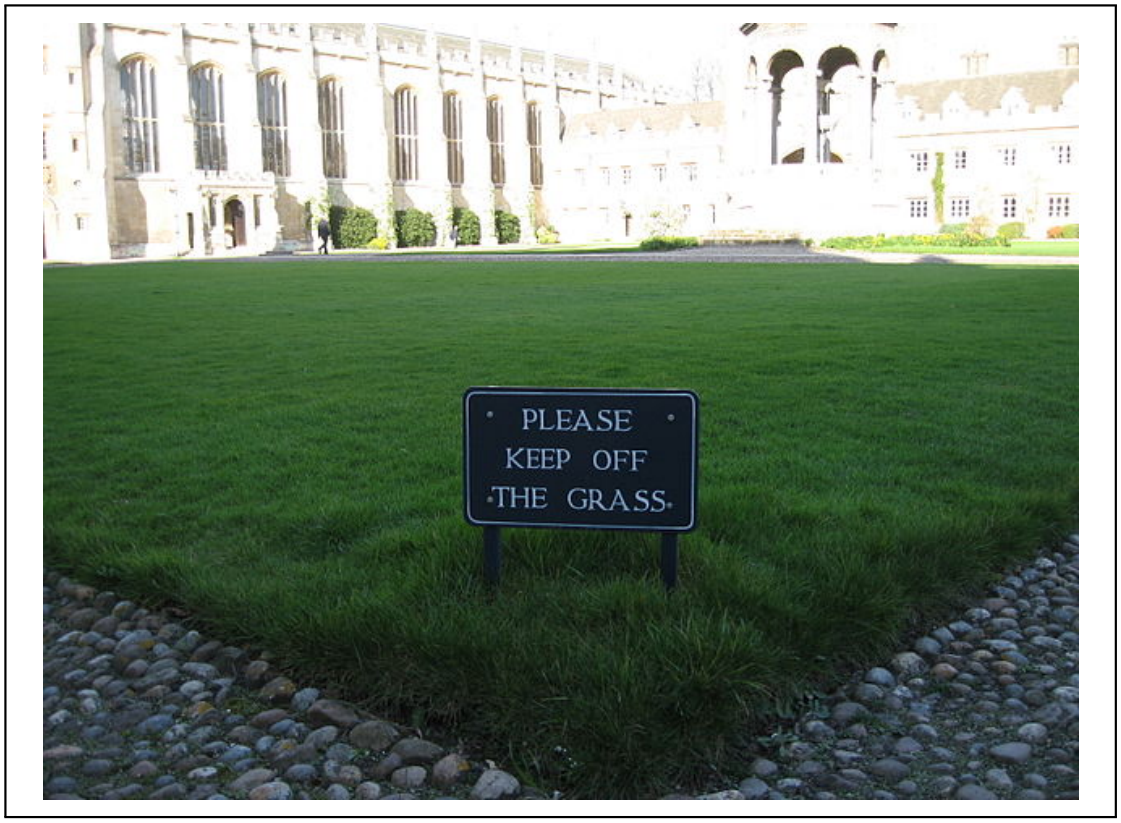
Risk, Artificial Intelligence and Discrete Geometry

Dr. Paul Larsen

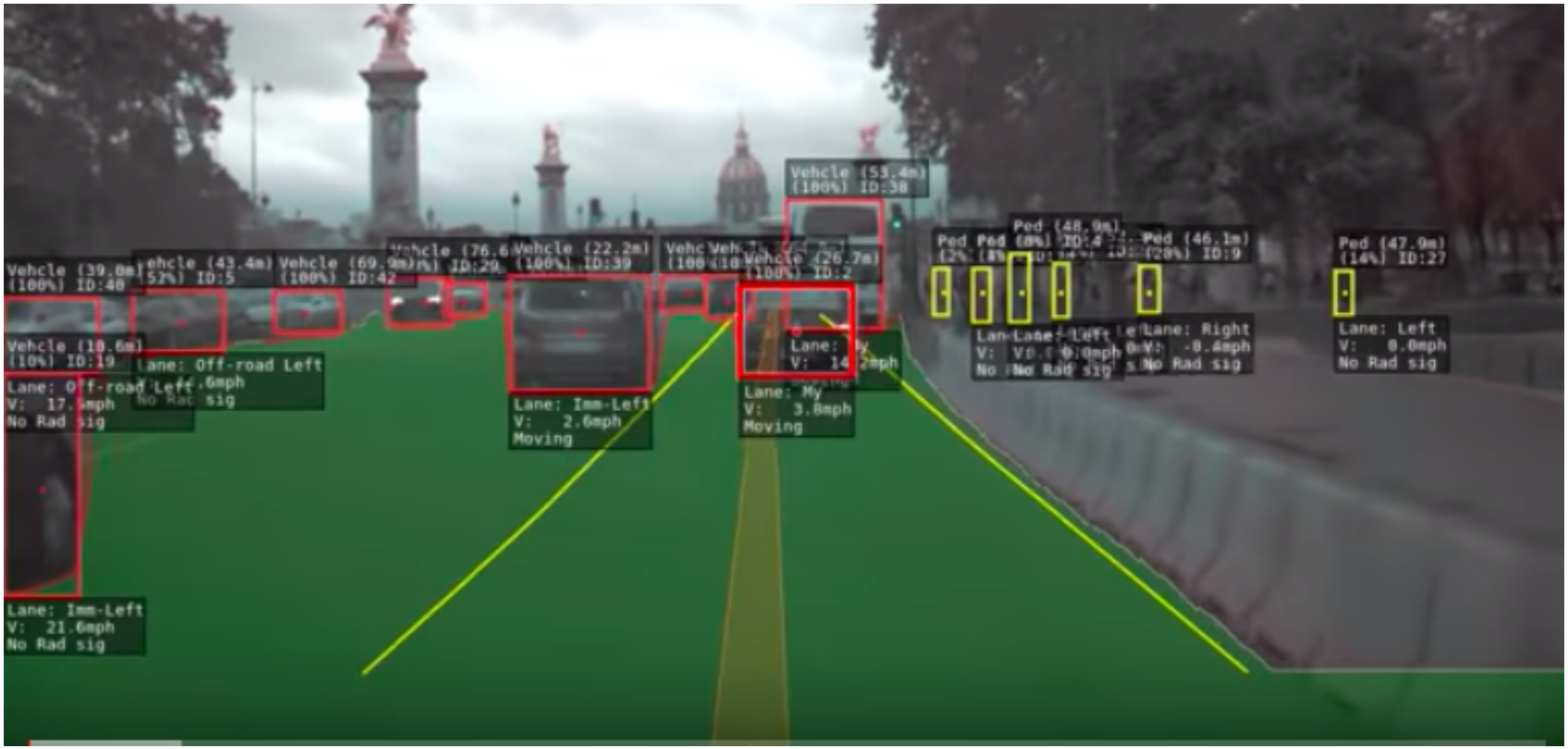
February 21, 2020

Human are amazing at navigating a world that does not make sense



Artificial Intelligence Does Not Exist

Artificial Intelligence Does Not Exist, But ...



Source: YouTube: greentheonly, Paris streets in the eyes of Tesla Autopilot

Artificial Intelligence Does Not Exist, But ...

Google Duplex: An AI System for Accomplishing Real-World Tasks Over the Phone

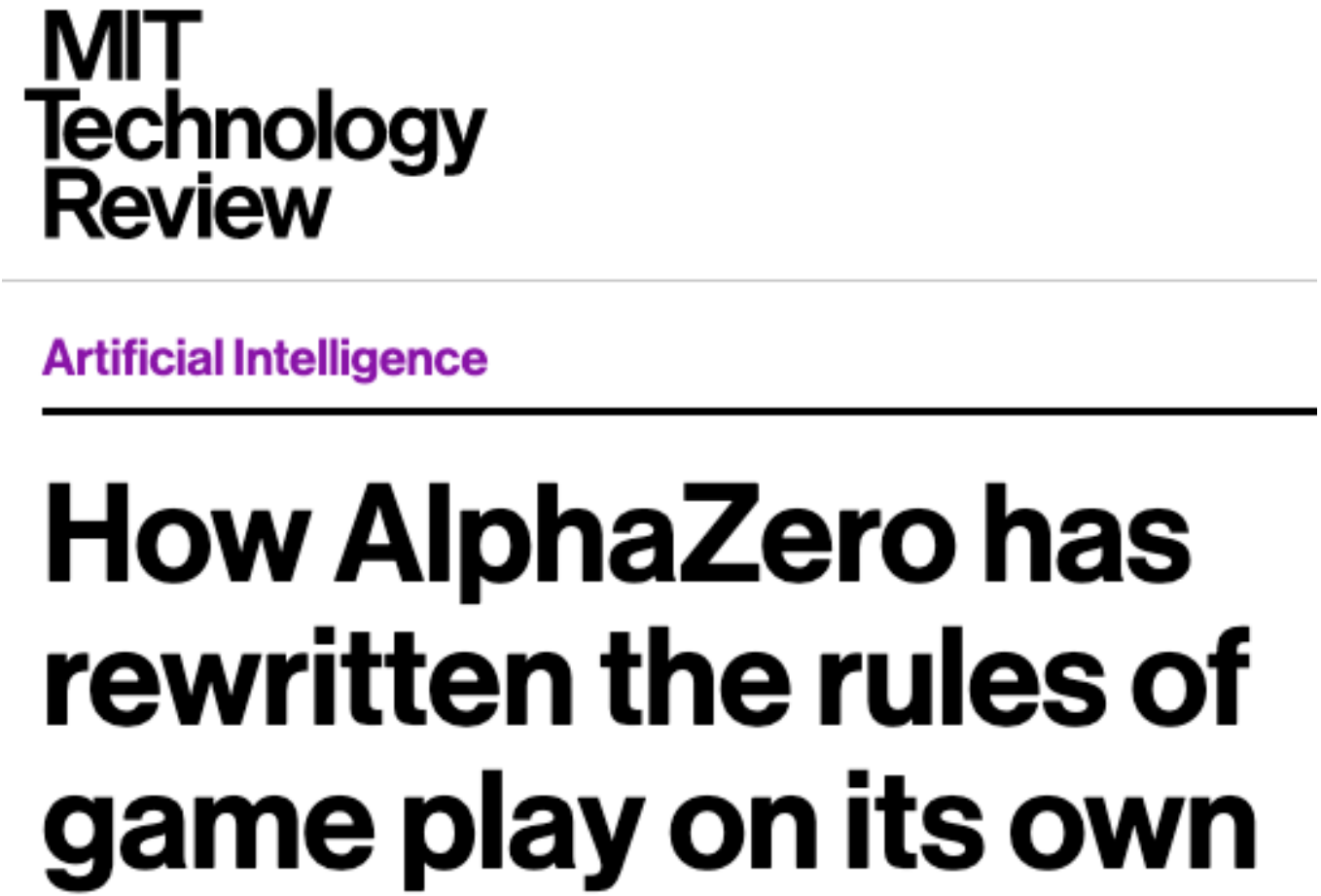
Tuesday, May 8, 2018

Posted by Yaniv Leviathan, Principal Engineer and Yossi Matias, Vice President, Engineering, Google

A long-standing goal of human-computer interaction has been to enable people to have a natural conversation with computers, as they would with each other. In recent years, we have witnessed a revolution in the ability of computers to understand and to generate natural speech, especially with the application of deep neural networks (e.g., Google voice search, WaveNet). Still, even with today's state of the art systems, it is often frustrating having to talk to stilted computerized voices that don't understand natural language. In particular, automated phone systems are still struggling to recognize simple words and commands. They don't engage in a conversation flow and force the caller to adjust to the system instead of the system adjusting to the caller.

Source: Google AI Blog

Artificial Intelligence Does Not Exist, But ...



Source: MIT Technology Review

AI Struggles With Context

The scientist named the population, after their distinctive horn, Ovid's Unicorn. These four-horned, silver-white unicorns were previously unknown to science.

Source:
https://www.lesswrong.com/posts/4AHXDwcGab5PhKhHT/humans-who-are-not-concentrating-are-not-general

AI Struggles With Bias

Facial Recognition Is Accurate, if You're a White Guy

By Steve Lohr

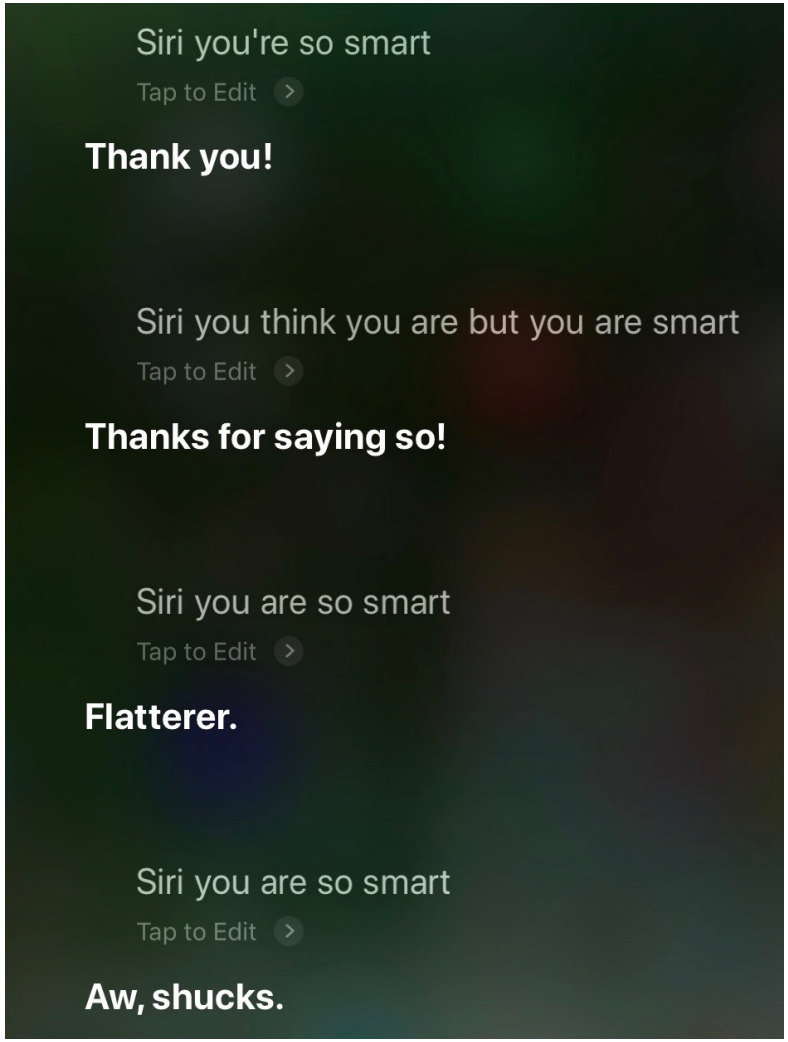
Feb. 9, 2018

Facial recognition technology is improving by leaps and bounds. Some commercial software can now tell the gender of a person in a photograph.

Source: New York Times
See also: James Vincent, The Verge, Google 'fixed' its racist algorithm by removing gorillas from its image-labeling tech

AI Struggles With Human Behavior

Siri Compliments

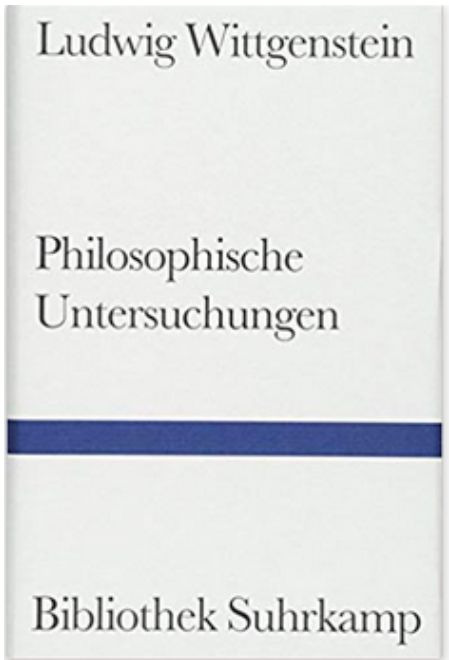


Workshop topics

1. Artificial intelligence for risk
2. Discrete geometry for artificial intelligence
3. Correlation and causation
4. Putting artificial intelligence into action

Workshop format

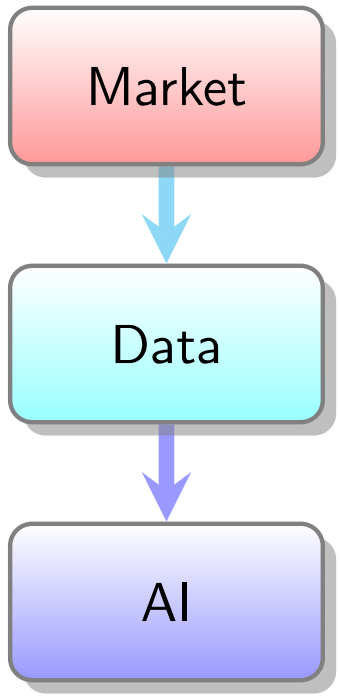
- For each topic, lecture + tutorial
- Emphasis on examples with data, thanks to python package fake-data-for-learning
- Light on non-mathematical definitions, thanks to



Artificial intelligence in insurance

- Underwriting: selling insurance, e.g. using a *technical pricing model* that predicts expected claims
- Claims: dealing with client loss events, e.g. using algorithmic *straight-through-processing* or *fraud detection*
- Operations: process efficiency

Good Market Understanding Trumps Technology



Model selection: theory sketch

Goal: select the best model for data to come, i.e. minimize *generalization error*.

Basic steps

- Split data set into train, validation and test sets
- For each model family (e.g. logistic regression, gradient boosted tree classifier), select features and optimize model parameters on training data
- For each model family, evaluate fitted models on validation set with chosen metric (e.g. recall, accuracy) and choose best
- To estimate generalization error, evaluate the best model on the test set.

Model selection example: hit rate for insurance quotes

Fake data generated with fake-data-for-learning

product_type	days	rating	hit
property	3	1	0
financial	2	1	0
financial	1	1	0
financial	0	0	1
financial	0	1	0

Variables

- product_type: Client line of business
- days: Number of days to generate quote
- rating: Binary indication of client risk
- hit: Binary, 1 for success (binding the quote), 0 for failure

Model selection example: lightning data

Fake data generated with fake-data-for-learning

city	year	lightning_strike
Ljutomer	1990	0
Turnišče	1990	0
Velenje	1990	0
Vrhnika	1990	0
Jesenice	1990	0

Variables

- city: Slovenian city
- year: calendar year AD
- lightning_strike: Number of lightning strike deaths per year

Model selection example: university admissions data

Fake data generated with fake-data-for-learning

gender	dept	comp	admission
0	1	1	1
0	1	1	0
1	0	0	0
0	0	0	1
0	1	1	1

Variables

- gender of applicant: 0 for male, 1 for female
- department: integer code for university department to which applicant applied
- comp: binary code for competitiveness of department, 1 for competitive
- admission: binary, 0 for rejection, 1 for acceptance

Model selection: practice sketch

- Market / business need trumps model selection theory
- Don't cede sense to statistics
- Regularize
- Use pipelines to avoid model leakage