Machine Learning in Physics and Astronomy

Kartheik Iyer, John Wu, Raghav Kunnawalkam Elayavalli
Rutgers University
SSPAR Oct 5th 2017

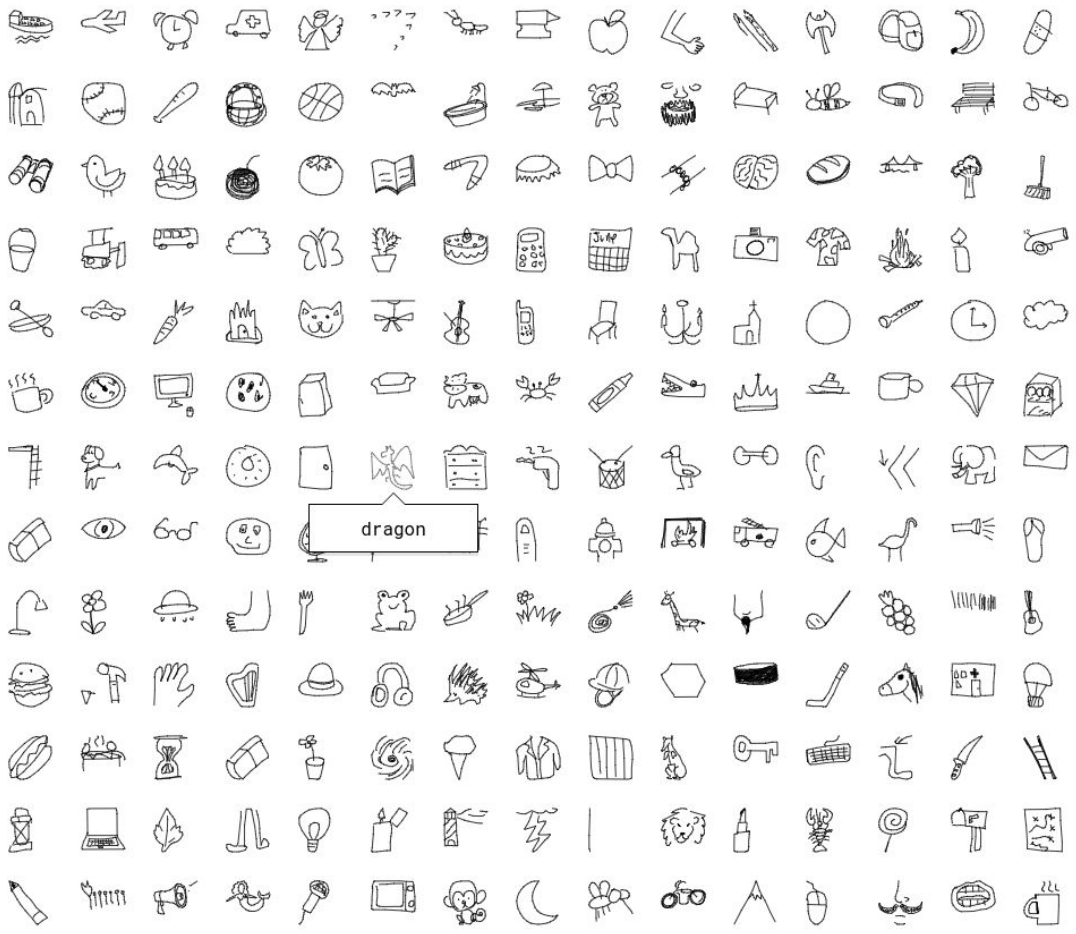
What is machine learning?

Dealing with incomplete or empirical physics. - the cutting edge is always unknown.

Dealing with an overload of data, often noisy, biased and incomplete.

Dealing with repeatable processes that can’t be described by simple linear relations.

Automating ourselves back into manual labor

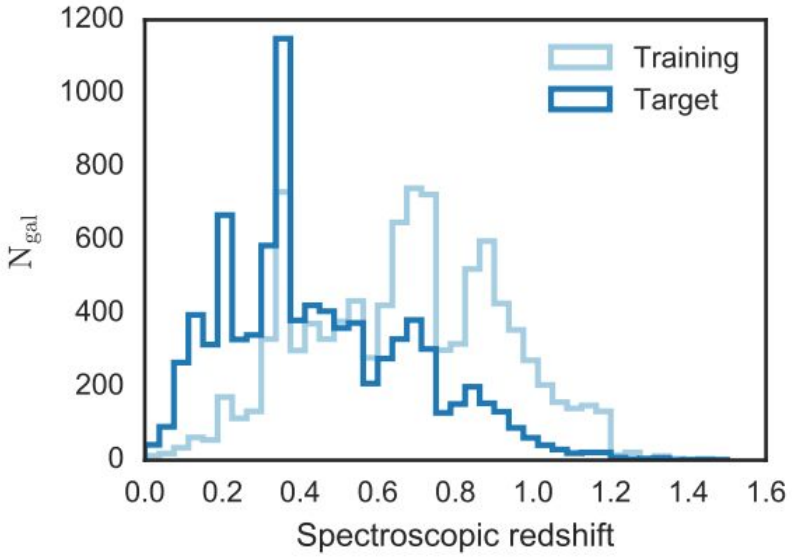


Picture from: https://quickdraw.withgoogle.com/data

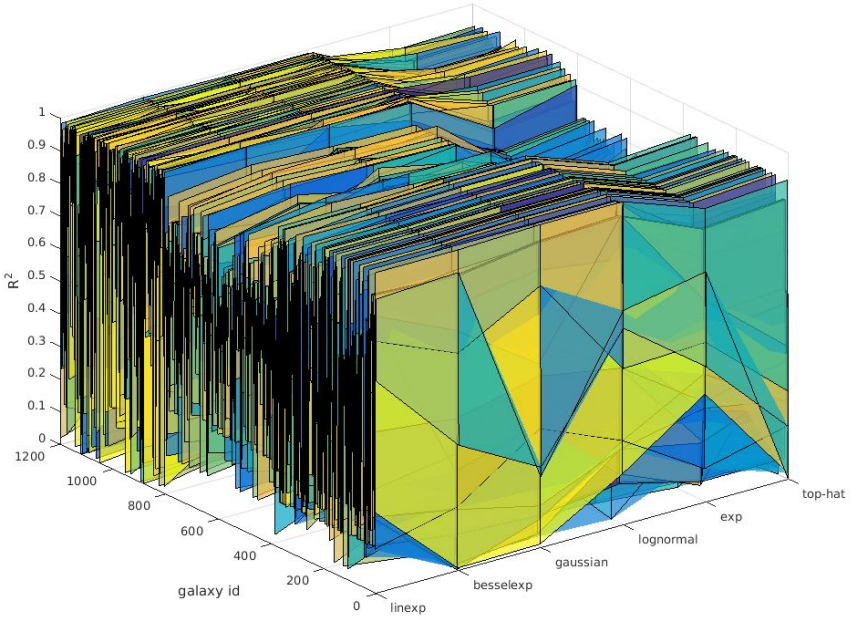
Why use ML? Do we need it in physics?

Galaxy spectra -> Stellar mass, Star Formation Rate, Redshift.. and more

Problems: highly nonlinear relations, increasingly degenerate as we go to older ages, noisy, spec-z distribution not representative of larger photo-z sample.



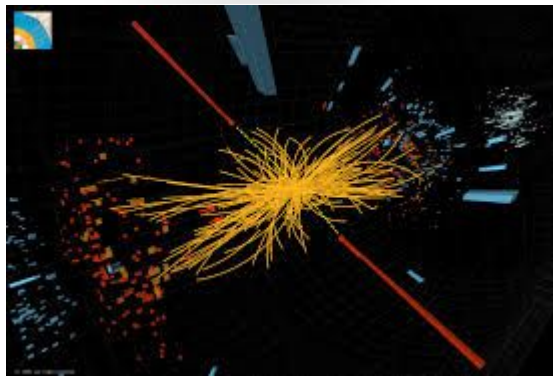
Phase transitions in complex systems often don't have analytic solutions. Additionally, simulating these systems often suffers from exponential growth of the space of possible configurations.



Do we need it in physics? - II

Techniques coming of age - proverbial black box starting to open …

Experiments at the LHC are essentially cameras - Producing pretty pictures



Datasets are really really huge and signal is very small

New physics is elusive! We are searching for something that we do not know what it looks like!

We want something thats faster, better and essentially new and doesnt involve grad students running code for a very long time !

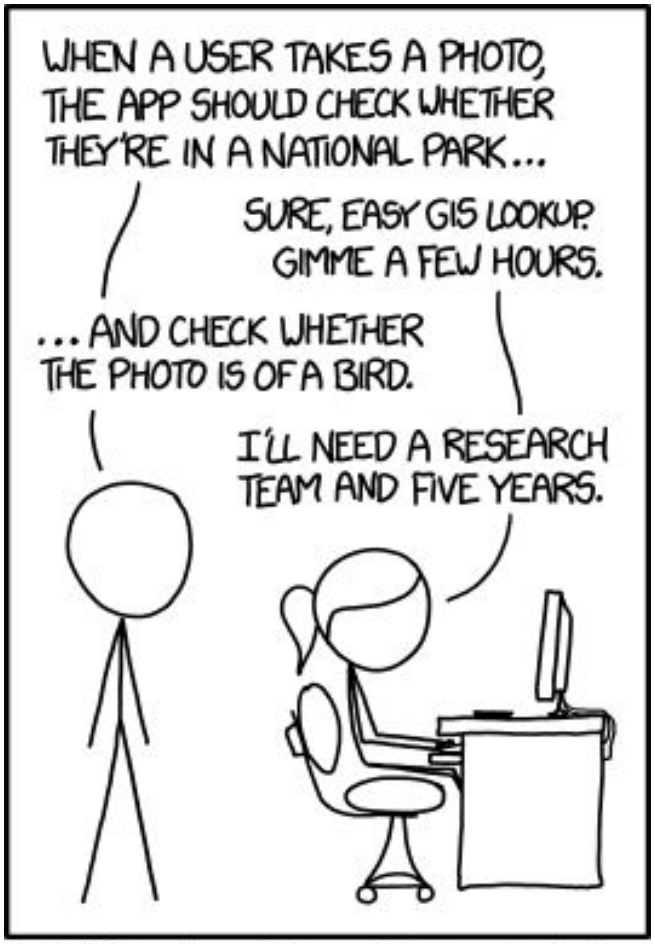
Might as well get comfortable with our future overlords

What is deep learning? (and why do we care?)

In cases with:

- Highly nonlinear problems
- Modeling time constraints
- A lack of knowledge about feature space
- The need for accurate forecasting without creating a complete model…

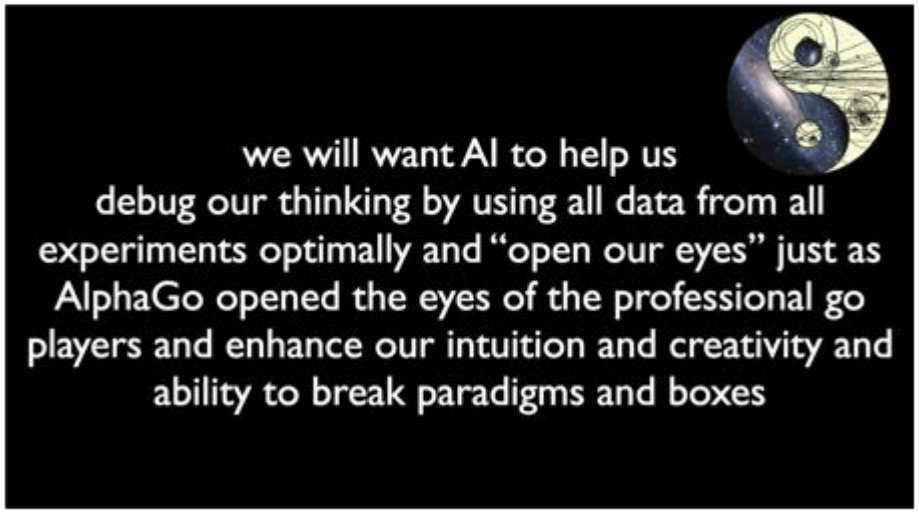
Build a network with many layers, that won't die when trained.



Technical vs practical machine learning

Machine learning is the subfield of computer science that, according to Arthur Samuel, gives "computers the ability to learn without being explicitly programmed."

Machine learning - Wikipedia
https://en.wikipedia.org/wiki/Machine_learning



Maria Spiropulu (Caltech)

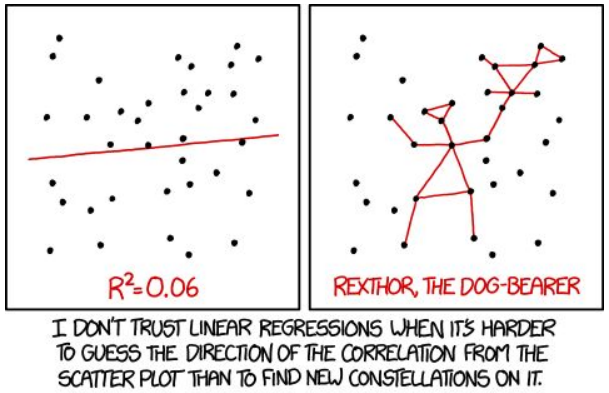
Two main class of problems we deal with -

Classification

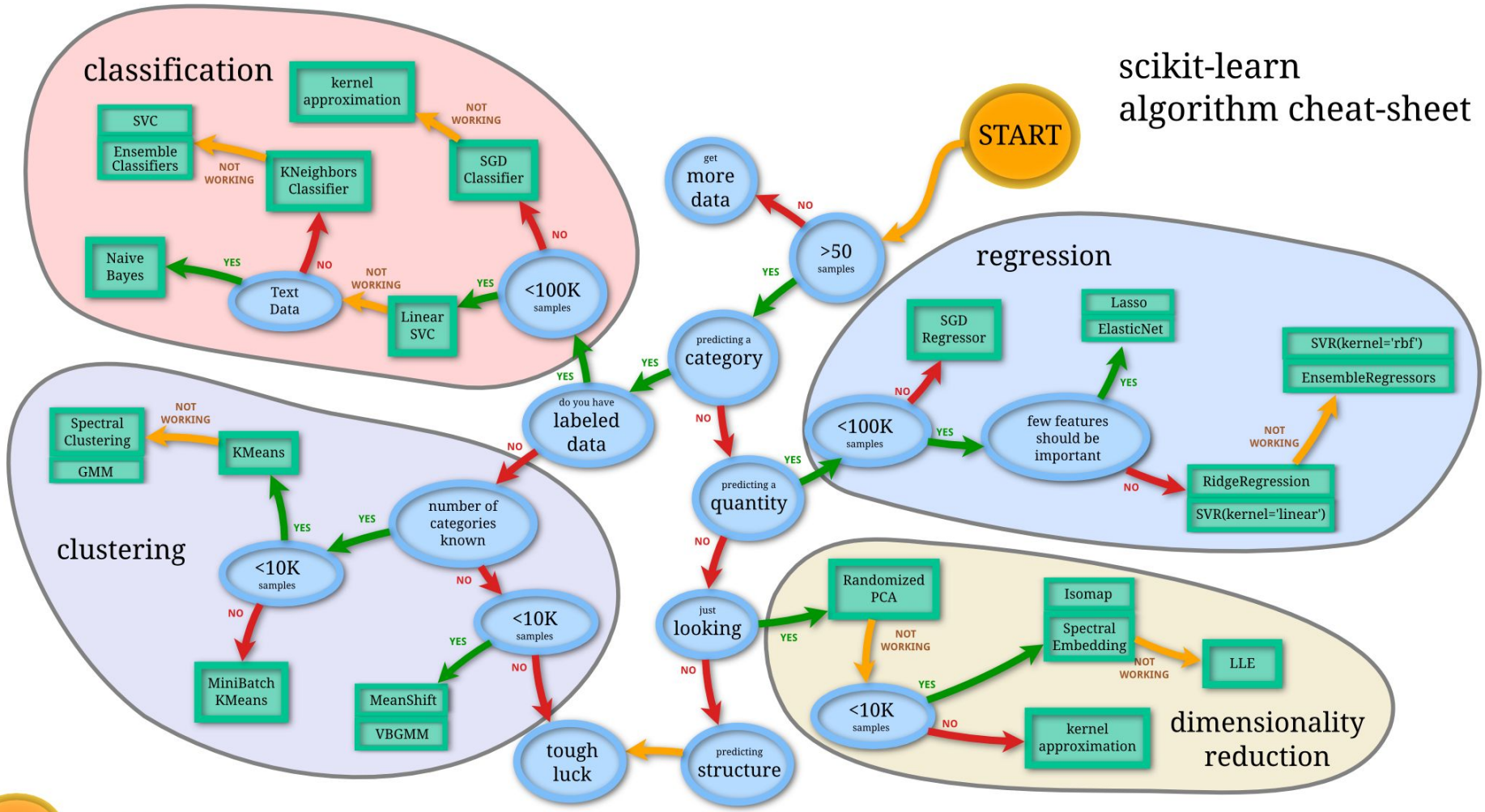
- Identify if an object belongs to one of N subgroups
- Divide objects into distinct classes and find the discriminating feature(s)
- Identify outliers / class of interest in a dataset

Regression

- Estimate the relation between observables and quantities of interest
- Both parametric (eg. fitting a line to data) and nonparametric (eg. splining / kriging)
- Interpolation and extrapolation
- Prediction and forecasting.



Resources [just google it] and Scikit-learn:::



Three terms: [Training, Testing, Validation]

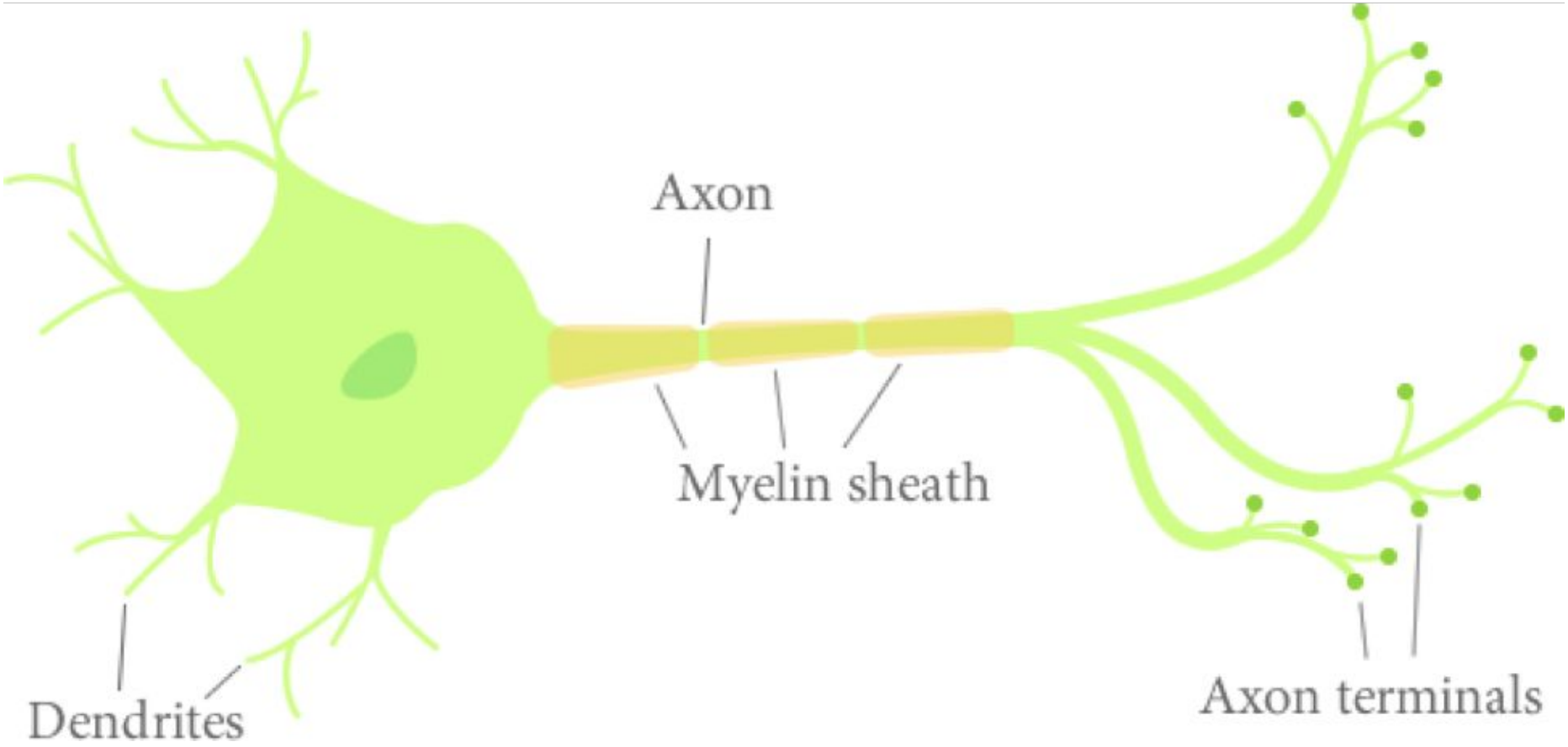
Training - giving (labeled or unlabeled) data to your method and letting it find a mapping between input and output variables

Validation - checking to see if this mapping still works when applied to data not in the training set. By being clever about this we can avoid overfitting - creating a mapping that describes the training data completely (noise and all) and nothing else.

Testing - after the training is done, this last piece of data is used to check if the mapping we've got works - determines the predictive power of the ML

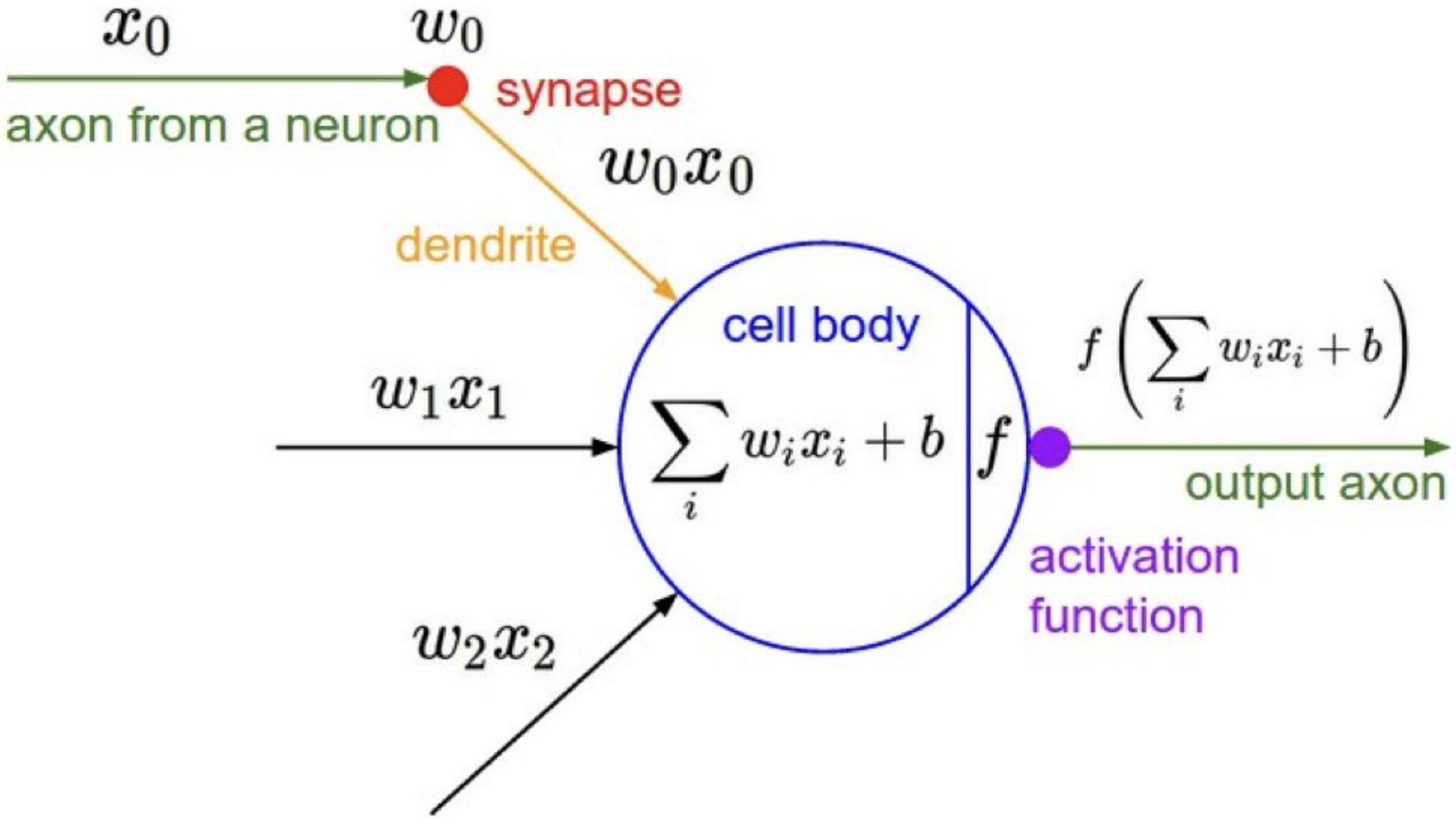
Now for some biology

Real Neuron



https://leonardoaraujosantos.gitbooks.io/artificial-inteligence/content/neural_networks.html

Artificial Neuron



https://leonardoaraujosantos.gitbooks.io/artificial-inteligence/content/neural_networks.html

Simple neural network

Single hidden layer with one output layer

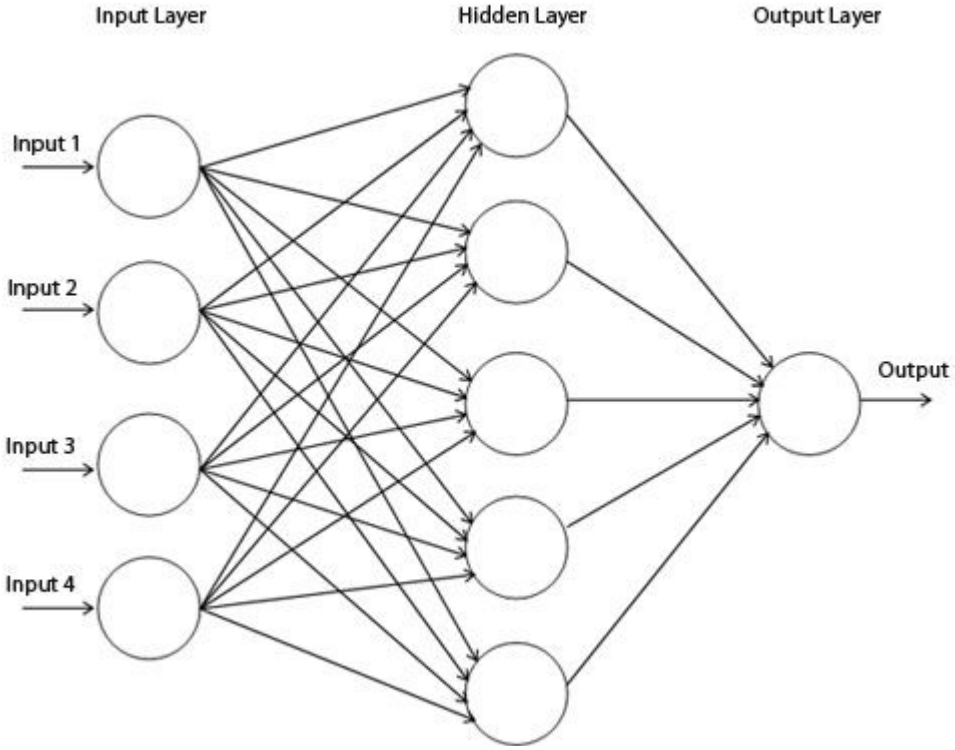
Fully connected - Each node in the hidden layer has an input from the input layer

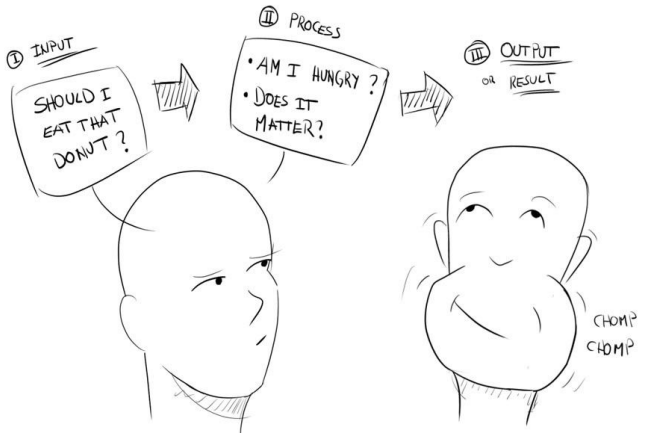
Total number of parameters : ?

4 x 5 + 5 + 5 + 1= 31 trainable parameters

Activation functions are dependent on your problem at hand.

- What is are you training against?
- Is your feature symmetric?
- Is it bounded? Binary?





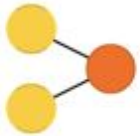
Comics from becoming human

A mostly complete chart of

Neural Networks

- Backfed Input Cell
- Input Cell
- Noisy Input Cell
- Hidden Cell
- Probablistic Hidden Cell
- Spiking Hidden Cell
- Output Cell
- Match Input Output Cell
- Recurrent Cell
- Memory Cell
- Different Memory Cell
- Kernel
- Convolution or Pool

Perceptron (P)

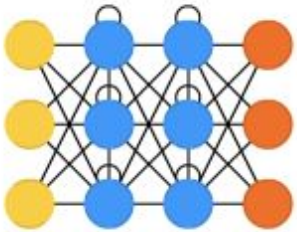


Feed Forward (FF)

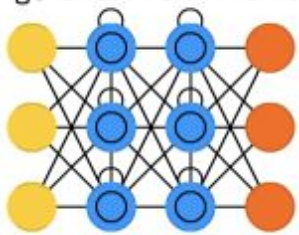
Radial Basis Network (RBF)

Deep Feed Forward (DFF)

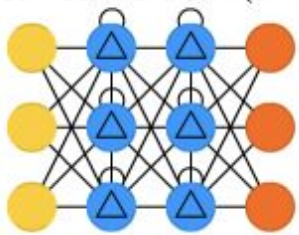
Recurrent Neural Network (RNN)



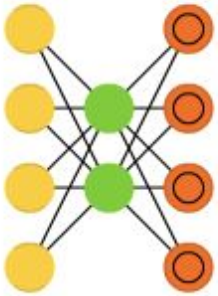
Long / Short Term Memory (LSTM)



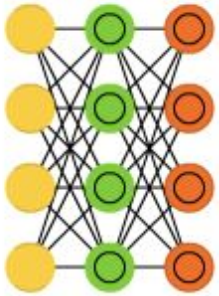
Gated Recurrent Unit (GRU)



Auto Encoder (AE)

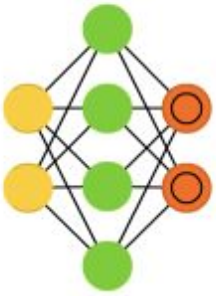


Variational AE (VAE)

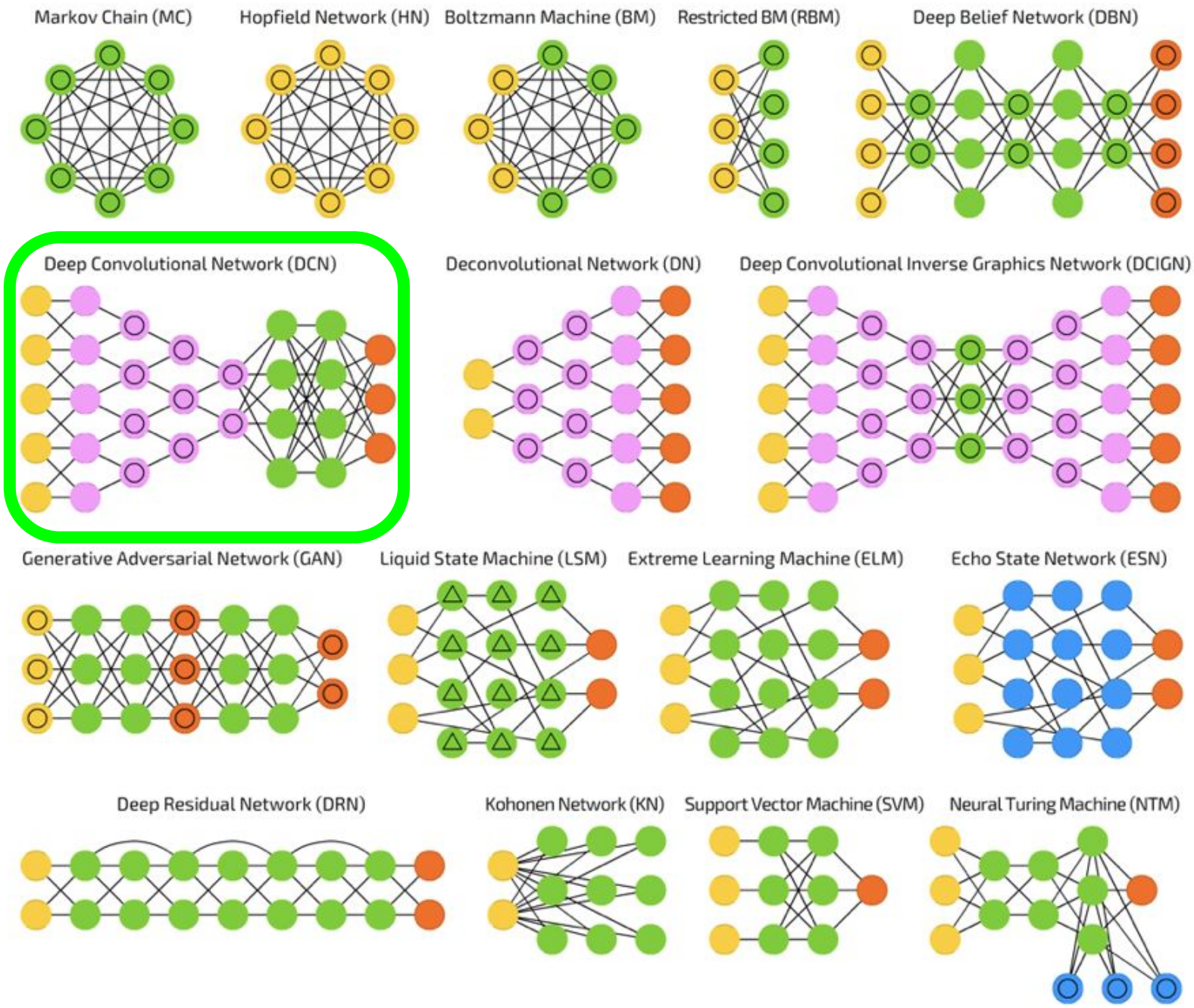


Denoising AE (DAE)

Sparse AE (SAE)



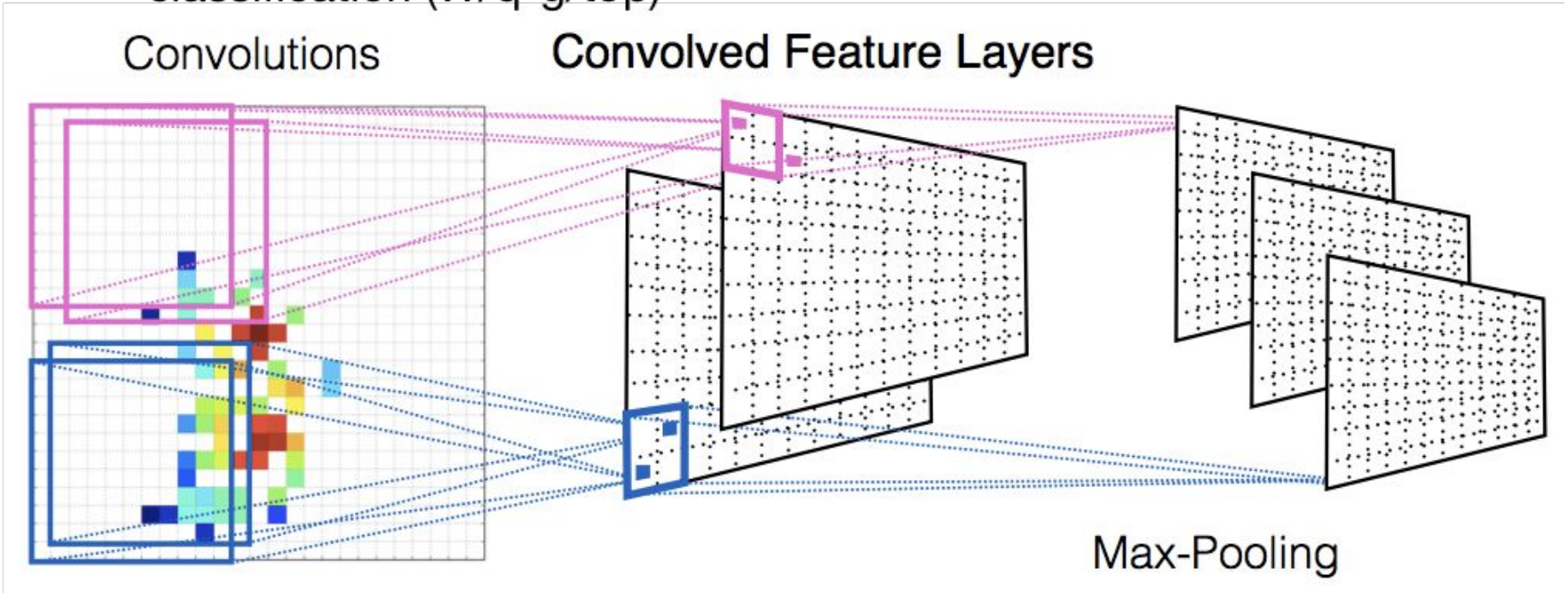
https://leonardoaraujosantos.gitbooks.io/artificial-inteligence/content/neural_networks.html


Markov Chain (MC)
Hopfield Network (HN)
Boltzmann Machine (BM)
Restricted BM (RBM)
Deep Belief Network (DBN)
Deep Convolutional Network (DCN)
Deconvolutional Network (DN)
Deep Convolutional Inverse Graphics Network (DCIGN)
Generative Adversarial Network (GAN)
Liquid State Machine (LSM)
Extreme Learning Machine (ELM)
Echo State Network (ESN)
Deep Residual Network (DRN)
Kohonen Network (KN)
Support Vector Machine (SVM)
Neural Turing Machine (NTM)

Setting up a DCNN

- Increasing usage amongst physicists for jets/calorimetry
- Specifically utilized in classification (W/q-g/top)

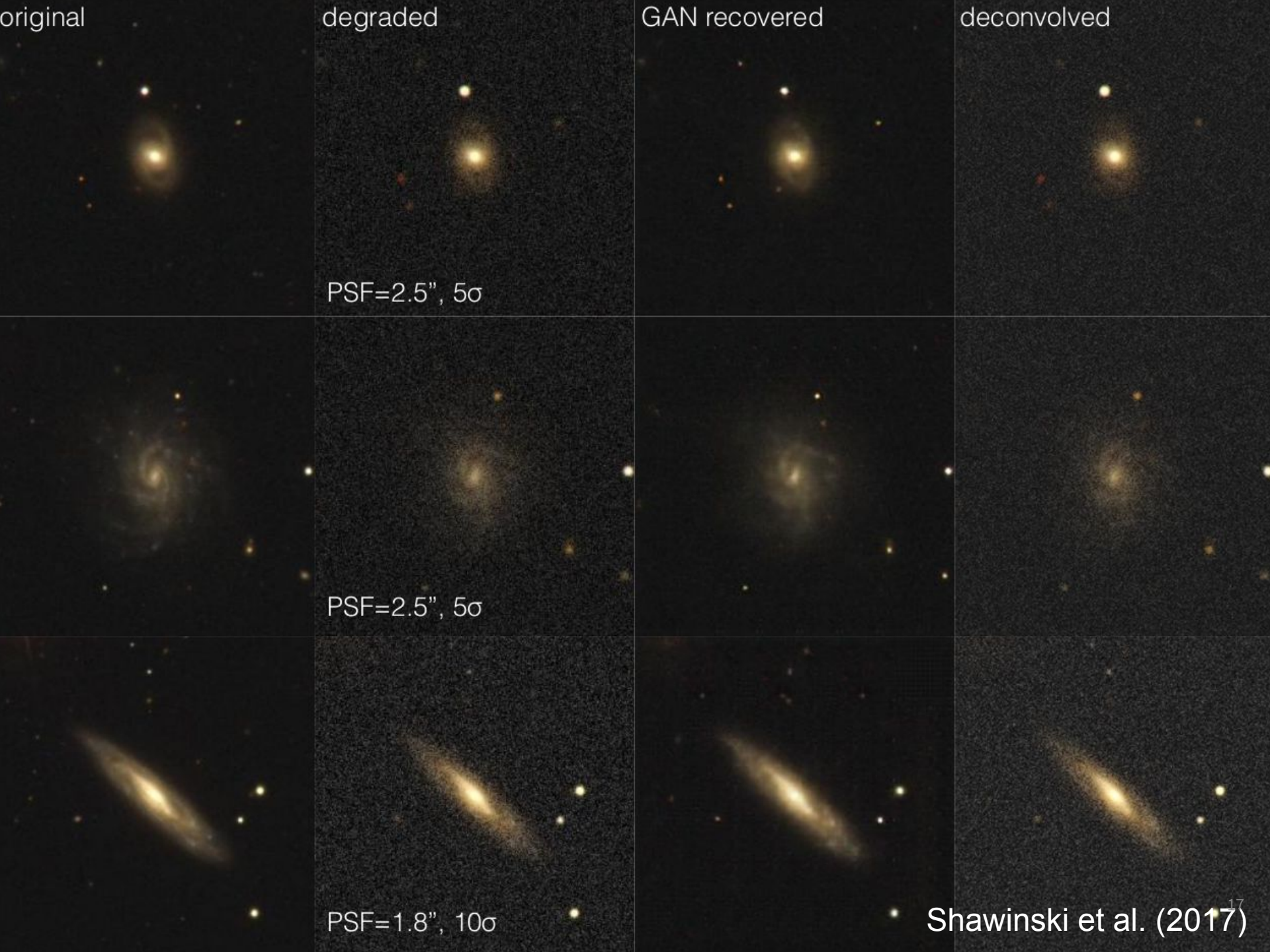
P. Baldi et al. 1603.09349 **(W-tagging)**
J. Barnard et al. 1609.00607 **(W-tagging)**
P. Komiske et al. 1612.01551 **(q/g-tagging)**
G. Kasieczka et al. 1701.08784 **(top-tagging)**



Predominantly used in astro and starting to gain popularity in HEX

Visual Example - How a DCNN actually works

http://scs.ryerson.ca/~aharley/vis/conv/

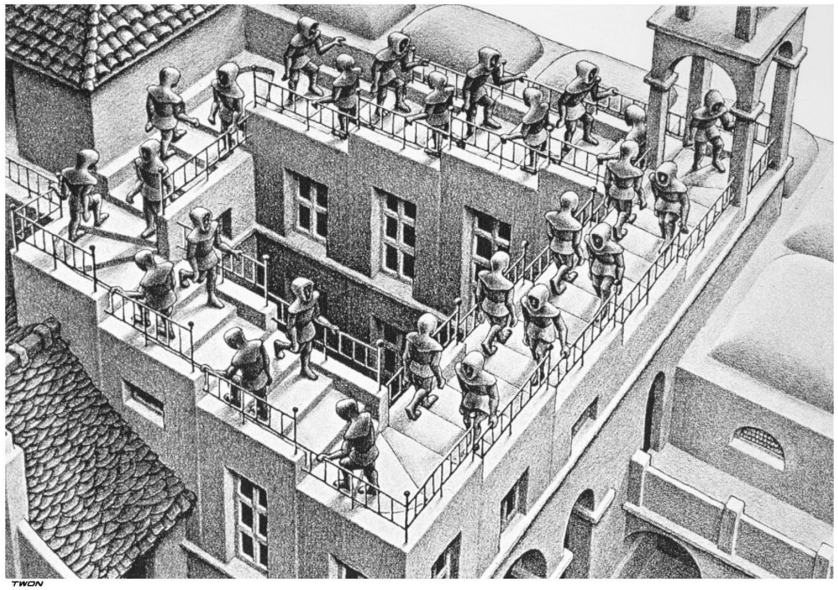


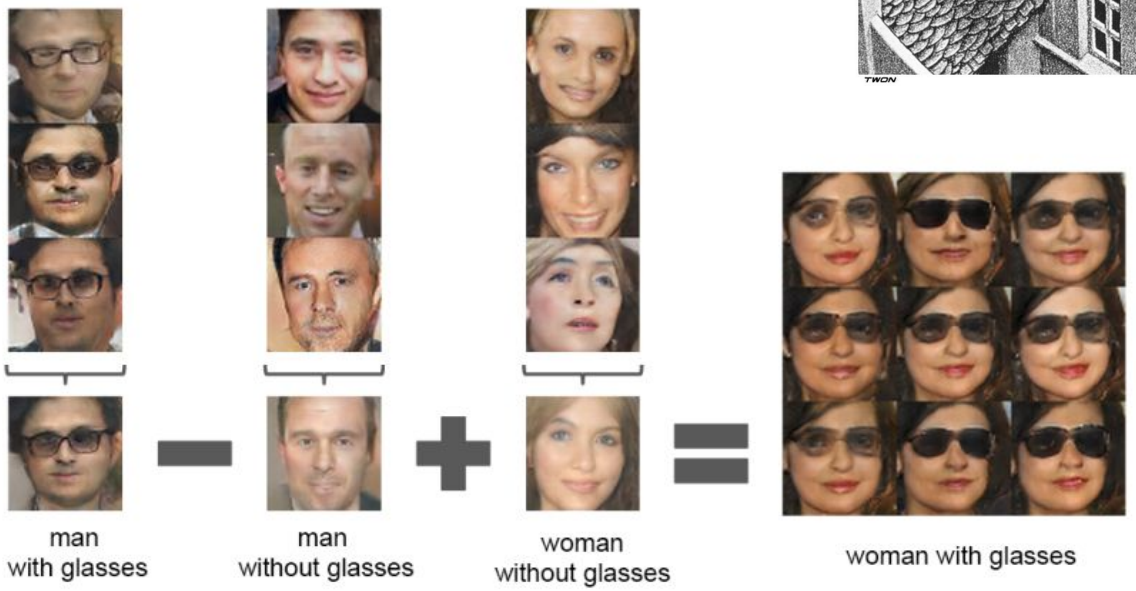
Shawinski et al. (2017)

Generative Adversarial Networks?

What if the cat and mouse game goes on forever? (model instabilities with oscillating solutions)

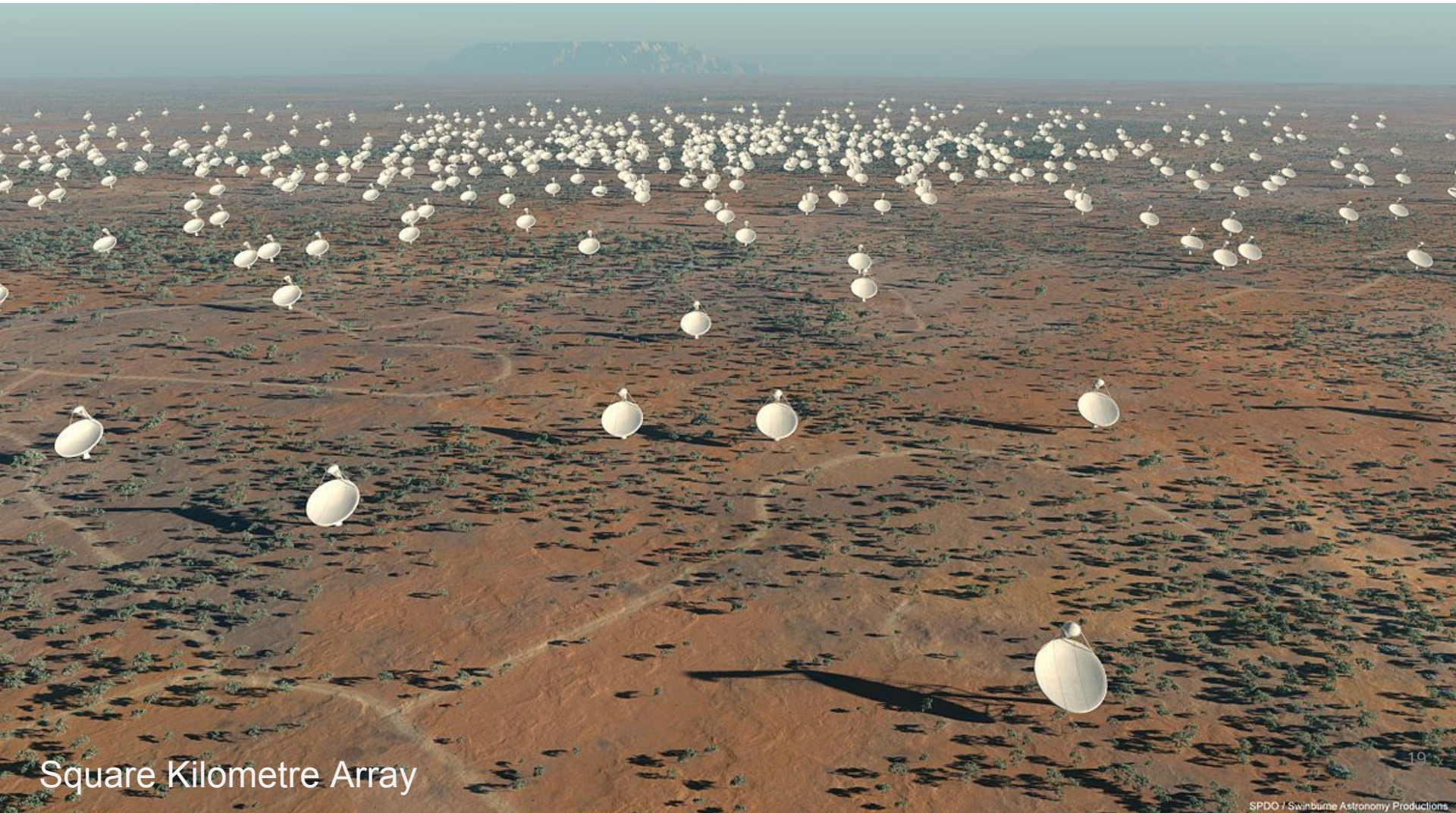
But they can still learn representations of, e.g., images, that can be rich in their own (linear) structure.



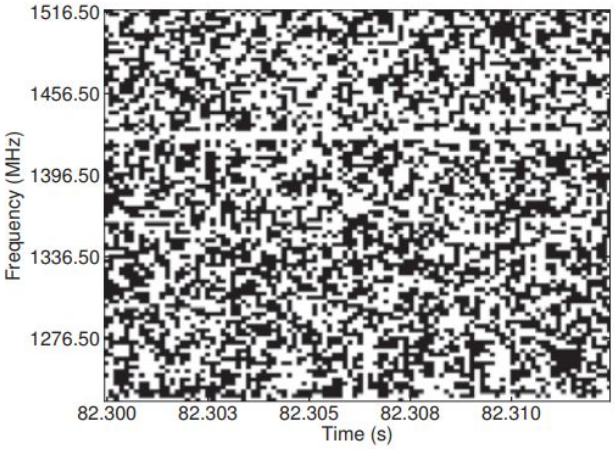


Radford et al. 2016

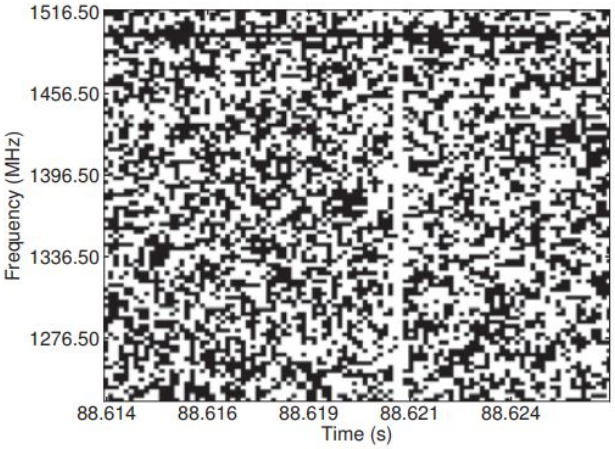
Radio frequency interference



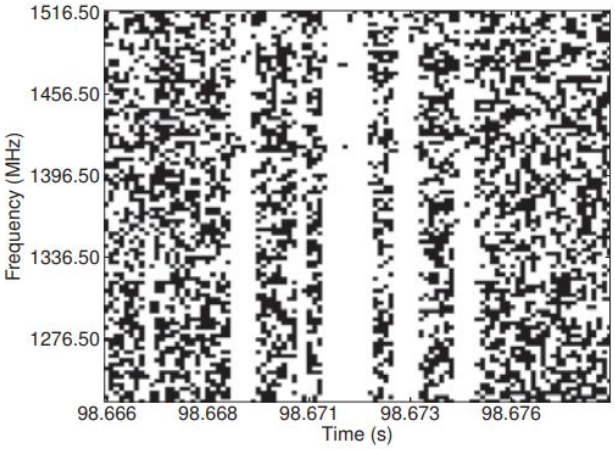
Radio frequency interference



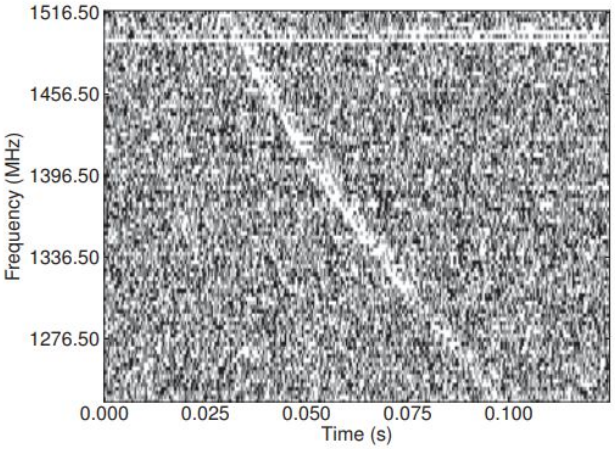
(a) Channelized RFI around 1425 MHz



(b) Broadband, short-duration RFI



(c) An RFI event consisting of several broadband bursts

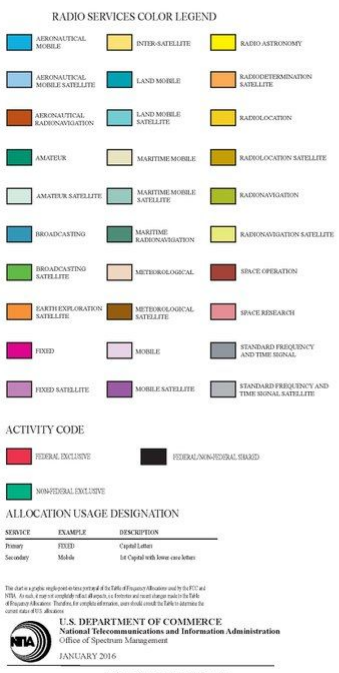


(d) A dispersed pulsar signal

Doran (2013)

UNITED STATES FREQUENCY ALLOCATIONS

THE RADIO SPECTRUM



This chart is a public-use poster-size portrayal of the Table of Frequency Allocations used by the FCC and NTIA. As a result, it may not completely reflect all aspects of, or footnotes and recent changes made to the Table of Frequency Allocations. Therefore, for complete information, users should consult the Table to determine the current status of U.S. allocations.

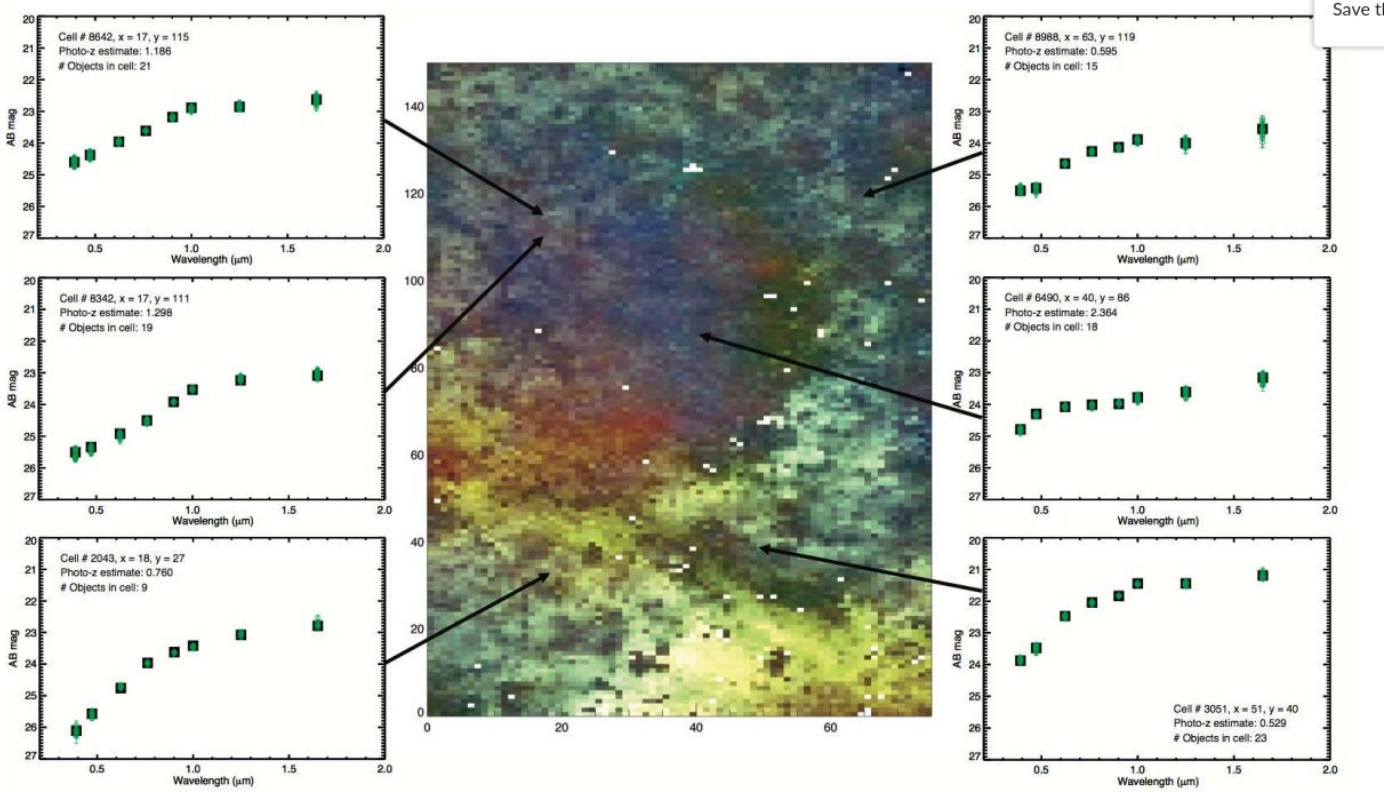
U.S. DEPARTMENT OF COMMERCE
National Telecommunications and Information Administration
Office of Spectrum Management

JANUARY 2016

Self organising maps

Kind of NN used to produce a low-dimensional representation of complex data.

Metric on the map is some kind of distance. Points close on the map are similar, points distant are dissimilar. Maps can be self-growing, elastic, conformal...



Picture from Masters et al.2015. ArXiv: 1509.03318

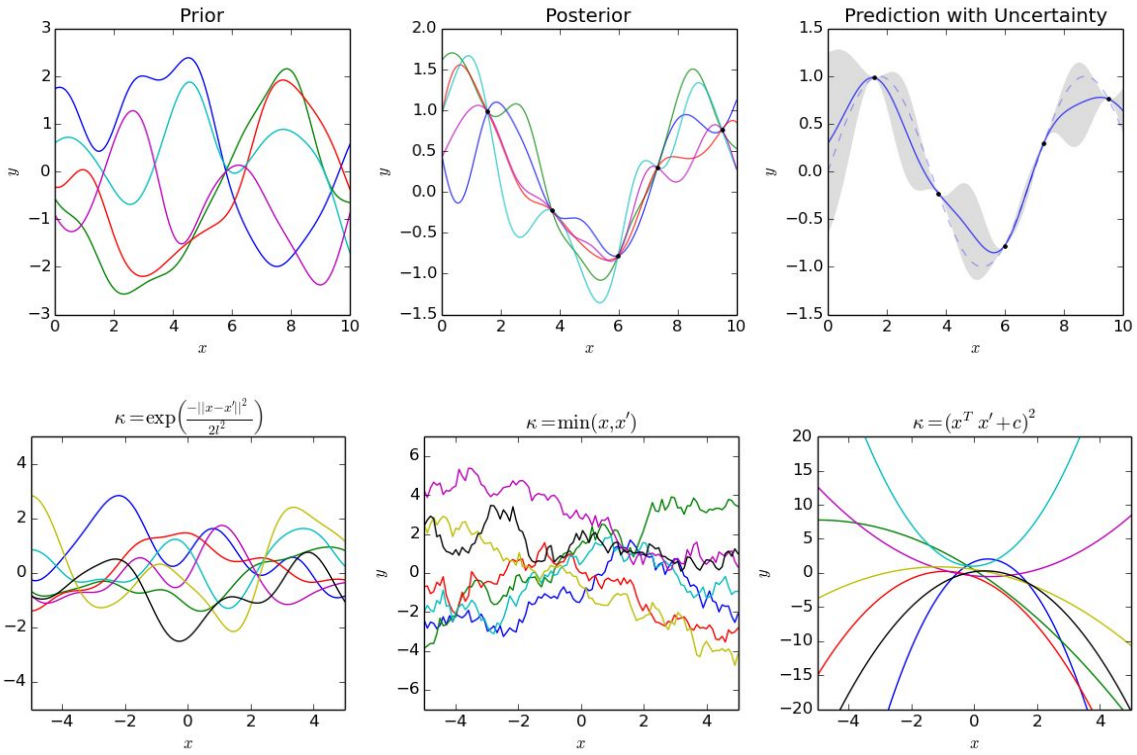
Gaussian Processes

Class of Kernel machines. + Lazy learning

‘Process’? - generalization of a probability distribution to functions.

Can control the process' stationarity, isotropy, smoothness and periodicity through its covariance function.

The prediction is not just an estimate for that point, but also has uncertainty information



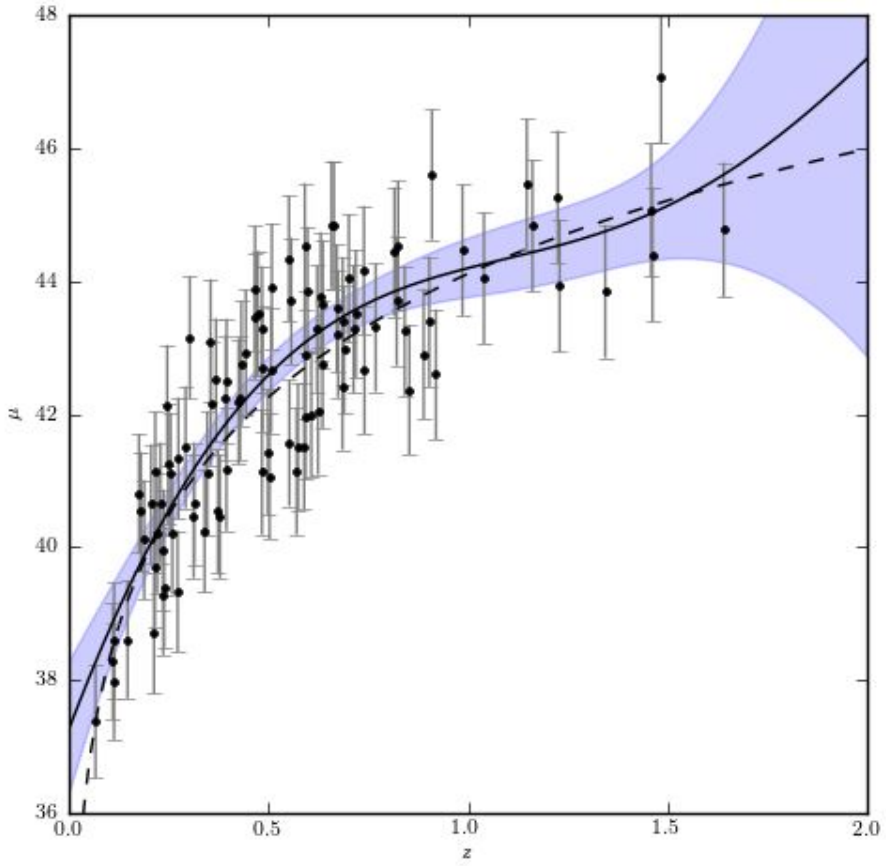
Gaussian Processes

Class of Kernel machines. + Lazy learning

‘Process’? - generalization of a probability distribution to functions.

Can control the process' stationarity, isotropy, smoothness and periodicity through its covariance function.

The prediction is not just an estimate for that point, but also has uncertainty information



Picture from: http://www.astroml.org/book_figures/chapter8/fig_gp_mu_z.html

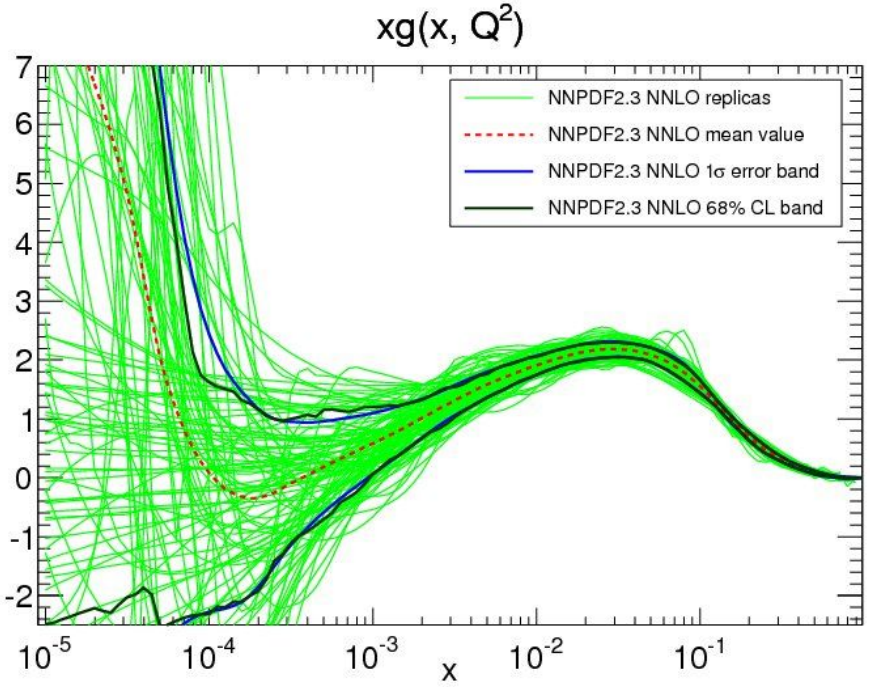
Uncertainties and error estimation:

More on uncertainties:

Using input uncertainties. - improve accuracy and prevent overfitting
Getting output uncertainties. - especially important in any prediction

Probabilistic methods
Dropout layers in neural networks.
Information entropy measures

and more… a convergence of statistics and ML

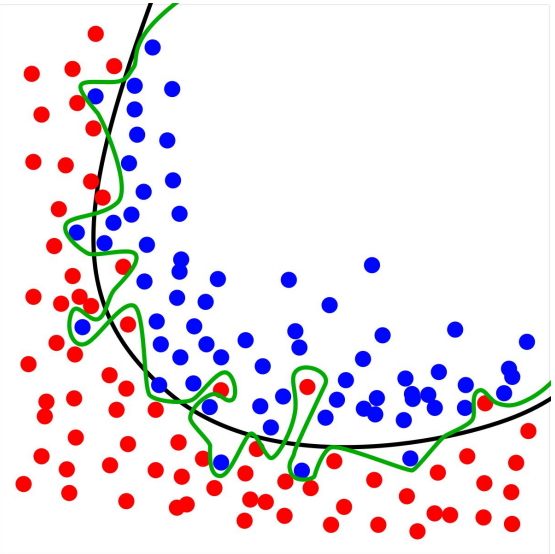


NNPDF - fits to deep inelastic data

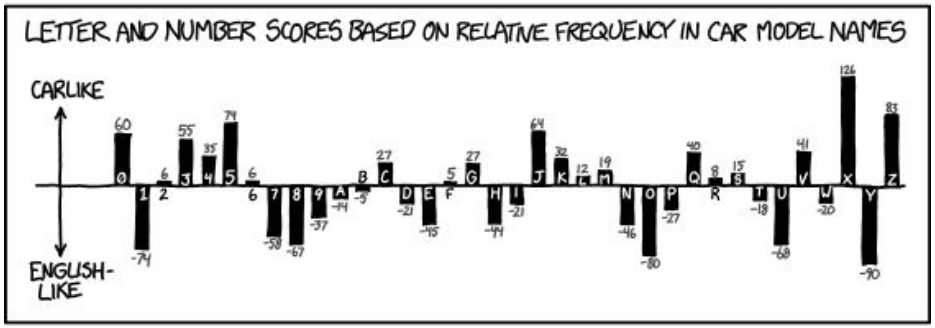
ML: Pitfalls to avoid

Know what training and test data you're working with.

- Missing data
- Unrepresentative distributions
- Outliers!
- Overfitting = your model sucks
- No free lunch theorem



CERTAIN LETTERS AND NUMBERS ARE USED DISPROPORTIONATELY OFTEN IN CAR MODELS COMPARED TO REGULAR TEXT.
(SEE: "REV-4 CR-X x3 G6 MAXX")



BASED ON THESE SCORES, HERE ARE A FEW SUGGESTIONS FOR CAR COMPANIES:
(WITH AVERAGE LETTER SCORES)

NAMES TO AVOID	POTENTIAL HITS
HONDA 2CHAINZ (-0.13)	HONDA 3CHAINZ (0.57)
MITSUBISHI FHQWHGADS (-0.62)	SUBARU ANDRE3000 (1.30)
KIA 49ANDGOTHY (-2.96)	SUZUKI SEXISM (1.82)
CHEVROLET NICEGUY (-3.09)	LINCOLN MARXISM (2.17)
OLDSMOBILE GOODWOOD (-4.44)	HYUNDAI CLIMAX (2.48)
INFINITI TOOTHY69 (-4.51)	PORSCHE ZIZEK9000 (3.06)
BMW OUTHOUSE (-4.85)	LEXUS 3x3CUTRIX (3.22)
VOLKSWAGEN WOODPONY 70H7 (-5.70)	ACURA PIZZAJAZZ (3.56)
CHRYSLER UH 10NO (-5.65)	FORD SIXAXLE 4x4 (3.95)
NISSAN DOODY (-5.84)	TOYOTA CERVIXXX (4.85)

What have we learnt?

Possibly nothing … (yet)

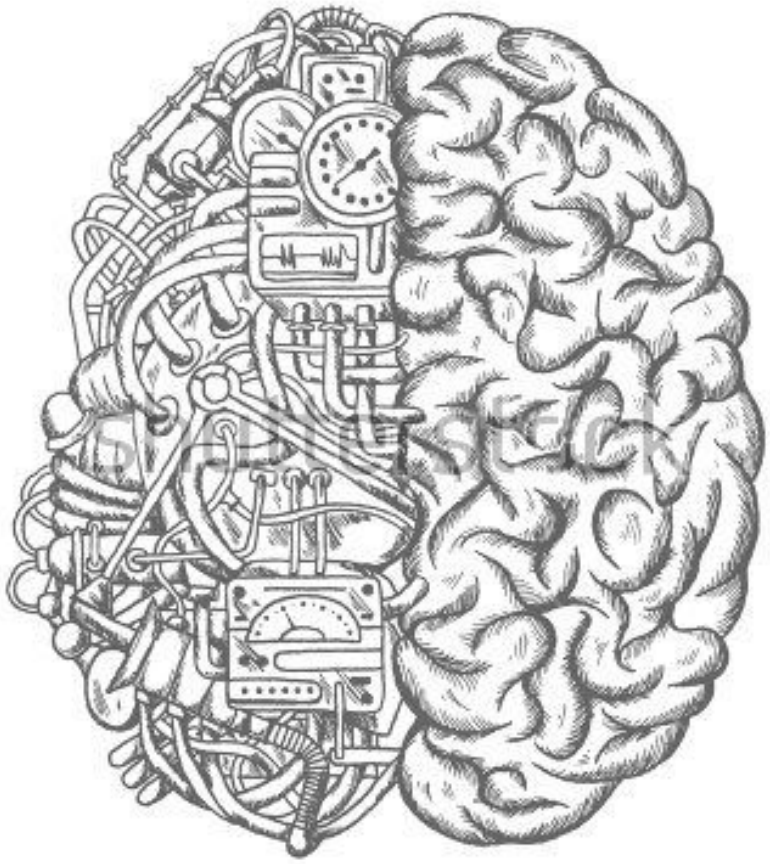
But this is very exciting and state of the art!

Relatively easy to download datasets and get started on your own fun project

Very active dev and user community - Easy to find stack exchange pages with SOLUTIONS on exactly the error you are seeing

Go and try it out!

Need more work here



Inference - in + ferus + ents

(part of) OLD ENGLISH | (wild) LATIN | (tree-hosts) QUENYA

- **Using the wild power of giant sentient trees to validate or invalidate conclusions based on logic and reasoning.**

Physics literature using ML techniques:

An automatic taxonomy of galaxy morphology using unsupervised machine learning

Alex Hocking (Hertfordshire), James E. Geach, Yi Sun, Neil Davey

(Submitted on 18 Sep 2017)

We present an unsupervised machine learning technique that automatically segments and labels galaxies in astronomical imaging surveys using only pixel data. Distinct from previous unsupervised machine learning approaches used in astronomy we use no pre-selection or pre-filtering of target galaxy type to identify galaxies that are similar. We demonstrate the technique on the HST Frontier Fields. By training the algorithm using galaxies from one field (Abell 2744) and applying the result to another (MACS0416.1-2403), we show how the algorithm can cleanly separate early and late type galaxies without any form of pre-directed training for what an 'early' or 'late' type galaxy is. We then apply the technique to the HST CANDELS fields, creating a catalogue of approximately 60,000 classifications. We show how the automatic classification groups galaxies of similar morphological (and photometric) type, and make the classifications public via a catalogue, a visual catalogue and galaxy similarity search. We compare the CANDELS machine-based classifications to human-based classifications from the Galaxy Zoo: CANDELS project. Although there is not a direct mapping between Galaxy Zoo and our hierarchical labelling, we demonstrate a good level of concordance between human and machine classifications. Finally, we show how the technique can be used to identify rarer objects and present new lensed galaxy candidates from the CANDELS imaging.

Physics literature using ML techniques:

Photometric Supernova Classification With Machine Learning

Michelle Lochner, Jason D. McEwen, Hiranya V. Peiris, Ofer Lahav, Max K. Winter

(Submitted on 2 Mar 2016 (v1), last revised 7 Sep 2016 (this version, v3))

Automated photometric supernova classification has become an active area of research in recent years in light of current and upcoming imaging surveys such as the Dark Energy Survey (DES) and the Large Synoptic Survey Telescope, given that spectroscopic confirmation of type for all supernovae discovered will be impossible. Here, we develop a multi-faceted classification pipeline, combining existing and new approaches. Our pipeline consists of two stages: extracting descriptive features from the light curves and classification using a machine learning algorithm. Our feature extraction methods vary from model-dependent techniques, namely SALT2 fits, to more independent techniques fitting parametric models to curves, to a completely model-independent wavelet approach. We cover a range of representative machine learning algorithms, including naive Bayes, k-nearest neighbors, support vector machines, artificial neural networks and boosted decision trees (BDTs). We test the pipeline on simulated multi-band DES light curves from the Supernova Photometric Classification Challenge. Using the commonly used area under the curve (AUC) of the Receiver Operating Characteristic as a metric, we find that the SALT2 fits and the wavelet approach, with the BDTs algorithm, each achieves an AUC of 0.98, where 1 represents perfect classification. We find that a representative training set is essential for good classification, whatever the feature set or algorithm, with implications for spectroscopic follow-up. Importantly, we find that by using either the SALT2 or the wavelet feature sets with a BDT algorithm, accurate classification is possible purely from light curve data, without the need for any redshift information.

Physics literature using ML techniques:

A Hybrid Ensemble Learning Approach to Star-Galaxy Classification

Edward J. Kim, Robert J. Brunner, Matias Carrasco Kind

(Submitted on 8 May 2015 (v1), last revised 14 Jul 2015 (this version, v2))

There exist a variety of star-galaxy classification techniques, each with their own strengths and weaknesses. In this paper, we present a novel meta-classification framework that combines and fully exploits different techniques to produce a more robust star-galaxy classification. To demonstrate this hybrid, ensemble approach, we combine a purely morphological classifier, a supervised machine learning method based on random forest, an unsupervised machine learning method based on self-organizing maps, and a hierarchical Bayesian template fitting method. Using data from the CFHTLenS survey, we consider different scenarios: when a high-quality training set is available with spectroscopic labels from DEEP2, SDSS, VIPERS, and VVDS, and when the demographics of sources in a low-quality training set do not match the demographics of objects in the test data set. We demonstrate that our Bayesian combination technique improves the overall performance over any individual classification method in these scenarios. Thus, strategies that combine the predictions of different classifiers may prove to be optimal in currently ongoing and forthcoming photometric surveys, such as the Dark Energy Survey and the Large Synoptic Survey Telescope.

Physics literature using ML techniques:

Estimating Extinction using Unsupervised Machine Learning

Stefan Meingast, Marco Lombardi, Joao Alves

(Submitted on 27 Feb 2017)

Dust extinction is the most robust tracer of the gas distribution in the interstellar medium, but measuring extinction is limited by the systematic uncertainties involved in estimating the intrinsic colors to background stars. In this paper we present a new technique, PNICER, that estimates intrinsic colors and extinction for individual stars using unsupervised machine learning algorithms. This new method aims to be free from any priors with respect to the column density and intrinsic color distribution. It is applicable to any combination of parameters and works in arbitrary numbers of dimensions. Furthermore, it is not restricted to color space. Extinction towards single sources is determined by fitting Gaussian Mixture Models along the extinction vector to (extinction-free) control field observations. In this way it becomes possible to describe the extinction for observed sources with probability densities. PNICER effectively eliminates known biases found in similar methods and outperforms them in cases of deep observational data where the number of background galaxies is significant, or when a large number of parameters is used to break degeneracies in the intrinsic color distributions. This new method remains computationally competitive, making it possible to correctly de-redden millions of sources within a matter of seconds. With the ever-increasing number of large-scale high-sensitivity imaging surveys, PNICER offers a fast and reliable way to efficiently calculate extinction for arbitrary parameter combinations without prior information on source characteristics. PNICER also offers access to the well-established NICER technique in a simple unified interface and is capable of building extinction maps including the NICEST correction for cloud substructure. PNICER is offered to the community as an open-source software solution and is entirely written in Python.

Physics literature using ML techniques:

Cosmological model discrimination with Deep Learning

Jorit Schmelzle, Aurelien Lucchi, Tomasz Kacprzak, Adam Amara, Raphael Sgier, Alexandre Réfrégier, Thomas Hofmann

(Submitted on 17 Jul 2017 (v1), last revised 18 Jul 2017 (this version, v2))

We demonstrate the potential of Deep Learning methods for measurements of cosmological parameters from density fields, focusing on the extraction of non-Gaussian information. We consider weak lensing mass maps as our dataset. We aim for our method to be able to distinguish between five models, which were chosen to lie along the σ_8 - Ω_m degeneracy, and have nearly the same two-point statistics. We design and implement a Deep Convolutional Neural Network (DCNN) which learns the relation between five cosmological models and the mass maps they generate. We develop a new training strategy which ensures the good performance of the network for high levels of noise. We compare the performance of this approach to commonly used non-Gaussian statistics, namely the skewness and kurtosis of the convergence maps. We find that our implementation of DCNN outperforms the skewness and kurtosis statistics, especially for high noise levels. The network maintains the mean discrimination efficiency greater than 85% even for noise levels corresponding to ground based lensing observations, while the other statistics perform worse in this setting, achieving efficiency less than 70%. datasets.This demonstrates the ability of CNN-based methods to efficiently break the σ_8 - Ω_m degeneracy with weak lensing mass maps alone. We discuss the potential of this method to be applied to the analysis of real weak lensing data and other

Physics literature using ML techniques:

Probability density estimation of photometric redshifts based on machine learning

Stefano Cavuoti, Massimo Brescia, Valeria Amaro, Civita Vellucci, Giuseppe Longo, Crescenzo Tortora

(Submitted on 12 Jun 2017)

Photometric redshifts (photo-z's) provide an alternative way to estimate the distances of large samples of galaxies and are therefore crucial to a large variety of cosmological problems. Among the various methods proposed over the years, supervised machine learning (ML) methods capable to interpolate the knowledge gained by means of spectroscopical data have proven to be very effective. METAPHOR (Machine-learning Estimation Tool for Accurate PHOtometric Redshifts) is a novel method designed to provide a reliable PDF (Probability density Function) of the error distribution of photometric redshifts predicted by ML methods. The method is implemented as a modular workflow, whose internal engine for photo-z estimation makes use of the MLPQNA neural network (Multi Layer Perceptron with Quasi Newton learning rule), with the possibility to easily replace the specific machine learning model chosen to predict photo-z's. After a short description of the software, we present a summary of results on public galaxy data (Sloan Digital Sky Survey - Data Release 9) and a comparison with a completely different method based on Spectral Energy Distribution (SED) template fitting.

Physics literature using ML techniques:

Improving galaxy morphology with machine learning

P. H. Barchi, F. G. da Costa, R. Sautter, T. C. Moura, D. H. Stalder, R. R. Rosa, R. R. de Carvalho

(Submitted on 18 May 2017)

This paper presents machine learning experiments performed over results of galaxy classification into elliptical (E) and spiral (S) with morphological parameters: concetration (CN), assimetry metrics (A3), smoothness metrics (S3), entropy (H) and gradient pattern analysis parameter (GA). Except concentration, all parameters performed a image segmentation pre-processing. For supervision and to compute confusion matrices, we used as true label the galaxy classification from GalaxyZoo. With a 48145 objects dataset after preprocessing (44760 galaxies labeled as S and 3385 as E), we performed experiments with Support Vector Machine (SVM) and Decision Tree (DT). Whit a 1962 objects balanced dataset, we applied K- means and Agglomerative Hierarchical Clustering. All experiments with supervision reached an Overall Accuracy OA >= 97%.

Physics literature using ML techniques:

Machine Learning of Explicit Order Parameters: From the Ising Model to SU(2) Lattice Gauge Theory

Sebastian Johann Wetzel, Manuel Scherzer

(Submitted on 16 May 2017)

We present a procedure for reconstructing the decision function of an artificial neural network as a simple function of the input, provided the decision function is sufficiently symmetric. In this case one can easily deduce the quantity by which the neural network classifies the input. The procedure is embedded into a pipeline of machine learning algorithms able to detect the existence of different phases of matter, to determine the position of phase transitions and to find explicit expressions of the physical quantities by which the algorithm distinguishes between phases. We assume no prior knowledge about the Hamiltonian or the order parameters except Monte Carlo-sampled configurations. The method is applied to the Ising Model and SU(2) lattice gauge theory. In both systems we deduce the explicit expressions of the known order parameters from the decision functions of the neural networks.

Physics literature using ML techniques:

Development of a Machine Learning Based Analysis Chain for the Measurement of Atmospheric Muon Spectra with IceCube

Tomasz Fuchs

(Submitted on 15 Jan 2017)

High-energy muons from air shower events detected in IceCube are selected using state of the art machine learning algorithms. Attributes to distinguish a HE-muon event from the background of low-energy muon bundles are selected using the mRMR algorithm and the events are classified by a random forest model. In a subsequent analysis step the obtained sample is used to reconstruct the atmospheric muon energy spectrum, using the unfolding software TRUEE. The reconstructed spectrum covers an energy range from 10^4GeV to 10^6GeV. The general analysis scheme is presented, including results using the first year of data taken with IceCube in its complete configuration with 86 instrumented strings.

Physics literature using ML techniques:

Rate Constants for Fine-Structure Excitations in O-H Collisions with Error Bars Obtained by Machine Learning

Daniel Vieira, Roman Krems

(Submitted on 8 Jan 2017)

We present an approach using a combination of coupled channel scattering calculations with a machine- learning technique based on Gaussian Process regression to determine the sensitivity of the rate constants for non-adiabatic transitions in inelastic atomic collisions to variations of the underlying adiabatic interaction potentials. Using this approach, we improve the previous computations of the rate constants for the fine-structure transitions in collisions of O(3Pj) with atomic H. We compute the error bars of the rate constants corresponding to 20 % variations of the ab initio potentials and show that this method can be used to determine which of the individual adiabatic potentials are more or less important for the outcome of different fine-structure changing collisions.

Physics literature using ML techniques:

What does a convolutional neural network recognize in the moon?

Daigo Shoji

(Submitted on 18 Aug 2017 (v1), last revised 21 Aug 2017 (this version, v2))

Many people see a human face or animals in the pattern of the maria on the moon. Although the pattern corresponds to the actual variation in composition of the lunar surface, the culture and environment of each society influence the recognition of these objects (i.e., symbols) as specific entities. In contrast, a convolutional neural network (CNN) recognizes objects from characteristic shapes in a training data set. Using CNN, this study evaluates the probabilities of the pattern of lunar maria categorized into the shape of a crab, a lion and a hare. If Mare Frigoris (a dark band on the moon) is included in the lunar image, the lion is recognized. However, in an image without Mare Frigoris, the hare has the highest probability of recognition. Thus, the recognition of objects similar to the lunar pattern depends on which part of the lunar maria is taken into account. In human recognition, before we find similarities between the lunar maria and objects such as animals, we may be persuaded in advance to see a particular image from our culture and environment and then adjust the lunar pattern to the shape of the imagined object.

Machine learning phases of matter

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

[1]*Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*

[2]*Department of Physics and Astronomy, University of Waterloo, Ontario, N2L 3G1, Canada*

Neural networks can be used to identify phases and phase transitions in condensed matter systems via supervised machine learning. Readily programmable through modern software libraries, we show that a standard feed-forward neural network can be trained to detect multiple types of order parameter directly from raw state configurations sampled with Monte Carlo. In addition, they can detect highly non-trivial states such as Coulomb phases, and if modified to a convolutional neural network, topological phases with no conventional order parameter. We show that this classification occurs within the neural network without knowledge of the Hamiltonian or even the general locality of interactions. These results demonstrate the power of machine learning as a basic research tool in the field of condensed matter and statistical physics.

Physics literature using ML techniques:

Estimating hydrogen sulfide solubility in ionic liquids using a machine learning approach

Ali Shafiei [a], Mohammad Ali Ahmadi [b] ✉, Seyed Hayan Zaheri [b], Alireza Baghban [c], Ali Amirfakhrian [d], Reza Soleimani [e]

Show more

https://doi.org/10.1016/j.supflu.2014.08.011 Get rights and content

Highlights

- An intelligent and simple-to-use approach for prediction H_2S solubility in various ionic liquids has been developed.
- Accurate, precise and extensive H_2S solubility in various ionic liquids data banks have been utilized.
- Statistical analysis was implemented to the outputs generated by PSO-ANN model.

Physics literature using ML techniques:

LETTER

doi:10.1038/nature14964

Conventional superconductivity at 203 kelvin at high pressures in the sulfur hydride system

A. P. Drozdov[1]*, M. I. Eremets[1]*, I. A. Troyan[1], V. Ksenofontov[2] & S. I. Shylin[2]

A superconductor is a material that can conduct electricity without resistance below a superconducting transition temperature, T_c. The highest T_c that has been achieved to date is in the copper oxide system[1]: 133 kelvin at ambient pressure[2] and 164 kelvin at high pressures[3]. As the nature of superconductivity in these materials is still not fully understood (they are not conventional superconductors), the prospects for achieving still higher transition temperatures by this route are not clear. In contrast, the Bardeen–Cooper–Schrieffer theory of conventional superconductivity gives a guide for achieving high T_c with no theoretical upper bound—all that is needed is a favourable combination of high-frequency phonons, strong electron–phonon coupling, and a high density of states[4]. These conditions can in principle be fulfilled for metallic hydrogen and covalent compounds dominated by hydrogen[5,6], as hydrogen atoms provide the necessary high-frequency phonon modes as well as the strong electron–phonon coupling. Numerous calculations support this idea and have predicted transition temperatures in the range 50–235 kelvin for many hydrides[7], but only a moderate T_c of 17 kelvin has been observed experimentally[8]. Here we investigate sulfur hydride[9], where a T_c of 80 kelvin has been predicted[10]. We find that this system transforms to a metal at a pressure of approximately 90 gigapascals. On cooling, we see signatures of superconductivity: a sharp drop of the resistivity to zero and a decrease of the transition temperature with magnetic field, with magnetic susceptibility measurements confirming a T_c of 203 kelvin. Moreover, a pronounced isotope shift of T_c in sulfur deuteride is suggestive of an electron–phonon mechanism of superconductivity that is consistent with the Bardeen–Cooper–Schrieffer scenario. We argue that the phase responsible for high-T_c superconductivity in this system is likely to be H_3S, formed from H_2S by decomposition under pressure. These findings raise hope for the prospects for achieving room-temperature superconductivity in other hydrogen-based materials.

superconductivity[6]. Similarly to pure hydrogen, they have high Debye temperatures. Moreover, heavier elements might be beneficial as they contribute to the low frequencies that enhance electron–phonon coupling. Importantly, lower pressures are required to metallize hydrides in comparison to pure hydrogen. Ashcroft's general idea was supported in numerous calculations[7,10] predicting high values of T_c for many hydrides. So far only a low T_c (~17 K) has been observed experimentally[8].

For the present study we selected H_2S, because it is relatively easy to handle and is predicted to transform to a metal and a superconductor at a low pressure $P \approx 100$ GPa with a high $T_c \approx 80$ K (ref. 10). Experimentally, H_2S is known as a typical molecular compound with a rich phase diagram[14]. At about 96 GPa, hydrogen sulphide transforms to a metal[15]. The transformation is complicated by the partial dissociation of H_2S and the appearance of elemental sulfur at $P > 27$ GPa at room temperature, and at higher pressures at lower temperatures[14]. Therefore, the metallization of hydrogen sulphide can be explained by elemental sulfur, which is known to become metallic above 95 GPa (ref. 16). No experimental studies of hydrogen sulphide are known above 100 GPa.

In a typical experiment, we performed loading and the initial pressure increase at temperatures of ~200 K; this is essential for obtaining a good sample (Methods). The Raman spectra of H_2S and D_2S were measured as the pressure was increased, and were in general agreement with the literature data[17,18] (Extended Data Fig. 1). The sample starts to conduct at $P \approx 50$ GPa. At this pressure it is a semiconductor, as shown by the temperature dependence of the resistance and pronounced photoconductivity. At 90–100 GPa the resistance drops further, and the temperature dependence becomes metallic. No photoconductive response is observed in this state. It is a poor metal—its resistivity at ~100 K is $\rho \approx 3 \times 10^{-5}$ ohm m at 110 GPa and $\rho \approx 3 \times 10^{-7}$ ohm m at ~200 GPa.

During the cooling of the metal at pressures of about 100 GPa

Physics literature using ML techniques:

Machine Learning Spatial Geometry from Entanglement Features

Yi-Zhuang You, Zhao Yang, Xiao-Liang Qi

(Submitted on 5 Sep 2017)

Motivated by the close relations of the renormalization group with both the holography duality and the deep learning, we propose that the holographic geometry can emerge from deep learning the entanglement feature of a quantum many-body state. We develop a concrete algorithm, call the entanglement feature learning (EFL), based on the random tensor network (RTN) model for the tensor network holography. We show that each RTN can be mapped to a Boltzmann machine, trained by the entanglement entropies over all subregions of a given quantum many-body state. The goal is to construct the optimal RTN that best reproduce the entanglement feature. The RTN geometry can then be interpreted as the emergent holographic geometry. We demonstrate the EFL algorithm on 1D free fermion system and observe the emergence of the hyperbolic geometry (AdS_3 spatial geometry) as we tune the fermion system towards the gapless critical point (CFT_2 point).

Physics literature using ML techniques:

The Fog of War: A Machine Learning Approach to Forecasting Weather on Mars

Daniele Bellutta

(Submitted on 26 Jun 2017)

For over a decade, scientists at NASA's Jet Propulsion Laboratory (JPL) have been recording measurements from the Martian surface as a part of the Mars Exploration Rovers mission. One quantity of interest has been the opacity of Mars's atmosphere for its importance in day-to-day estimations of the amount of power available to the rover from its solar arrays. This paper proposes the use of neural networks as a method for forecasting Martian atmospheric opacity that is more effective than the current empirical model. The more accurate prediction provided by these networks would allow operators at JPL to make more accurate predictions of the amount of energy available to the rover when they plan activities for coming sols.

Physics literature using ML techniques:

A hybrid supervised/unsupervised machine learning approach to solar flare prediction

Federico Benvenuto, Michele Piana, Cristina Campi, Anna Maria Massone

(Submitted on 21 Jun 2017)

We introduce a hybrid approach to solar flare prediction, whereby a supervised regularization method is used to realize feature importance and an unsupervised clustering method is used to realize the binary flare/no-flare decision. The approach is validated against NOAA SWPC data.

Physics literature using ML techniques:

Real-time detection of transients in OGLE-IV with application of machine learning

Jakub Klencki, Łukasz Wyrzykowski

(Submitted on 22 Jan 2016)

The current bottleneck of transient detection in most surveys is the problem of rejecting numerous artifacts from detected candidates. We present a triple-stage hierarchical machine learning system for automated artifact filtering in difference imaging, based on self-organizing maps. The classifier, when tested on the OGLE-IV Transient Detection System, accepts ~ 97 % of real transients while removing up to ~ 97.5 % of artifacts.

Physics literature using ML techniques:

CaloGAN: Simulating 3D High Energy Particle Showers in Multi-Layer Electromagnetic Calorimeters with Generative Adversarial Networks

---

Michela Paganini[a,b], **Luke de Oliveira**[a], **and Benjamin Nachman**[a]

[a] *Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA, 94720, USA*
[b] *Department of Physics, Yale University, New Haven, CT 06520, USA*

E-mail: michela.paganini@yale.edu, lukedeoliveira@lbl.gov, bnachman@cern.ch

Physics literature using ML techniques:

Classification without labels:
Learning from mixed samples in high energy physics

Eric M. Metodiev,[a] **Benjamin Nachman,**[b] **and Jesse Thaler**[a]

[a] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[b] *Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

E-mail: metodiev@mit.edu, bpnachman@lbl.gov, jthaler@mit.edu

Physics literature using ML techniques:

Pileup Mitigation with Machine Learning (PUMML)

Patrick T. Komiske,[a] **Eric M. Metodiev,**[a] **Benjamin Nachman,**[b] **Matthew D. Schwartz**[c]

[a] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[b] *Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[c] *Department of Physics, Harvard University, Cambridge, MA 02138, USA*

E-mail: pkomiske@mit.edu, metodiev@mit.edu, bpnachman@lbl.gov, schwartz@physics.harvard.edu

Physics literature using ML techniques:

Jet-Images – Deep Learning Edition

Luke de Oliveira,[a] Michael Kagan,[b] Lester Mackey,[c] Benjamin Nachman,[b] and Ariel Schwartzman[b]

[a] *Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA 94305, USA*
[b] *SLAC National Accelerator Laboratory, Stanford University, 2575 Sand Hill Rd, Menlo Park, CA 94025, U.S.A.*
[a] *Department of Statistics, Stanford University, Stanford, CA 94305, USA*

E-mail: lukedeo@stanford.edu, mkagan@cern.ch, lmackey@stanford.edu, bnachman@cern.ch, sch@slac.stanford.edu

Physics literature using ML techniques:

Deep learning in color: towards automated quark/gluon jet discrimination

Patrick T. Komiske,[a] **Eric M. Metodiev,**[a] **and Matthew D. Schwartz**[b]

[a] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[b] *Department of Physics, Harvard University, Cambridge, MA 02138, USA*

E-mail: pkomiske@mit.edu, metodiev@mit.edu, schwartz@physics.harvard.edu

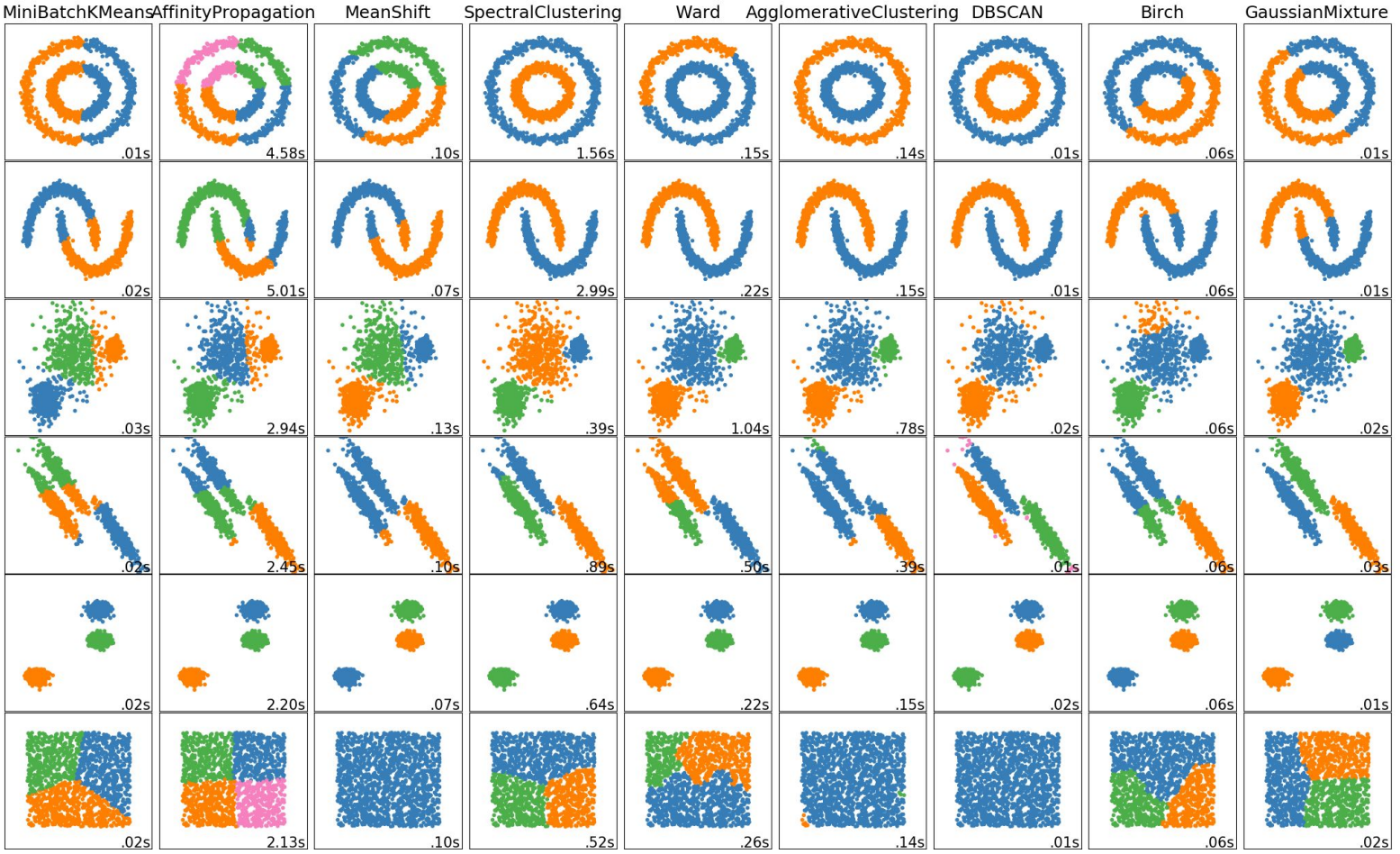
ABSTRACT: Artificial intelligence offers the potential to automate challenging data-processing tasks in collider physics. To establish its prospects, we explore to what extent deep learning with convolutional neural networks can discriminate quark and gluon jets better than observables designed by physicists. Our approach builds upon the paradigm that a jet can be treated as an image, with intensity given by the local calorimeter deposits. We supplement this construction by adding color to the images, with red, green and blue intensities given by the transverse momentum in charged particles, transverse momentum in neutral particles, and pixel-level charged particle counts. Overall, the deep networks match or outperform traditional jet variables. We also find that, while various simulations produce different quark and gluon jets, the neural networks are surprisingly insensitive to these differences, similar to traditional observables. This suggests that the networks can extract robust physical information from imperfect simulations.

Some scenarios:

Given a large amount of data...
- Is this email about your qualifier spam?
- Can you fit a line to this data? Is this the best line to fit?
- Can I extrapolate beyond my current measured values? With what confidence?
- Can I remove contaminants from my data?

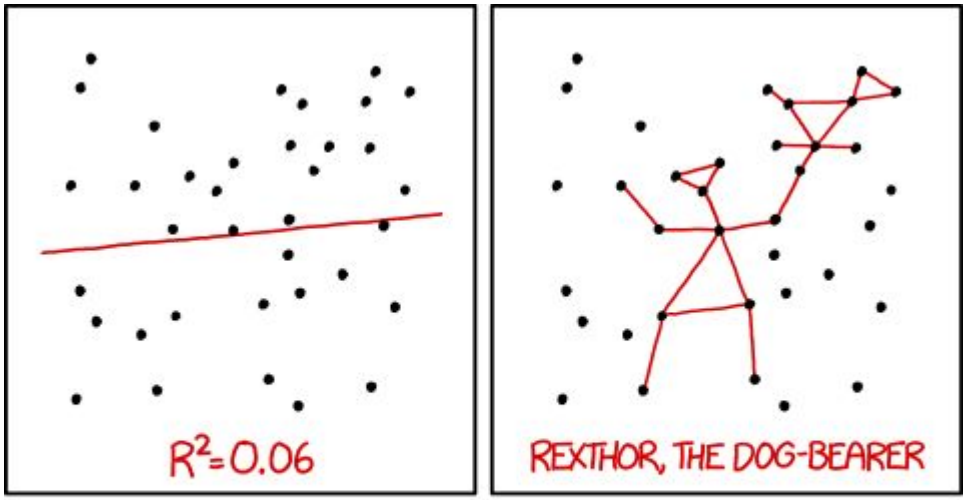
- What can I say about my data? (are there groups? Interactions? structure?)
- Can I somehow use citizen science?

Clustering [1 slide]



What is machine learning?

Automating ourselves back into manual labor



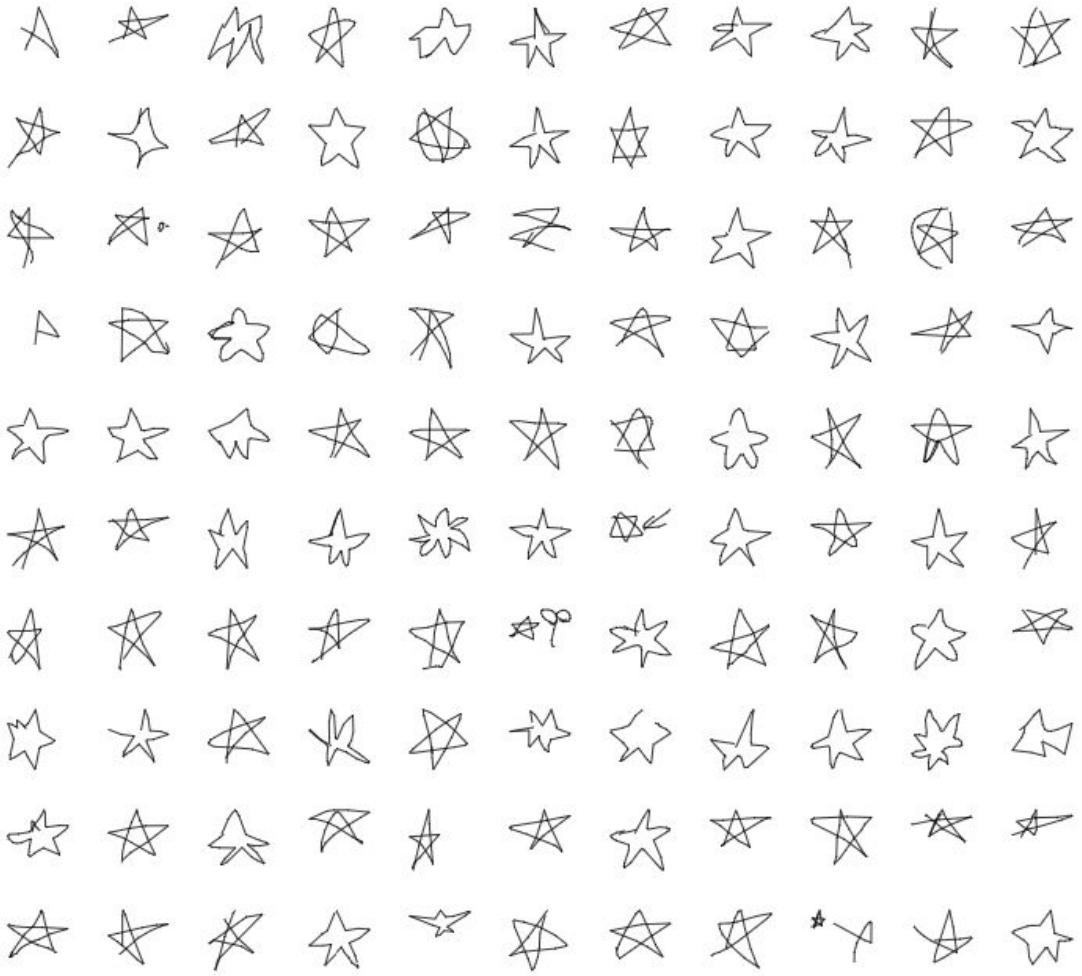
What is machine learning?

Dealing with incomplete or empirical physics. - the cutting edge is always unknown.

Dealing with an overload of data, often noisy, biased and incomplete.

Dealing with repeatable processes that can't be described by simple linear relations.

Automating ourselves back into manual labor

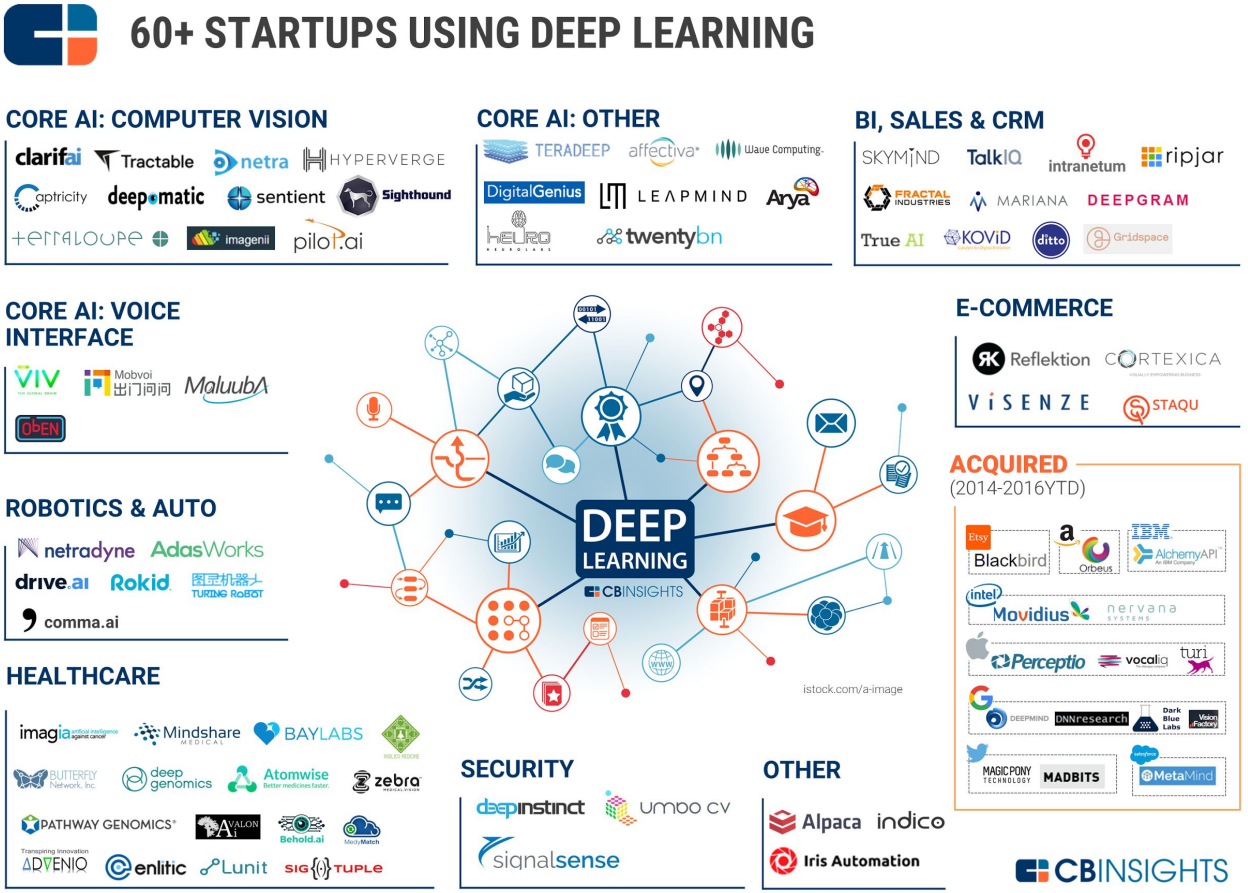


Picture from: https://quickdraw.withgoogle.com/data

What is deep learning? (and why do we care?)

Possibly remove this slide, or at least replace it with something more relevant to us physics people.

Yeah this is going straight to backup



outline

~15 mins / person

What is machine learning? - in current timeframe - a way of hiding our ignorance of how intelligence works. [algorithms vs models]

2 kinds of ML: supervised - provide a model to train with (classification, regression) and unsupervised - find N things (Derp learning, RNNs, CNN, RBMs..)

Dive into classification, regression, more abstractions ...NNs and come what may.
What do we talk about? - regression, trees, RF, k-NN, bayes, curse of dimensionality, gaussian mixture models

The vices of ML - overfitting, blindly trusting your ML results, error estimation, training variance, dealing with noisy data / contaminants, computational complexity, demotivating chess /go players.

Papers/quiz

Resources

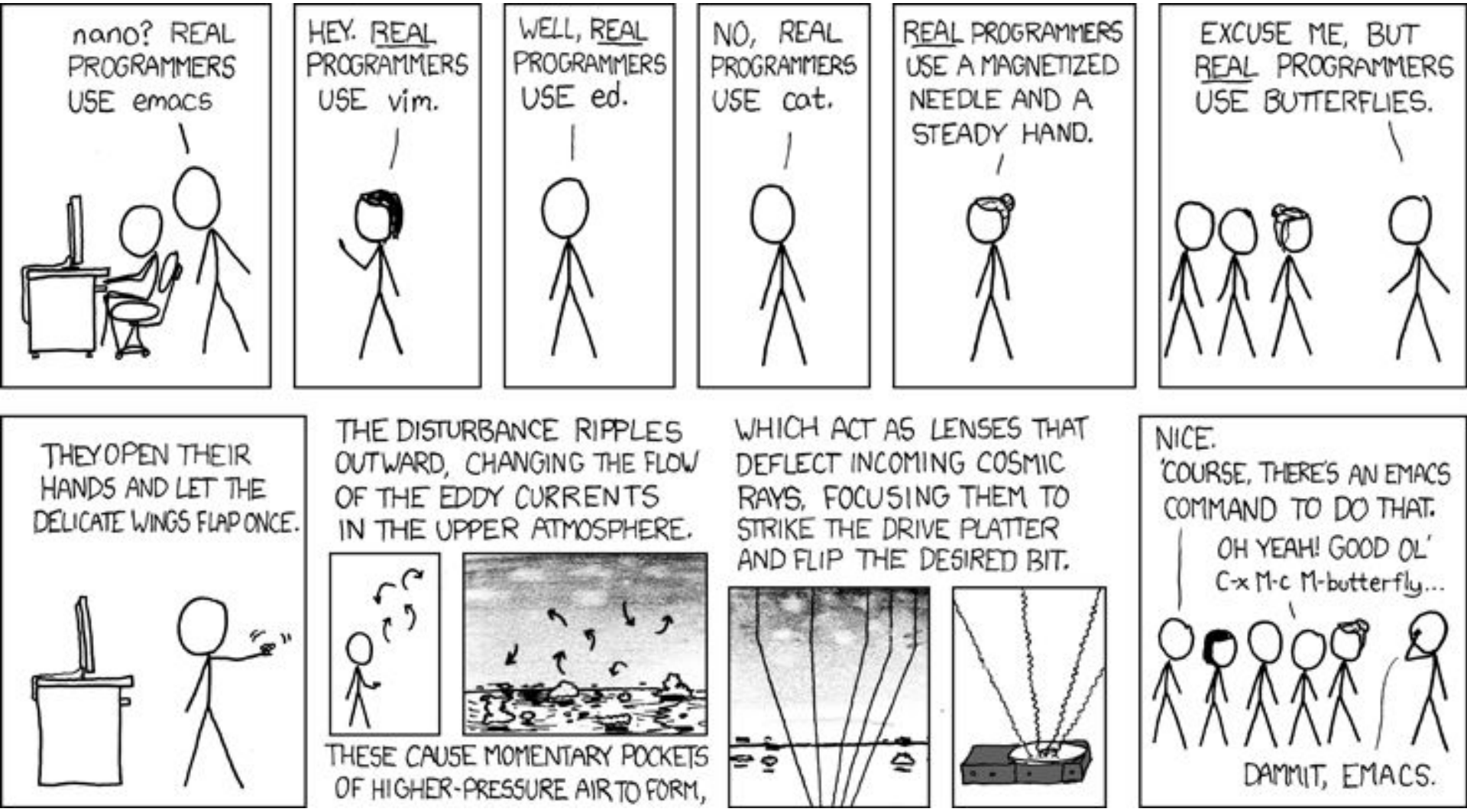
[move either to the end of the talk or right after the introduction]

Mnist
Scikit-learn
Theano/Tensorflow/Keras…
AstroML
/r/datasets, /r/dataisbeautiful …

Raghav, can you add more places to start off with CNNs? I haven't added any of those yet apart from the representative Theano etc.

Ive only used Keras and its easy to get started on. I have not used CNNs in other places actually…

Emacs all the way...



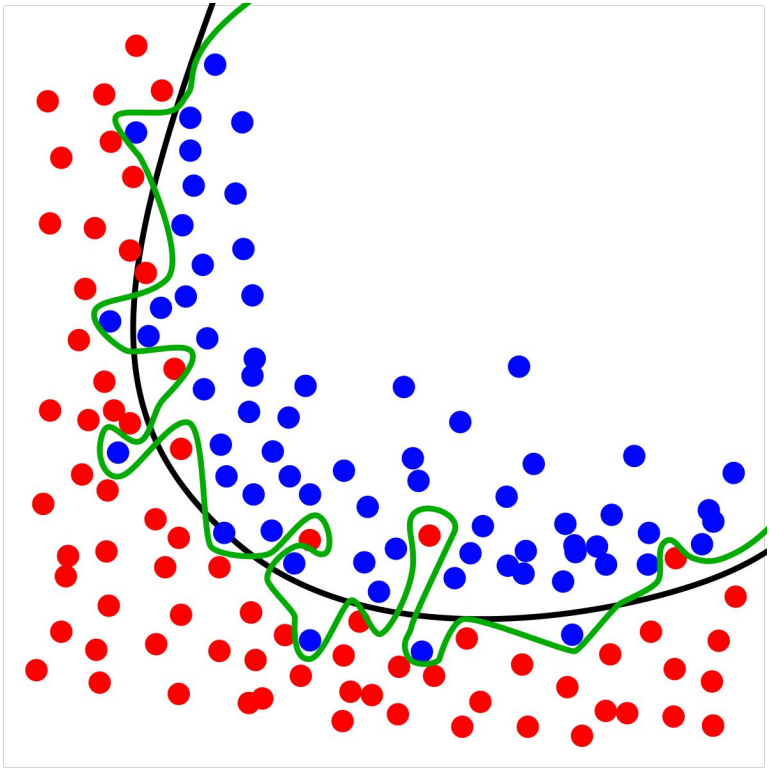
Physics literature using ML techniques:

What is machine learning? Chang+16, https://arxiv.org/abs/1709.10106v1
http://www.nature.com/nphys/journal/v13/n5/full/nphys4053.html
ML for physicists course at BU: http://physics.bu.edu/~pankajm/PY895-ML.html
Astronomy and Particle physics in general have a ton of data, so lots of papers there…
Condensed matter is also starting to catch on

ML: Pitfalls to avoid

Know what training and test data you're working with.

- Missing data
- Unrepresentative distributions
- Outliers!
- Overfitting = your model sucks
- No free lunch theorem



Mixture models [possibly backup]

Probabilistic models useful for identifying components of an observed distribution.

Gaussian mixture models often used to separate fuzzy data.

Used in combination with other methods (MCMC, SVD, Spectral methods) to boost speed and/or accuracy.