

# **Cogs 9**

# **Discussion Section**

**FA22 Week 7**  
**Will McCarthy**

# Upcoming due dates

**Friday, Nov 11th** Assignment 3

**Thursday, Nov 17th** Reading Quiz 5

# **This week: Machine Learning**

**Reading 5(a): Nicholas Diakopoulos, 2016, Accountability in Algorithmic Decision Making**

**Hand-written character classification demo**

## Reading 5(a): Nicholas Diakopoulos, 2016, Accountability in Algorithmic Decision Making

“A view from computational journalism.”

Algorithms make a lot of decisions...

Those decisions have consequences...

But algorithms make mistakes.

This is a problem in “virtually all strands of industry and government”.

How should we hold algorithmically informed decisions **accountable** to the public?

You can't really hold an algorithm accountable:

Many places for **humans to influence algorithms**: “criteria choices, optimization functions, training data, and the semantics of categories...”

# Kinds of algorithmic decision making

**Prioritizing:** ordering entities (search results, newsfeeds, products, etc.)

**Classification:** putting entities (& people) in groups (good vs bad loan recipient)

**Association:** relationships between entities (search autocomplete)

**Filtering:** exclusion of information (censorship)

For each of these, describe a scenario/ technology where an algorithm could make a decision that have undesirable consequences.

# Government vs. Private Sector Accountability

**Elected governments should be “accountable to their citizenry”**

“algorithms are largely unregulated now, and they are indeed exercising power over individuals or policies in a way that in some cases (for example, hidden government watch lists) lacks any accountability whatsoever.”

“We, the governed, should find it unacceptable there is no transparency or even systematic benchmarking and evaluation of these forecasts, given the important policy decisions they feed.”

**Corporations don’t have to be**

“may sometimes be impelled to act through social pressure”

Government vs. Private Sector respond to different pressures

# ***Transparency* leads to accountability**

**“Transparency can be a mechanism that facilitates accountability, one that we should demand from government and exhort from industry.”**

Transparency could lead to a competitive disadvantage

But transparency need not be absolute (e.g. not entire source code)

What could be made more transparent?

# An algorithmic transparency standard

**Human involvement:** goal, purpose, and intent of the algorithm. Who was involved? => public responsibility/ accountability

**Data:** accuracy, completeness, uncertainty, timeliness, representativeness. How was it gathered?

**The model:** what algorithm is used? inputs? weights?

**Inferencing:** how successful is the model (at some standard measure). Margin of error? Where do errors come from?

**Algorithmic presence:** algorithm or not?



# Criticisms of this article

Transparency framework seems sensible, but vague

Transparency obviously relevant but not well motivated

What else could lead to accountability besides transparency?

Distinction between Government and Private Sector is relevant but not discussed much going forward

Scope of this article is unclear. Journalism? Or bigger than that?

# **Demo: handwritten digit classification**

# Goal: recognize handwritten digits

**Input:**



**Model:**

?

**Output:**

0

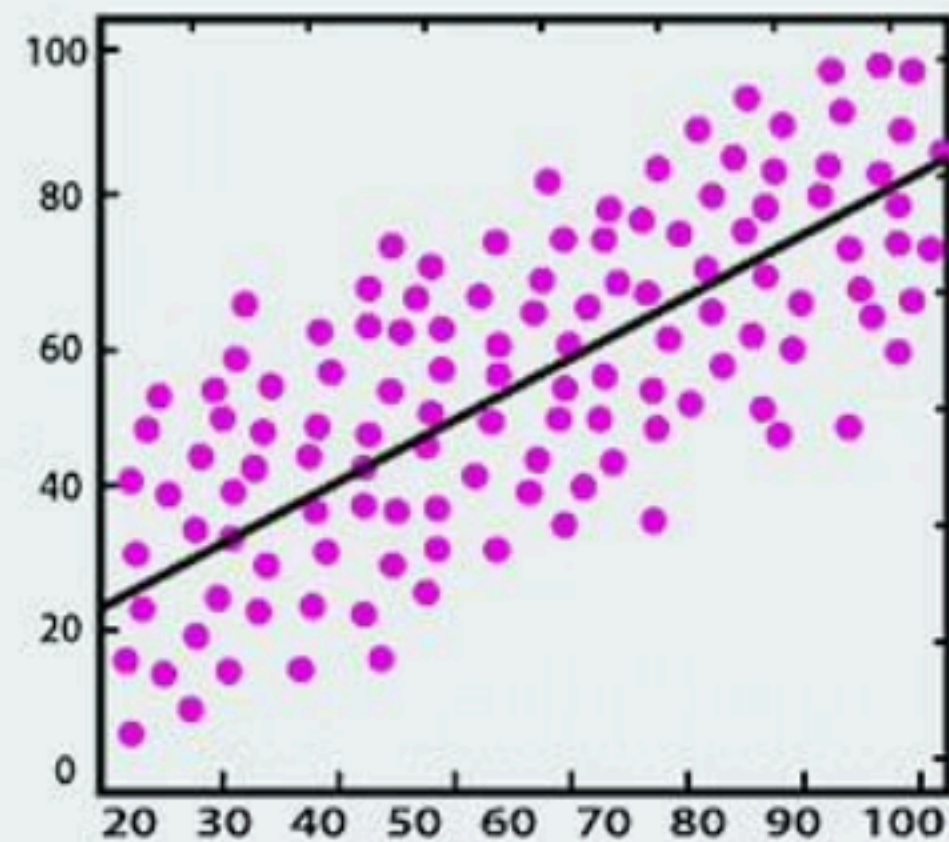
1

2

3

# What kind of task is this?

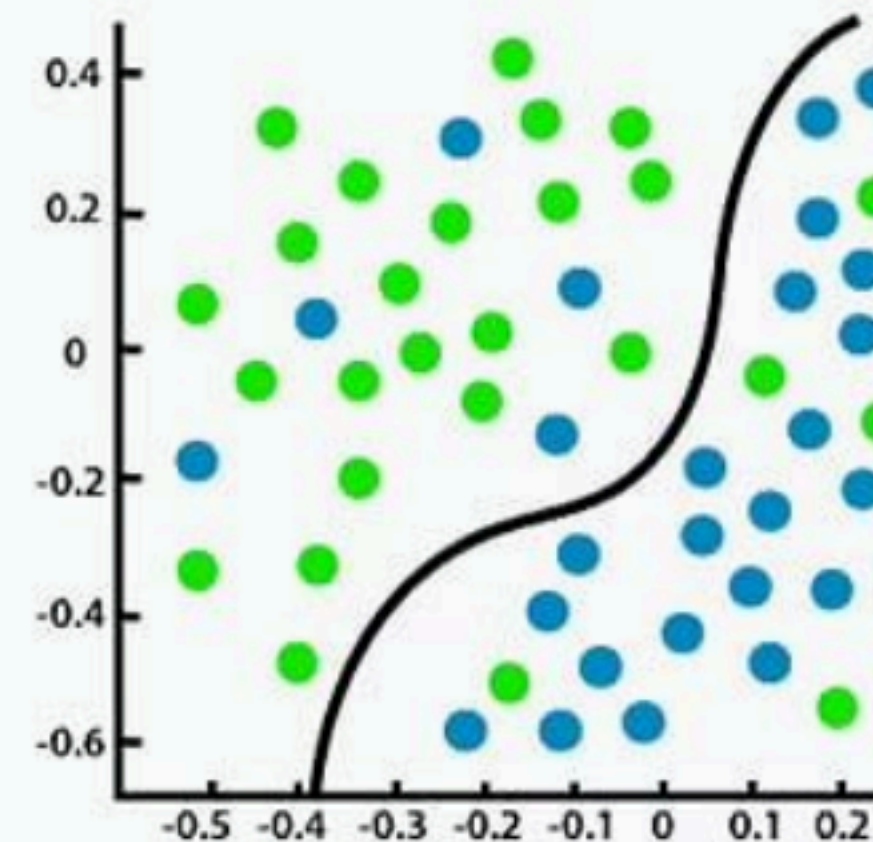
Predicting a  
Continuous Value



Regression

*versus*

Predicting a Class



Classification

# Goal: recognize handwritten digits

**Input:**



**Model:**

**Classifier**

**Output:**

0 1 2 3

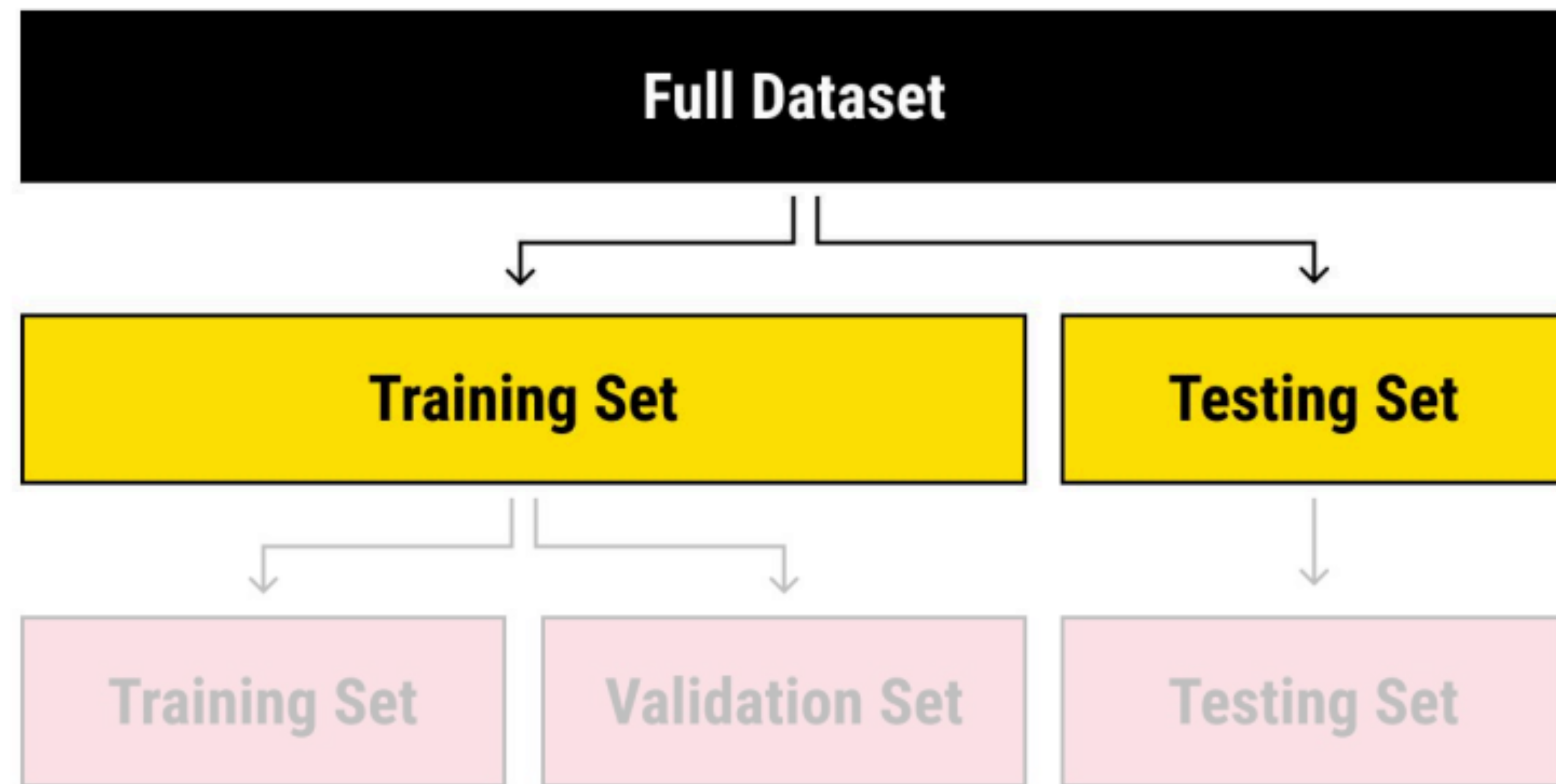
# Predictive culture (cf. Donoho/ Breiman)

“If the prediction is good, who cares what the model is?” - Prof. Shannon

Great, so long as we actually know that the prediction is actually good!

How do we ensure this?

Testing



# **Demo: handwritten digit classification**

**Group work / questions**



# Future Readings

5(b): Julia Angwin, et al., 2016, Machine Bias