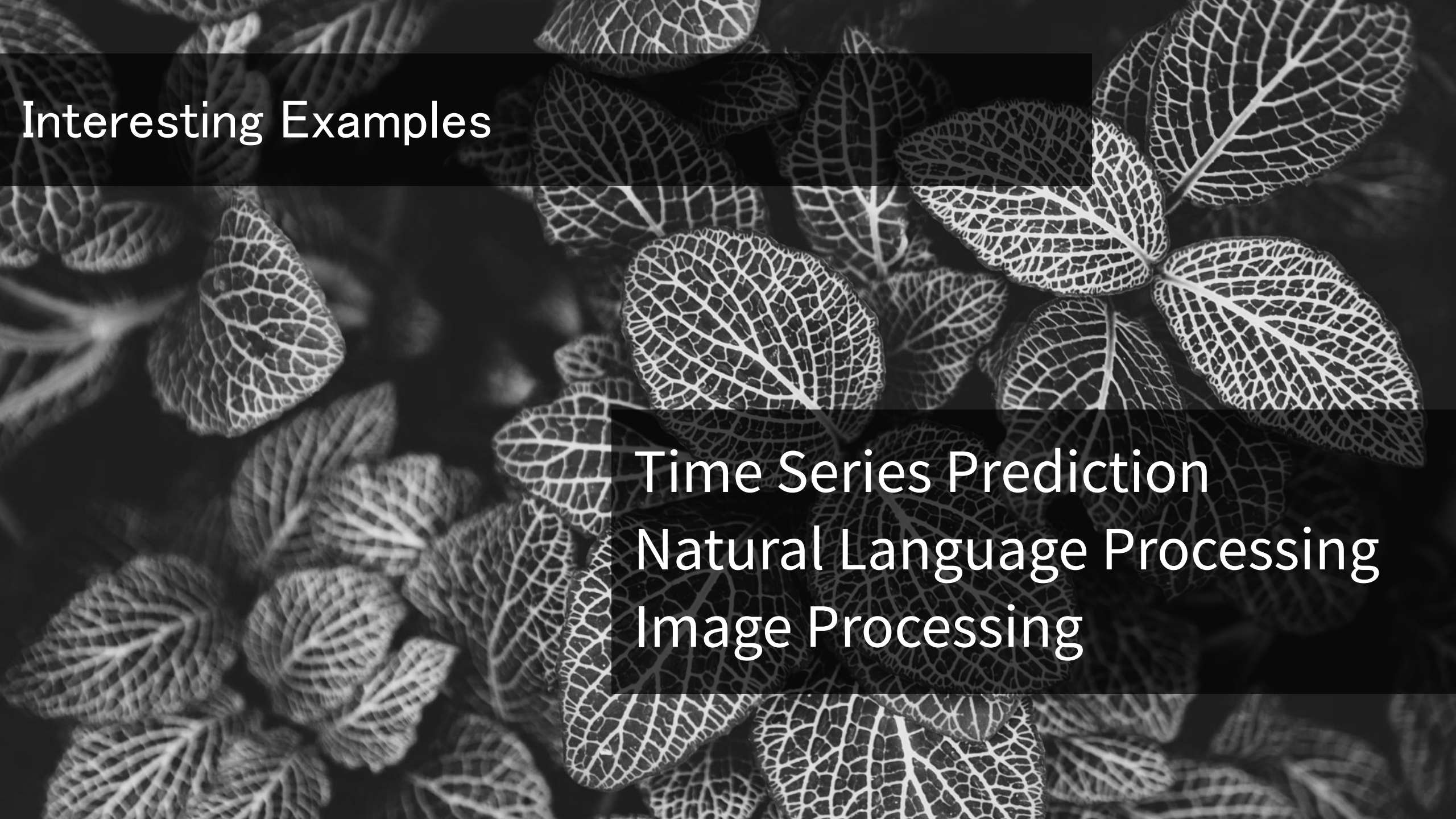
Solving Real World Machine Learning Problems

Questions

We will see singularity happening


Questions
AI is going to change the world

Interesting Examples



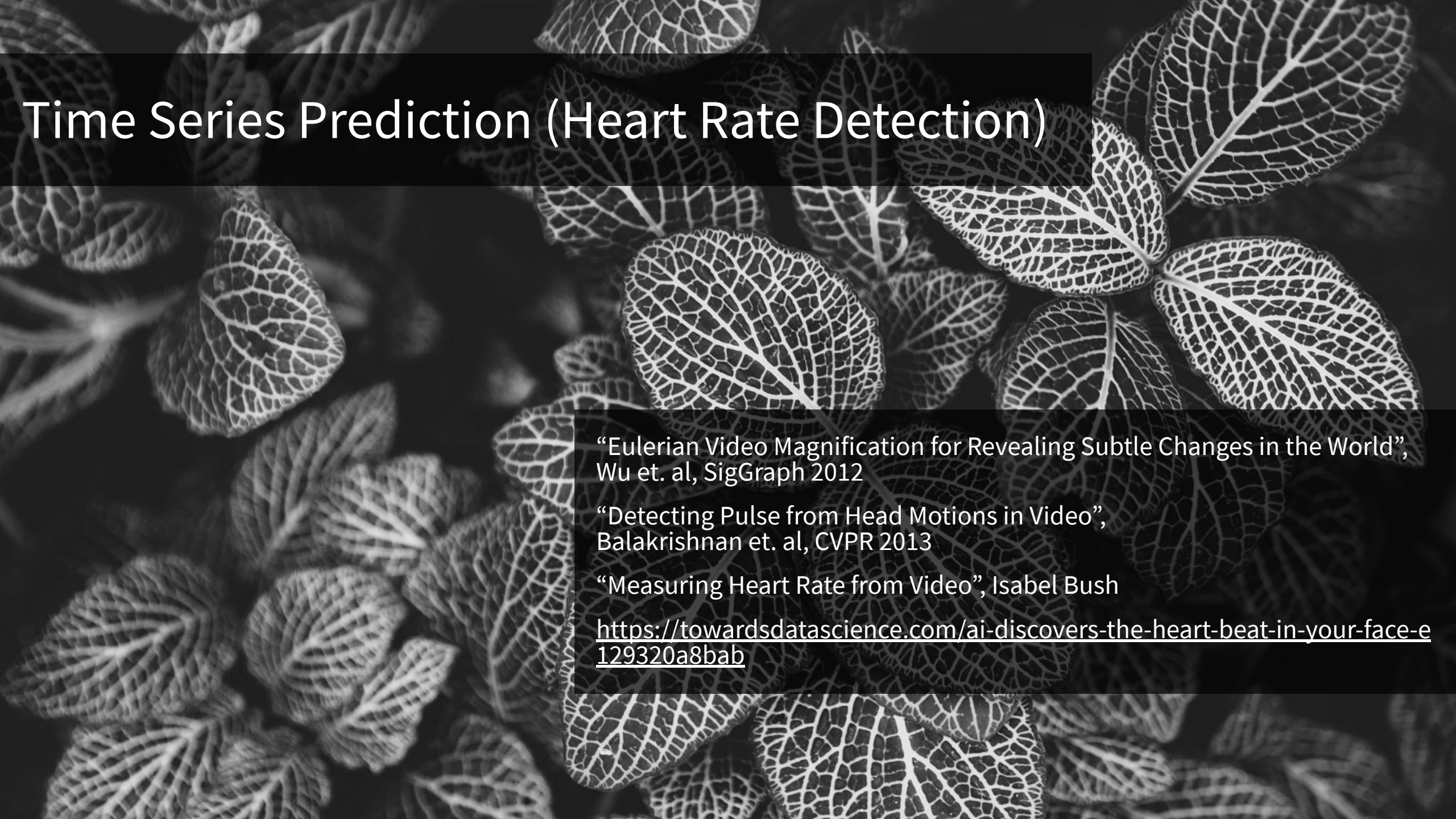
Time Series Prediction
Natural Language Processing
Image Processing

Time Series Examples

Heart Rate Detection from Video Streams

Stock Market Prediction

Time Series Prediction (Heart Rate Detection)



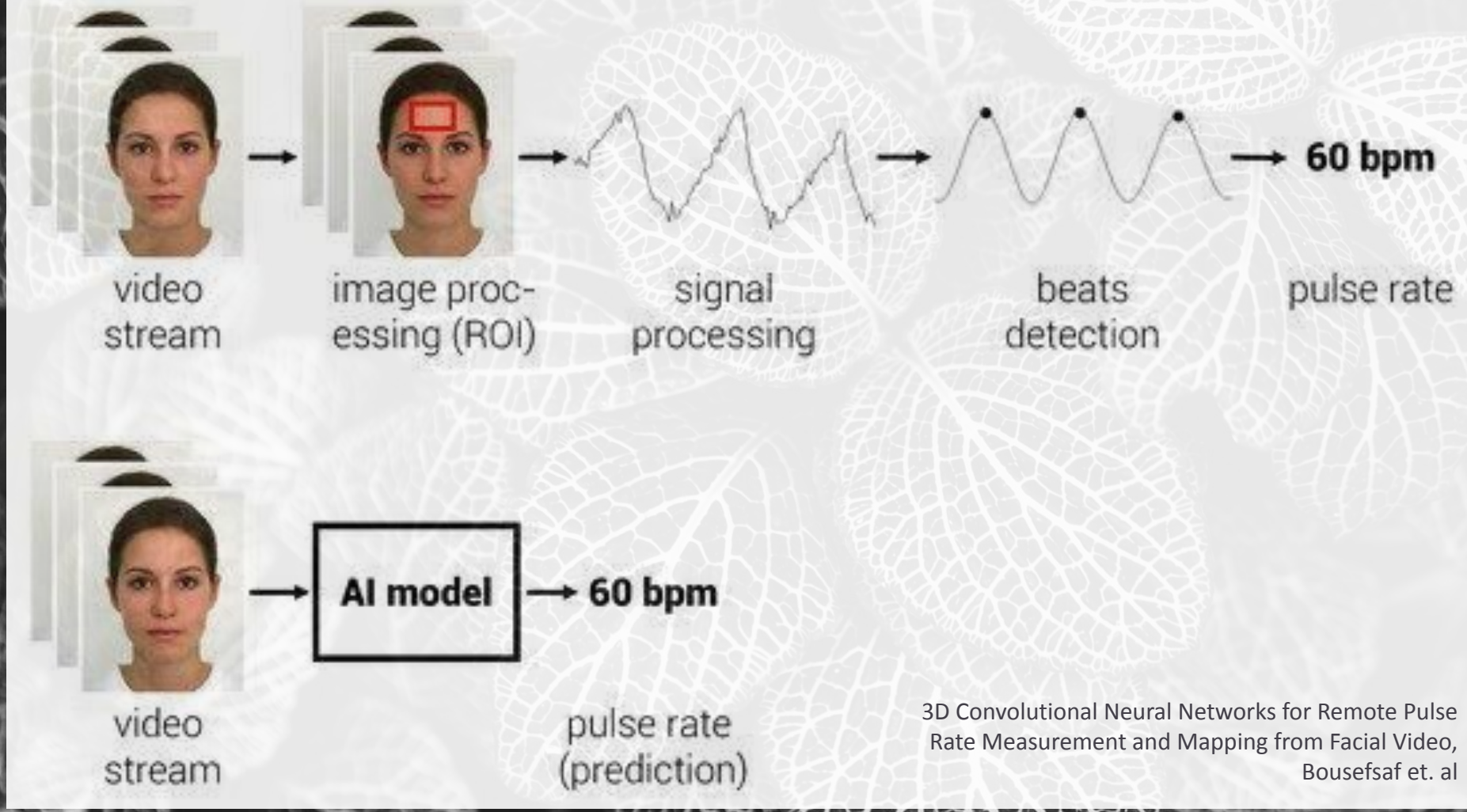
"Eulerian Video Magnification for Revealing Subtle Changes in the World", Wu et. al, SigGraph 2012

"Detecting Pulse from Head Motions in Video", Balakrishnan et. al, CVPR 2013

"Measuring Heart Rate from Video", Isabel Bush

https://towardsdatascience.com/ai-discovers-the-heart-beat-in-your-face-e129320a8bab

Time Series Prediction (Heart Rate Detection)



3D Convolutional Neural Networks for Remote Pulse Rate Measurement and Mapping from Facial Video, Bousefsaf et. al

Time Series Prediction

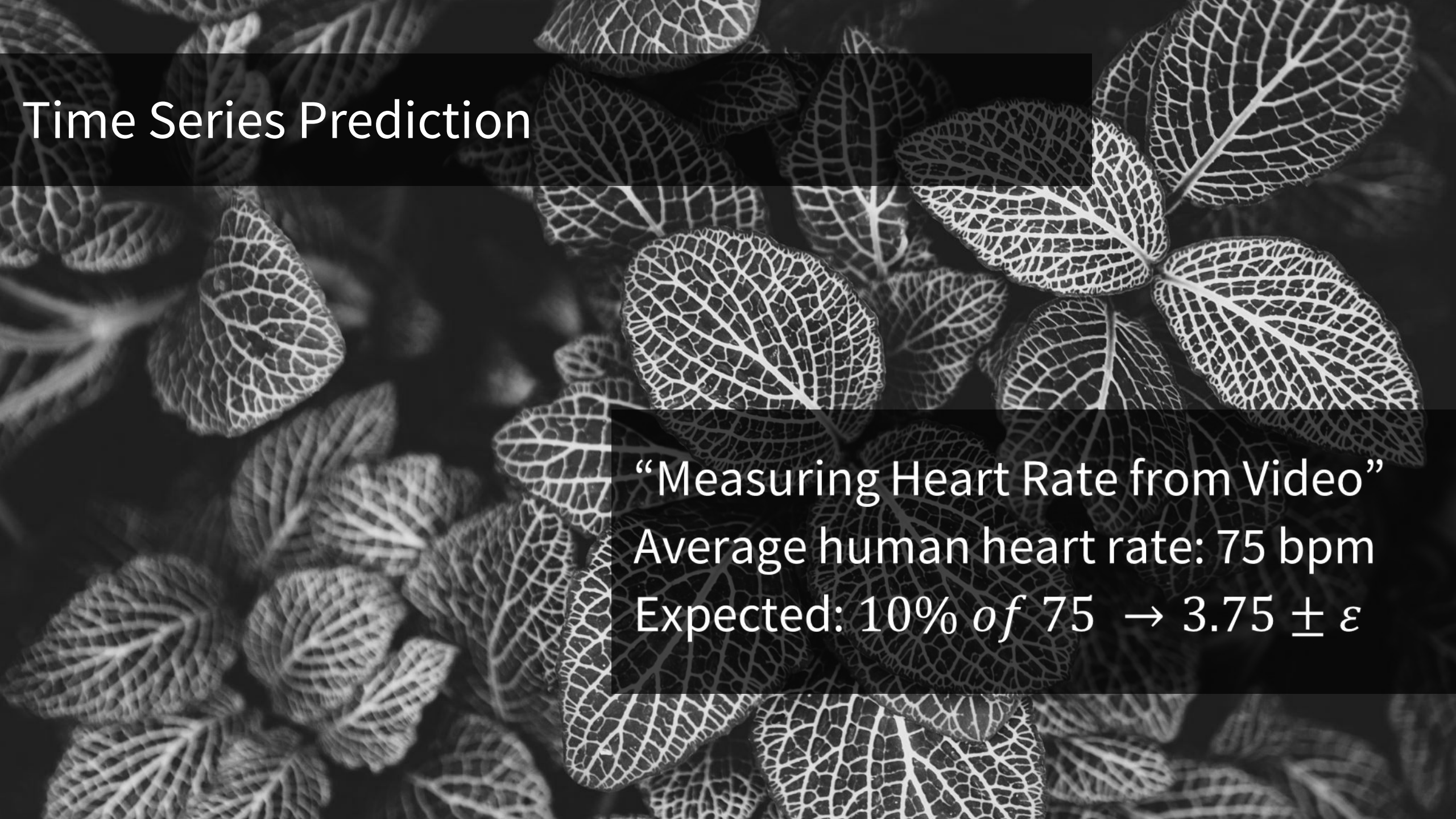
For videos of still faces, the calculated heart rate error was 3.4 ± 0.6 bpm. For videos of subjects tilting, turning, and shifting their faces, the calculated heart rate error was 2.0 ± 1.6 bpm.

“Measuring Heart Rate from Video”

Time Series Prediction

These errors were calculated using only inlier measurements, which were defined to be measurements within 10% of the reference heart rate.

"Measuring Heart Rate from Video"



Time Series Prediction

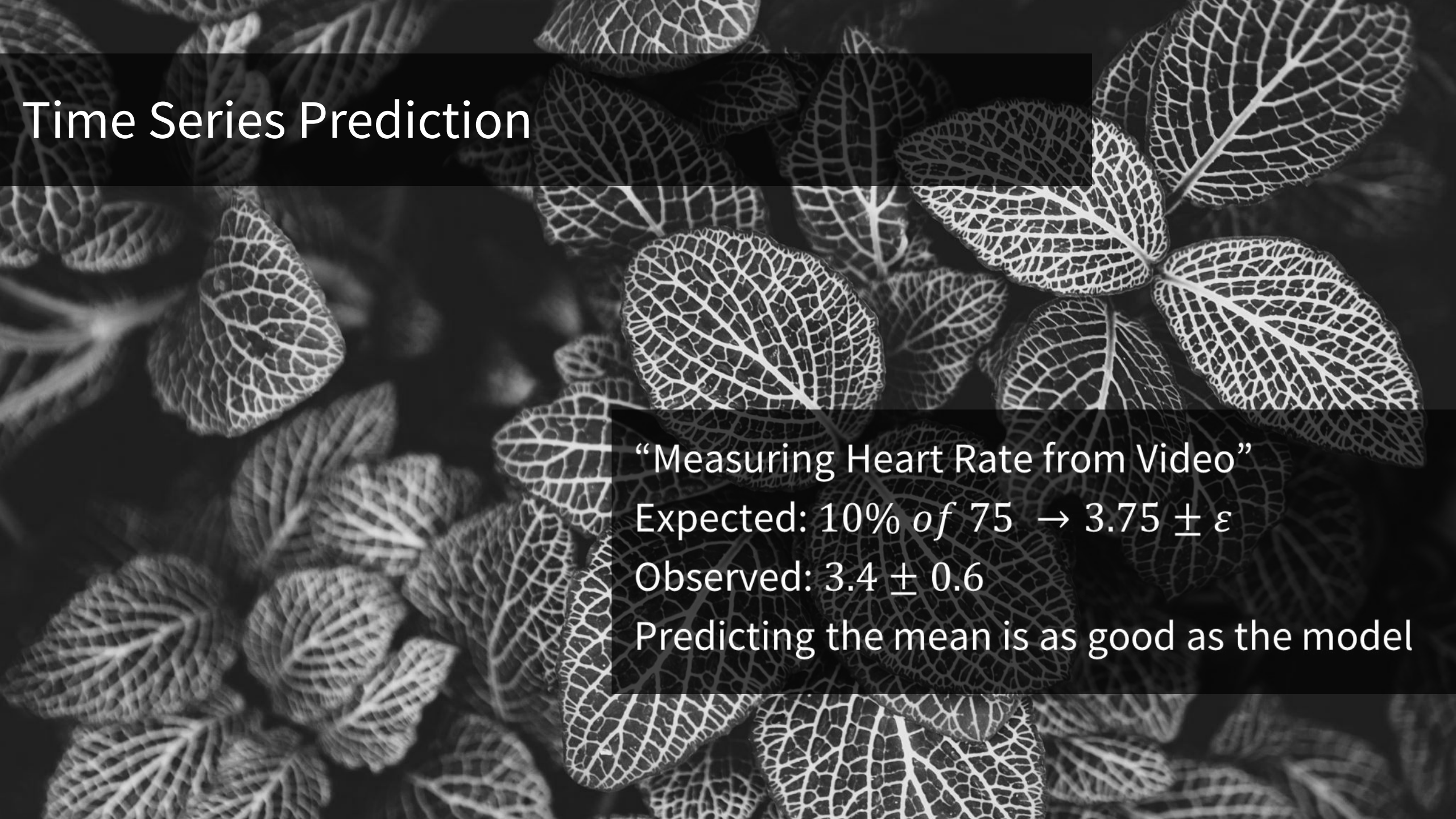
"Measuring Heart Rate from Video"
Average human heart rate: 75 bpm
Expected: $10\%\ of\ 75\ \rightarrow 3.75 \pm \varepsilon$

Time Series Prediction

"Measuring Heart Rate from Video"

Expected: $10\%\ of\ 75\ \rightarrow 3.75 \pm \varepsilon$

Observed: 3.4 ± 0.6



Time Series Prediction

"Measuring Heart Rate from Video"

Expected: $10\%\ of\ 75\ \rightarrow 3.75 \pm \varepsilon$

Observed: 3.4 ± 0.6

Predicting the mean is as good as the model

Time Series Prediction

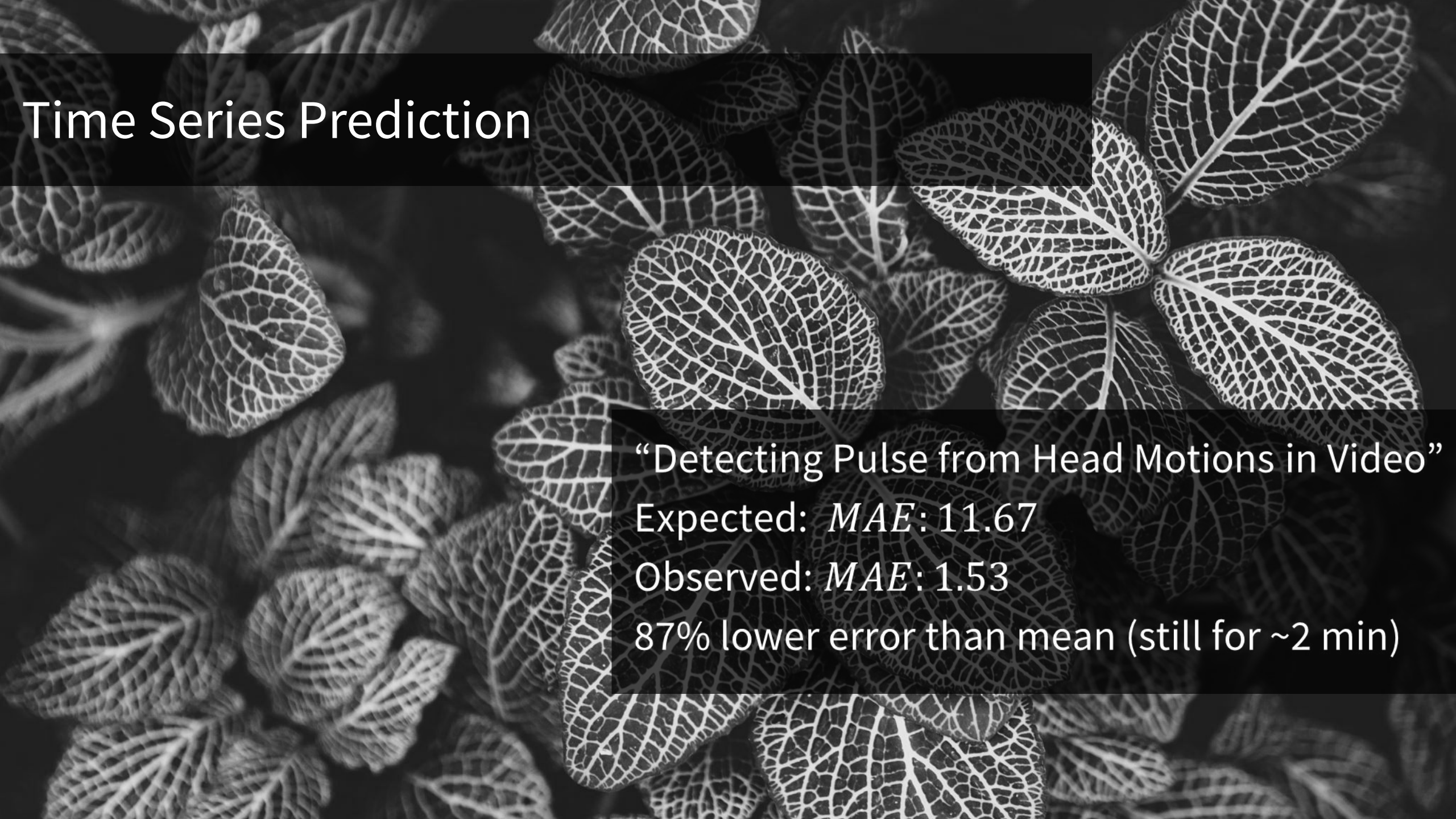
“Measuring Heart Rate from Video”

Choose a good performance metric

Measure against expected error

With a reasonable base line estimator e.g.,
Mean or Class Distribution

Time Series Prediction



"Detecting Pulse from Head Motions in Video"

Expected: MAE: 11.67

Observed: MAE: 1.53

87% lower error than mean (still for ~2 min)

Time Series Prediction

“Detecting Pulse from Head Motions in Video”

Is the way data needs to be collected feasible?

Time Series Prediction

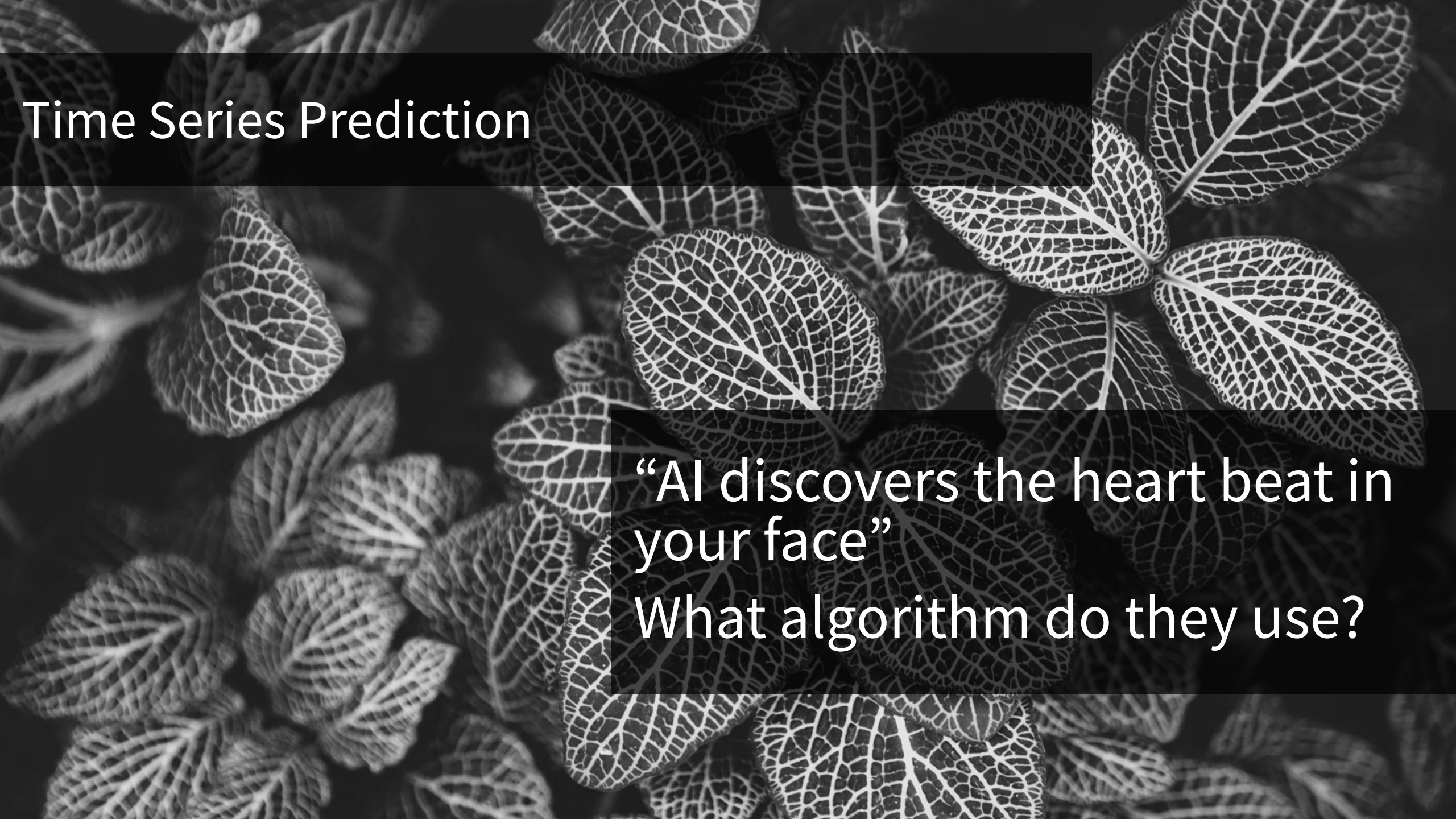
"AI discovers the heart beat in your face"

Time Series Prediction

“AI discovers the heart beat in your face”

https://towardsdatascience.com/ai-discovers-the-heart-beat-in-your-face-e129320a8bab

Time Series Prediction



"AI discovers the heart beat in your face"

What algorithm do they use?

Time Series Prediction

“AI discovers the heart beat in your face”
What algorithm do you think they use?
ICA and PCA

Time Series Prediction

“AI discovers the heart beat in your face”

Use the most straight forward algorithm.
Easier to explain, easier to use, easier to debug

Time Series Prediction

Noise in Noise out

If your Signal has less than 0.5 dB a machine learning algorithm will not solve it for you.

Design a clean signal and test what amount of noise is manageable. Can you reasonably assume this data is available for your problem?

Time Series Prediction (Stock Market)

Stock Market Prediction

Big investment banks moving closer to Wall Street / Frankfurt / London Stock exchange because the faster connection gives them an edge.

Jay Tuck, AI will kill us

Time Series Prediction (Stock Market)

Stock Market Prediction

Banks moving closer to Wallstreet to have a speed edge.

Possible origin in the book “Flash Boys”:
https://www.laphamsquarterly.org/time/not-millisecond-spare

Michael Lewis

Time Series Prediction (Stock Market)

Stock Market Prediction

Bountiful Trust, Kansas City

I was not able to find any evidence this company ever existed

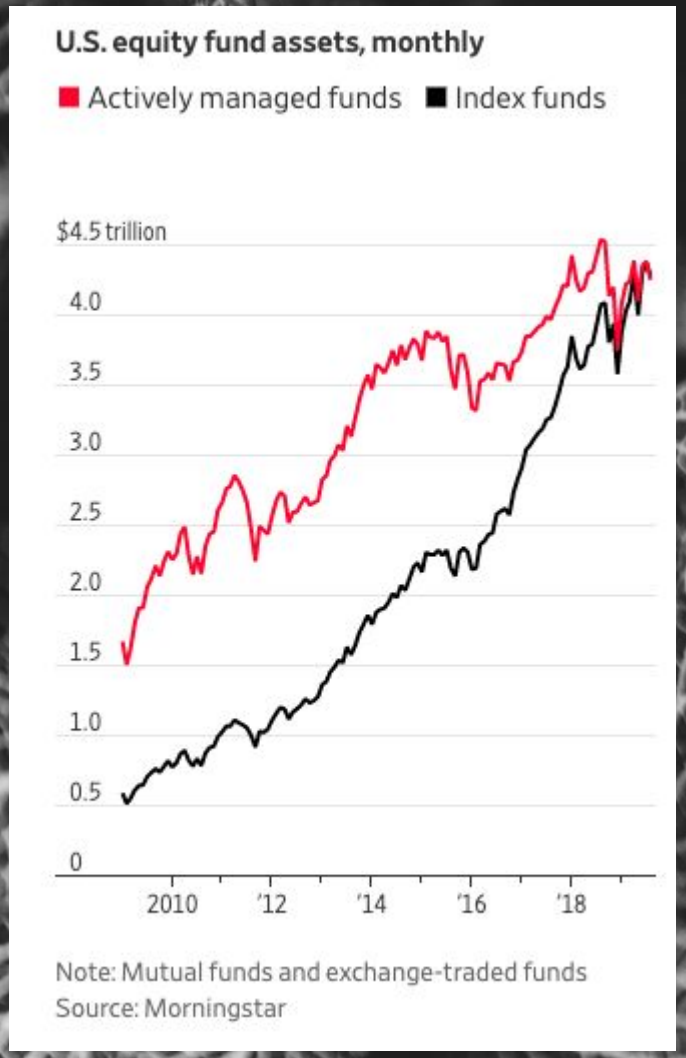
Time Series Prediction (Stock Market)

Stock Market Prediction

Another interesting point with all this:

Managed funds do not outperform the market in the long run

Time Series Prediction (Stock Market)



Stock Market Prediction

Another interesting point with all this:

Managed funds do not outperform the market in the long run

Time Series Prediction

“Stock Market”

Just because someone taunts a solution does not mean it is really a better solution to your challenge

Time Series Prediction (Stock Market)

In 2010 these algorithms burned 50 billion in less than a second for exactly the following reason:

Parent: If one of your friends snorted too much cocaine and jumped in front of a train would you?

Kid: o.O ⋯ what, Dad?

Artificial Intelligence: $\begin{bmatrix} 1.0 \\ 0.0 \end{bmatrix}$

Time Series Prediction (Stock Market)

In 2010 these algorithms burned 50 billion in less than a second for exactly the following reason:

Parent: If one of your friends snorted too much cocaine and jumped in front of a train would you?

Kid: o.O ⋯ what, Dad?

Artificial Intelligence: $\begin{bmatrix} 1.0 \\ 0.0 \end{bmatrix} \rightarrow$ 

Time Series Prediction (Stock Market)

Stock Market #2

What are possible negative consequences of your system failing?

Detour: Bias

Algorithmic Bias

Bias: Redlining

a discriminatory practice that puts services (financial and otherwise) out of reach for residents of certain areas

Community Reinvestment Act was passed in 1977 to help prevent redlining

Reintroducing it through ML?

Bias: Data

Parole

Should an algorithm decide about parole?

Compas gives judges “recommendations”

https://en.wikipedia.org/wiki/COMPAS (software)

And is mot likely biased:

https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm

Bias

Bias

What are possible negative consequences of your system working as intended?

https://dash.harvard.edu/bitstream/handle/1/33746041/2017-07_responsivecommunities_2.pdf


NLP
GPT-3
Google assist

NLP: GPT-3

GPT-3 too early to talk about
Enjoy slides with lemon and salt

NLP: GPT-3

The Guardian let GPT-3 write an op-ed piece

They edited 8 versions into one article

https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3

NLP: GPT-3

Marcus et. al**: You poured yourself a glass of cranberry juice, but then you absentmindedly poured about a teaspoon of grape juice into it. It looks okay. You try sniffing it, but you have a bad cold, so you can't smell anything. You are very thirsty. So **you drink it.

GPT-3: You are now dead.

Marcus et. al

NLP: GPT-3

So what is GPT-3?

Lets try to understand the strength of the tool we see

Simulacrum or smart clippy?

NLP: GPT-3

Reasonable Applications

Quick translation aide

Smart auto complete / spelling corrector

Support writing papers / stories

Innovative game controller / better NPCs

NLP: GPT-3

ML is a useful tool if one has
Reasonable Assumptions
Incremental Improvement

NLP: Google Assist

NLP: Google Assist

Google I/O 2018

https://www.youtube.com/watch?v=D5VN56jQMWM

Google I/O 2019

Sundar Pichai: “great feedback”

NLP: Google Assist

Take a look how it works and see if you can spot what it does:

https://www.youtube.com/watch?v=DGwALqd1YxQ

It also uses a lot of human in the loop

NLP: GPT-3

Is there a tool that is used for this already?

Would a higher input variance help the tool?

Can you find errors and delegate to a human?

Image Analysis: Cancer Detection

Study: Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study

Funding: Lunit

Lunit mission statement: ” With AI, we aim to make data-driven medicine the new standard of care. …”

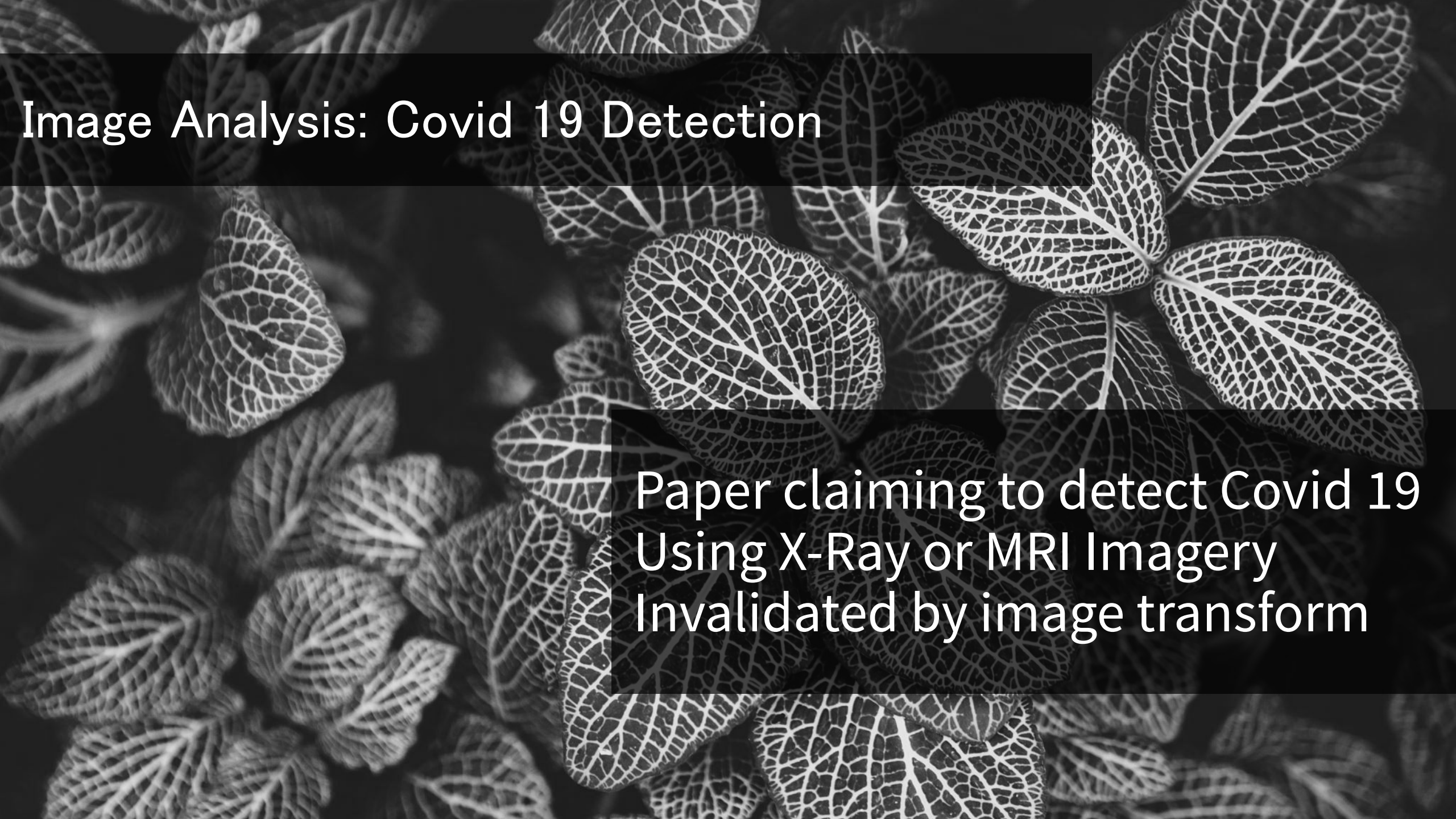


Image Analysis: Covid 19 Detection

Paper claiming to detect Covid 19 Using X-Ray or MRI Imagery Invalidated by image transform

Image Analysis: Car Identifier

Decide make of car
0.99 Cohens Kappa (scary)

Image Analysis: Car Identifier

Why so good?

Solving Real World Machine Learning Problems

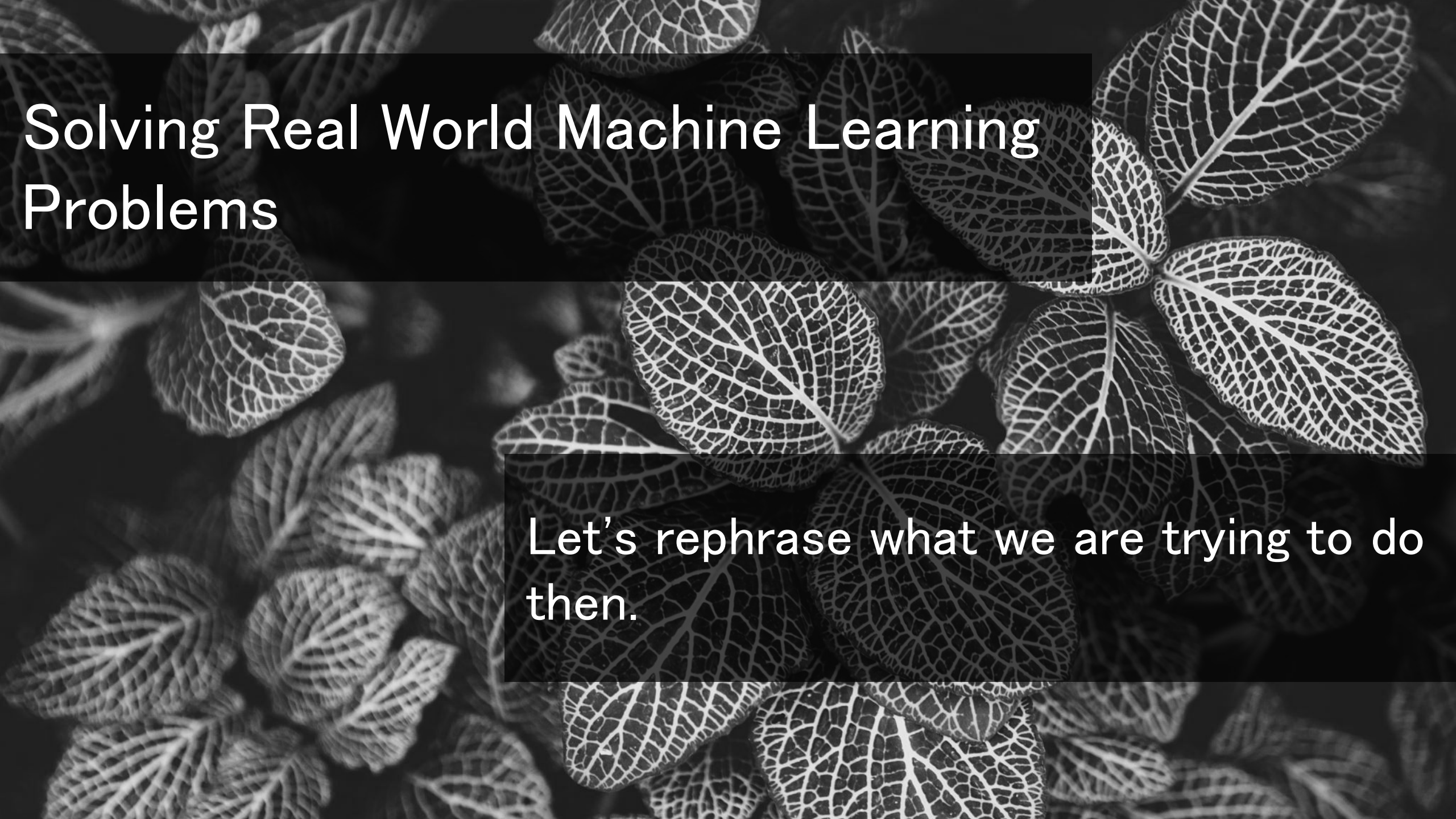

Solving Real World Machine Learning Problems
Think about the problem
And how a solution would look like

Solving Real World Machine Learning Problems

If one thinks they have a silver bullet,

Solving Real World Machine Learning Problems

every problem becomes a werewolf.



Solving Real World Machine Learning Problems

Let’s rephrase what we are trying to do then.

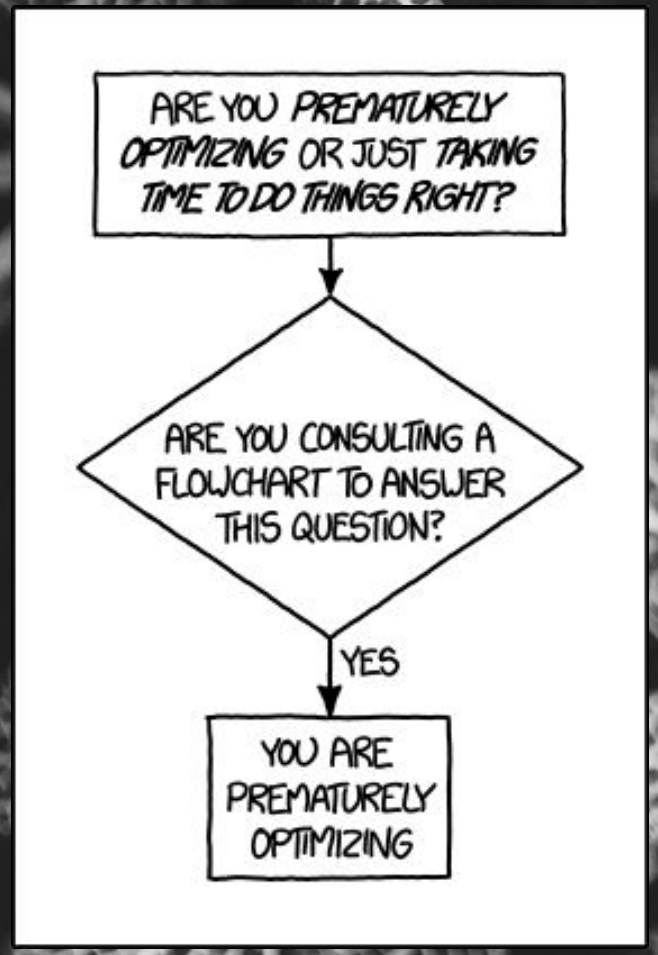
Solving Real Problems*

* Using Computing Machinery

Let’s rephrase what we are trying to do then.

Solving Real Problems*

* Using Computing Machinery



The algorithm is the last piece of the puzzle.

“Premature optimization is the root cause of all evil.”

Problem Formalization

To understand the problem try solving it yourself.

Data Collection

Collect a small data set yourself!

Build or find the tools you need to collect the data at scale.

Data Collection

How much data do we need?

How to collect that data?

What is the price tag? Still worth it?

Wizard of Oz

Select some examples from your collection.

Write a description how to solve the problem.

Let your teammates solve the examples only using your description.

Wizard of Oz

Gives you an idea of data quality and general task feasibility.


Performance Metric
Baseline Implementation for Comparison

Performance Metric

Histogram and threshold
Using the class distribution
Predicting the mean


Performance Metric
Replace your baseline algorithm
Test performance changes