Artificial Unintelligence (Cap. 7)

ML & Titanic (Kaggle)

Prof. Sandra Avila

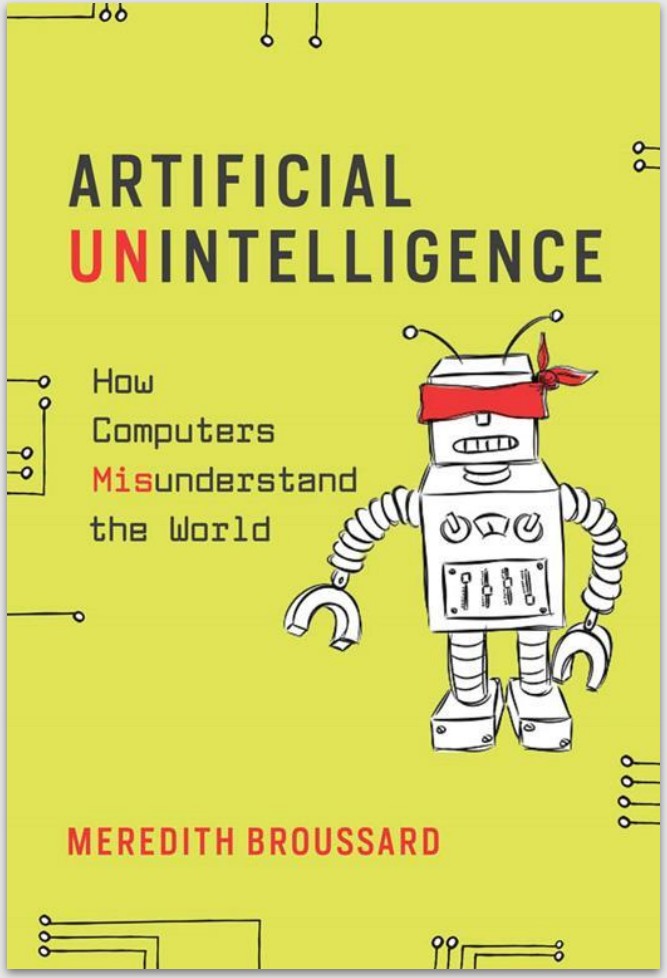
Institute of Computing (IC/Unicamp)

@sandraavilabr

MC886/MO444, August 23, 2022

Today's Agenda

- What is Machine Learning?
- Titanic Disaster



Book Chapter

7: Machine Learning: The DL on ML

Published: 2018

What is Machine Learning?

“Machine Learning is the science (and art) of programming computers so they can learn from data”.

[Aurélien Géron, 2019]

“Computing the capacity of a computer to learn from experience, i.e. to modify its processing on the basis of newly acquired information”.

[Oxford English Dictionary, 2000]

“A branch of artificial intelligence concerned with the construction of programs that learn from experience.”

[Oxford’s A Dictionary of Computing, 2008]

“Machine learning is about learning some properties of data set and applying them to new data.”

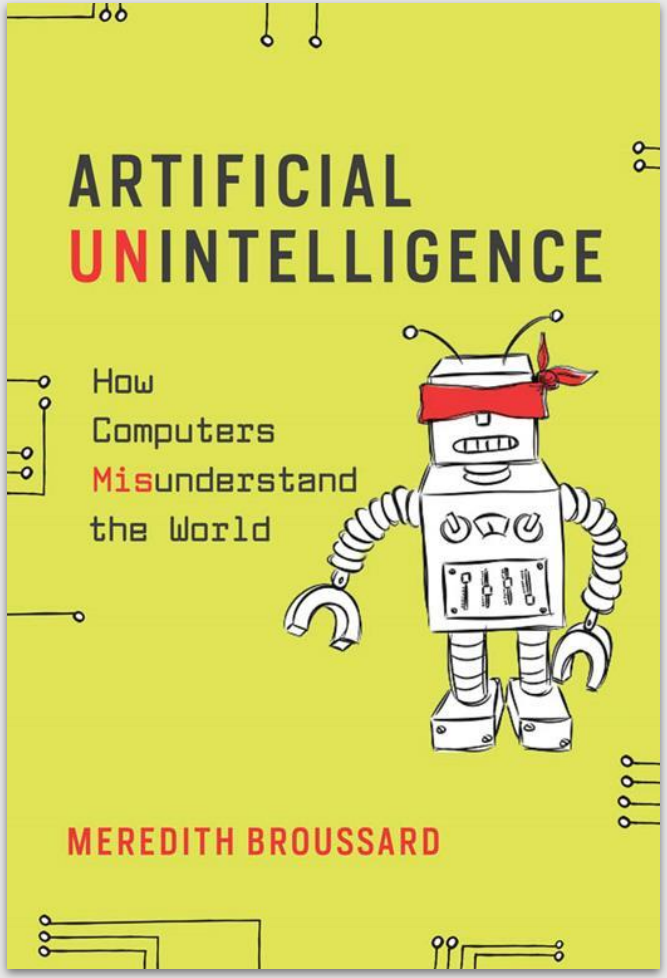
"We say that a machine learns with respect to a particular task T, performance metric P, and type of experience E, if the systems reliably improves its performance P at task T, following experience E."

[Tom M. Mitchell, 1997]

When a machine “learns”, it means that the machine has become more accurate at performing a single, specific task according to a specific metric that a person has defined.

Today's Agenda

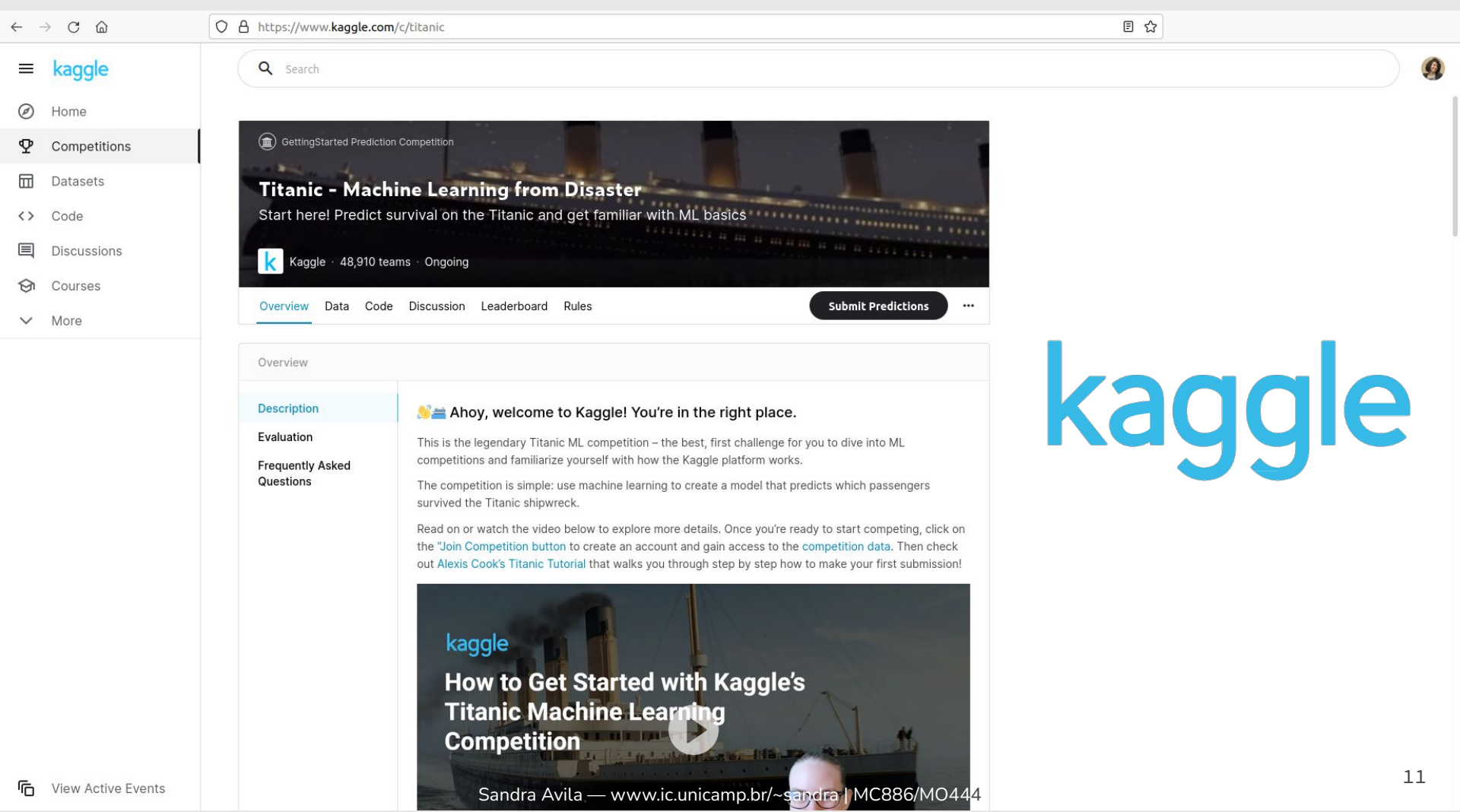
- What is Machine Learning?
- Titanic Disaster



Book Chapter

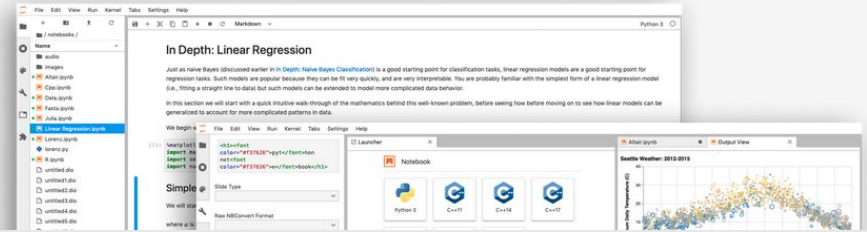
7: Machine Learning: The DL on ML

Published: 2018


https://www.kaggle.com/c/titanic
kaggle
Search
Home
Competitions
Datasets
Code
Discussions
Courses
More
View Active Events
GettingStarted Prediction Competition
Titanic - Machine Learning from Disaster
Start here! Predict survival on the Titanic and get familiar with ML basics
Kaggle · 48,910 teams · Ongoing
Overview
Data
Code
Discussion
Leaderboard
Rules
Submit Predictions
Overview
Description
Evaluation
Frequently Asked Questions
Ahoy, welcome to Kaggle! You're in the right place.
This is the legendary Titanic ML competition – the best, first challenge for you to dive into ML competitions and familiarize yourself with how the Kaggle platform works.
The competition is simple: use machine learning to create a model that predicts which passengers survived the Titanic shipwreck.
Read on or watch the video below to explore more details. Once you're ready to start competing, click on the "Join Competition button to create an account and gain access to the competition data. Then check out Alexis Cook's Titanic Tutorial that walks you through step by step how to make your first submission!
kaggle
How to Get Started with Kaggle's Titanic Machine Learning Competition

kaggle

Project Jupyter exists to develop open-source software, open-standards, and services for interactive computing across dozens of programming languages.



JupyterLab: Jupyter's Next-Generation Notebook Interface

JupyterLab is a web-based interactive development environment for Jupyter notebooks, code, and data. JupyterLab is flexible: configure and arrange the user interface to support a wide range of workflows in data science, scientific computing, and machine learning. JupyterLab is extensible and modular: write plugins that add new components and integrate with existing ones.

Welcome To Colaboratory

File Edit View Insert Runtime Tools Help

Share

+ Code + Text Copy to Drive

Connect Editing

What is Colaboratory?

Colaboratory, or "Colab" for short, allows you to write and execute Python in your browser, with

- Zero configuration required
- Free access to GPUs
- Easy sharing

Whether you're a **student**, a **data scientist** or an **AI researcher**, Colab can make your work easier. Watch Introduction to Colab to learn more, or just get started below!

https://colab.research.google.com

Getting started

The document you are reading is not a static web page, but an interactive environment called a **Colab notebook** that lets you write and execute code.

For example, here is a **code cell** with a short Python script that computes a value, stores it in a variable, and prints the result:

```
[ ] seconds_in_a_day = 24 * 60 * 60
    seconds_in_a_day
```

```
86400
```

To execute the code in the above cell, select it with a click and then either press the play button to the left of the code, or use the keyboard shortcut "Command/Ctrl+Enter". To edit the code, just click the cell and start editing.

Variables that you define in one cell can later be used in other cells:

```
[ ] seconds_in_a_week = 7 * seconds_in_a_day
    seconds_in_a_week
```

```
604800
```

https://youtu.be/Jq4aKxaoLGM

intel progra{m}aria

MÃO NA MASSA COM MACHINE LEARNING

PRIMEIROS PASSOS COM MACHINE LEARNING

-

MÓDULO 03

FERNANDA WANDERLEY

CIENTISTA DE DADOS

NEURALMED

Sprint >>> progra{m}aria

https://www.fast.ai/2021/08/16/eleven-videos

11 Short Videos About AI Ethics

Written: 16 Aug 2021 by *Rachel Thomas*

I made a playlist of 11 short videos (most are 6-13 mins long) on Ethics in Machine Learning. This is from my ethics lecture in Practical Deep Learning for Coders v4. I thought these short videos would be easier to watch, share, or skip around.

What are Ethics and Why do they Matter? Machine Learning Edition: Through 3 key case studies, I cover how people can be harmed by machine learning gone wrong, why we as machine learning practitioners should care, and what tech ethics are.

All machine learning systems need ways to identify & address mistakes. It is crucial that all machine learning systems are implemented with ways to correctly surface and correct mistakes, and to provide recourse to those harmed.

The Problem with Metrics, Feedback Loops, and Hypergrowth: Overreliance on metrics is a core problem both in the field of machine learning and in the tech industry more broadly. As Goodhart's Law tells us, when a measure becomes the target, it ceases to be a good measure, yet the incentives of venture capital push companies in this direction. We see out-of-control feedback loops, widespread gaming of metrics, and

© fast.ai 2021. All rights reserved.