

# Parallel session High Tech



# From Visual Inspection To Causal Discovery

April 2022, Nijmegen

*Ali Bahramisharif*



# OUTLINE

---

INTRODUCTION

CAUSALITY AND EXPLAINABILITY

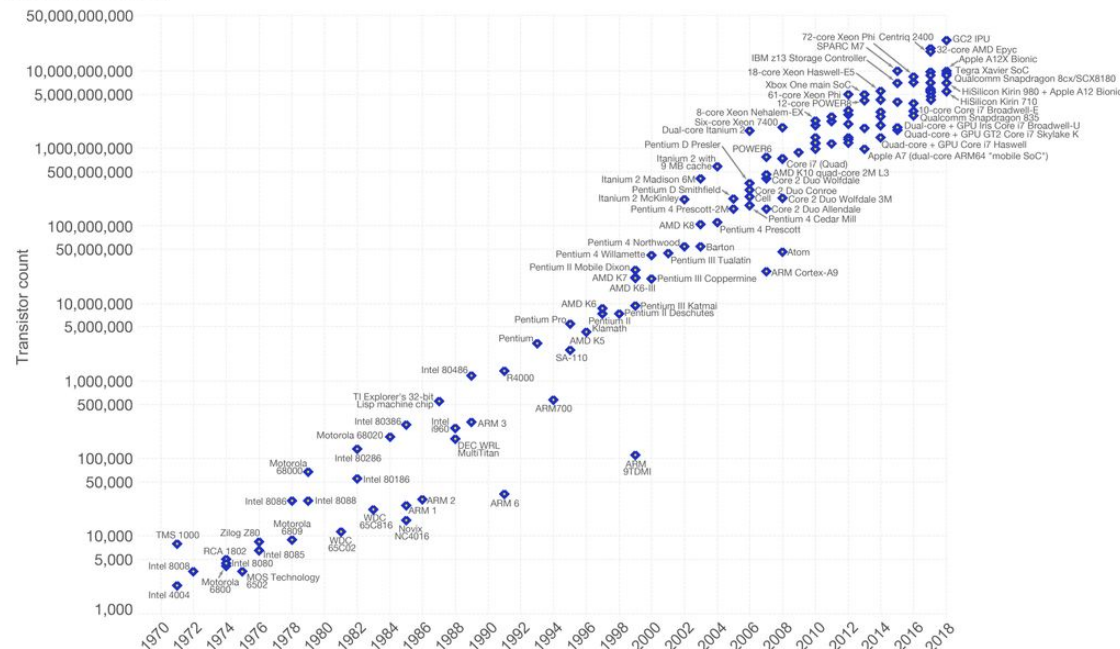
VISION AND INSPECTION

COOL STUFF

CONCLUSION

## Moore's Law – The number of transistors on integrated circuit chips (1971-2018)

Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important as other aspects of technological progress – such as processing speed or the price of electronic products – are linked to Moore's law.



Data source: Wikipedia ([https://en.wikipedia.org/wiki/Transistor\\_count](https://en.wikipedia.org/wiki/Transistor_count))  
The data visualization is available at OurWorldinData.org. There you find more visualizations and research on this topic.

Licensed under CC-BY-SA by the author Max Roser.

The number of transistors in a dense integrated circuit doubles about every two years.



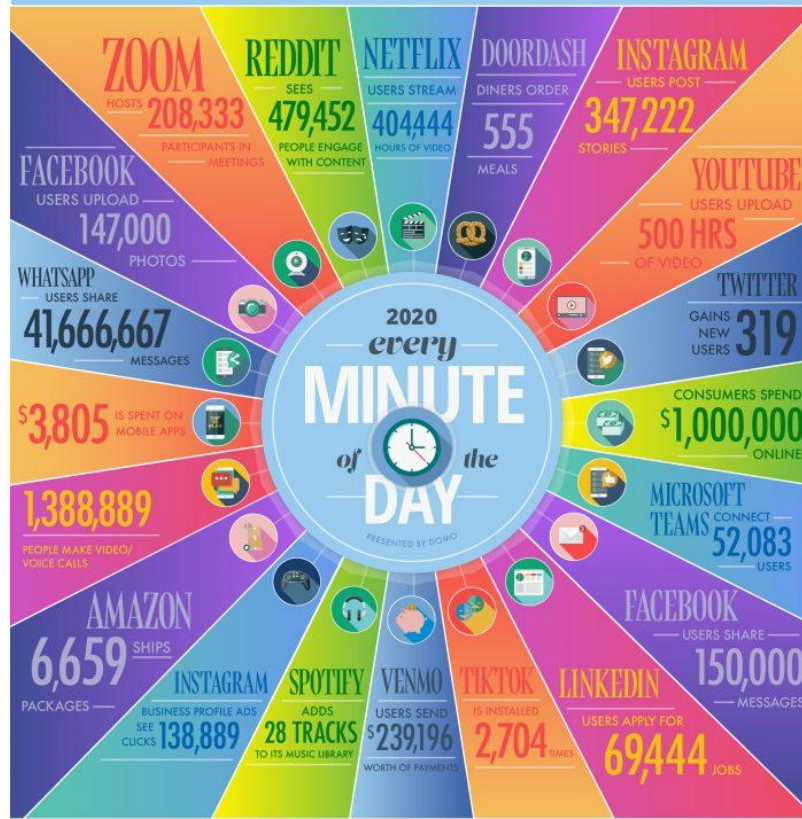
# DATA NEVER SLEEPS 8.0

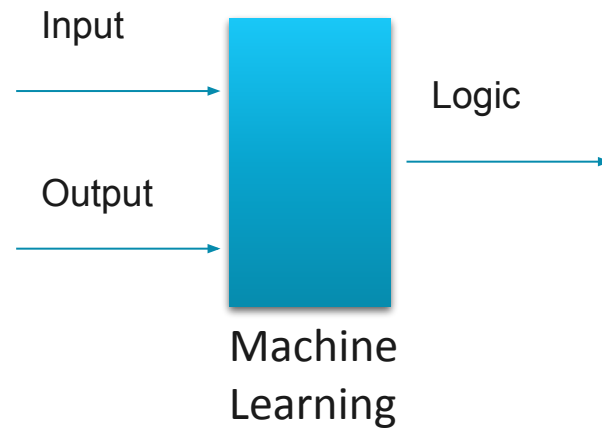
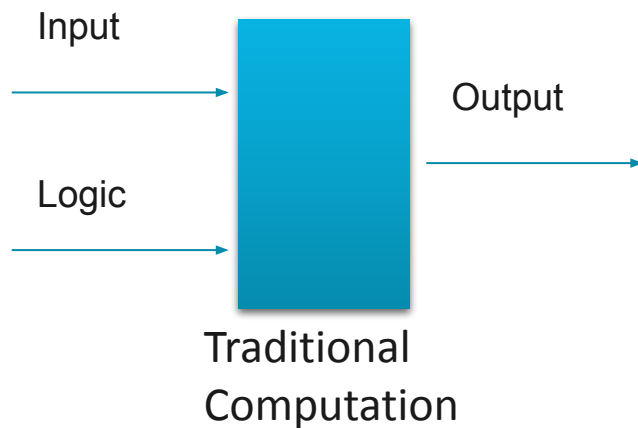
How much data is generated *every minute*?

In 2020, the world changed fundamentally—and so did the data that makes the world go round. As COVID-19 swept the globe, nearly every aspect of life—from work to working out—moved online, and people depended more and more on apps and the Internet to socialize, educate and entertain ourselves. Before quarantine, just 15% of Americans worked from home. Now over half do. And that's not the only big shift. In our 8th edition of Data Never Sleeps, we bring you the latest stats on how much data is being created in every digital minute—a trend that shows no sign of stopping.



MACHINE2LEARN

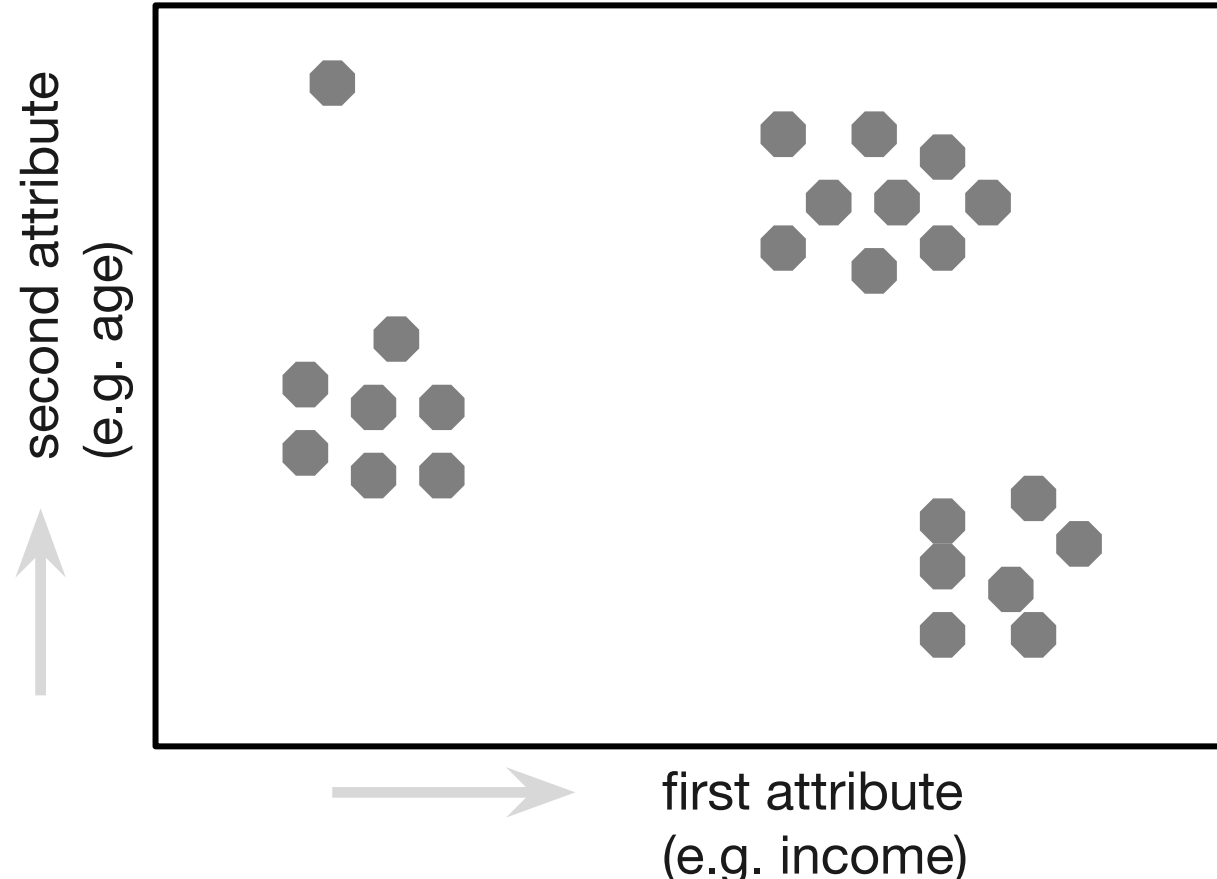




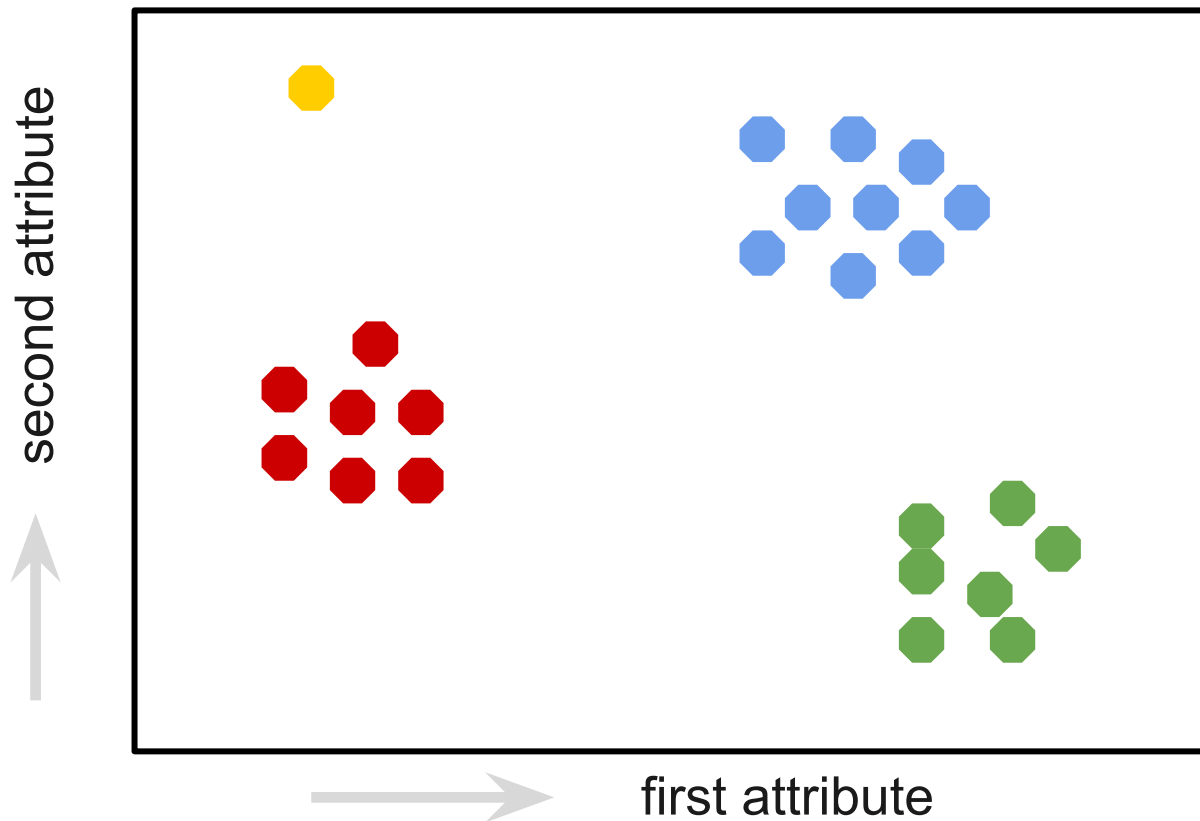
## Attributes

## Objects

<i>Tid</i>	Refund	Marital Status	Taxable Income	Cheat
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes









THIS IS YOUR MACHINE LEARNING SYSTEM?

YUP! YOU POUR THE DATA INTO THIS BIG PILE OF LINEAR ALGEBRA, THEN COLLECT THE ANSWERS ON THE OTHER SIDE.

WHAT IF THE ANSWERS ARE WRONG?

JUST STIR THE PILE UNTIL THEY START LOOKING RIGHT.



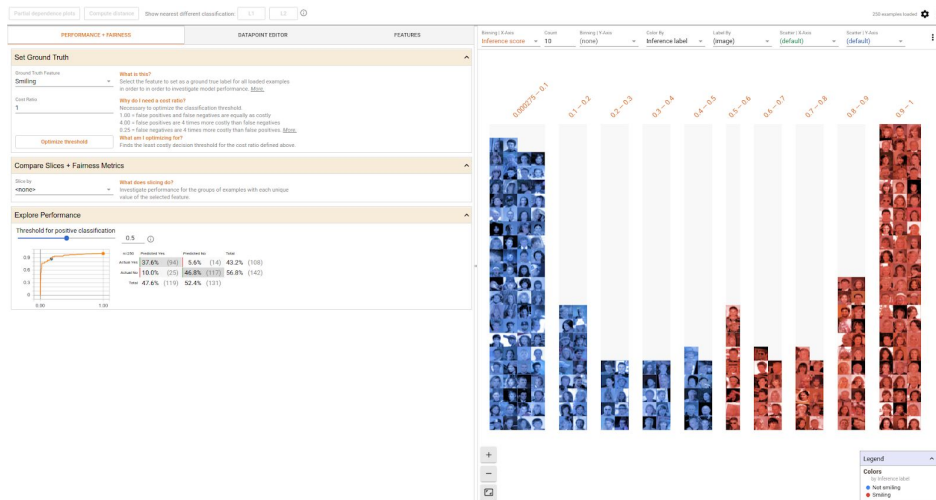
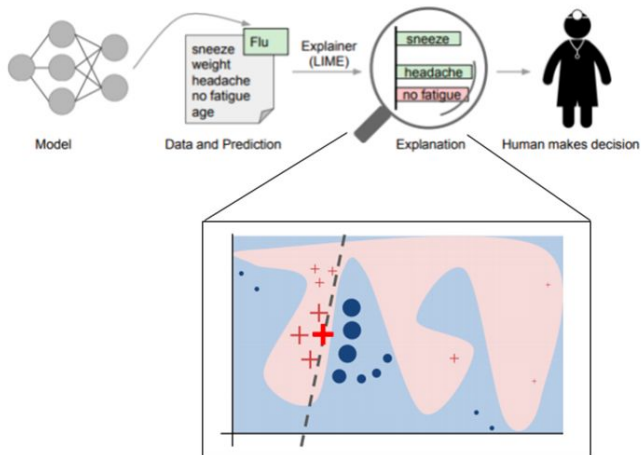
# EXPLAINABILITY

## “Why Should I Trust You?” Explaining the Predictions of Any Classifier

Marco Tulio Ribeiro  
University of Washington  
Seattle, WA 98105, USA  
marcotcr@cs.uw.edu

Sameer Singh  
University of Washington  
Seattle, WA 98105, USA  
sameer@cs.uw.edu

Carlos Guestrin  
University of Washington  
Seattle, WA 98105, USA  
guestrin@cs.uw.edu



## Covariates

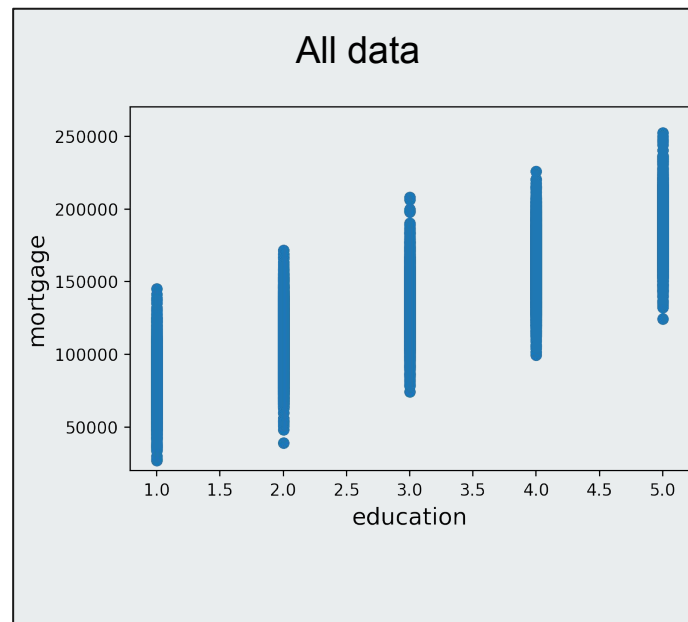
Marital  
Status

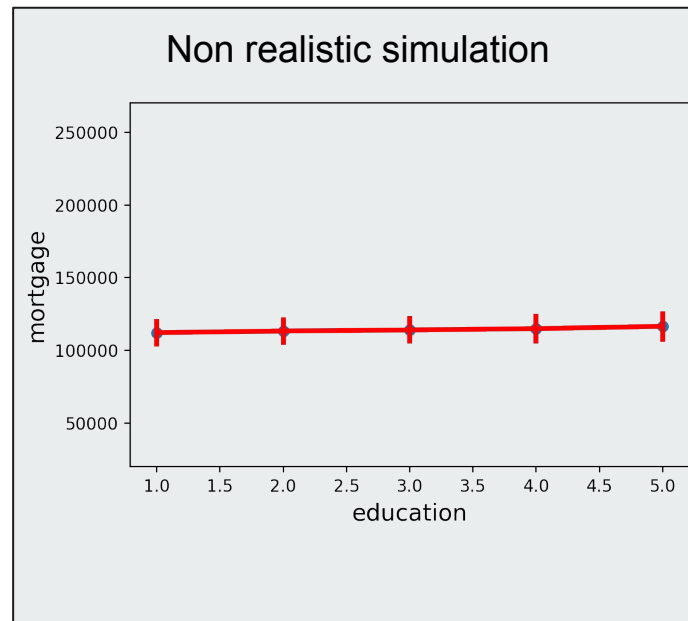
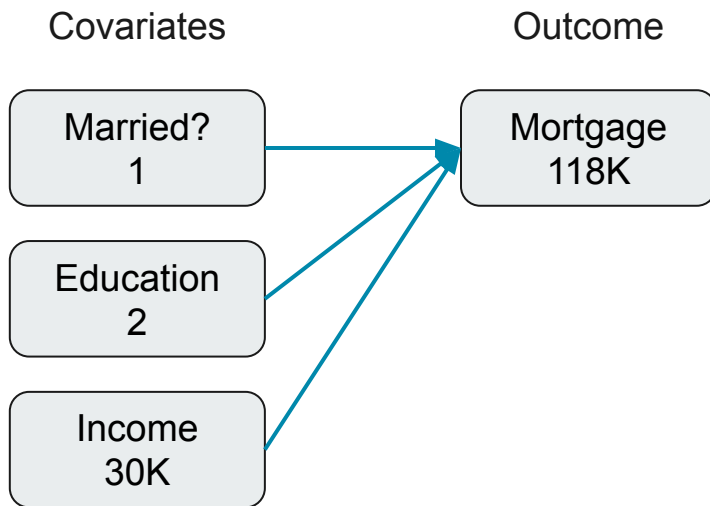
Education

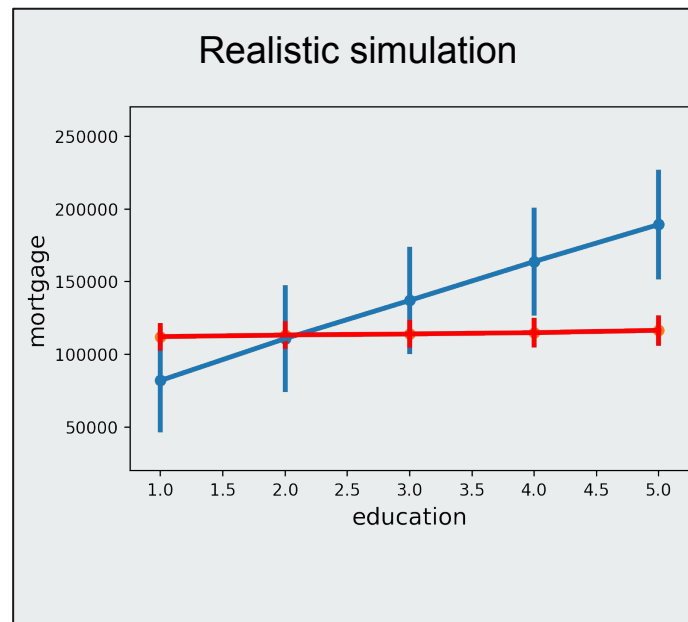
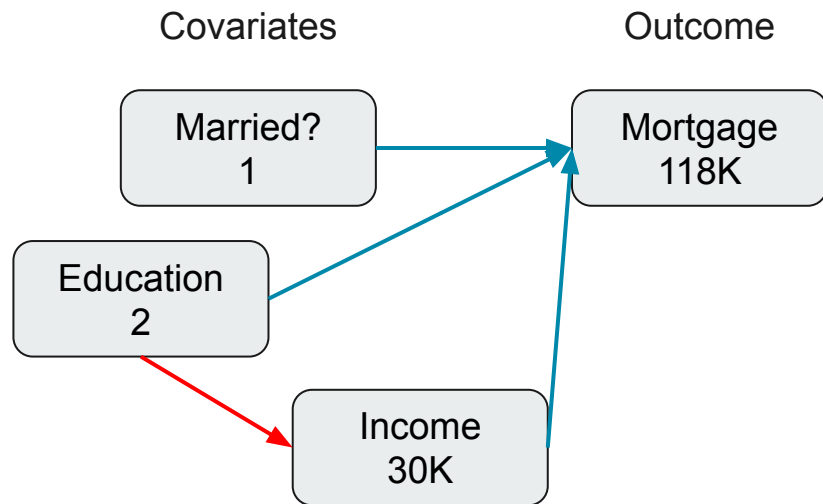
Income

## Outcome

Mortgage  
Amount







## Explain

Features values

Feature	Value
x1	-5
x2	2.01593
x3	2.21147

Showing 1 to 3 of 3 entries

Prediction probabilities



Explain graphs



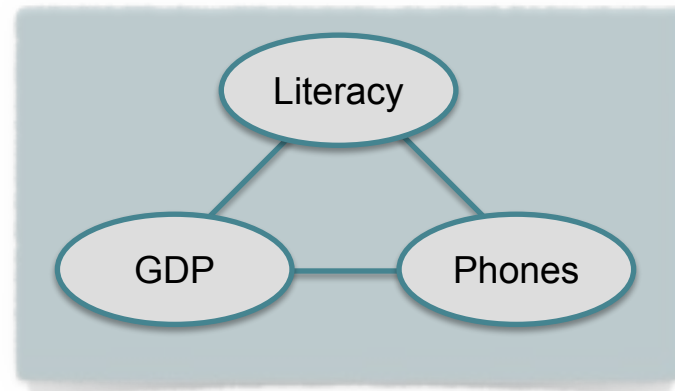
All rights reserved. No part of this document may be reproduced in any material form without the written permission of Machine2Learn.

# CAUSAL DISCOVERY

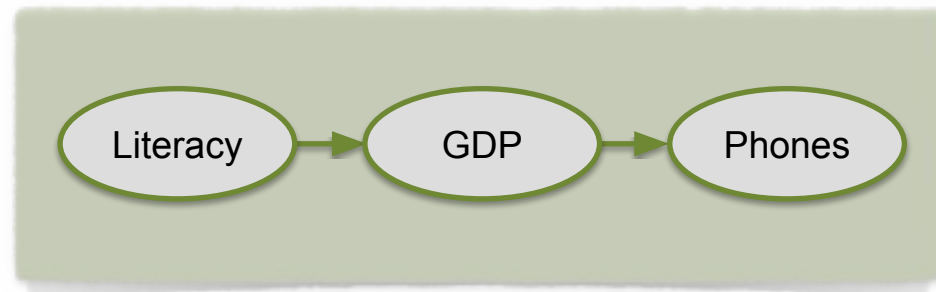


Literacy (%)	GDP (\$ per capita)	Phones (per 1000)
36.0	700.0	3.2
86.5	4500.0	71.2
70.0	6000.0	78.1
97.0	8000.0	259.5
100.0	19000.0	497.2
42.0	1900.0	7.8
95.0	8600.0	460.0
89.0	11000.0	549.9
97.1	11200.0	220.4
98.6	3500.0	195.7
97.0	28000.0	516.1
100.0	29000.0	565.5
98.0	30000.0	452.2
97.0	3400.0	137.1
95.6	16700.0	460.6
89.1	16900.0	281.3
43.1	1900.0	7.3
97.4	15700.0	481.9
99.6	6100.0	319.1
98.0	29100.0	462.6
94.1	4900.0	115.7
40.9	1100.0	9.7
98.0	36000.0	851.4
42.2	1300.0	14.3
87.2	2400.0	71.9
NaN	6100.0	215.4
79.8	9000.0	80.5
86.4	7600.0	225.3
97.8	16000.0	506.5
93.9	18600.0	237.2
...	...	...

Regular  
statistics



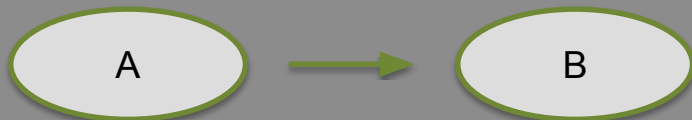
Causal  
discovery





# CAUSAL RELATIONS

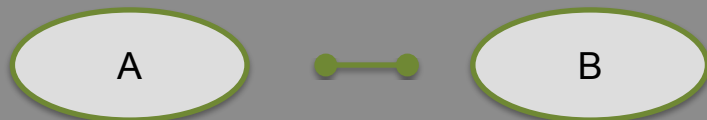
**A causes B**



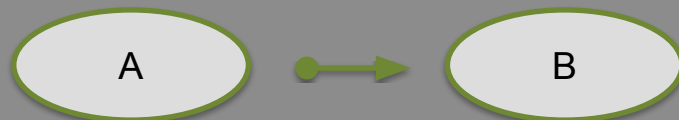
**Selection bias**



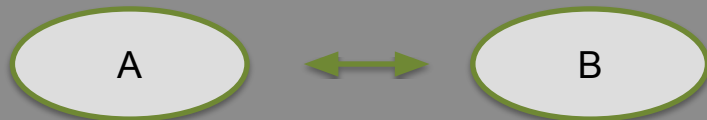
**Correlation between A and B**



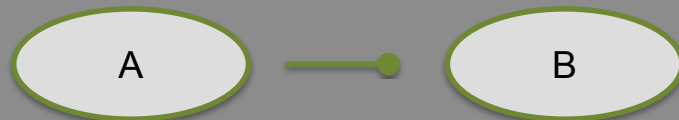
**B does not cause A**

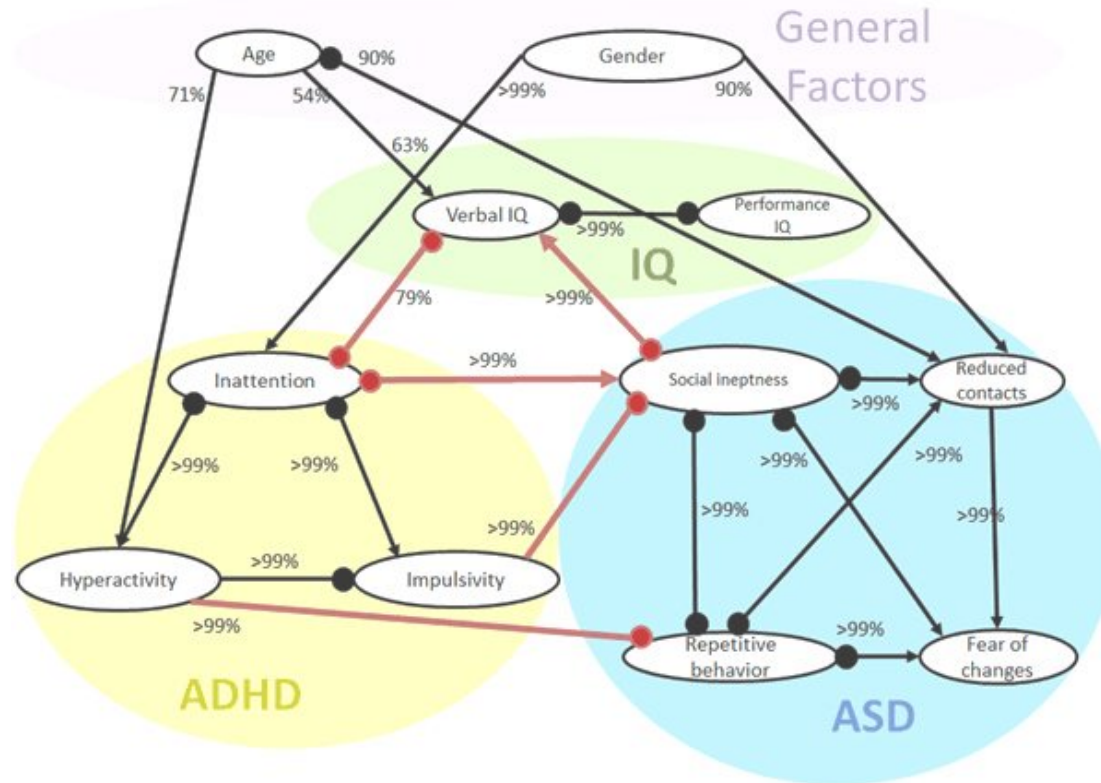


**Hidden common cause**



**A causes B / selection bias**



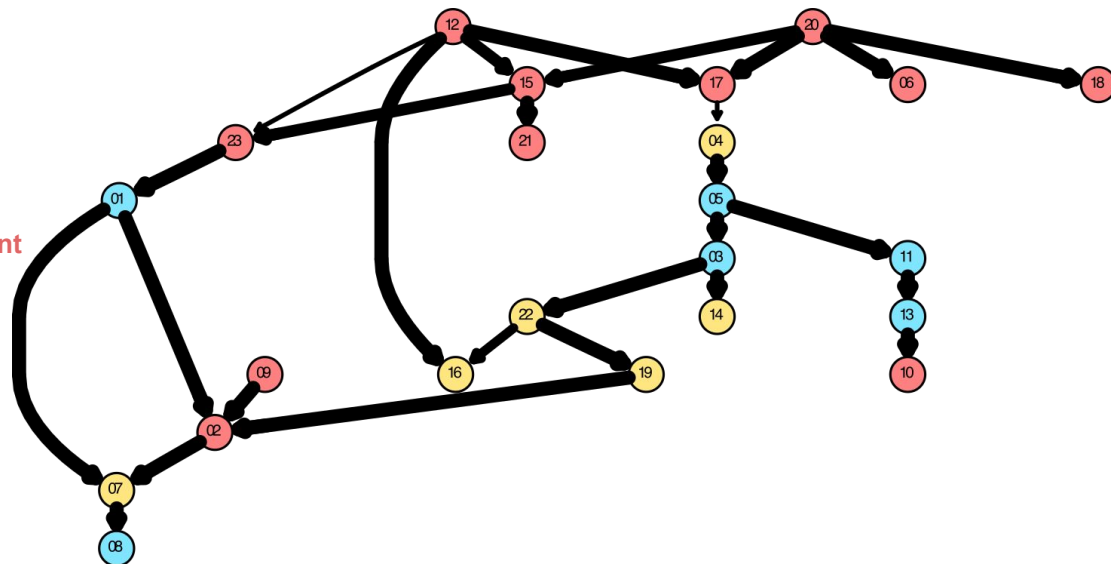


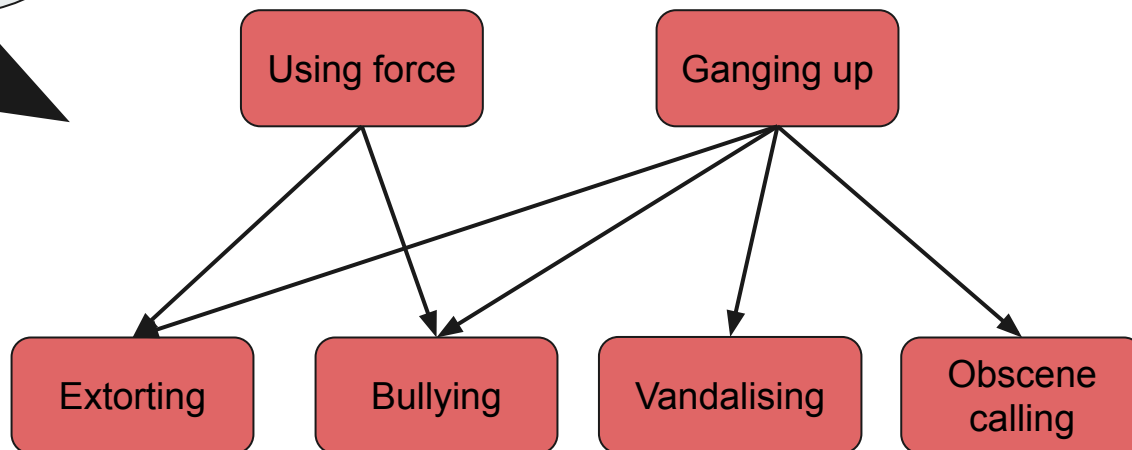
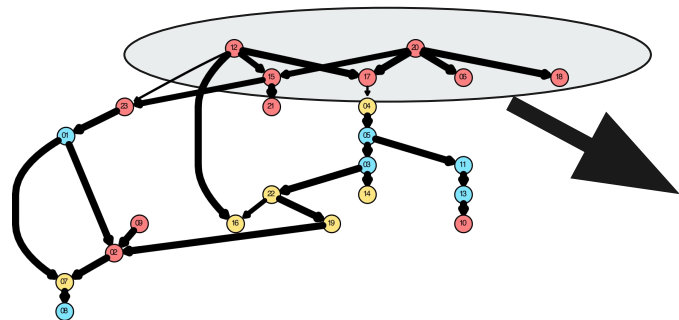


# QUESTIONNAIRE



1. Yelled at others when they have annoyed you
2. Had fights with others to show who was on top
3. Reacted angrily when provoked by others
4. Taken things from other students
5. Gotten angry when frustrated
6. Vandalized something for fun
7. Had temper tantrums
8. Damaged things because you felt mad
9. Had a gang fight to be cool
10. Hurt others to win a game
11. Become angry or mad when you don't get your way
12. Used physical force to get others to do what you want
13. Gotten angry or mad when you lost a game
14. Gotten angry when others threatened you
15. Used force to obtain money or things from others
16. Felt better after hitting or yelling at someone
17. Threatened and bullied someone
18. Made obscene phone calls for fun
19. Hit others to defend yourself
20. Gotten others to gang up on someone else
21. Carried a weapon to use in a fight
22. Gotten angry or mad or hit others when teased
23. Yelled at others so they would do things for you





# VISION AND INSPECTION

# M2L IMPLEMENTATION CYCLE

Identify and React

M2L Studio Classify



Learn and Optimise

M2L Studio Customisation



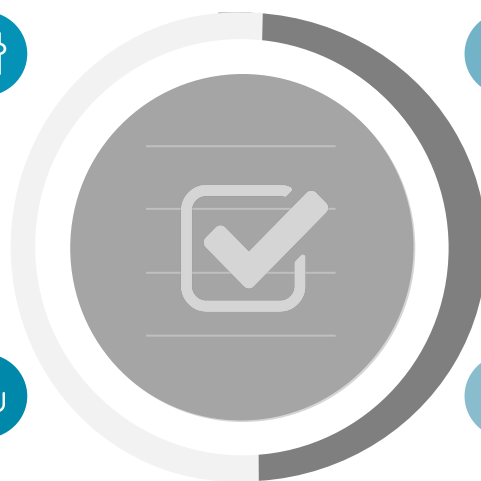
Detect and Collect

M2L Studio Annotation



Support

M2L Team



# AKZONOBEL



Problem Description

Current State



# PROBLEM DESCRIPTION



MACHINE2LEARN



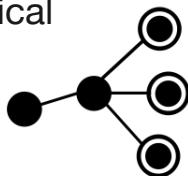
Every year, new cars with **new colors** are introduced.



AkzoNobel provides the **paint to repair** these cars when drivers have an accident.



Up until recent years, all **prediction** calculations were based on physical models.

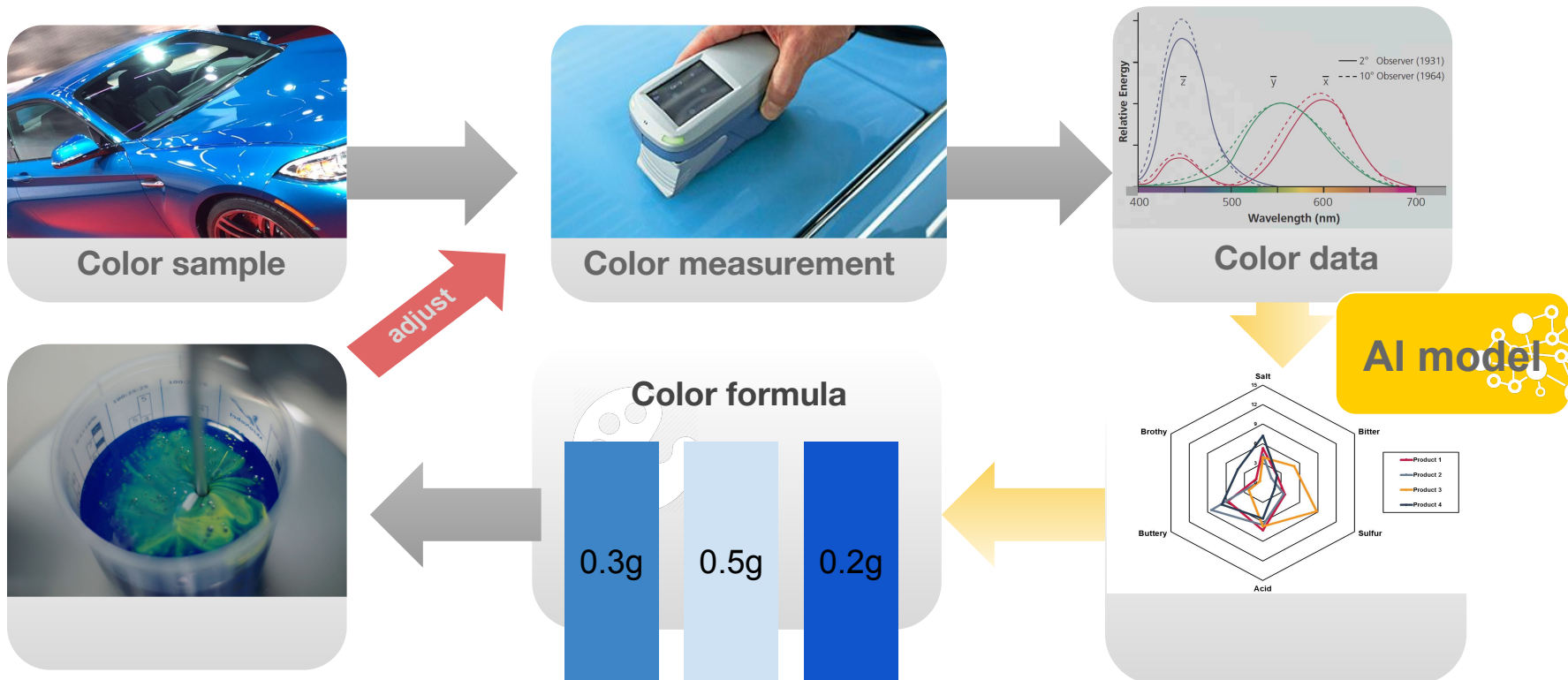


In the past, it sometimes took up to **two years to get a car color available** for the market.



All rights reserved. No part of this document may be reproduced in any material form without the written permission of Machine2Learn.

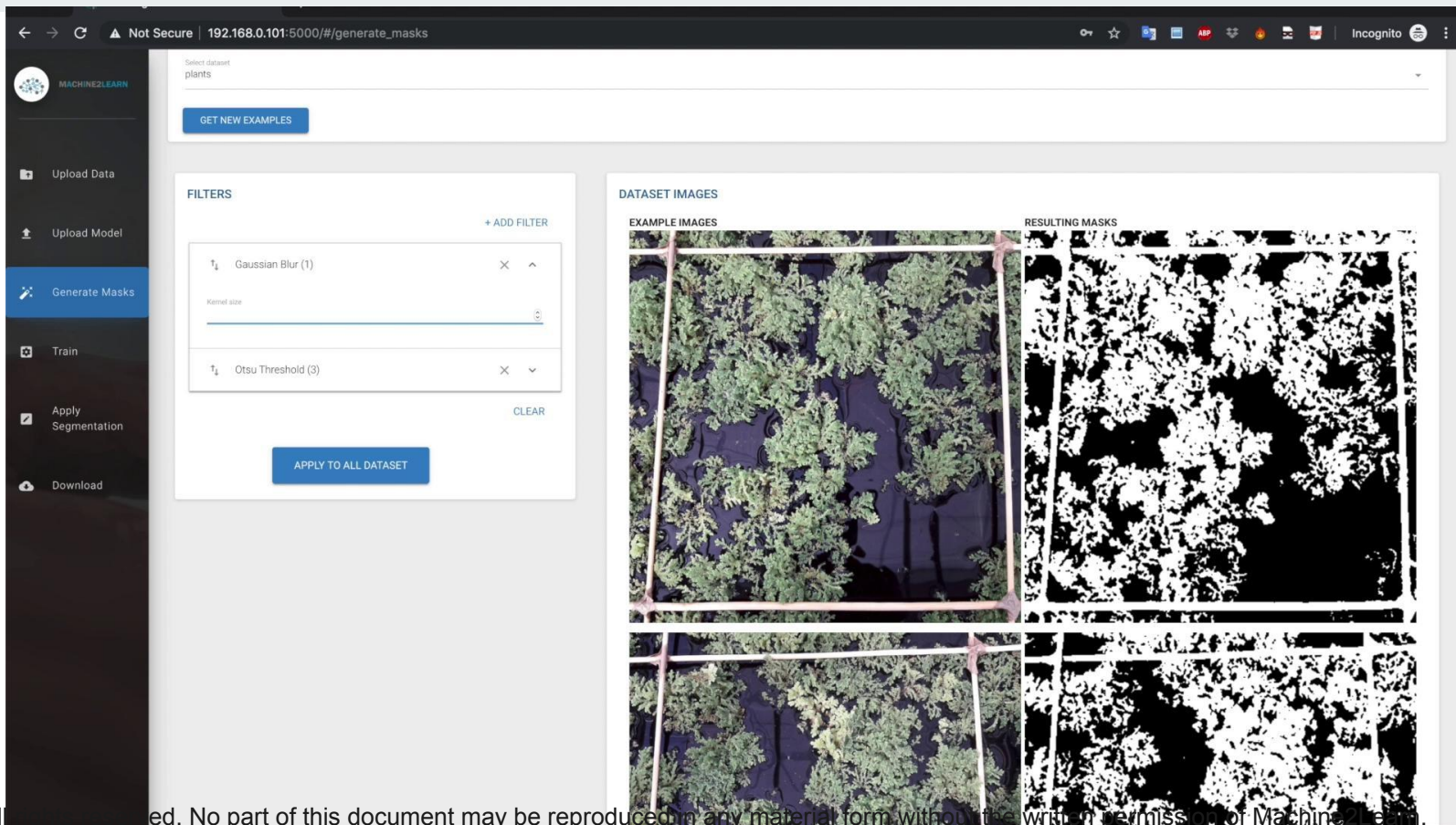
# CURRENT STATE



All rights reserved. No part of this document may be reproduced in any material form without the written permission of Machine2Learn.

# COOL STUFF

# GREEN SEGMENTATION



The screenshot displays the Machine2Learn web application interface for generating segmentation masks. The browser address bar shows the URL `192.168.0.101:5000/#/generate_masks`. The interface includes a sidebar with navigation options: Upload Data, Upload Model, Generate Masks (highlighted), Train, Apply Segmentation, and Download. The main content area is divided into several sections:

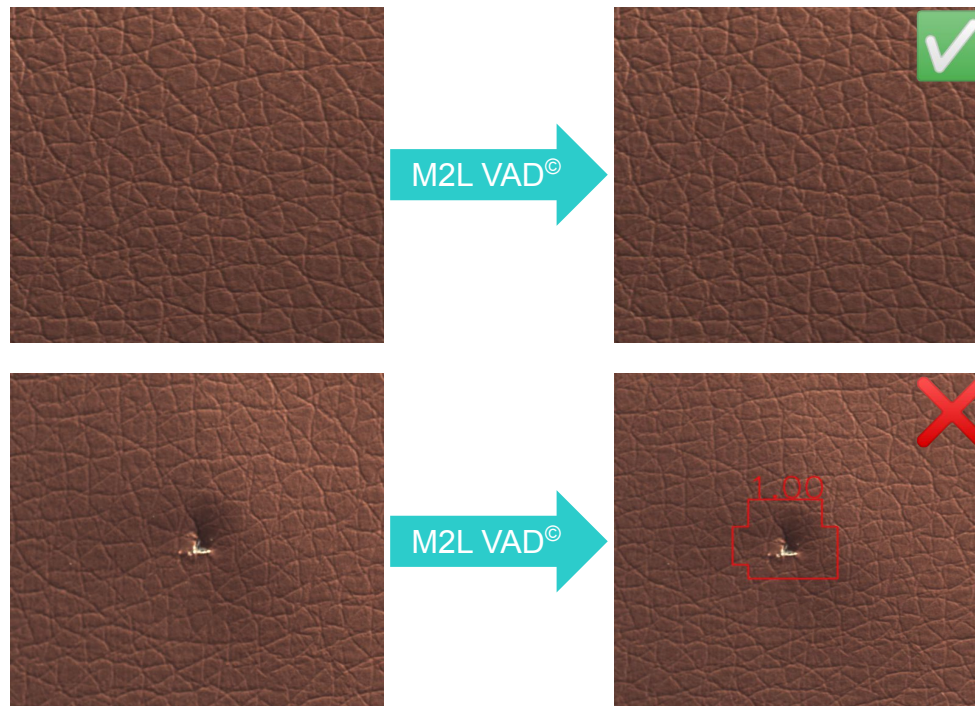
- Select dataset:** A dropdown menu currently shows 'plants', with a 'GET NEW EXAMPLES' button below it.
- FILTERS:** A panel on the left containing two active filters: 'Gaussian Blur (1)' and 'Otsu Threshold (3)'. It includes a 'Kernal size' slider, a 'CLEAR' button, and an 'APPLY TO ALL DATASET' button.
- DATASET IMAGES:** A section titled 'EXAMPLE IMAGES' showing a grid of aerial photos of a forest.
- RESULTING MASKS:** A section showing the corresponding binary segmentation masks for the example images, where green areas are white and other areas are black.

All rights reserved. No part of this document may be reproduced in any material form without the written permission of Machine2Learn.

# VISUAL ANOMALY DETECTION




MACHINE2LEARN



<http://anomaly.machine2learn.com>

# ANOMALY PREVENTION


The screenshot displays the CelSian 'Gates & Solar' interface for anomaly prevention. It is divided into two main steps on the left sidebar: 'STEP 1: Data' and 'STEP 2: Model'.  
**STEP 1: Data**  
Instructions: 'Select a dataset available and click Load to start.'  
A dropdown menu shows 'data.csv' as the selected dataset.  
A blue 'LOAD' button is present.  
**STEP 2: Model**  
Instructions: 'Select the columns of the list to feed your model and click Train to start training.'  
Five sensors are listed with checkboxes:  
- sensor 1: ☒  
- sensor 2: ☒  
- sensor 3: ☐  
- sensor 4: ☒  
- sensor 5: ☐  
An orange 'TRAIN' button is at the bottom of the sidebar.  
In the center of the main workspace, a green box indicates 'Training ... please wait' with a progress bar that is approximately 75% full.  
The footer of the interface shows '©2019 — MACHINE2LEARN'.




MACHINE2LEARN

## Tabular Data Generatator

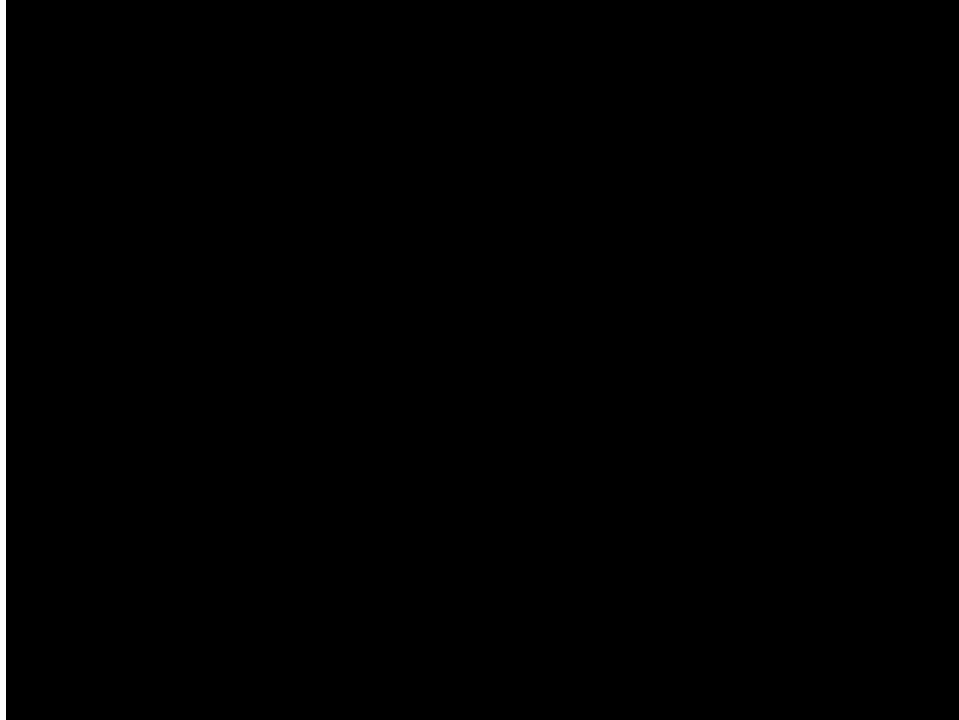
Upload your .csv file, train a model and  
generate new samples!

 Upload data (.csv)

NANS ALLOWED



© 2020 – Machine2Learn





# MUSIC GENRE TRANSFER

# IMAGE STYLE TRANSFER



MACHINE2LEARN



Content Musical Piece		Style (Musical Piece)		Output Musical Piece
Metallica's Master of Puppets	+	Beethoven's 5th Symphony	=	Beethoven's Master of Puppets?
OR				
Stevie Wonder's Isn't She Lovely	+	Hip-hop	=	Hip-hoppy Isn't She Lovely?

# METALLICA-ENTER SANDMAN JAZZ



# EXAMPLE EUROPOP TO SOUL

---

Source



Transfer



All rights reserved. No part of this document may be reproduced in any material form without the written permission of Machine2Learn.

# TAKE HOME MESSAGES

Deep learning is cool,  
especially for  
computer vision.

Explainability /  
transparency is  
essential for gaining  
trust.

Causal inference if  
you care about the  
effect of  
interventions.

# Empowering Smart Industry



**MACHINE2LEARN**

[machine2learn.com](http://machine2learn.com)

**+31 20 21 171 41**

**Bijlmerdreef 101**

**1102 BP, Amsterdam**

[info@machine2learn.nl](mailto:info@machine2learn.nl)



# Parallel session High Tech

