The exciting applications of machine learning in science

Shravan Hanasoge
Tata Institute of Fundamental Research

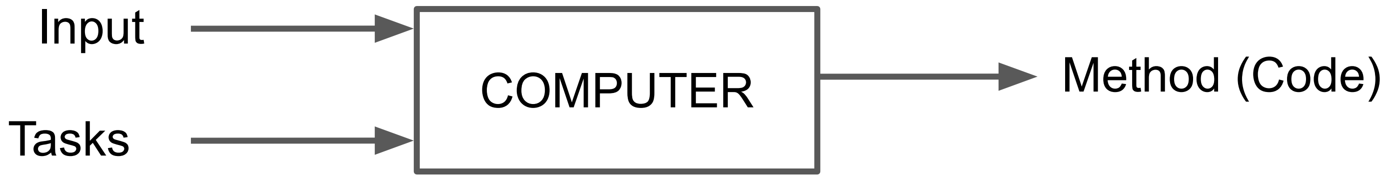
Machine learning is the most exciting technique in contemporary science

- ➔ “Learning is any process by which a system improves performance from experience.” - Herbert Simon (father of AI)
- ➔ “Machine Learning: Field of study that gives computers the ability to learn without being explicitly programmed.” -Arthur Samuel (1959)
- ➔ Tom Mitchell (1998): Machine Learning is the study of algorithms that
  - ◆ improve their performance P
  - ◆ at some task T
  - ◆ with experience E

Machines evolve and improve

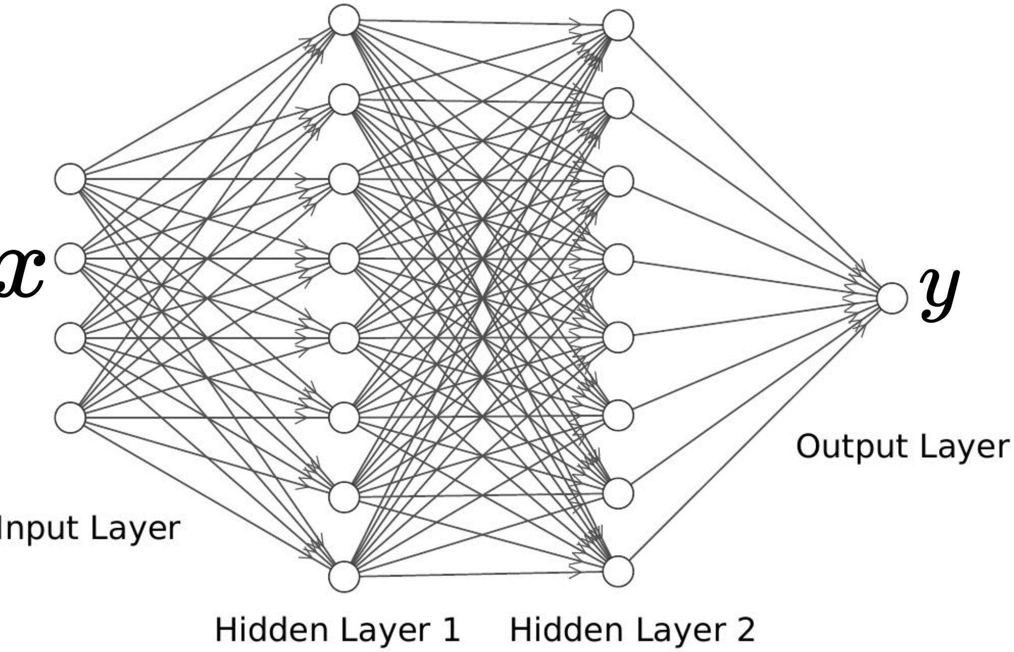
Traditional programming:

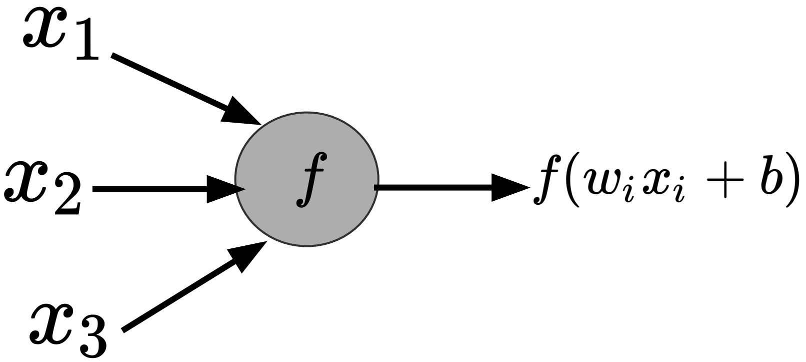
Machine Learning:



The machine learns whenever it changes its parameters based on its inputs or in response to external information in such a manner that its expected future performance on given tasks improves.

Neural Networks and deep learning can build complex regression models



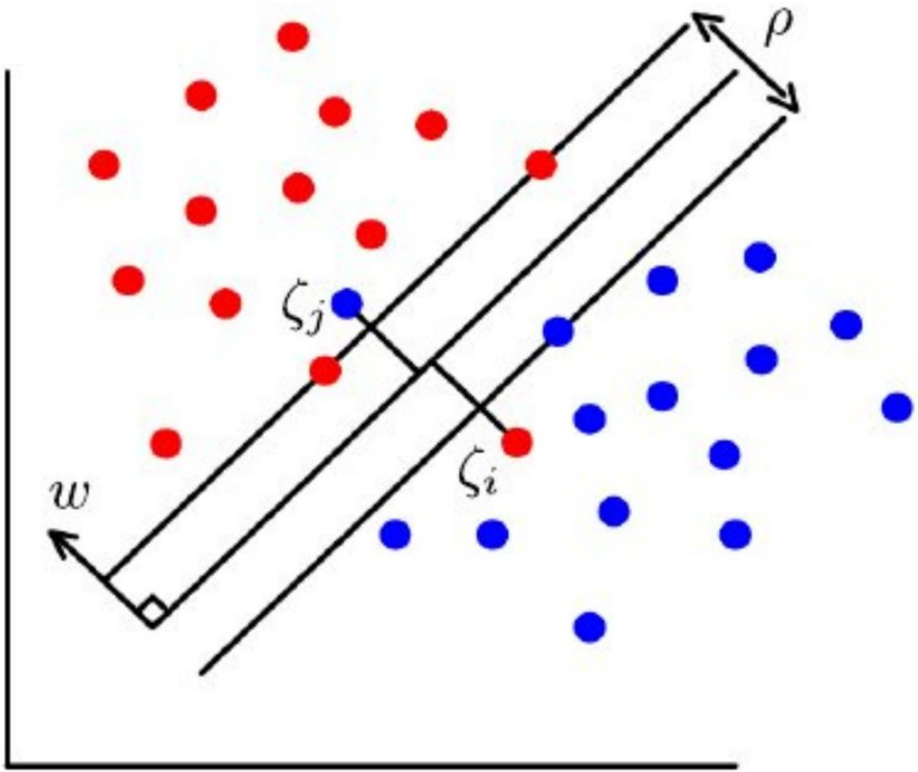


Weights and biases $\theta = \{w, b\}$

$Loss\ Function : L_\theta\left(y, y^{pred}\right)$

Weights and biases are obtained using gradient descent to minimize the *loss*.

Support Vector Machines are excellent classifiers

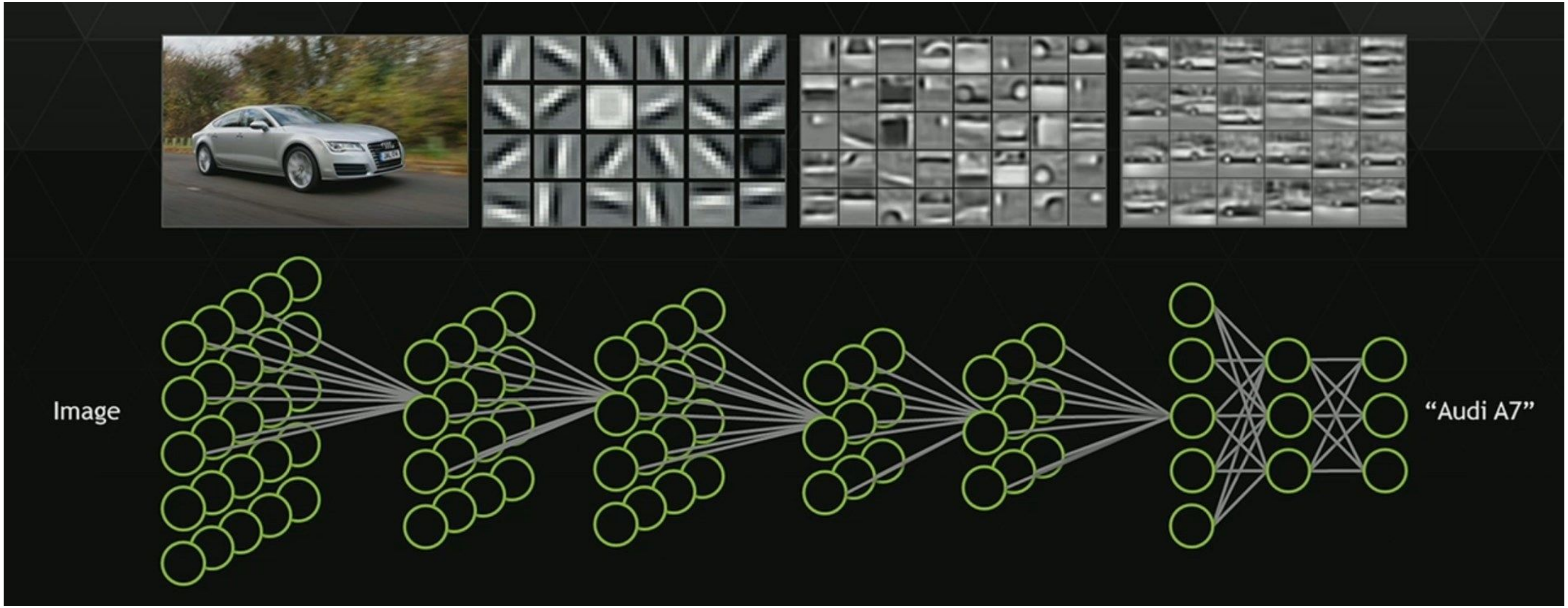


Maximize width of margin separating the two classes.

$$width\,of\,margin\,\rho = (1/2||w||)$$

$$min\left(||w||^2 + C_{blue}\sum_{i_{blue}}\zeta_{i_{blue}} + C_{red}\sum_{i_{red}}\zeta_{i_{red}}\right)$$

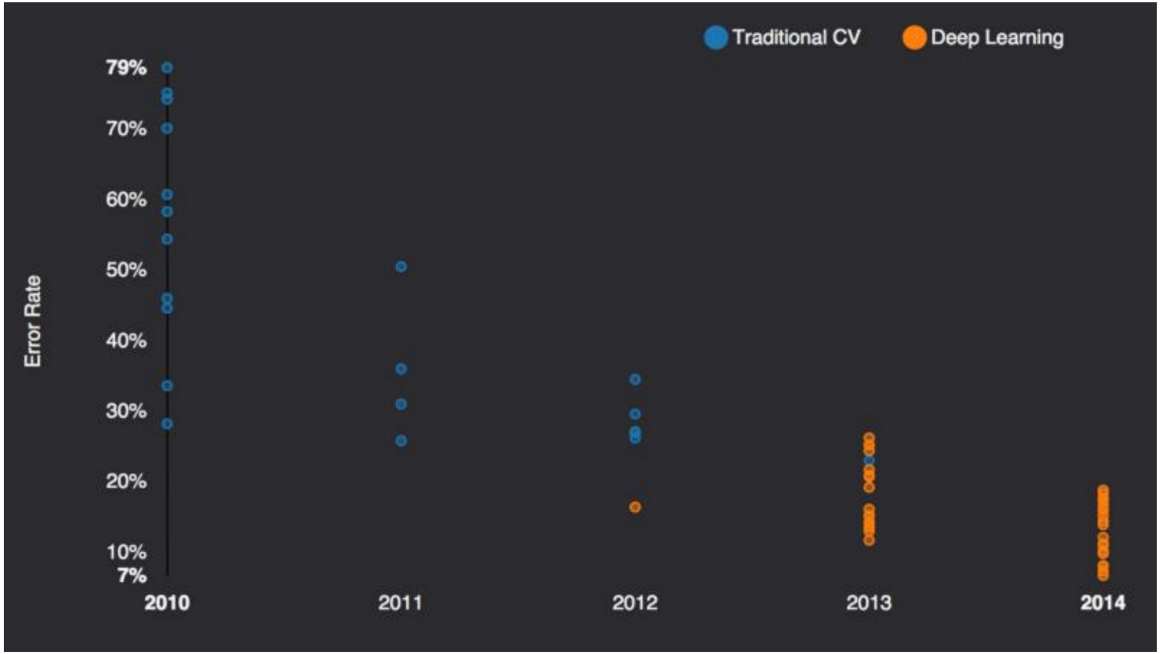
Convolutional Neural Networks perform high-quality image recognition



The goal of this talk is to briefly go over some notable results, challenges and techniques

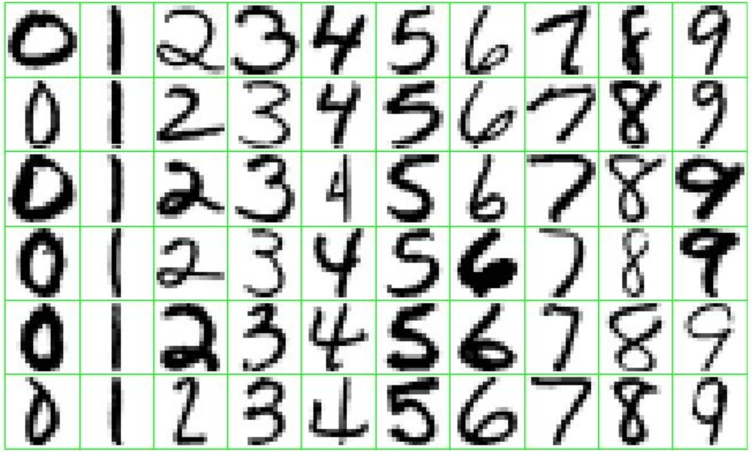
- ➔ What machines are capable of
  - ◆ Examples of remarkable applications
- ➔ Types of algorithms
- ➔ Challenges in interpretation
- ➔ A specific science application
- ➔ Future directions

GPUs have revolutionized machine learning



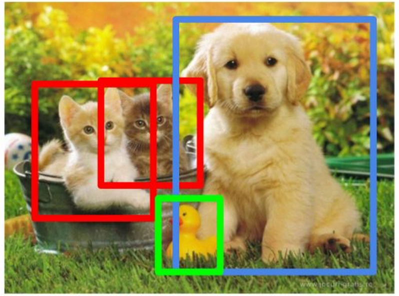
AI performance at ImageNet competition

Machines are capable of many tasks ...

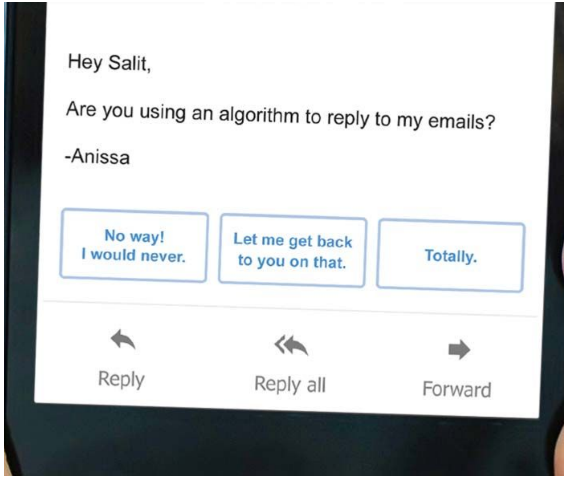


Classification

Object Detection



CAT, DOG, DUCK



Text Generation

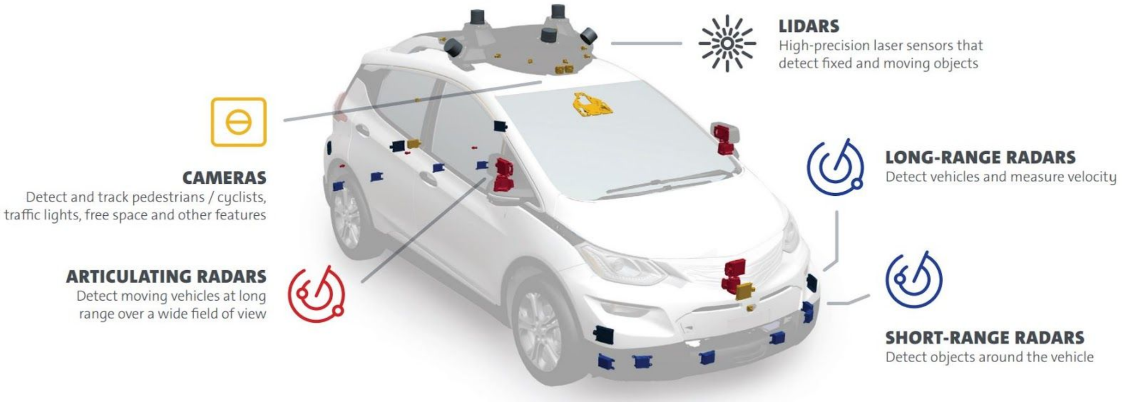
Image Generation

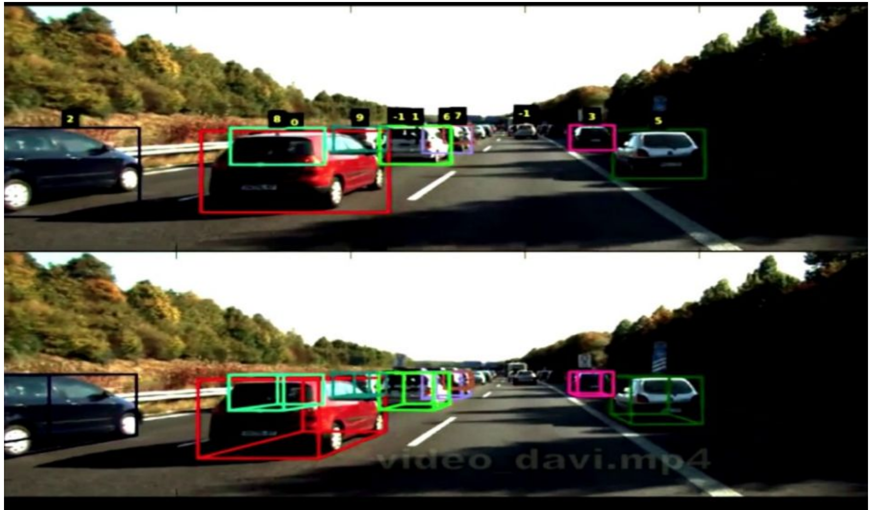
Photograph → Monet | Van Gogh | Cezanne | Ukiyo-e

...some of which are simply too complex for conventional techniques

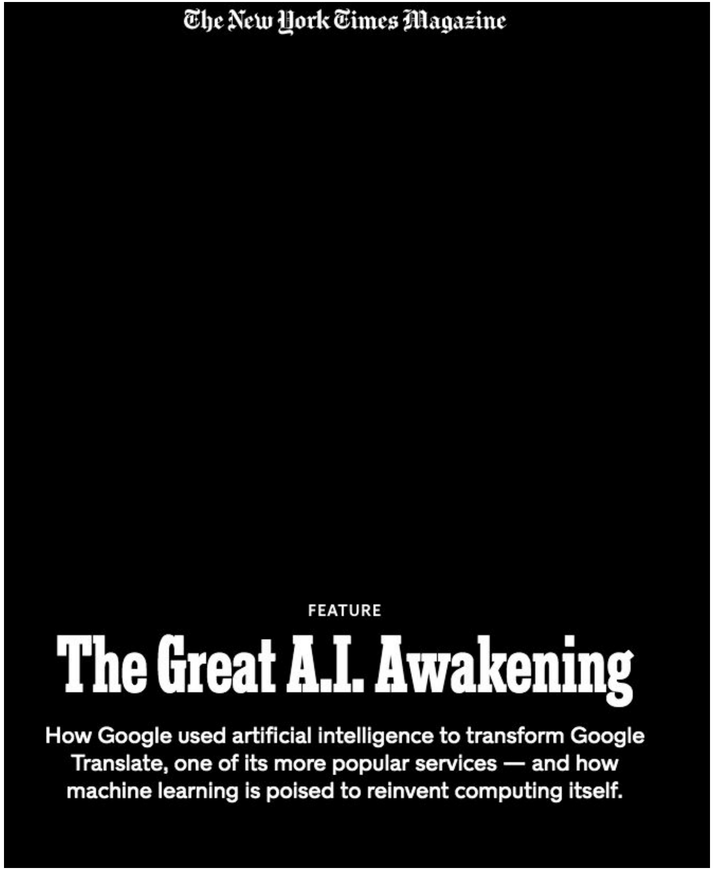
- ➔ Some tasks can only be defined by giving examples i.e., we might be able to specify (input,output) pairs but not a precise relationship between them.
- ➔ ***Data Mining*** Important relationships and correlations are hidden within large piles of data.
- ➔ ***Big Data*** The amount of knowledge available about certain tasks might be too large for explicit encoding by humans.
- ➔ ***Unknown/dynamic environment*** It is difficult to comprehensively incorporate all characteristics of the working environment at the design time.

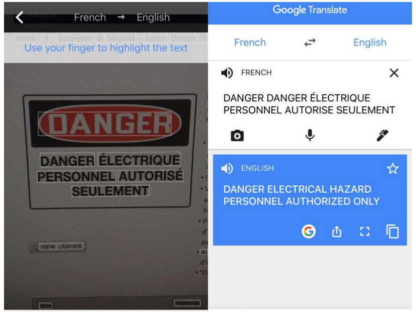
Cars can now drive themselves

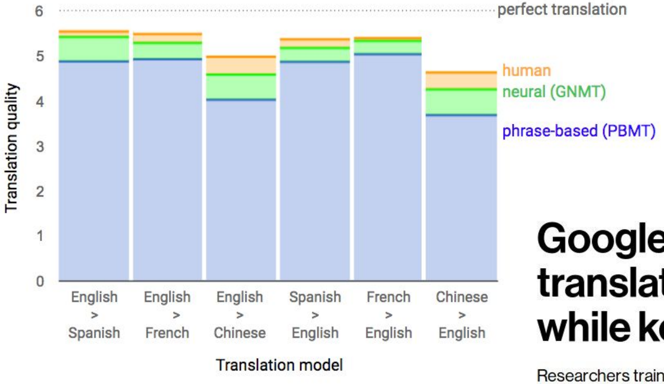




Machines translate languages with high fidelity







Google's AI can now translate your speech while keeping your voice

Researchers trained a neural network to map audio "voiceprints" from one language to another.

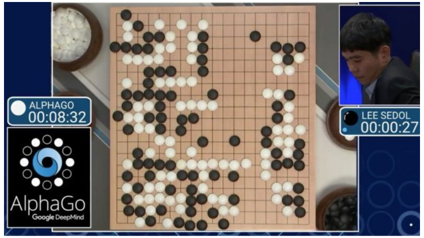
by **Karen Hao** May 20, 2019

Machines turn creative and competitive

```
PANDARUS:
Alas, I think he shall be come approached and the day
When little srain would be attain'd into being never fed,
And who is but a chain and subjects of his death,
I should not sleep.

Second Senator:
They are away this miseries, produced upon my soul,
Breaking and strongly should be buried, when I perish
The earth and thoughts of many states.

DUKE VINCENTIO:
Well, your wit is in the care of side and that.

Second Lord:
They would be ruled after this chamber, and
my fair nues begun out of the fact, to be conveyed,
Whose noble souls I'll have the heart of the wars.

Clown:
Come, sir, I will make did behold your worship.

VIOLA:
I'll drink it.
```

Shakespeare Styled Text

Reinforcement Learning



INTERNATIONAL

'Like A God,' Google A.I. Beats Human Champ Of Notoriously Complex Go Game

May 23, 2017 · 1:38 PM ET

COLIN DWYER

SHARE

REPORT

A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play

David Silver[1,2,*,†], Thomas Hubert[1,*], Julian Schrittwieser[1,*], Ioannis Antonoglou[1], Matthew Lai[1], Arthur Guez[1], Marc Lancto...

+ See all authors and affiliations

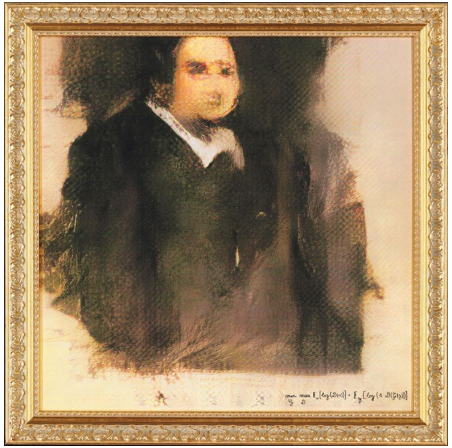
Science 07 Dec 2018:
Vol. 362, Issue 6419, pp. 1140-1144
DOI: 10.1126/science.aar6404

Article | Figures & Data | Info & Metrics | eLetters | PDF

One program to rule them all

Computers can beat humans at increasingly complex games, including chess and Go. However, these programs are typically constructed for a particular game, exploiting its properties, such as the symmetries of the board on which it is played. Silver *et al.* developed a program called AlphaZero, which taught itself to play Go, chess, and shogi (a Japanese version of chess) (see the Editorial, and the Perspective by Campbell). AlphaZero managed to beat state-of-the-art programs specializing in these three games. The ability of AlphaZero to adapt to various game rules is a notable step toward achieving a general game-playing system.

Science, this issue p. 1140; see also pp. 1087 and 1118



Artwork by Generative Adversarial Networks (GANs)

Machine learning enables new science

nature physics

LETTERS

PUBLISHED ONLINE: 13 FEBRUARY 2017 | DOI: 10.1038/NPHYS4035

Machine learning phases of matter

Juan Carrasquilla[1*] and Roger G. Melko[1,2]

THE ASTRONOMICAL JOURNAL, 155:94 (21pp), 2018 February https://doi.org/10.3847/1538-3881/aa9e09

© 2018. The American Astronomical Society.

OPEN ACCESS

CrossMark

Identifying Exoplanets with Deep Learning: A Five-planet Resonant Chain around Kepler-80 and an Eighth Planet around Kepler-90

Christopher J. Shallue[1] and Andrew Vanderburg[2,3,4]

[1] Google Brain, 1600 Amphitheatre Parkway, Mountain View, CA 94043, USA; shallue@google.com

[2] Department of Astronomy, The University of Texas at Austin, 2515 Speedway, Stop C1400, Austin, TX 78712, USA

[3] Harvard–Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

Received 2017 September 19; revised 2017 November 13; accepted 2017 November 20; published 2018 January 30

MOTHERBOARD

TECH BY VICE

AI Trained on Old Scientific Papers Makes Discoveries Humans Missed

Scientists used machine learning to reveal new scientific knowledge hidden in old research papers.

PHYSICAL REVIEW X **8**, 031084 (2018)

Reinforcement Learning with Neural Networks for Quantum Feedback

Thomas Fösel, Petru Tighineanu, and Talitha Weiss

Max Planck Institute for the Science of Light, Staudtstraße 2, 91058 Erlangen, Germany

Florian Marquardt

Max Planck Institute for the Science of Light, Staudtstraße 2, 91058 Erlangen, Germany and Physics Department, University of Erlangen-Nuremberg, Staudtstraße 5, 91058 Erlangen, Germany

nature medicine

LETTERS

https://doi.org/10.1038/s41591-018-0335-9

Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence

nature

International journal of science

Letter | Published: 03 July 2019

Unsupervised word embeddings capture latent knowledge from materials science literature

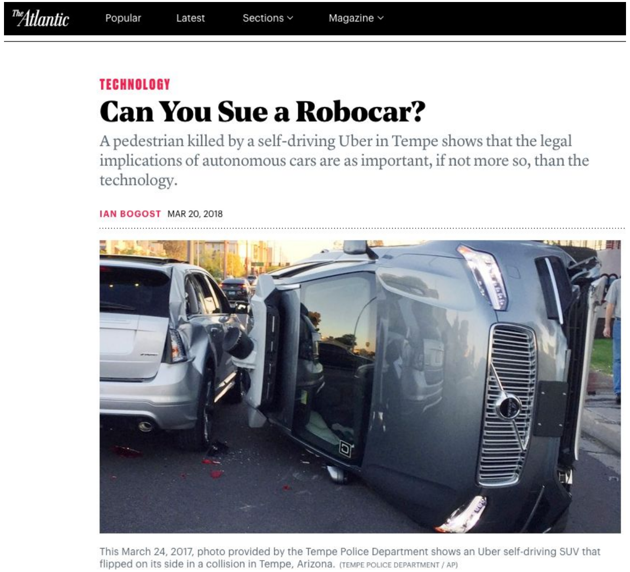
Vahe Tshitoyan, John Dagdelen, Leigh Weston, Alexander Dunn, Ziqin Rong, Olga Kononova, Kristin A. Persson, Gerbrand Ceder & Anubhav Jain

Nature **571**, 95–98 (2019) | Download Citation

There are many types of ML algorithms

- Supervised Learning - Labelled Data
  - Classification, regression
  - Semi-supervised learning
  - Weakly supervised learning
- Unsupervised Learning - Unlabelled data
  - Clustering, anomaly detection, generation
  - Self-supervised learning
- Reinforcement Learning

Machine interpretation is a major challenge


The Atlantic   Popular   Latest   Sections   Magazine

TECHNOLOGY

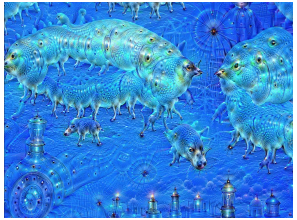
Can You Sue a Robocar?

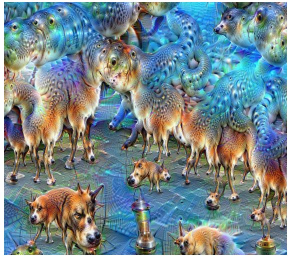
A pedestrian killed by a self-driving Uber in Tempe shows that the legal implications of autonomous cars are as important, if not more so, than the technology.

IAN BOGOST   MAR 20, 2018

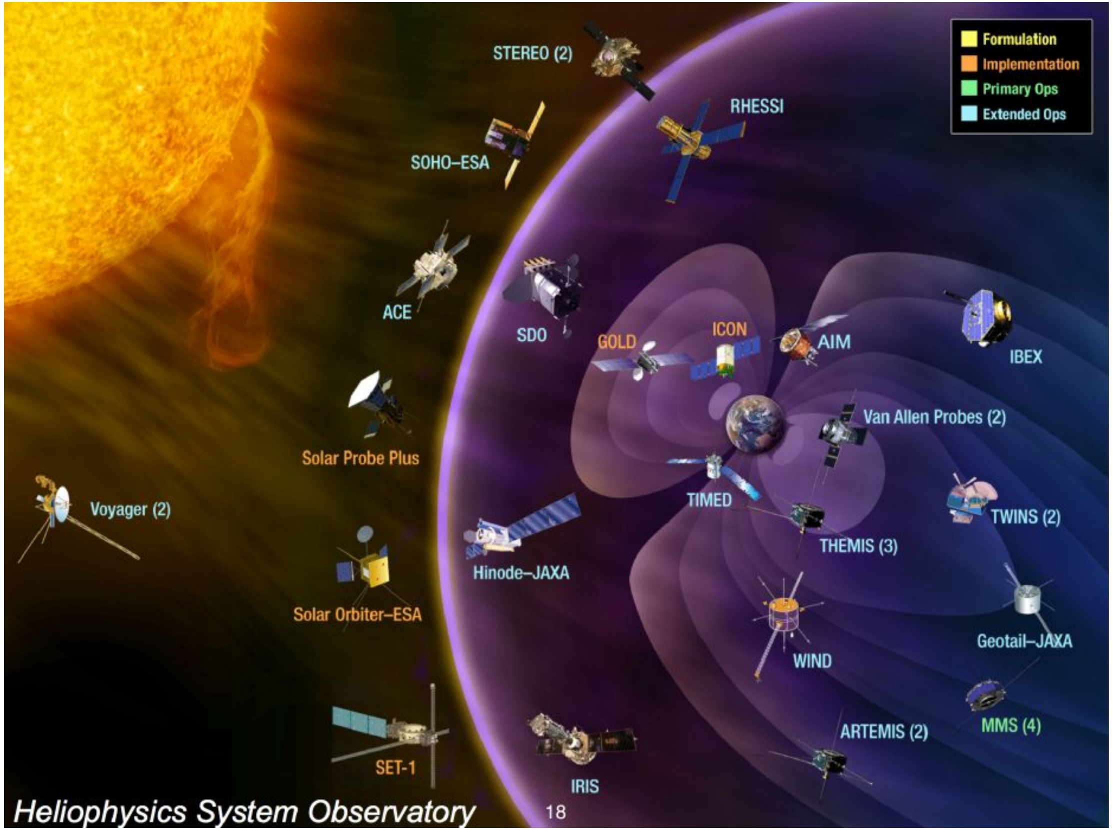
This March 24, 2017, photo provided by the Tempe Police Department shows an Uber self-driving SUV that flipped on its side in a collision in Tempe, Arizona. (TEMPE POLICE DEPARTMENT / AP)

Google deep dream's hallucinatory images



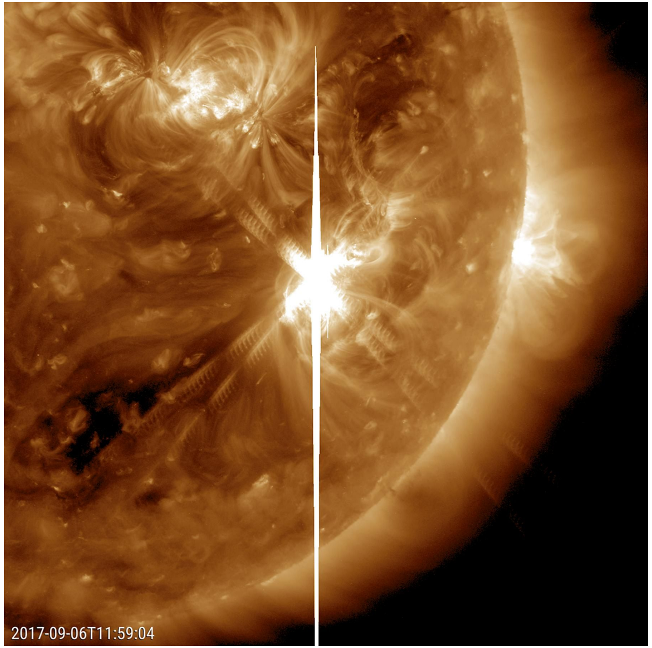


Solar missions are exciting - Big Data, big opportunities.



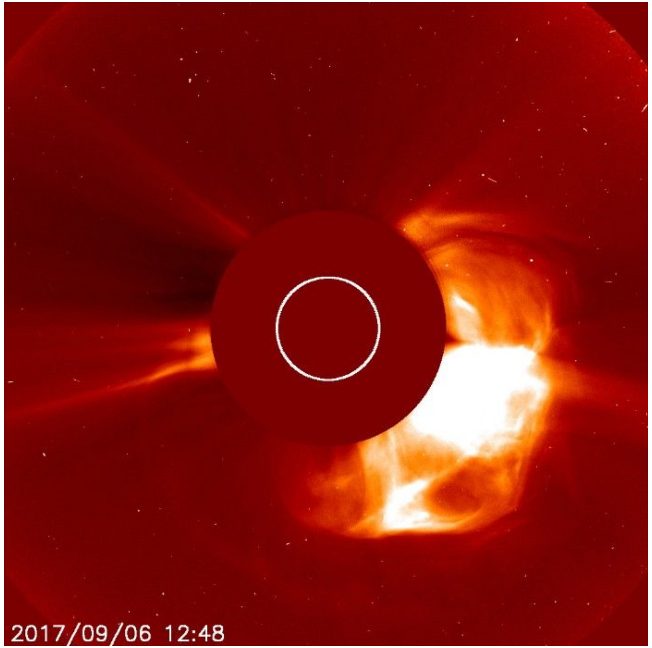
Solar magnetic activity drives space weather.

Solar Flares



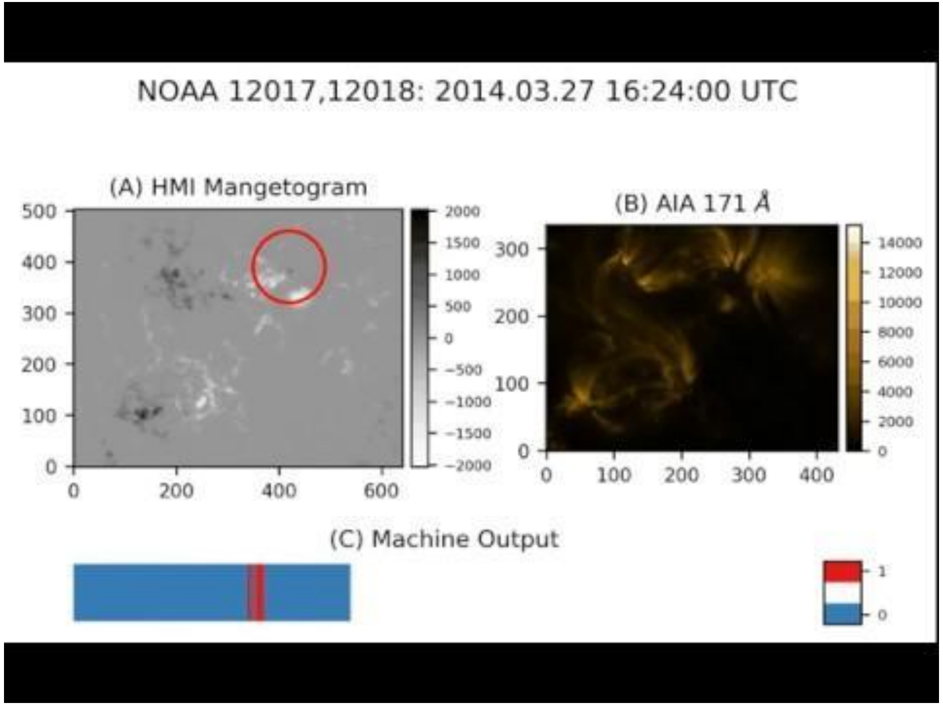
No warning

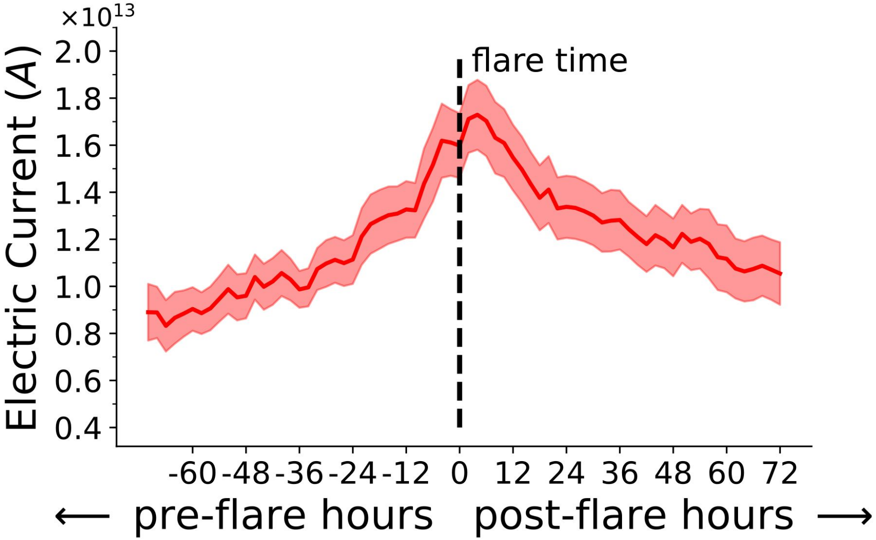
Coronal Mass Ejections



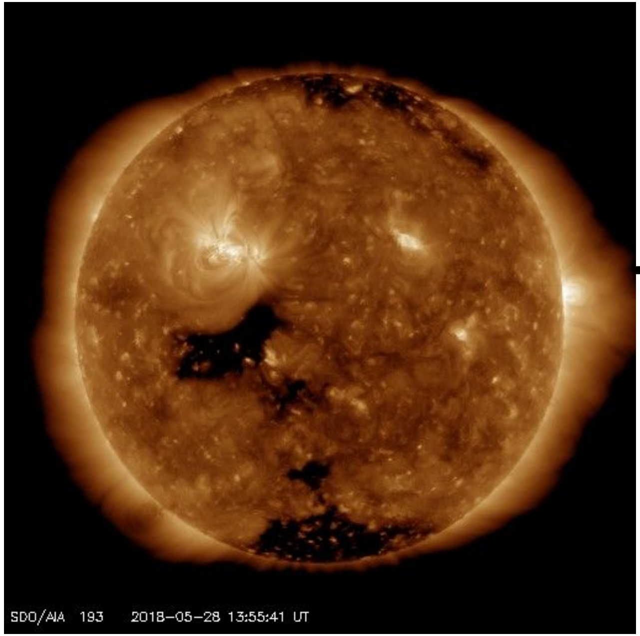
~20 hour warning

ML for identifying flare-productive magnetic regions.





Forecasting solar wind properties at L1.



SDO/AIA 193

X

~ 2 - 4 days

L1

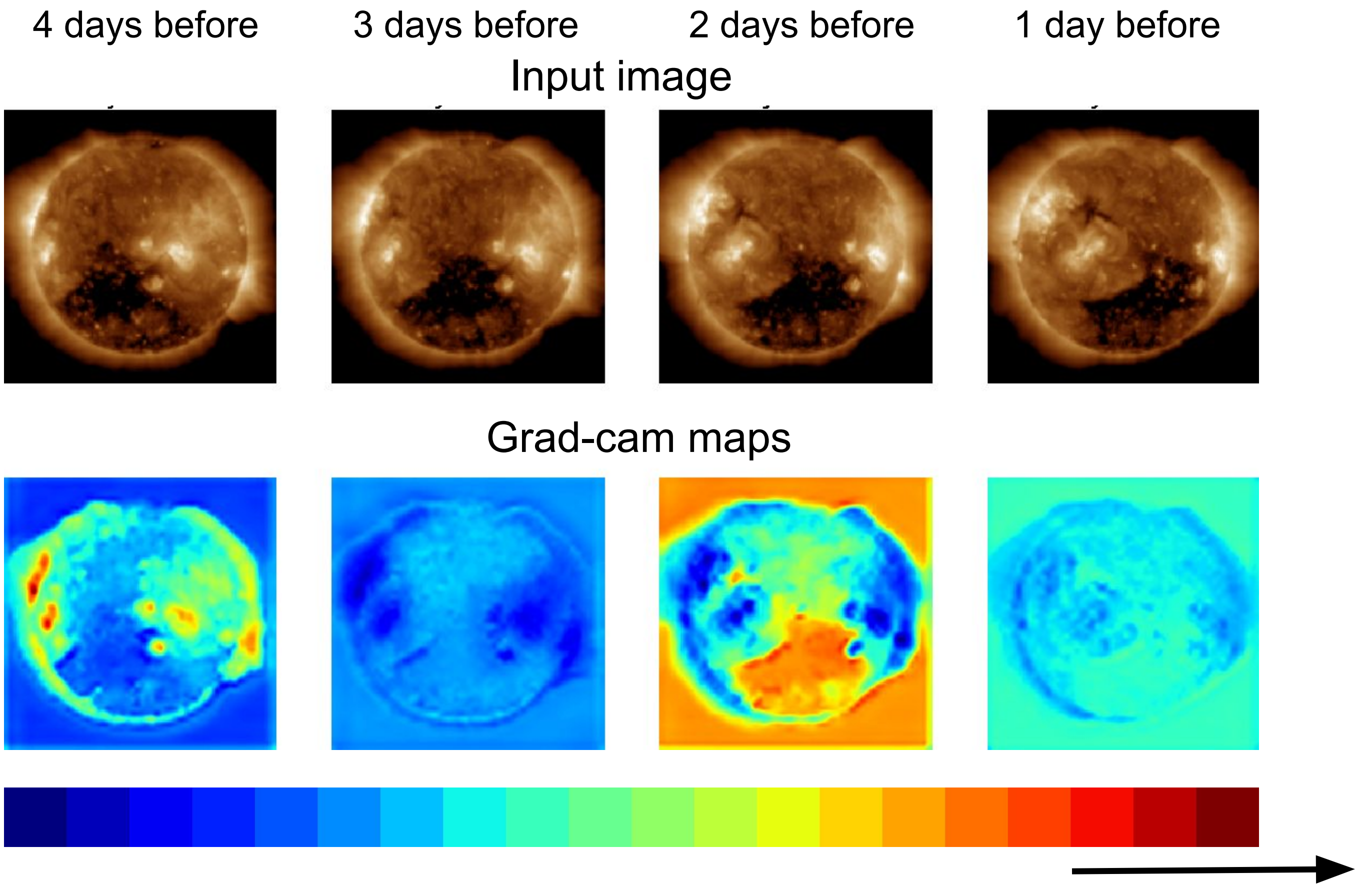
~ 30 min - 1 hr

Solar wind

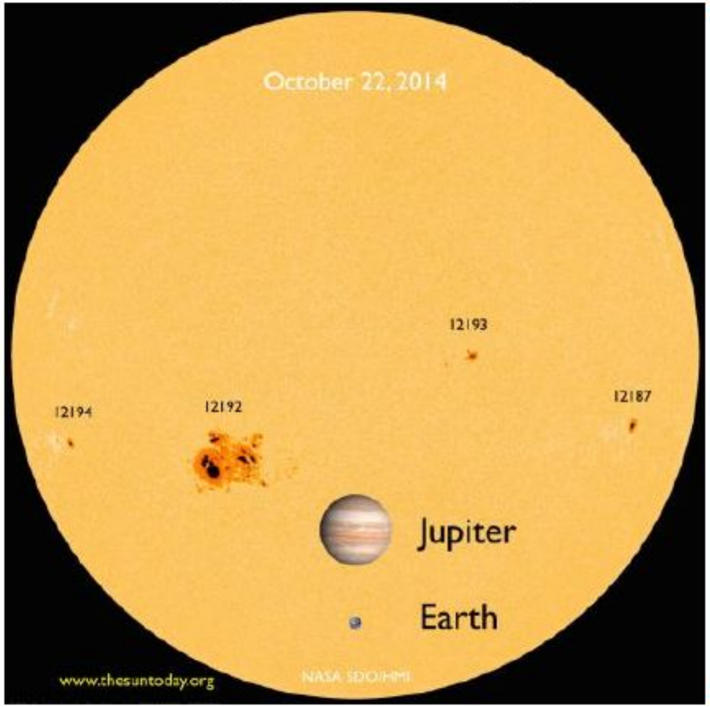
- ❑ Speed
- ❑ Flux
- ❑ Magnetic field
- •
- •
- •

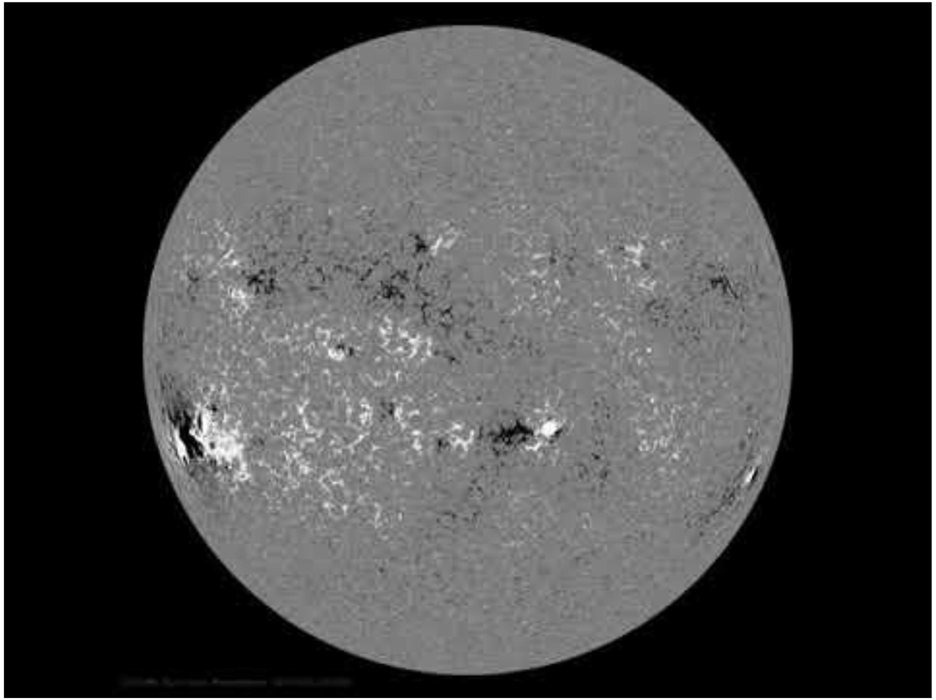
Y

WindNet recognizes origin of fast and slow solar wind.



ML may help appreciate magnetic region emergence



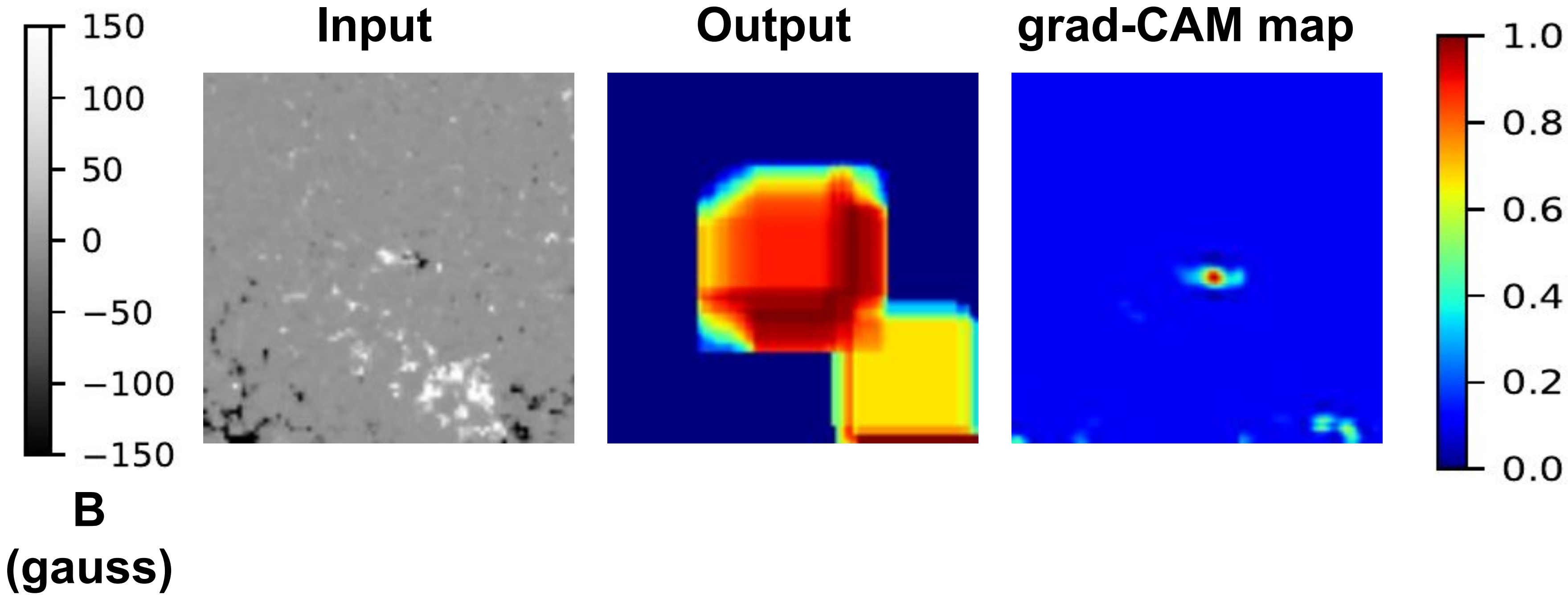


Future of fundamental ML science

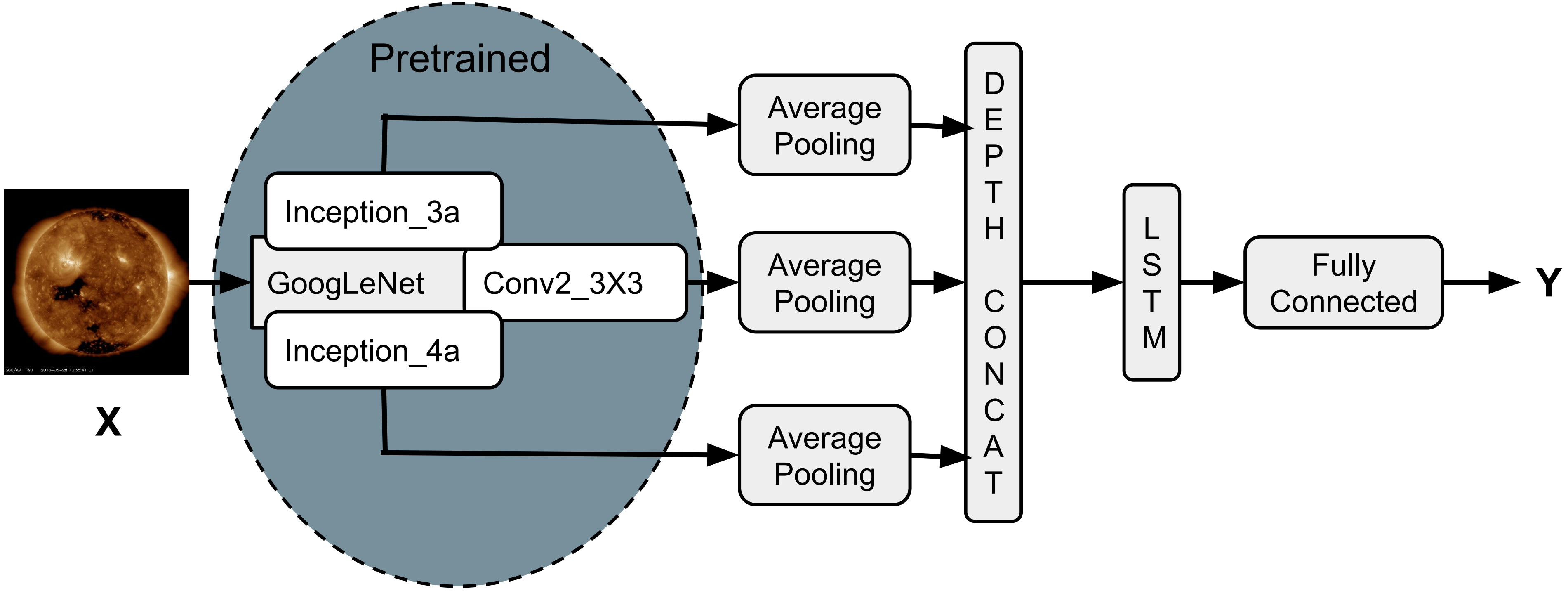
- ➔ Unsupervised Learning - Unlabelled data
  - ◆ Clustering, anomaly detection, generation
  - ◆ Self-supervised learning
- ➔ Explainable AI

Self-driving cars.

Machine interpretation is challenging.



WindNet for forecasting solar wind properties.



Backpropagation

$$z_j^n \equiv w_{ji}^n a_i^{n-1} + b_j^n \qquad a_j^n = f(z_j^n) \qquad \delta_j^n \equiv \partial L / \partial z_j^n$$

$$\left(\partial L / \partial w_{ji}^n\right) = \delta_j^n a_i^{n-1} \qquad \left(\partial L / \partial b_j^n\right) = \delta_j^n$$

For final layer: $$\delta_j^N = \left(\partial L / \partial a_j^N\right) f'\left(z_j^N\right)$$

For intermediate layer:
$$\begin{aligned}\delta_j^{n-1} &= \left(\partial L / \partial a_j^{n-1}\right) f'\left(z_j^n\right) \\ &= \left(\partial z_j^n / \partial a_j^{n-1}\right)\left(\partial L / \partial z_j^n\right) f'\left(z_j^n\right) \\ &= w_{ij}^n \delta_i^n f'\left(z_j^n\right)\end{aligned}$$

Generative Adversarial Networks (GANs)

Algorithm 1 Minibatch stochastic gradient descent training of generative adversarial nets. The number of steps to apply to the discriminator, k, is a hyperparameter. We used $k = 1$, the least expensive option, in our experiments.

for number of training iterations **do**

for k steps **do**

- Sample minibatch of m noise samples $\{\boldsymbol{z}^{(1)}, \dots, \boldsymbol{z}^{(m)}\}$ from noise prior $p_g(\boldsymbol{z})$.
- Sample minibatch of m examples $\{\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(m)}\}$ from data generating distribution $p_{\text{data}}(\boldsymbol{x})$.
- Update the discriminator by ascending its stochastic gradient:

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^{m} \left[\log D\left(\boldsymbol{x}^{(i)}\right) + \log\left(1 - D\left(G\left(\boldsymbol{z}^{(i)}\right)\right)\right) \right].$$

end for

- Sample minibatch of m noise samples $\{\boldsymbol{z}^{(1)}, \dots, \boldsymbol{z}^{(m)}\}$ from noise prior $p_g(\boldsymbol{z})$.
- Update the generator by descending its stochastic gradient:

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^{m} \log\left(1 - D\left(G\left(\boldsymbol{z}^{(i)}\right)\right)\right).$$

end for

The gradient-based updates can use any standard gradient-based learning rule. We used momentum in our experiments.