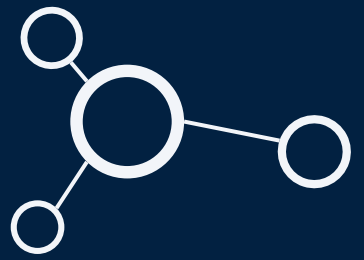
DOMINO

Ethics and Bias in Data Science

DAVID BLOCH | London Meetup |MARCH 2020

Agenda

Ethics and Bias in Data Science



- Introduction
    - Defining Ethics and what it means to us as Data Scientists
    - The difference between Ethics and Governance
- Example data science use-cases with ethical impacts
- Managing for Bias within Data Science
- Building ethics and bias training into your Data Science curriculum
- Conclusion.

© 2019 Domino Data Lab, Inc.

Introduction - Ethics in Data Science

Defining Ethics

ethics [eth-iks]

SEE SYNONYMS FOR *ethics* ON THESAURUS.COM

noun

(*used with a singular or plural verb*) a system of moral principles:

the ethics of a culture.

(*used with a plural verb*) the rules of conduct recognized in respect to a particular class of human actions or a particular group, culture, etc.:

medical ethics; Christian ethics.

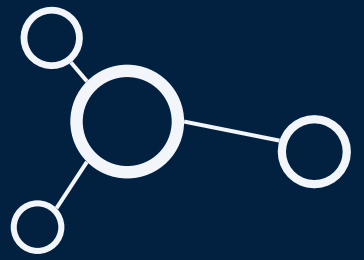
(*used with a plural verb*) moral principles, as of an individual:

His ethics forbade betrayal of a confidence.

(*used with a singular verb*) that branch of philosophy dealing with values relating to human conduct, with respect to the rightness and wrongness of certain actions and to the goodness and badness of the motives and ends of such actions.Compare axiological ethics, deontological ethics.

Introduction - Ethics in Data Science

Making Ethics Real in Data Science



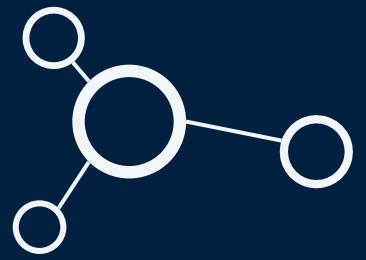
ethics (in Data Science)

- A set of shared values that help us differentiate what our customers, main stakeholders and business view as right or wrong.
- Provides a set of principles and parameters around how data can be collected, captured and used inside the modern enterprise.
- Helps to build transparency and promote trust in the way data science practices operate inside modern businesses.

© 2019 Domino Data Lab, Inc.

Ethics vs Governance

What are ethical decisions vs regulatory decisions?



ethics

- Leading question - “Should we do this?”
- Ensures adherence to company and customer values.
- Reduces reputational risk to the organization.
- Concluding Question - “If this was in the paper on Monday morning, how would our customers feel?”

governance

- Leading question - “Are we allowed to do this?”
- Ensures compliance with regulatory standards,
- Reduces legal and governance risk to the organization.
- Concluding Question - “Are we in breach of any governance, regulation or legal standards in doing this?”

© 2019 Domino Data Lab, Inc.

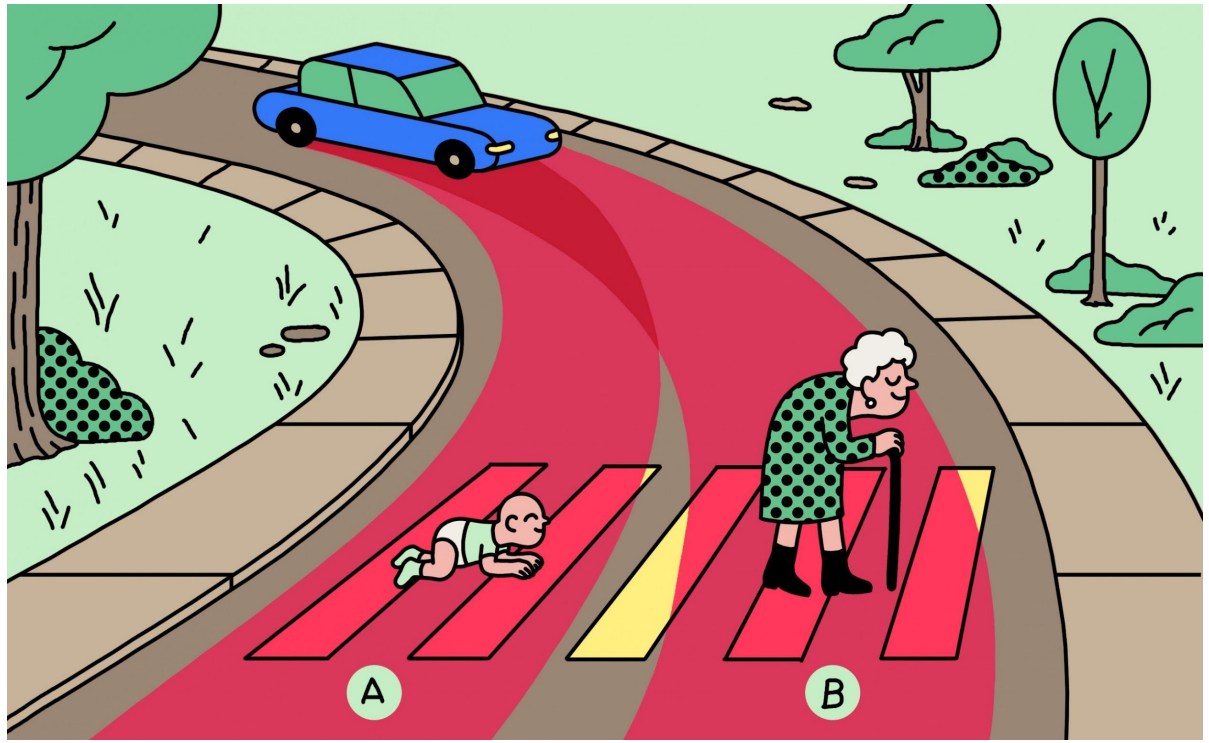
DOMINO

Example cases of ethics in Data Science

Models making decisions that are life or death

The Trolley Problem

- Often used as an example when discussing autonomous vehicles.
- Should you include provisions in your software to make a decision around the value of life?
- Current Example: Assessing expected life for COVID-19 patients before offering treatment



© 2019 Domino Data Lab, Inc.

DOMINO

Example cases of ethics in Data Science

Models being used as a form of policing or governance

Hyper Personalization

- Social Credit Scoring - A form of hyper personalization with influencer network analysis
- The ability to combine multiple datasets to build ensemble models that assess the “value” of a person by any given metric
- Other Examples: Facebook advertising platform using proximity + influencer interests to advertise to you

© 2019 Domino Data Lab, Inc.

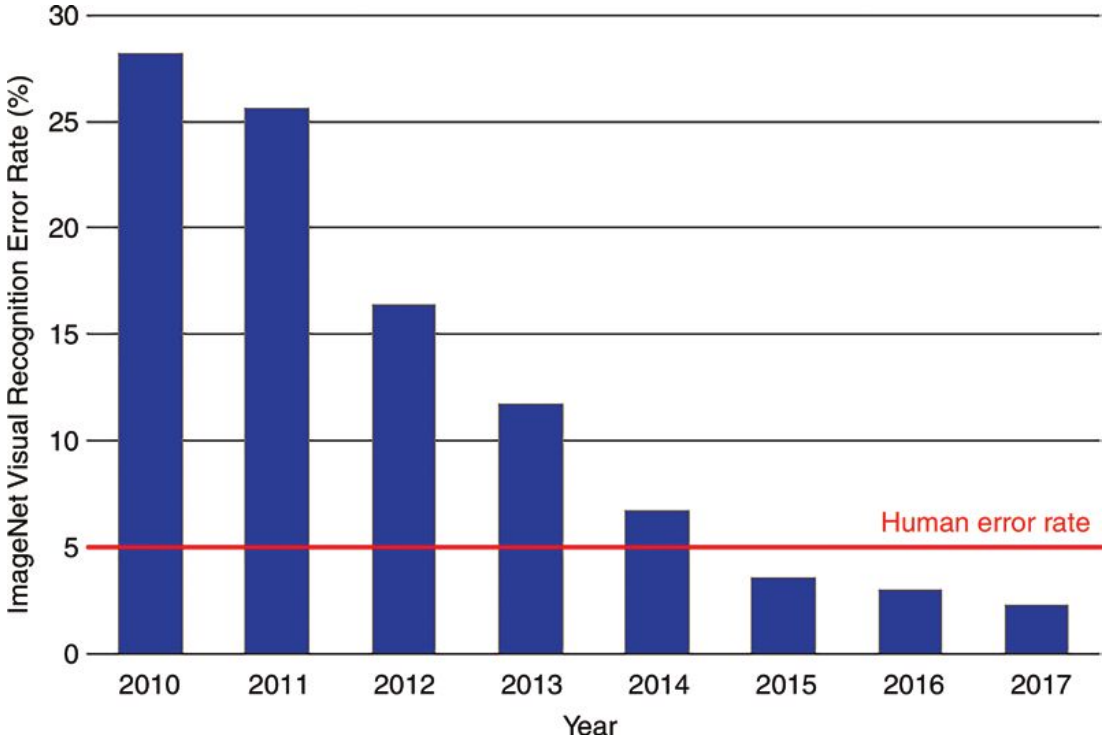
DOMINO

Example cases of ethics in Data Science

Models being used to assist in diagnosis of health conditions, growth rates of plants or pest control

Diagnosis and Image Recognition

- Image Classification and NLP techniques being used to make diagnoses in healthcare sector
- ImageNet rate of classification shows that machines make fewer errors in recognizing visual patterns, however we still trust humans more than machines
- Other Examples: Crop growth, identification of pests in forests and facial recognition



© 2019 Domino Data Lab, Inc.

DOMINO

Example cases of ethics in Data Science

Models being used to identify performance markers of people

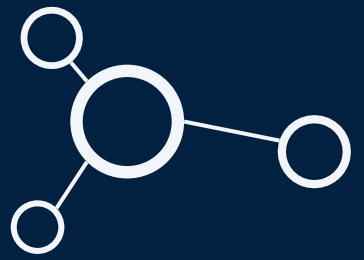
Assessing People Performance

- Using data capture and collection alongside ensemble modelling techniques to identify performance, behavior and attendance of employees
- Leading organizations such as Google have become KPI obsessed about how talent work and operate
- Other Examples: Wearables being utilized in construction sites to understand active periods vs rest periods

© 2019 Domino Data Lab, Inc.

Guiding Questions to be ethically aware

Quick goalposts to understand ethical implications of actions



Guiding Questions

- If someone did this with our own personal information, how would we feel?
- If this ended up in the Monday morning paper, would there be any reputational impact to the work we're doing?
- Are we able to be fully transparent with our customers and/or key stakeholders about the way that we're using models?
- Have we included the ability for people to opt-in or opt-out?

Introduction - Bias in Data Science

Defining Bias

bias [bahy-*uhs*]

SEE SYNONYMS FOR *bias* ON THESAURUS.COM

noun

a particular tendency, trend, inclination, feeling, or opinion, especially one that is preconceived or unreasoned:

illegal bias against older job applicants;

the magazine's bias toward art rather than photography;

our strong bias in favor of the idea.

unreasonably hostile feelings or opinions about a social group; prejudice:

accusations of racial bias.

an oblique or diagonal line of direction, especially across a woven fabric.

Statistics. a systematic as opposed to a random distortion of a statistic as a result of sampling procedure.

Introduction - Ethics in Data Science

Making Bias Real in Data Science

Assessing Bias (in Data Science)

- Identify potential unintended consequences of a model due to biases introduced through datasets, selective sampling, modelling techniques and false-positive conclusions.

TECH | AMAZON | ARTIFICIAL INTELLIGENCE

Amazon reportedly scraps internal AI recruiting tool that was biased against women

The secret program penalized applications that contained the word "women's"

By James Vincent | @jvincent | Oct 10, 2018, 7:09am EDT

SHARE

DOMINO

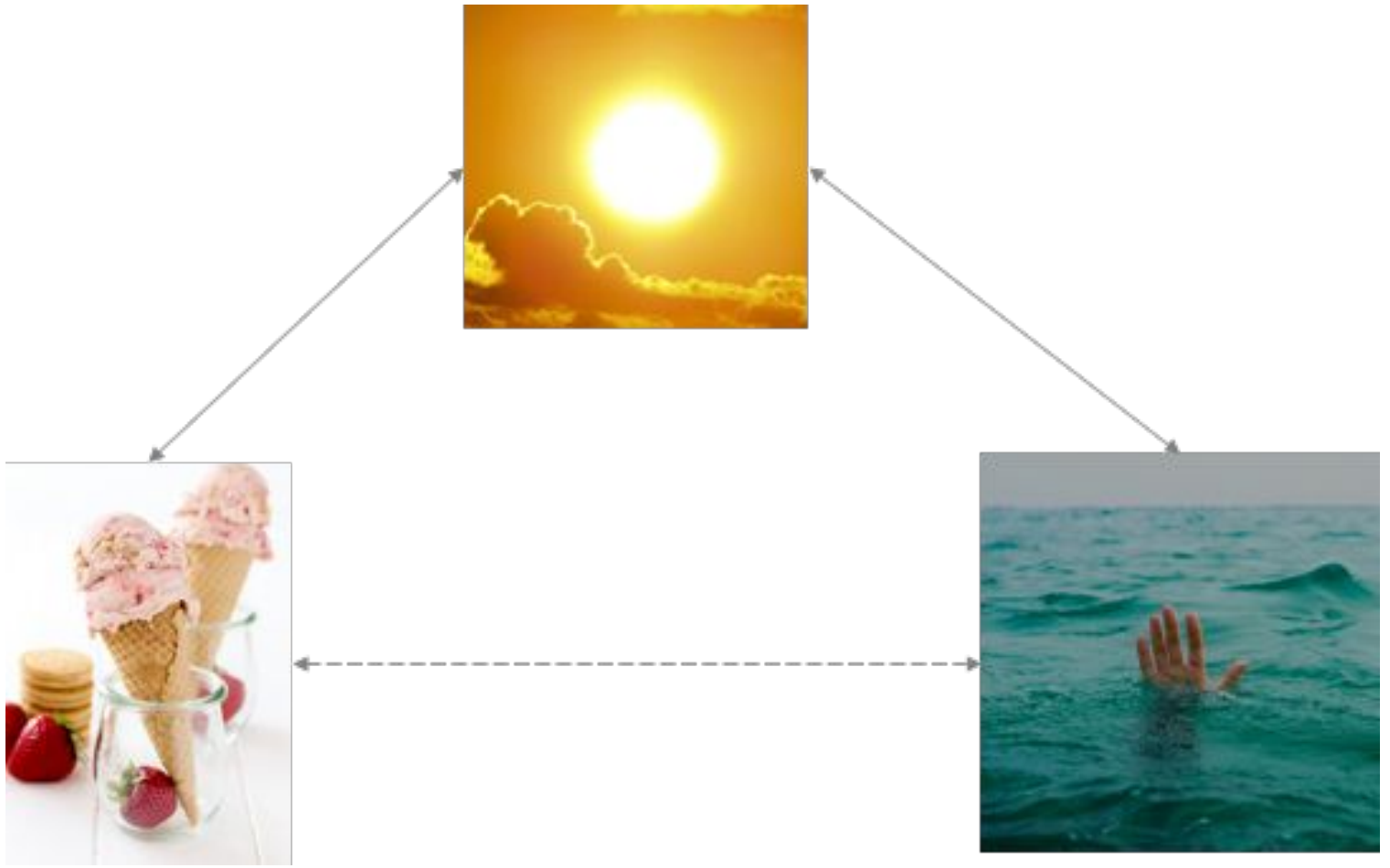
Types of Bias - Confirmation Bias

Only selecting examples that agree with your existing opinions

DOMINO

Types of Bias - Confounding Bias

Identifying correlation that may not directly apply to causation

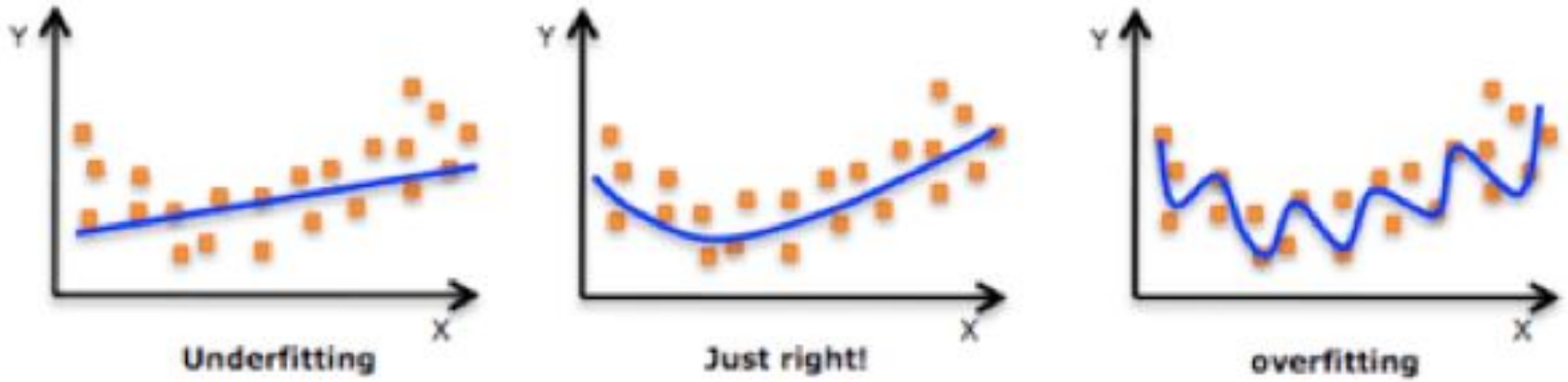


© 2019 Domino Data Lab, Inc.

DOMINO

Types of Bias - Overfitting

Building a model that is perfect for the training dataset, but not applicable to the real world



© 2019 Domino Data Lab, Inc.

DOMINO

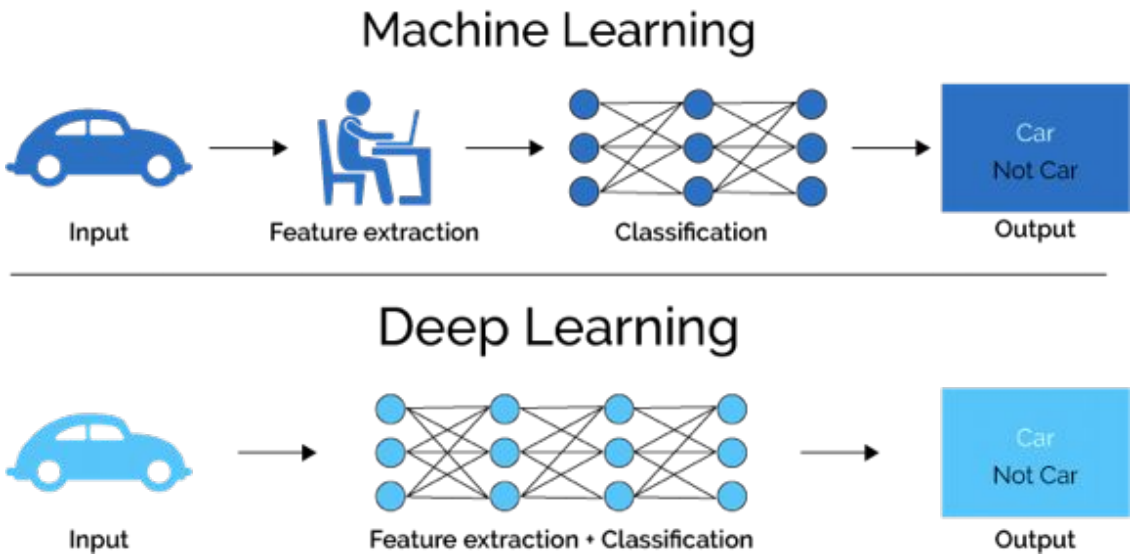
The Consequences of Deep Learning

Deep Learning removes people from the loop in variable selection, meaning more biases can be introduced

The need for explainability in Deep Learning

- Deep learning techniques allow data scientists to apply algorithms to massively deep and massively wide datasets, meaning that variable selection via feature extraction is not required.

- Due to this, the need to explain why a model predicts what it predicts becomes more important in order to correctly assess and manage for bias (not to mention apply the results of a model)



DOMINO

Key Questions for Bias

If we can't explain it, should we use it?

- Machine Learning, Deep Learning and other AI techniques open up many different possibilities for improving classification, categorization and prediction engines….

- The risks of these tools and techniques is that they become further removed from human selection and understanding, therefore new techniques must be used to explain the output - i.e "Why this is being predicted?"

© 2019 Domino Data Lab, Inc.

Key Questions for Bias

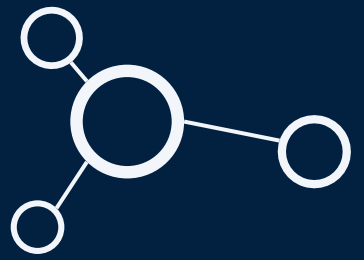
If we can't use it, should we include it?

- While many demographic indicators can lead to positive correlation in building prediction engines, indicators such as gender, age or ethnicity cannot be used in most situations.

- Identifying which variables cannot be leveraged once a model has been generated can help identify what is for the purposes of information only, versus actionable insights a business can use for decisions.

Conclusion: Final Thoughts

Ethics and Bias in Data Science



- Establishing governance and ethics boards inside organizations are helpful, but need to be backed by establishing ethics and bias training in your data scientist curriculum.

- Data Governance and Compliance boards should be setup to make quick decisions around “Should we do this?”

- Ethics and Bias boards should be setup to make quick decisions around “Could we do this?”

- The role of a Data Science leader should include regular refreshers on unintended consequences and the need to build explainability into each technique being utilized to build models

© 2019 Domino Data Lab, Inc.

Resources

Additional reading

- Ethical Charter for Data Science - Actuaries.org.uk
  https://www.actuaries.org.uk/system/files/field/document/An%20Ethical%20Charter%20for%20Date%20Science%20WEB%20FINAL.PDF

- Princeton University - Dialogue and Case Studies on AI Ethics use-cases
  https://aiethics.princeton.edu/case-studies/

- Towards Data Science - The Ethics of Data Science Blog
  https://towardsdatascience.com/the-ethics-of-data-science-e3b1828affa2

- Towards Data Science - Understanding Data Bias
  https://towardsdatascience.com/survey-d4f168791e57

- The Civis Journal - Reducing Bias and Ensuring Fairness in Data Science
  https://medium.com/civis-analytics/reducing-bias-and-ensuring-fairness-in-data-science-424ded3badbb

© 2019 Domino Data Lab, Inc.

DOMINO

Thank you

David Bloch
Data Science Evangelist

email: david.bloch@dominodatalab.com
LinkedIn: https://www.linkedin.com/in/dbloch