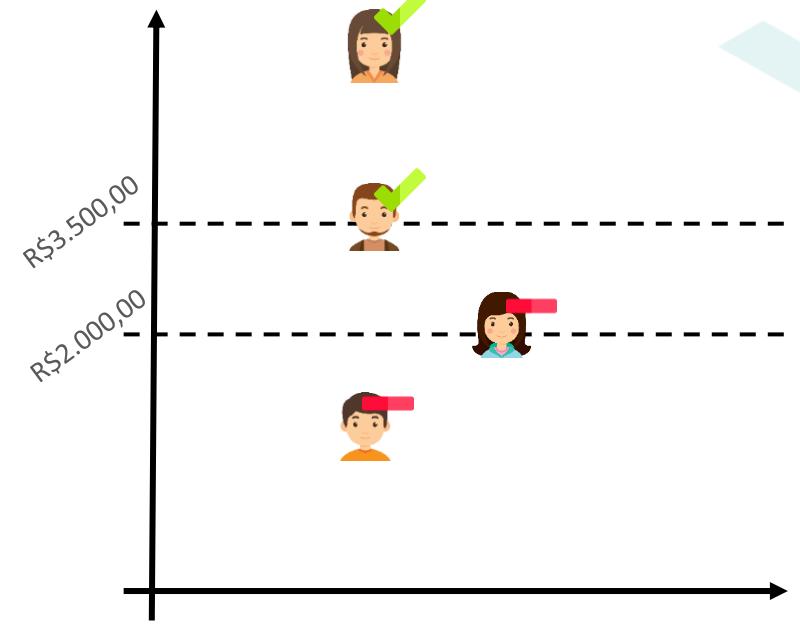
Nome	Renda	Investidor?
João	R\$3.500,00	Sim ✓
Luiz	R\$1.500,00	Não -
Maria	R\$5.000,00	Sim ✓
Ana	R\$2.000,00	Não -

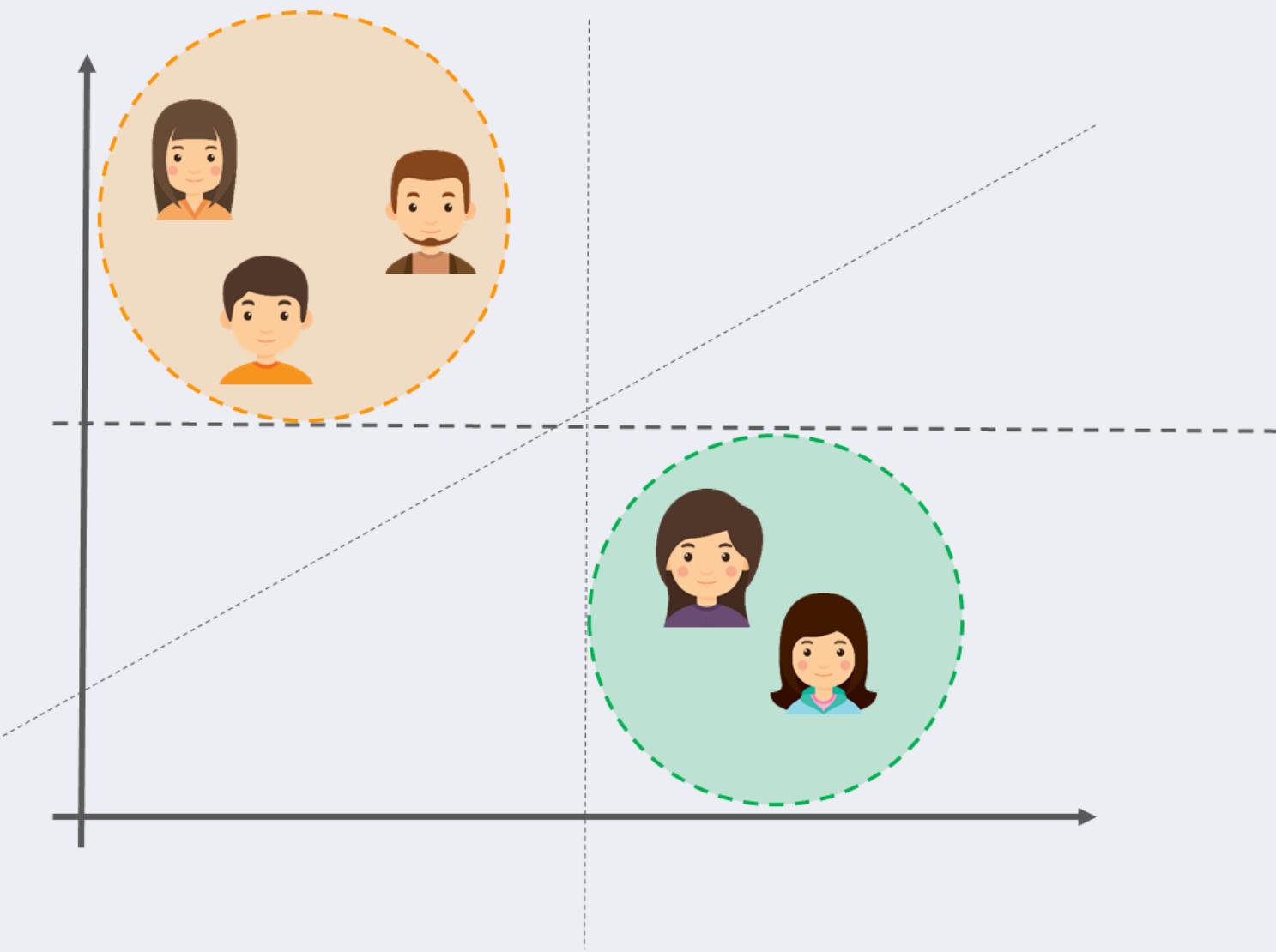


Podemos inserir novos
exemplos

Inteligência Artificial

Nome	Renda	Investidor?
 João	R\$3.500,00	Sim 
 Luiz	R\$1.500,00	Não 
 Maria	R\$5.000,00	Sim 
 Ana	R\$2.000,00	Não 
 José	R\$1.735,00	???



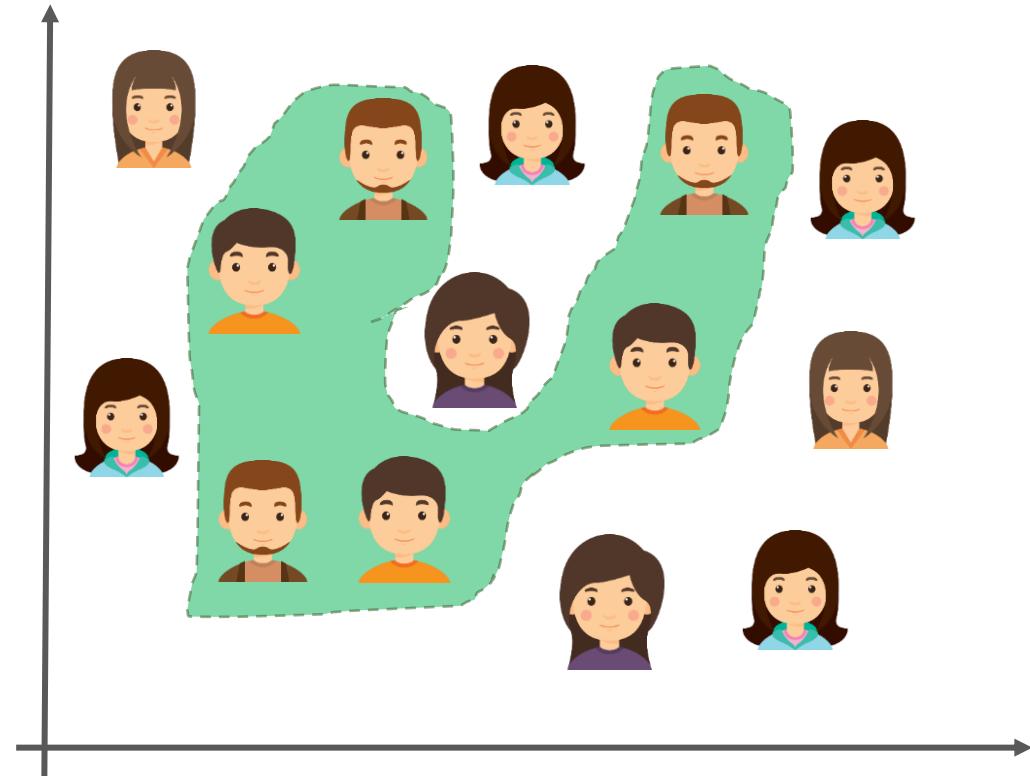


Afinal, como funciona a Inteligência Artificial?

Nome	Renda	Patrimônio	Residência	Já teve investimento	Cliente a quanto tempo	Escolaridade	...	Investidor?
 João	R\$4.500,00	R\$ 1 M	Vl. Mariana	Sim	2 anos	Médio	...	Não 
 Luiz	R\$2.500,00	R\$ 250 Mil	Jabaquara	Não	5 anos	Superior	...	Não 
 Maria	R\$6.000,00	R\$ 1.5 M	Morumbi	Não	6 anos	Médio	...	Sim 
 Ana	R\$7.000,00	R\$ 800 Mil	Jardins	Sim	3 anos	Superior	...	Sim 
 José	R\$5.000,00	R\$ 1.2 M	Conceição	Não	2 anos	Superior	...	??? 

Afinal, como funciona a Inteligência Artificial?

Mas os problemas do mundo real são mais complexos...

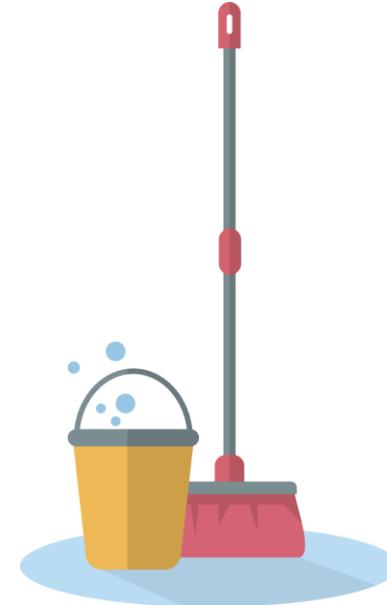








What about a dog and a mop? Easy, right?



Not so fast...



Not so fast...



Entender não é tão simples

Google me mostre fotos de raças cachorro exceto beagles

Q Tudo Imagens Compras Vídeos Notícias Mais Definições Ferramentas

filhote raças caninas

Beagle (Beagle Inglês ou famoso Snoopy...) cachorrogato.com.br

Beagle – Raças Caninas - Raças de Cac... petvale.com.br

476 x 238

Beagle (Beagle Inglês ou famoso Snoopy) - Raças de... cachorrogato.com.br

Beagle | Raças de cães | Royal Canin royalcanin.pt

Beagle - Blog do Cachorro blogdocachorro.com.br

Nós, os Cachorros - N... nososcachorros.blogspot.com

Beagle: O cachorro cantor | Au au au!!! caninablog.wordpress.com

Beagle: tudo sobre a raça em um guia co... arbolez.com

Raças: Beagle | BitCão bitcao.com.br

Introdução à Machine Learning

	Peso	Altura
 Pessoa 1	80 kg	163
 Pessoa 2	85 kg	168
 Pessoa 3	90 kg	175
 Pessoa 4	95 kg	188

Introdução à Machine Learning

	Peso	Altura
	Pessoa 1	80 kg
	Pessoa 2	85 kg
	Pessoa 3	90 kg
	Pessoa 4	95 kg

$$\hat{y} = \beta_0 + \beta_1 X_1$$

Introdução à Machine Learning

	Peso	Altura
	Pessoa 1	80 kg
	Pessoa 2	85 kg
	Pessoa 3	90 kg
	Pessoa 4	95 kg

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$\hat{y} = 17 + 1,8 \times 163$$

Introdução à Machine Learning

	Peso	Altura
	Pessoa 1	80 kg
	Pessoa 2	85 kg
	Pessoa 3	90 kg
	Pessoa 4	95 kg

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$\hat{y} = 310,4$$

Introdução à Machine Learning

	Peso	Altura
	Pessoa 1	80 kg 163
	Pessoa 2	85 kg 168
	Pessoa 3	90 kg 175
	Pessoa 4	95 kg 188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$\hat{y} = 310,4$$

Introdução à Machine Learning

	Peso	Altura
 Pessoa 1	80 kg	163
 Pessoa 2	85 kg	168
 Pessoa 3	90 kg	175
 Pessoa 4	95 kg	188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$MSE = (80 - 310,4)^2$$

$$\hat{y} = 310,4$$

Introdução à Machine Learning

	Peso	Altura
 Pessoa 1	80 kg	163
 Pessoa 2	85 kg	168
 Pessoa 3	90 kg	175
 Pessoa 4	95 kg	188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$MSE = (80 - 310,4)^2$$

$$\hat{y} = 310,4$$

$$MSE = 53.084,16$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163
Pessoa 2	85 kg	168
Pessoa 3	90 kg	175
Pessoa 4	95 kg	188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

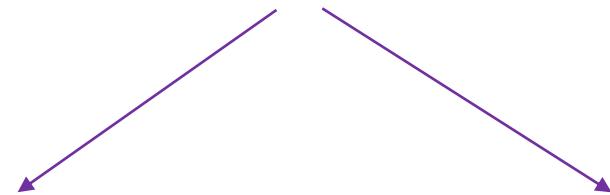
$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$MSE = (80 - 310,4)^2$$

$$\hat{y} = 310,4$$

$$MSE = 53.084,16$$



Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163
Pessoa 2	85 kg	168
Pessoa 3	90 kg	175
Pessoa 4	95 kg	188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

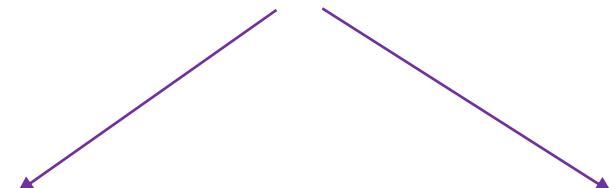
$$MSE = (80 - 310,4)^2$$

$$\hat{y} = 310,4$$

$$MSE = 53.084,16$$

$$\hat{y} = 20 + 2,1 \times 163$$

$$\hat{y} = 14 + 1,5 \times 163$$



Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163
Pessoa 2	85 kg	168
Pessoa 3	90 kg	175
Pessoa 4	95 kg	188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$MSE = (80 - 310,4)^2$$

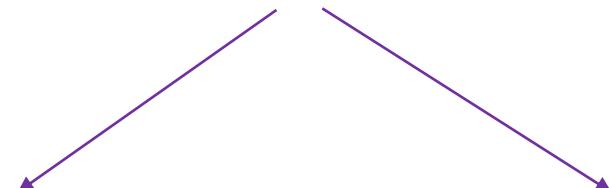
$$\hat{y} = 310,4$$

$$MSE = 53.084,16$$

$$\hat{y} = 20 + 2,1 \times 163$$

$$\hat{y} = 362,3$$

$$\hat{y} = 14 + 1,5 \times 163$$



Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163
Pessoa 2	85 kg	168
Pessoa 3	90 kg	175
Pessoa 4	95 kg	188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$MSE = (80 - 310,4)^2$$

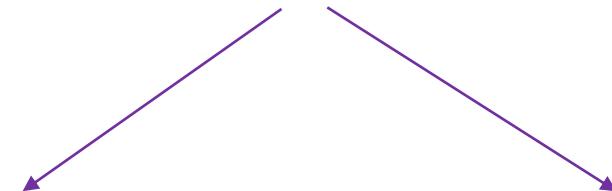
$$\hat{y} = 310,4$$

$$MSE = 53.084,16$$

$$\hat{y} = 20 + 2,1 \times 163$$

$$\hat{y} = 362,3$$

$$MSE = (80 - 362,3)^2$$



Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163
Pessoa 2	85 kg	168
Pessoa 3	90 kg	175
Pessoa 4	95 kg	188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$MSE = (80 - 310,4)^2$$

$$\hat{y} = 310,4$$

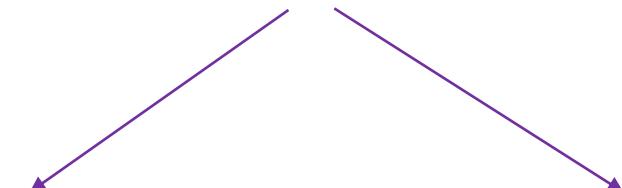
$$MSE = 53.084,16$$

$$\hat{y} = 20 + 2,1 \times 163$$

$$\hat{y} = 362,3$$

$$MSE = (80 - 362,3)^2$$

$$MSE = 79.693,29$$



$$\hat{y} = 14 + 1,5 \times 163$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163
Pessoa 2	85 kg	168
Pessoa 3	90 kg	175
Pessoa 4	95 kg	188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$MSE = (80 - 310,4)^2$$

$$\hat{y} = 310,4$$

$$MSE = 53.084,16$$

$$\hat{y} = 20 + 2,1 \times 163$$

$$\hat{y} = 362,3$$

$$\hat{y} = 14 + 1,5 \times 163$$

$$\hat{y} = 258,5$$

$$MSE = (80 - 362,3)^2$$

$$MSE = 79.693,29$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163
Pessoa 2	85 kg	168
Pessoa 3	90 kg	175
Pessoa 4	95 kg	188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$MSE = (80 - 310,4)^2$$

$$\hat{y} = 310,4$$

$$MSE = 53.084,16$$

$$\hat{y} = 20 + 2,1 \times 163$$

$$\hat{y} = 362,3$$

$$\hat{y} = 14 + 1,5 \times 163$$

$$\hat{y} = 258,5$$

$$MSE = (80 - 362,3)^2$$

$$MSE = 79.693,29$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163
Pessoa 2	85 kg	168
Pessoa 3	90 kg	175
Pessoa 4	95 kg	188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$MSE = (80 - 310,4)^2$$

$$\hat{y} = 310,4$$

$$MSE = 53.084,16$$

$$\hat{y} = 20 + 2,1 \times 163$$

$$\hat{y} = 362,3$$

$$\hat{y} = 14 + 1,5 \times 163$$

$$\hat{y} = 258,5$$

$$MSE = (80 - 362,3)^2$$

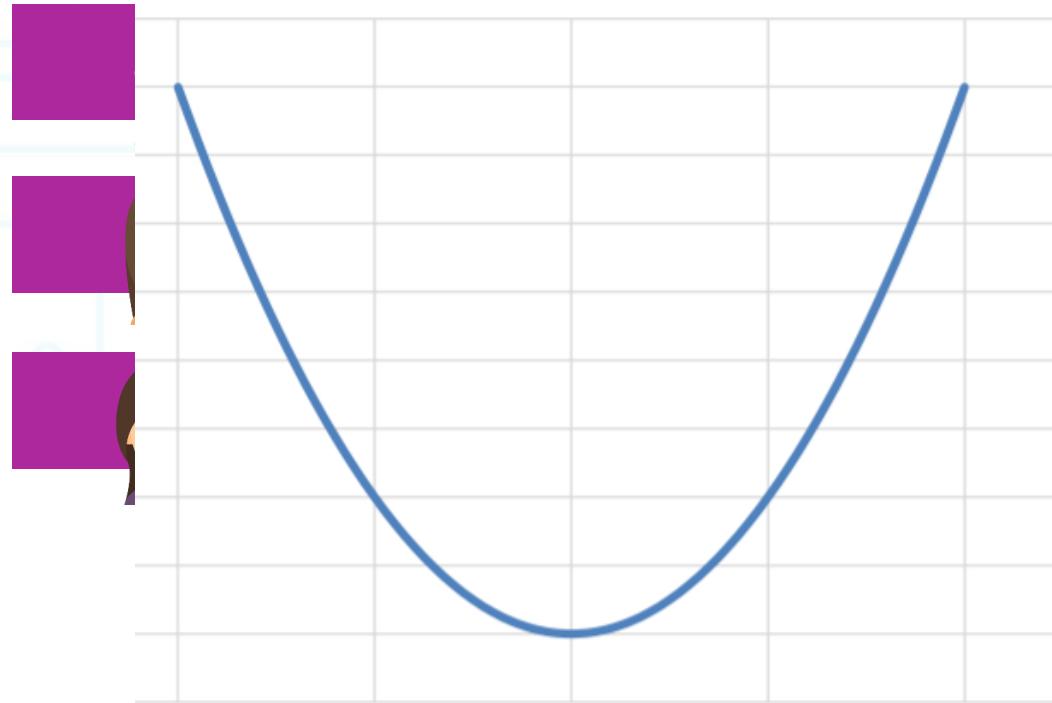
$$MSE = 79.693,29$$

$$MSE = (80 - 258,5)^2$$

$$MSE = 31.862,25$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163



$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$MSE = (80 - 310,4)^2$$

$$\hat{y} = 310,4$$

$$MSE = 53.084,16$$

$$\hat{y} = 20 + 2,1 \times 163$$

$$\hat{y} = 362,3$$

$$\hat{y} = 14 + 1,5 \times 163$$

$$\hat{y} = 258,5$$

$$MSE = (80 - 362,3)^2$$

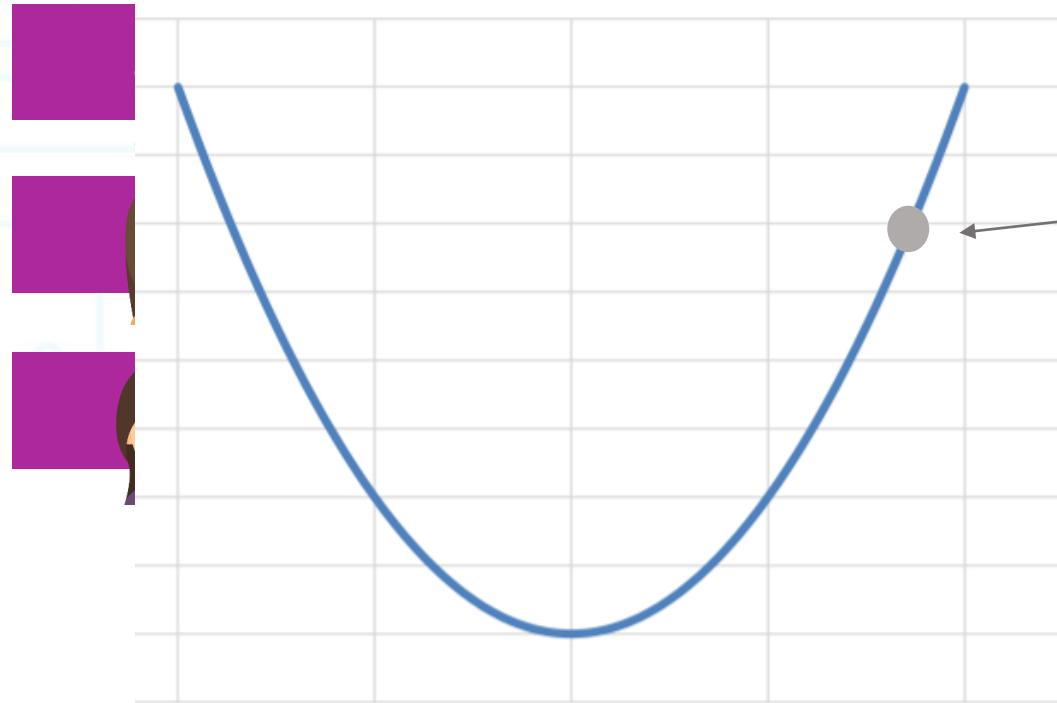
$$MSE = 79.693,29$$

$$MSE = (80 - 258,5)^2$$

$$MSE = 31.862,25$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163



$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$MSE = (80 - 310,4)^2$$

$$\hat{y} = 310,4$$

$$MSE = 53.084,16$$

$$\hat{y} = 20 + 2,1 \times 163$$

$$\hat{y} = 362,3$$

$$MSE = (80 - 362,3)^2$$

$$MSE = 79.693,29$$

$$\hat{y} = 14 + 1,5 \times 163$$

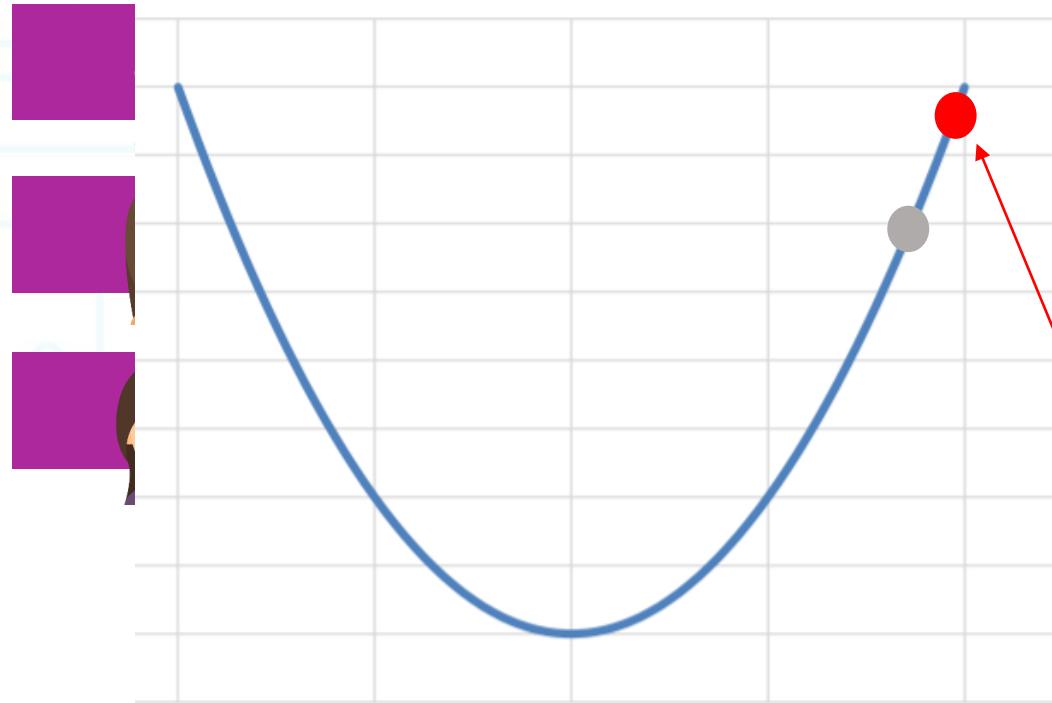
$$\hat{y} = 258,5$$

$$MSE = (80 - 258,5)^2$$

$$MSE = 31.862,25$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163



$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$MSE = (80 - 310,4)^2$$

$$\hat{y} = 310,4$$

$$MSE = 53.084,16$$

$$\hat{y} = 20 + 2,1 \times 163$$

$$\hat{y} = 362,3$$

$$\hat{y} = 362,3$$

$$\hat{y} = 258,5$$

$$MSE = (80 - 362,3)^2$$

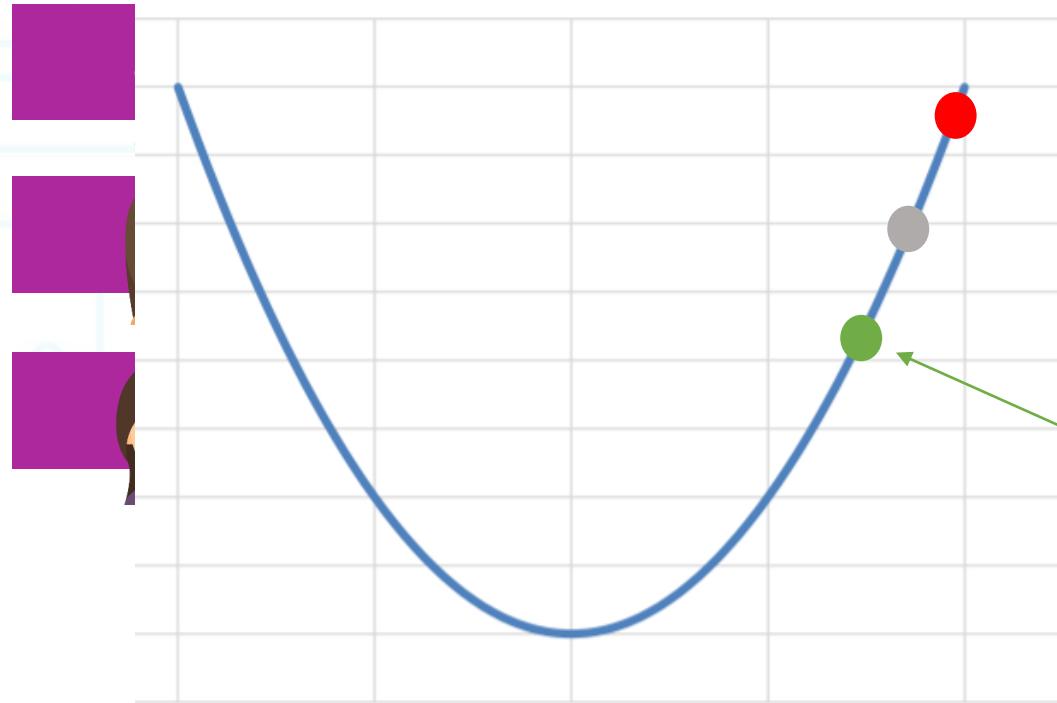
$$MSE = (80 - 258,5)^2$$

$$MSE = 79.693,29$$

$$MSE = 31.862,25$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163



$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$MSE = (80 - 310,4)^2$$

$$\hat{y} = 310,4$$

$$MSE = 53.084,16$$

$$\hat{y} = 20 + 2,1 \times 163$$

$$\hat{y} = 14 + 1,5 \times 163$$

$$\hat{y} = 362,3$$

$$\hat{y} = 258,5$$

$$MSE = (80 - 362,3)^2$$

$$MSE = (80 - 258,5)^2$$

$$MSE = 79.693,29$$

$$MSE = 31.862,25$$

Introdução à Machine Learning

	Peso	Altura	
Pessoa 1	80 kg	163	←
Pessoa 2	85 kg	168	
Pessoa 3	90 kg	175	
Pessoa 4	95 kg	188	

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$MSE = (80 - 310,4)^2$$

$$\hat{y} = 310,4$$

$$MSE = 53.084,16$$

$$\hat{y} = 14 + 1,5 \times 163$$

$$MSE = (80 - 258,5)^2$$

$$\hat{y} = 258,5$$

$$MSE = 31.862,25$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163

Pessoa 2	85 kg	168
----------	-------	-----

Pessoa 3	90 kg	175
----------	-------	-----

Pessoa 4	95 kg	188
----------	-------	-----

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$MSE = (80 - 310,4)^2$$

$$\hat{y} = 310,4$$

$$MSE = 53.084,16$$

$$\hat{y} = 14 + 1,5 \times 163$$

$$MSE = (80 - 258,5)^2$$

$$\hat{y} = 258,5$$

$$MSE = 31.862,25$$

$$\hat{y} = 10 + 1,1 \times 168$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163

Pessoa 2	85 kg	168
----------	-------	-----

Pessoa 3	90 kg	175
----------	-------	-----

Pessoa 4	95 kg	188
----------	-------	-----

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$MSE = (80 - 310,4)^2$$

$$\hat{y} = 310,4$$

$$MSE = 53.084,16$$

$$\hat{y} = 14 + 1,5 \times 163$$

$$MSE = (80 - 258,5)^2$$

$$\hat{y} = 258,5$$

$$MSE = 31.862,25$$

$$\hat{y} = 10 + 1,1 \times 168$$

$$\hat{y} = 194,8$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163

Pessoa 2	85 kg	168	←	
----------	-------	-----	---	--

Pessoa 3	90 kg	175
----------	-------	-----

Pessoa 4	95 kg	188
----------	-------	-----

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$MSE = (80 - 310,4)^2$$

$$\hat{y} = 310,4$$

$$MSE = 53.084,16$$

$$\hat{y} = 14 + 1,5 \times 163$$

$$MSE = (80 - 258,5)^2$$

$$\hat{y} = 258,5$$

$$MSE = 31.862,25$$

$$\hat{y} = 10 + 1,1 \times 168$$

$$MSE = (85 - 194,8)^2$$

$$\hat{y} = 194,8$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163

Pessoa 2	85 kg	168	←	
----------	-------	-----	---	--

Pessoa 3	90 kg	175
----------	-------	-----

Pessoa 4	95 kg	188
----------	-------	-----

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 17 + 1,8 \times 163$$

$$MSE = (80 - 310,4)^2$$

$$\hat{y} = 310,4$$

$$MSE = 53.084,16$$

$$\hat{y} = 14 + 1,5 \times 163$$

$$MSE = (80 - 258,5)^2$$

$$\hat{y} = 258,5$$

$$MSE = 31.862,25$$

$$\hat{y} = 10 + 1,1 \times 168$$

$$MSE = (85 - 194,8)^2$$

$$\hat{y} = 194,8$$

$$MSE = 12.056,04$$

Introdução à Machine Learning

	Peso	Altura
	Pessoa 1	80 kg 163
	Pessoa 2	85 kg 168
	Pessoa 3	90 kg 175
	Pessoa 4	95 kg 188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$



Introdução à Machine Learning

	Peso	Altura
	Pessoa 1	80 kg 163
	Pessoa 2	85 kg 168
	Pessoa 3	90 kg 175
	Pessoa 4	95 kg 188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$\hat{y} = 147$$



Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163
Pessoa 2	85 kg	168
Pessoa 3	90 kg	175
Pessoa 4	95 kg	188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$MSE = (90 - 147)^2$$

$$\hat{y} = 147$$



Introdução à Machine Learning

	Peso	Altura
	Pessoa 1	80 kg 163
	Pessoa 2	85 kg 168
	Pessoa 3	90 kg 175
	Pessoa 4	95 kg 188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$MSE = (90 - 147)^2$$

$$\hat{y} = 147$$

$$MSE = 3.249$$



Introdução à Machine Learning

	Peso	Altura
	Pessoa 1	80 kg 163
	Pessoa 2	85 kg 168
	Pessoa 3	90 kg 175
	Pessoa 4	95 kg 188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$MSE = (90 - 147)^2$$

$$\hat{y} = 147$$

$$MSE = 3.249$$

$$\hat{y} = 3 + 0,5 \times 188$$



Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163
Pessoa 2	85 kg	168
Pessoa 3	90 kg	175
Pessoa 4	95 kg	188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$MSE = (90 - 147)^2$$

$$\hat{y} = 147$$

$$MSE = 3.249$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$\hat{y} = 97$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163
Pessoa 2	85 kg	168
Pessoa 3	90 kg	175
Pessoa 4	95 kg	188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$MSE = (90 - 147)^2$$

$$\hat{y} = 147$$

$$MSE = 3.249$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$



Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163

Pessoa 2	85 kg	168
----------	-------	-----

Pessoa 3	90 kg	175
----------	-------	-----

Pessoa 4	95 kg	188
----------	-------	-----

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$MSE = (90 - 147)^2$$

$$\hat{y} = 147$$

$$MSE = 3.249$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

$$MSE = 4$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163

Pessoa 2	85 kg	168
----------	-------	-----

Pessoa 3	90 kg	175
----------	-------	-----

Pessoa 4	95 kg	188
----------	-------	-----

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$MSE = (90 - 147)^2$$

$$\hat{y} = 147$$

$$MSE = 3.249$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

$$MSE = 4$$

$$\hat{y} = 1 + 0,2 \times 188$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163
Pessoa 2	85 kg	168
Pessoa 3	90 kg	175
Pessoa 4	95 kg	188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$MSE = (90 - 147)^2$$

$$\hat{y} = 147$$

$$MSE = 3.249$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

$$MSE = 4$$

$$\hat{y} = 1 + 0,2 \times 188$$

$$\hat{y} = 38,6$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163

Pessoa 2	85 kg	168
----------	-------	-----

Pessoa 3	90 kg	175
----------	-------	-----

Pessoa 4	95 kg	188
----------	-------	-----

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$MSE = (90 - 147)^2$$

$$\hat{y} = 147$$

$$MSE = 3.249$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

$$MSE = 4$$

$$\hat{y} = 1 + 0,2 \times 188$$

$$MSE = (95 - 38,6)^2$$

$$\hat{y} = 38,6$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163

Pessoa 2	85 kg	168
----------	-------	-----

Pessoa 3	90 kg	175
----------	-------	-----

Pessoa 4	95 kg	188
----------	-------	-----

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$MSE = (90 - 147)^2$$

$$\hat{y} = 147$$

$$MSE = 3.249$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

$$MSE = 4$$

$$\hat{y} = 1 + 0,2 \times 188$$

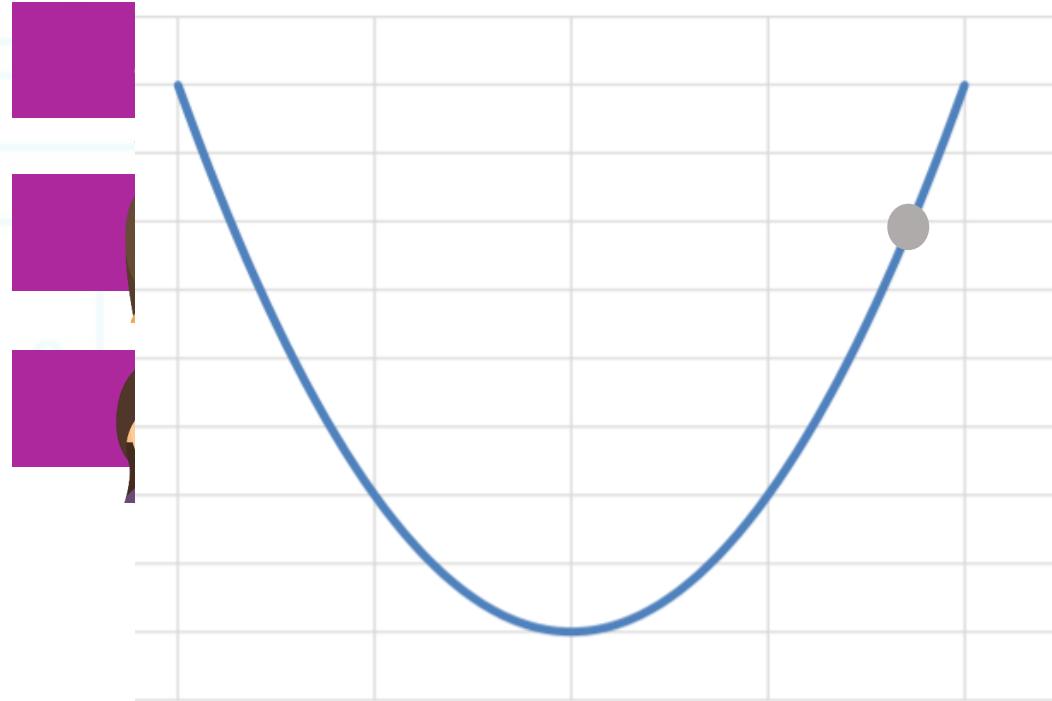
$$MSE = (95 - 38,6)^2$$

$$\hat{y} = 38,6$$

$$MSE = 3.180,97$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163



$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$MSE = (90 - 147)^2$$

$$\hat{y} = 147$$

$$MSE = 3.249$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

$$MSE = 4$$

$$\hat{y} = 1 + 0,2 \times 188$$

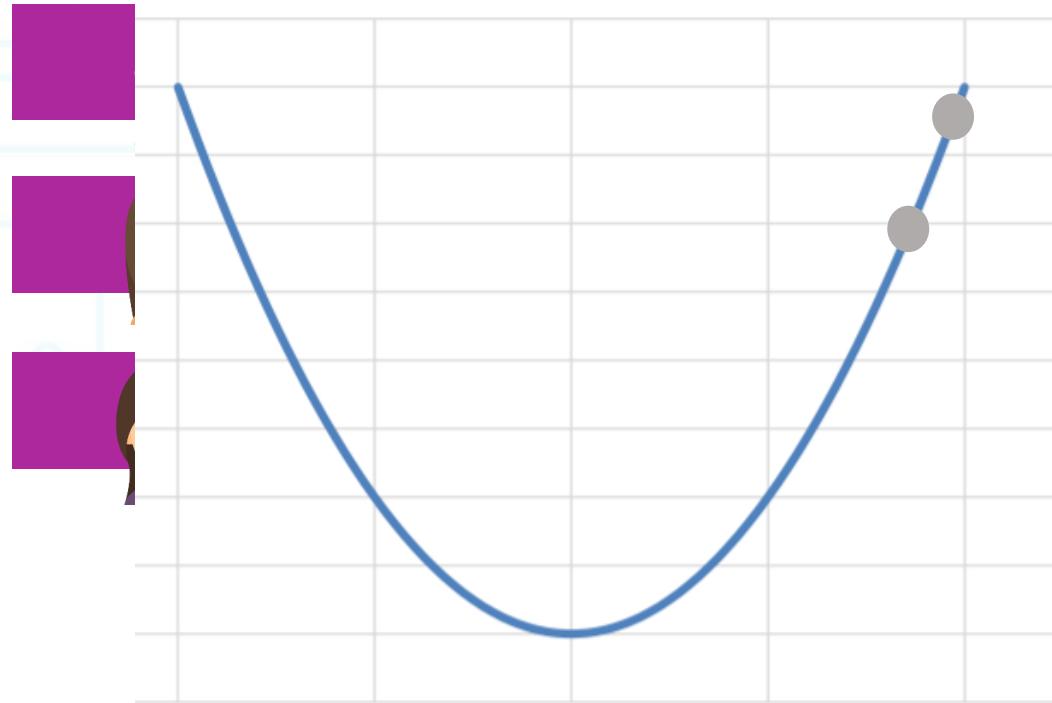
$$MSE = (95 - 38,6)^2$$

$$\hat{y} = 38,6$$

$$MSE = 3.180,97$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163



$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$MSE = (90 - 147)^2$$

$$\hat{y} = 147$$

$$MSE = 3.249$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

$$MSE = 4$$

$$\hat{y} = 1 + 0,2 \times 188$$

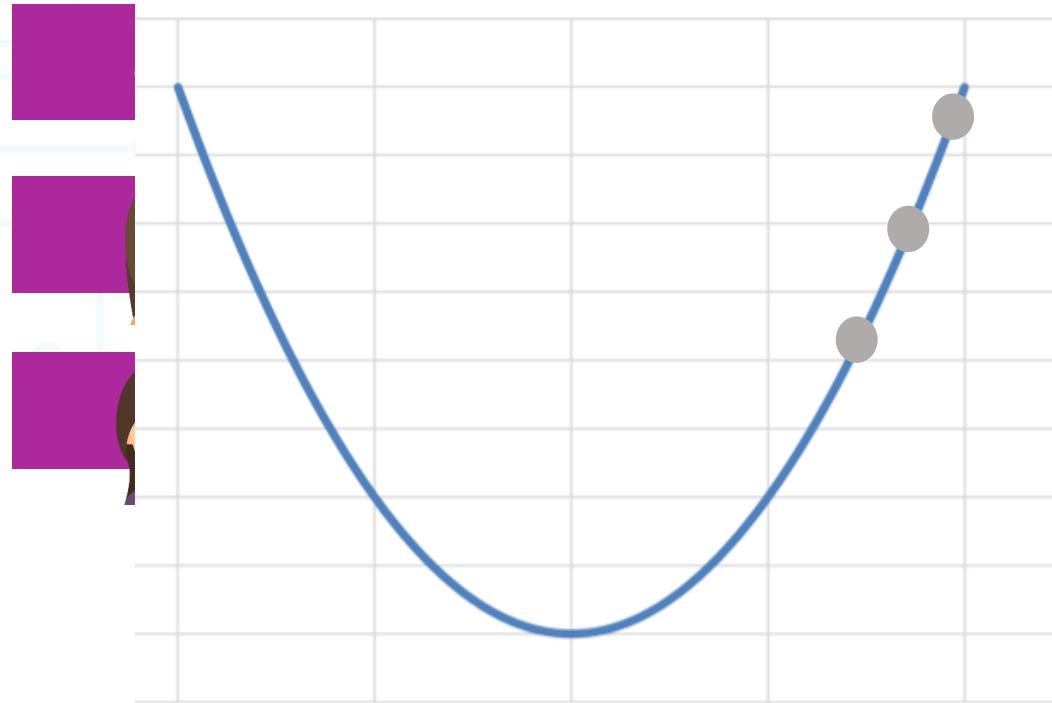
$$MSE = (95 - 38,6)^2$$

$$\hat{y} = 38,6$$

$$MSE = 3.180,97$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163



$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$MSE = (90 - 147)^2$$

$$\hat{y} = 147$$

$$MSE = 3.249$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

$$MSE = 4$$

$$\hat{y} = 1 + 0,2 \times 188$$

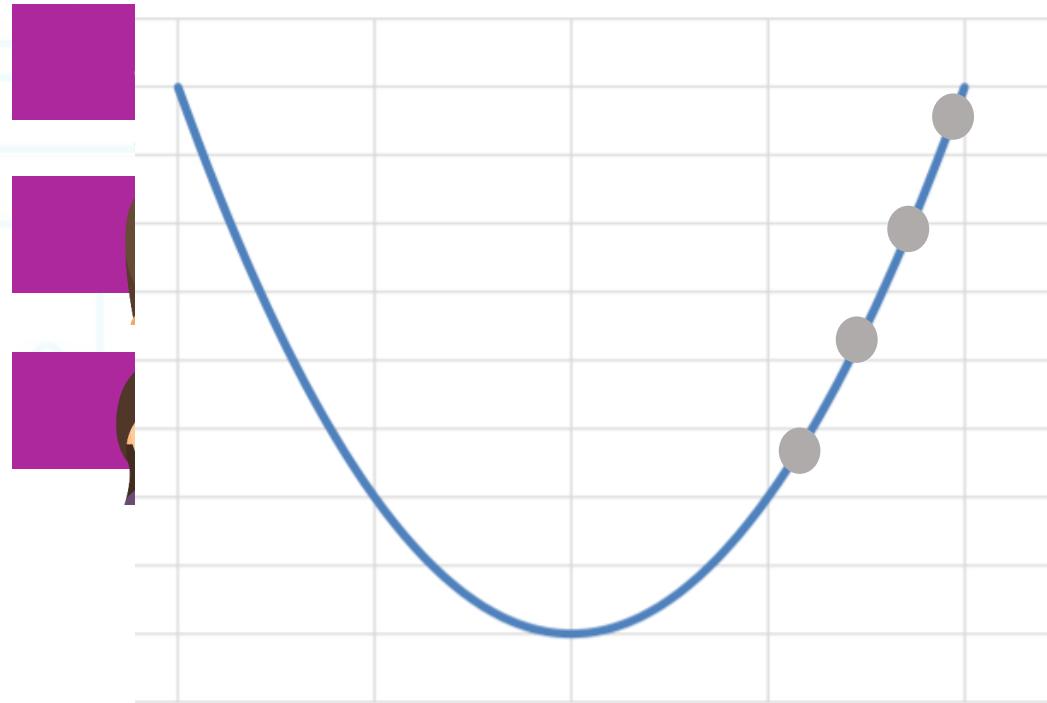
$$MSE = (95 - 38,6)^2$$

$$\hat{y} = 38,6$$

$$MSE = 3.180,97$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163



$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$MSE = (90 - 147)^2$$

$$\hat{y} = 147$$

$$MSE = 3.249$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

$$MSE = 4$$

$$\hat{y} = 1 + 0,2 \times 188$$

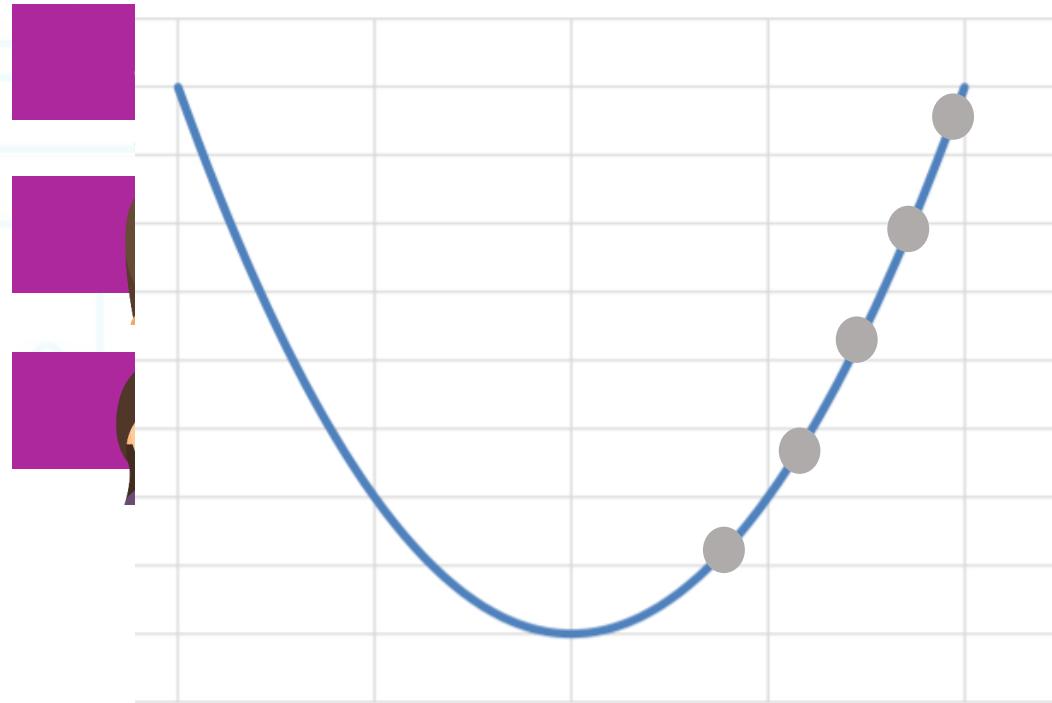
$$MSE = (95 - 38,6)^2$$

$$\hat{y} = 38,6$$

$$MSE = 3.180,97$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163



$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$MSE = (90 - 147)^2$$

$$\hat{y} = 147$$

$$MSE = 3.249$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

$$MSE = 4$$

$$\hat{y} = 1 + 0,2 \times 188$$

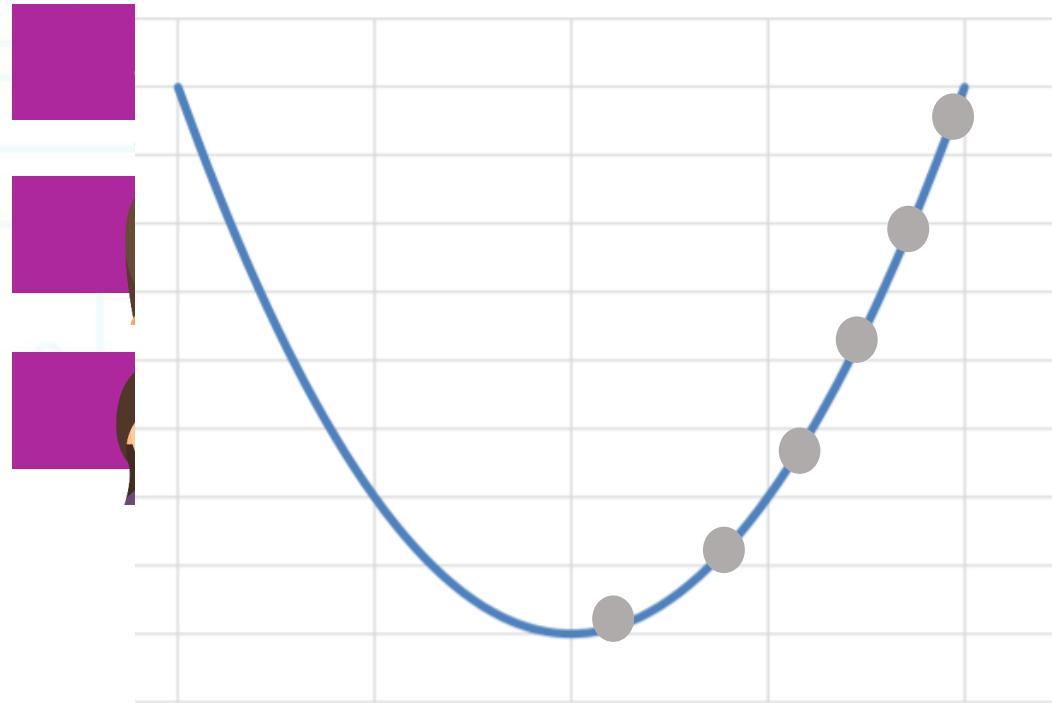
$$MSE = (95 - 38,6)^2$$

$$\hat{y} = 38,6$$

$$MSE = 3.180,97$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163



$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$MSE = (90 - 147)^2$$

$$\hat{y} = 147$$

$$MSE = 3.249$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

$$MSE = 4$$

$$\hat{y} = 1 + 0,2 \times 188$$

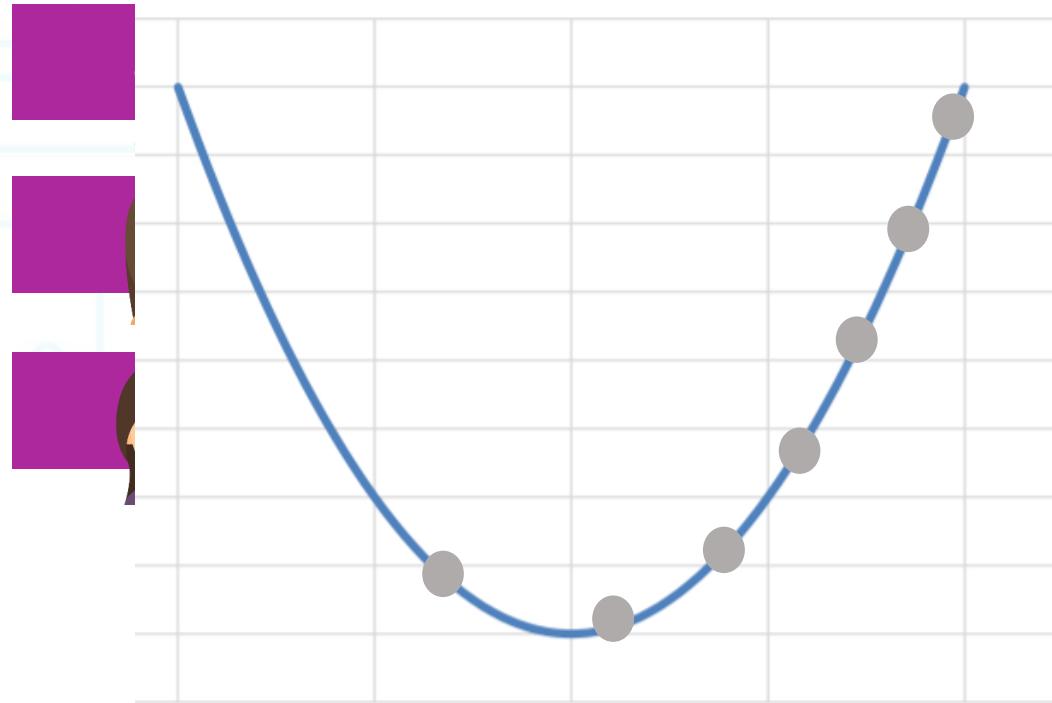
$$MSE = (95 - 38,6)^2$$

$$\hat{y} = 38,6$$

$$MSE = 3.180,97$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163



$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$MSE = (90 - 147)^2$$

$$\hat{y} = 147$$

$$MSE = 3.249$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

$$MSE = 4$$

$$\hat{y} = 1 + 0,2 \times 188$$

$$MSE = (95 - 38,6)^2$$

$$\hat{y} = 38,6$$

$$MSE = 3.180,97$$

Introdução à Machine Learning

	Peso	Altura
	Pessoa 1	80 kg 163
	Pessoa 2	85 kg 168
	Pessoa 3	90 kg 175
	Pessoa 4	95 kg 188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 7 + 0,8 \times 175$$

$$MSE = (90 - 147)^2$$

$$\hat{y} = 147$$

$$MSE = 3.249$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

$$MSE = 4$$

$$\hat{y} = 1 + 0,2 \times 188$$

$$MSE = (95 - 38,6)^2$$

$$\hat{y} = 38,6$$

$$MSE = 3.180,97$$

Introdução à Machine Learning

	Peso	Altura
	Pessoa 1	80 kg 163
	Pessoa 2	85 kg 168
	Pessoa 3	90 kg 175
	Pessoa 4	95 kg 188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

$$MSE = 4$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163
Pessoa 2	85 kg	168
Pessoa 3	90 kg	175
Pessoa 4	95 kg	188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

$$MSE = 4$$

$$\hat{y} = 84,5 \quad MSE = 20,25$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163
Pessoa 2	85 kg	168
Pessoa 3	90 kg	175
Pessoa 4	95 kg	188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

$$MSE = 4$$

$$\hat{y} = 84,5$$

$$MSE = 20,25$$

$$\hat{y} = 87$$

$$MSE = 4$$



Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163
Pessoa 2	85 kg	168
Pessoa 3	90 kg	175
Pessoa 4	95 kg	188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

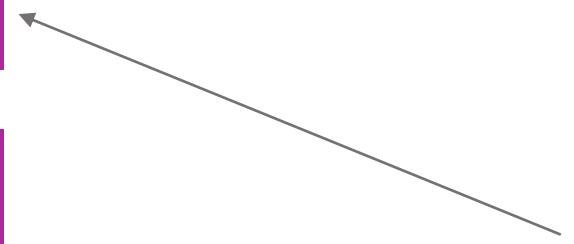
$$\hat{y} = 97$$

$$MSE = 4$$

$$\hat{y} = 84,5 \quad MSE = 20,25$$

$$\hat{y} = 87 \quad MSE = 4$$

$$\hat{y} = 90,5 \quad MSE = 0,25$$



Introdução à Machine Learning

	Peso	Altura
	Pessoa 1	80 kg 163
	Pessoa 2	85 kg 168
	Pessoa 3	90 kg 175
	Pessoa 4	95 kg 188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

$$MSE = 4$$

$$\hat{y} = 84,5$$

$$MSE = 20,25$$

$$\hat{y} = 87$$

$$MSE = 4$$

$$\hat{y} = 90,5$$

$$MSE = 0,25$$

$$\hat{y} = 97$$

$$MSE = 4$$



Introdução à Machine Learning

	Peso	Altura
	Pessoa 1	80 kg 163
	Pessoa 2	85 kg 168
	Pessoa 3	90 kg 175
	Pessoa 4	95 kg 188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

$$MSE = 4$$

$$\hat{y} = 84,5 \quad MSE = 20,25$$

$$\hat{y} = 87 \quad MSE = 4$$

$$\hat{y} = 90,5 \quad MSE = 0,25$$

$$\hat{y} = 97 \quad MSE = 4$$

$$MSE = 7,125$$

Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163
Pessoa 2	85 kg	168
Pessoa 3	90 kg	175
Pessoa 4	95 kg	188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{y} = 3 + 0,5 \times 188$$

$$MSE = (95 - 97)^2$$

$$\hat{y} = 97$$

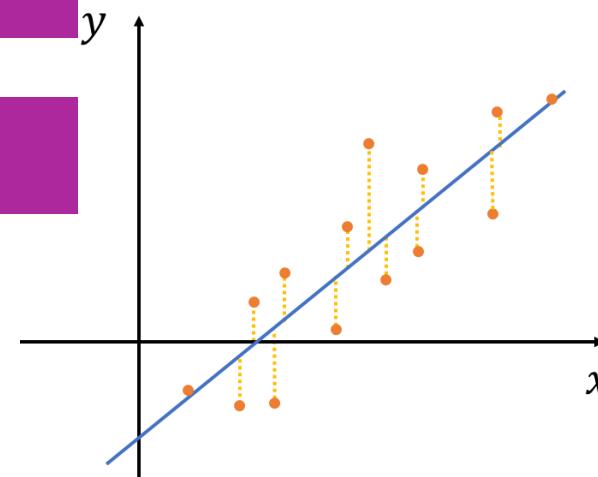
$$MSE = 4$$

$$\hat{y} = 84,5 \quad MSE = 20,25$$

$$\hat{y} = 87 \quad MSE = 4$$

$$\hat{y} = 90,5 \quad MSE = 0,25$$

$$\hat{y} = 97 \quad MSE = 4$$



$$MSE = 7,125$$

Introdução à Machine Learning

	Peso	Altura
	Pessoa 1	80 kg 163
	Pessoa 2	85 kg 168
	Pessoa 3	90 kg 175
	Pessoa 4	95 kg 188

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\beta_0 = 3$$

$$\beta_1 = 0,5$$

$$MSE = 7,125$$

$$\hat{y} = 3 + 0,5 X_1$$

Introdução à Machine Learning

	Peso	Altura
	Pessoa 1	80 kg 163
	Pessoa 2	85 kg 168
	Pessoa 3	90 kg 175
	Pessoa 4	95 kg 188
	Pessoa 5	?? kg 158

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\beta_0 = 3$$

$$\beta_1 = 0,5$$

$$MSE = 7,125$$

$$\hat{y} = 3 + 0,5 X_1$$

Introdução à Machine Learning

	Peso	Altura
	Pessoa 1	80 kg 163
	Pessoa 2	85 kg 168
	Pessoa 3	90 kg 175
	Pessoa 4	95 kg 188
	Pessoa 5	?? kg 158

$$\hat{y} = \beta_0 + \beta_1 X_1$$

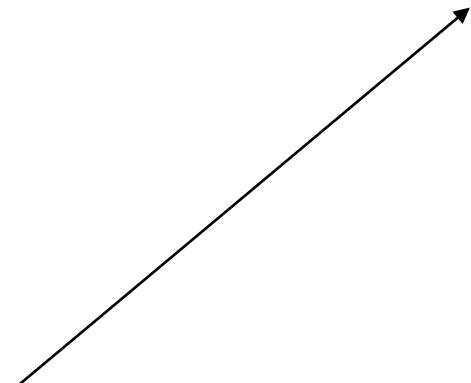
$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\beta_0 = 3$$

$$\beta_1 = 0,5$$

$$MSE = 7,125$$

$$\hat{y} = 3 + 0,5 X_1$$



Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163
Pessoa 2	85 kg	168
Pessoa 3	90 kg	175
Pessoa 4	95 kg	188
Pessoa 5	?? kg	158

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

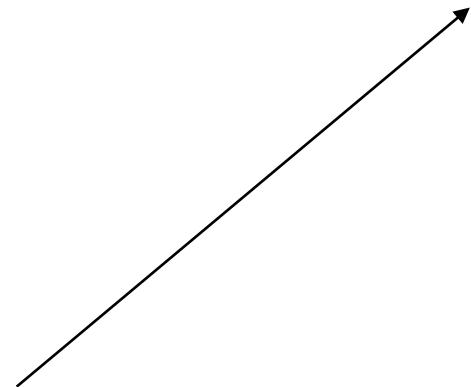
$$\beta_0 = 3$$

$$\beta_1 = 0,5$$

$$MSE = 7,125$$

$$\hat{y} = 3 + 0,5 X_1$$

$$\hat{y} = 3 + 0,5 \times 158$$



Introdução à Machine Learning

	Peso	Altura
Pessoa 1	80 kg	163
Pessoa 2	85 kg	168
Pessoa 3	90 kg	175
Pessoa 4	95 kg	188
Pessoa 5	82 kg	158

$$\hat{y} = \beta_0 + \beta_1 X_1$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\beta_0 = 3$$

$$\beta_1 = 0,5$$

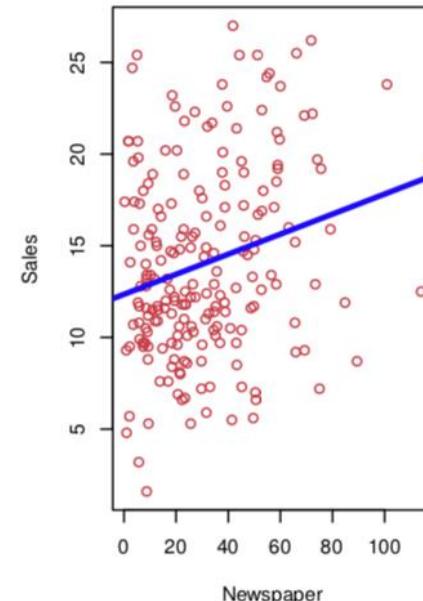
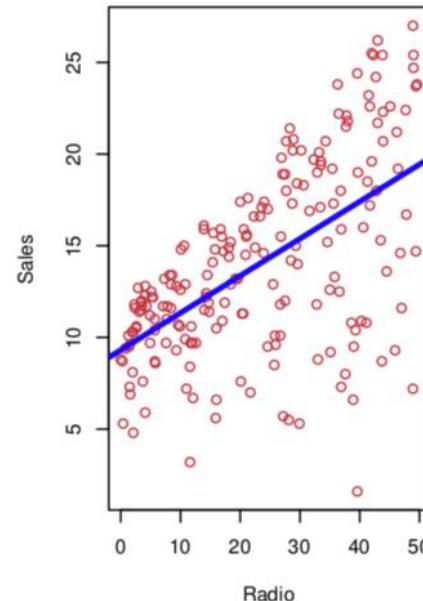
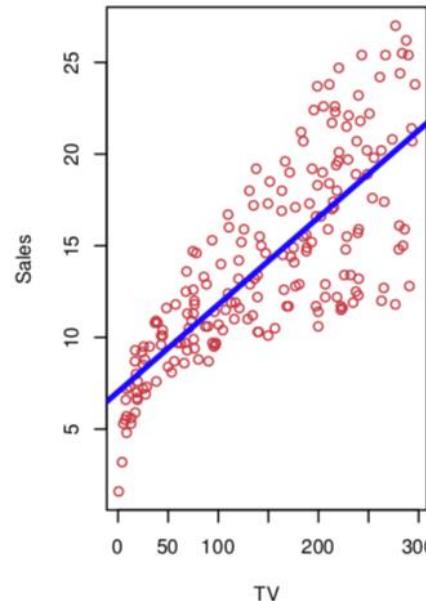
$$MSE = 7,125$$

$$\hat{y} = 3 + 0,5 X_1$$

$$\hat{y} = 3 + 0,5 \times 158$$

Regressão Linear

- Há alguma relação entre aumento de vendas e propaganda?
- Qual mídia contribui mais para as vendas?



Régressão Linear

- Abordagem supervisionada simples
- Assume uma dependência linear entre a variável resposta Y e os valores X_1, X_2, \dots, X_p
- Assume-se o modelo:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$$

Sendo $\beta_0, \beta_1, \dots, \beta_p$ coeficientes aprendidos pelo modelo

Régressão Linear

- Para o exemplo do slide 5, foi obtida a seguinte equação do hiperplano:

$$Vendas = 2,939 + 0,046 \times TV + 0,189 \times radio + 0.01 \times Jornal$$

Régressão Linear

- Para o exemplo do slide 5, foi obtida a seguinte equação do hiperplano:

$$Vendas = 2,939 + 0,046 \times TV + 0,189 \times radio + 0.01 \times Jornal$$

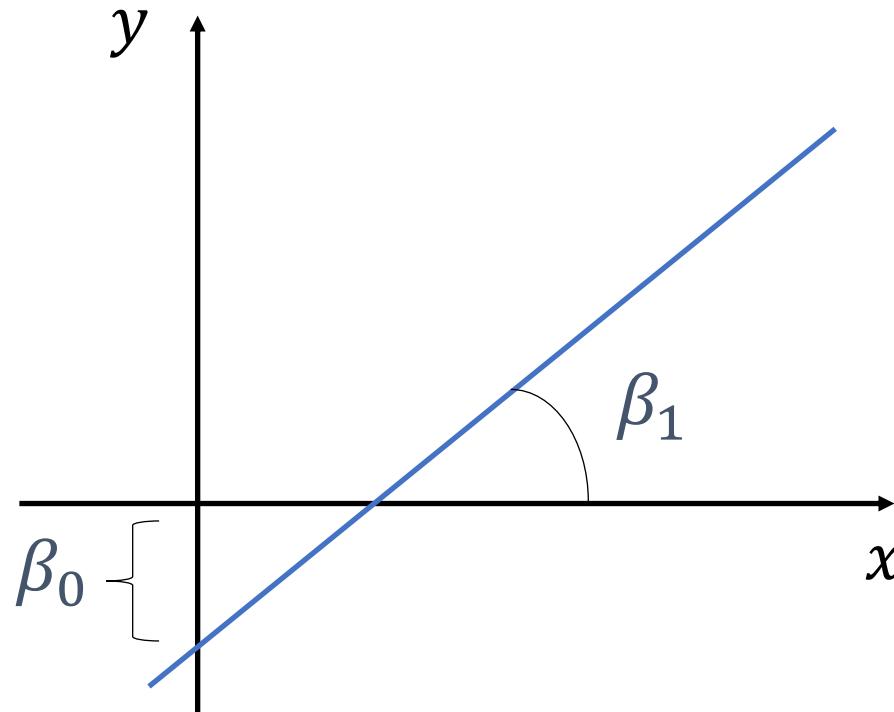
- caso nenhuma ação de propaganda seja feita as vendas serão de 2,939
- Mantendo todos os outros valores constantes, a cada uma unidade aumentada nas ações por TV, as vendas aumentam 0,046
- A influência da utilização de jornal é quase nula

Rregressão Linear

Reta

- β_0 : deslocamento
- β_1 : inclinação

$$y = \beta_0 + \beta_1 x$$

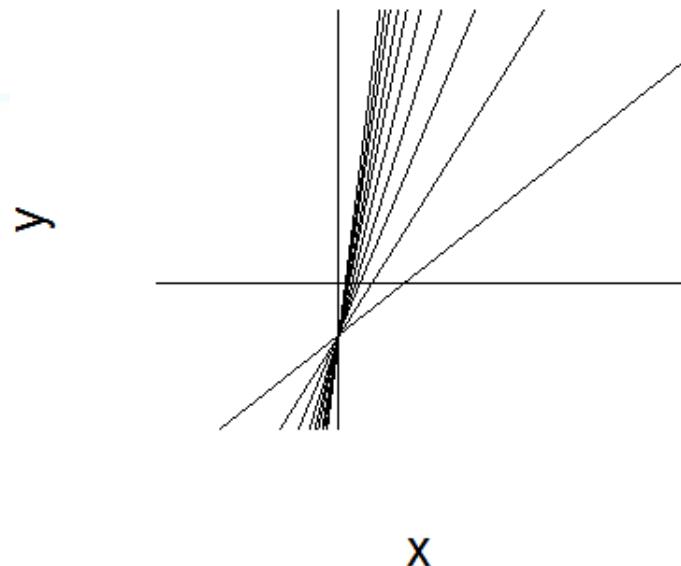


Regressão Linear

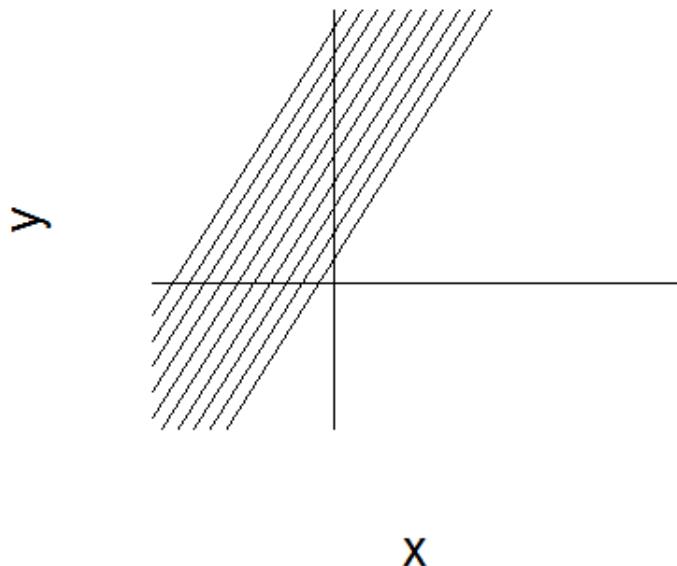
Reta

$$y = \beta_0 + \beta_1 x$$

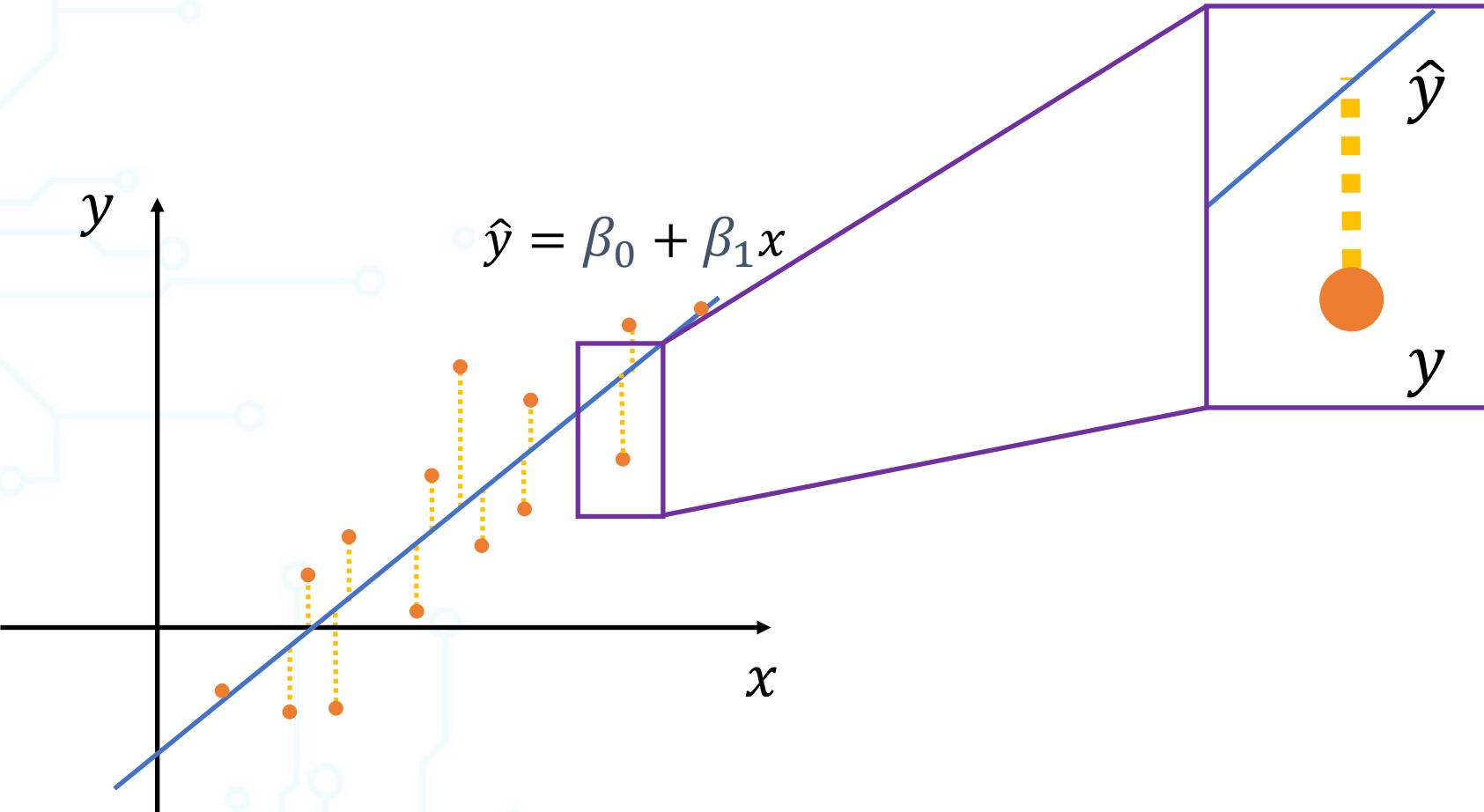
β_0 fixo; β_1 variável



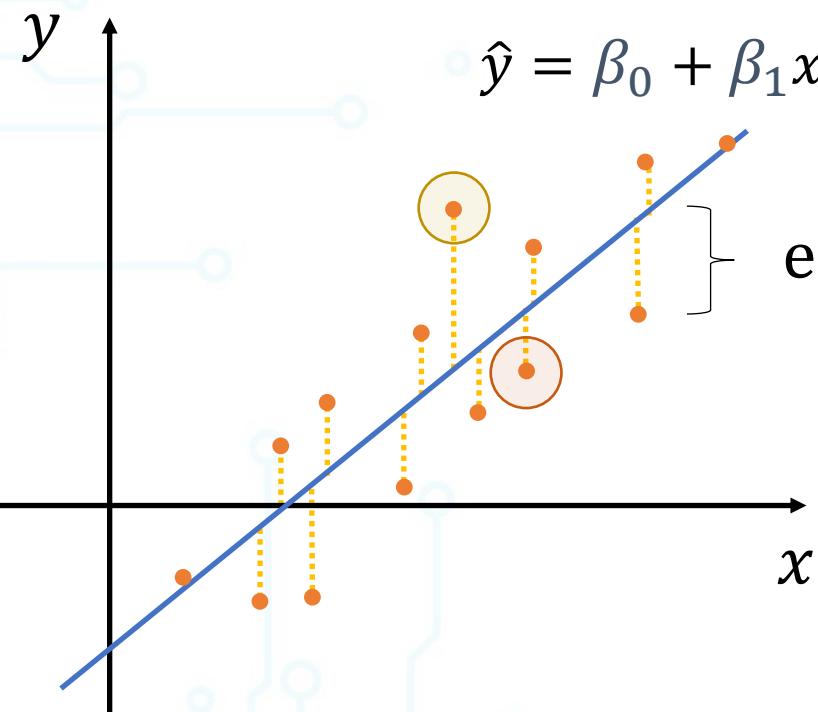
β_0 variável; β_1 fixo



Regressão Linear



Regressão Linear



$$e = y - \hat{y}$$

$$e = y - (\beta_0 + \beta_1 x)$$

Positivo

$$y - \hat{y} > 0$$

Negativo

$$y - \hat{y} < 0$$

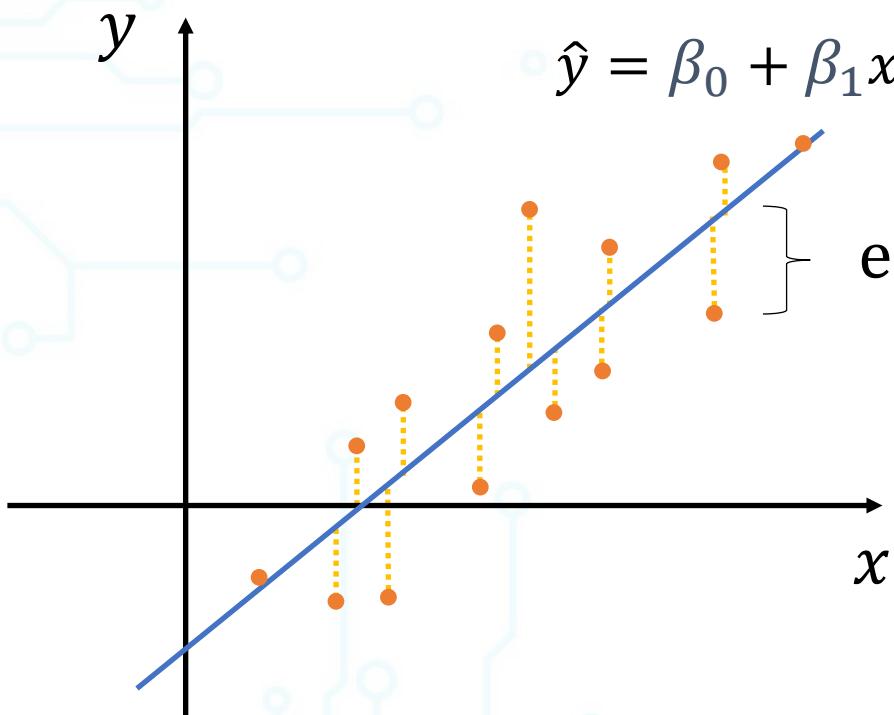
Solução?

$$(y - \hat{y})^2$$

$$|y - \hat{y}|$$



Regressão Linear



Como avaliar o erro total?

Erro quadrático
médio
R (MSE)

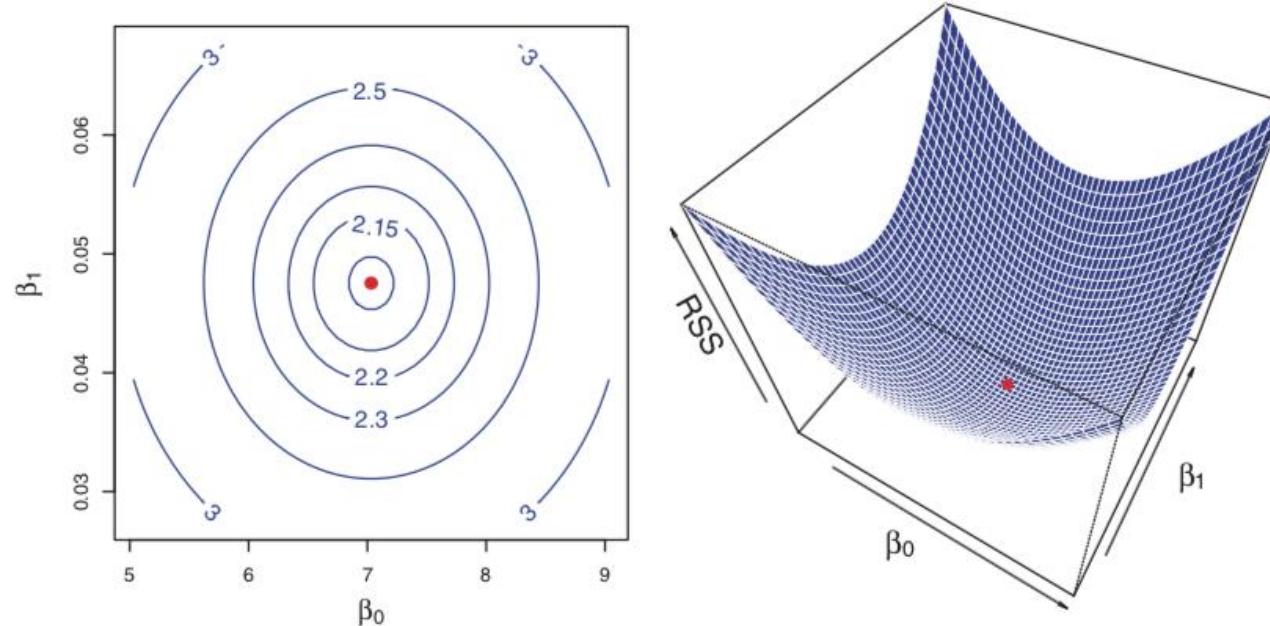
$$\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Erro absoluto
médio
(MAE)

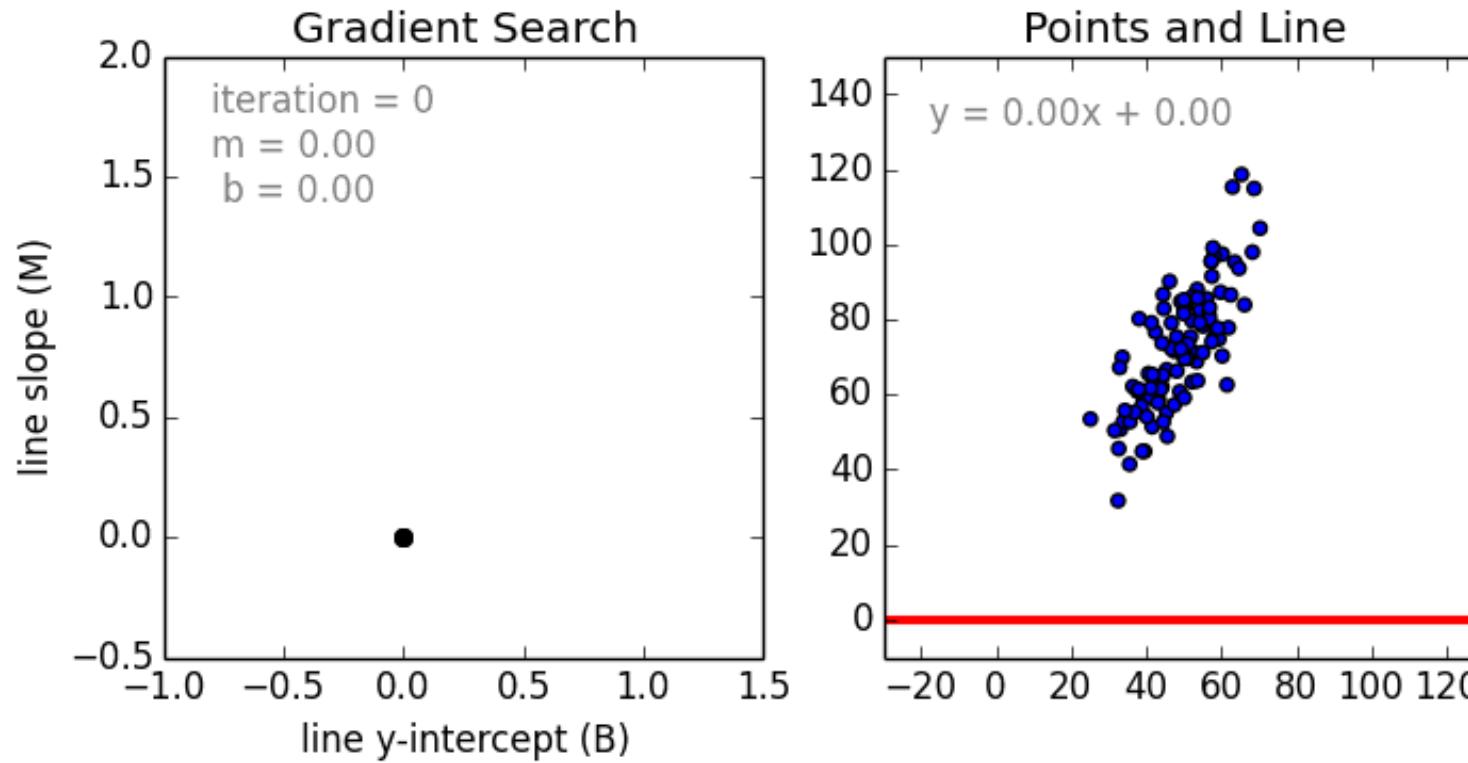
$$\frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Regressão Linear

- Como a função RMSE é convexa, é possível encontrar o valor mínimo por meio de algoritmos de otimização



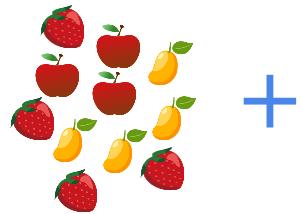
Regressão Linear



Types of machine learning algorithms | Supervised learning

Training

Labeled inputs



Training

Labels

Apple

Mango

Strawberry



Supervised machine learning algorithm

Trained model as output



$f(x)$

A trained ML model that maps inputs to outputs

Prediction

Inputs

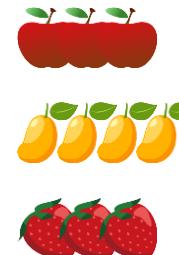


Processing



$f(x)$ trained model

Prediction



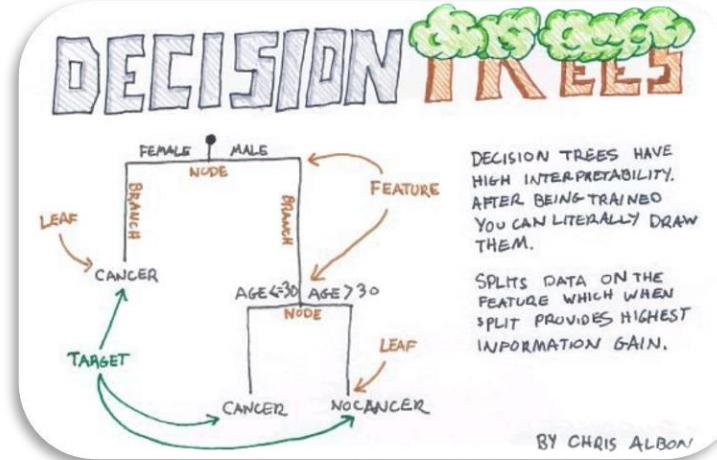
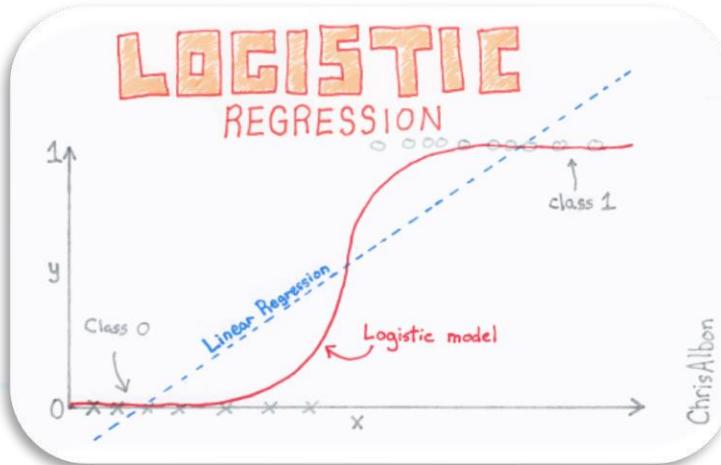
Labels

Apple

Mango

Strawberry

Classificação



GAUSSIAN NAIVE BAYES CLASSIFIER

"Gaussian" because this is a normal distribution

$$P(\text{class} | \text{data}) = \frac{P(\text{data} | \text{class}) \times p(\text{class})}{P(\text{data})}$$

We don't calculate this in naive bayes classifiers

This is our prior belief

Chris Albon

KNN NEIGHBORHOOD SIZE

Small \downarrow K = Low Bias, High Variance

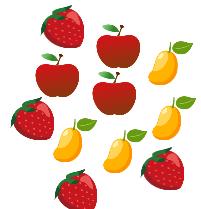
LARGE \uparrow K = High Bias, Low Variance

BY CHRIS ALBON

Types of machine learning algorithms | Unsupervised learning

Training

Labeled inputs



Training

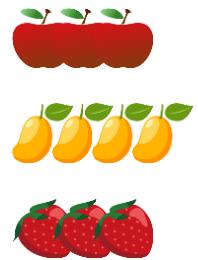


Unsupervised machine learning algorithm

Trained model as output



$f(x)$
Model: Learns clusters or patterns in data to detect anomalies



Prediction

Inputs



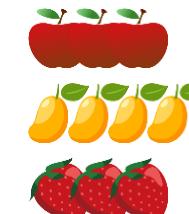
Processing



$f(x)$ trained model

Prediction

Known clusters

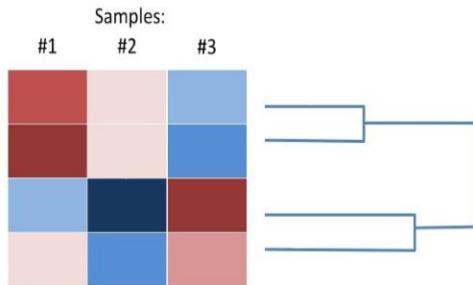


Anomalies



Agrupamento

Hierarchical Clustering, Clearly Explained!!!!



AGGLOMERATIVE CLUSTERING

All observations start as their own cluster. Clusters meeting some criteria are merged. This process is repeated, growing clusters until some end point is reached.

Chris Albon

k-MEANS CLUSTERING

1. k centerpoints are randomly initialized.
2. Observations are assigned to the closest centerpoint.
3. Centerpoints are moved to the center of their members.
4. Repeat steps 2 and 3 until no observation changes membership in step 2.

Chris Albon

DBSCAN

DBSCAN looks for densely packed observations and makes no assumptions about the number or shape of clusters.

1. A random observation, x_i , is selected
2. If x_i has a minimum of close neighbors, we consider it part of a cluster.
3. Step 2 is repeated recursively for all of x_i 's neighbors, then neighbors' neighbors etc... These are the cluster's core members.
4. Once Step 3 runs out of observations, a new random point is chosen

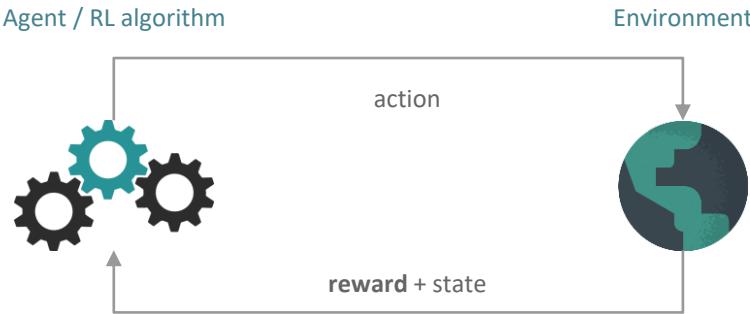
Afterwards, observations not part of a core are assigned to a nearby cluster or marked as outliers.

Chris Albon

Types of machine learning algorithms | Reinforcement learning

Training

Agent interacting with environment



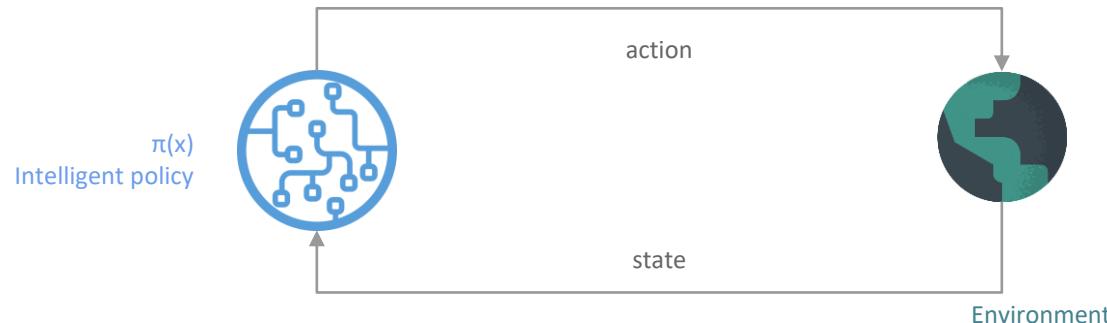
Intelligent Policy



$\pi(x)$
Intelligent policy that understands
environment

Policy in action

Trained agent / policy intelligently interacting with the environment



Thanks !



Vinicius Fernandes Caridá

vfcarida@gmail.com

