Novel uses of machine learning

Tom Charnock

Machine learning

Current projects and collaborations

Machine learning

What is machine learning?

Predict an outcome based on algorithm learned from data.

Neural networks can be universal representors.

Machine learning

What is machine learning?

Predict an outcome based on algorithm learned from data.

Neural networks can be universal representors.

Why do I study machine learning?

Beginning to thrive in astronomy.

We have loads of data!

Machine learning

What is machine learning?

Predict an outcome based on algorithm learned from data.

Neural networks can be universal representors.

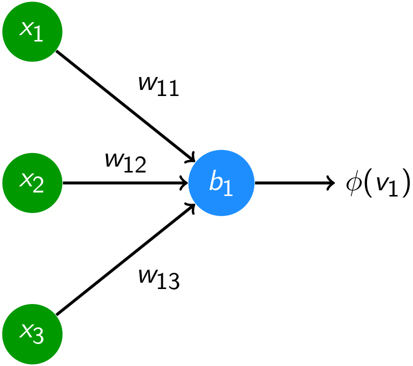
Why do I study machine learning?

Beginning to thrive in astronomy.

We have loads of data!

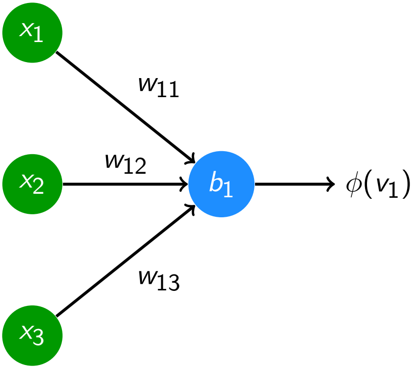
How do neural networks work?

How do neural networks work?



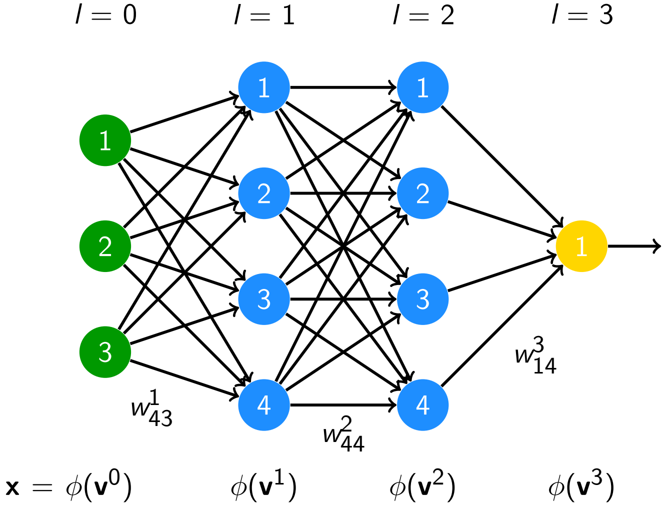
$$v_j^l = \sum_i w_{ji}^l x_i^l + b_j^l$$

How do neural networks work?



$\phi(v_i^l)$ is a $\mathrm{non-linear\ activation\ function}$

How do neural networks work?



How do neural networks work?

Loss function, λ

Calculate difference between prediction and label.

How do neural networks work?

Loss function, λ

Calculate difference between prediction and label.

Backpropagate to find $\dfrac{\partial\lambda}{\partial\mathbf{w}^l}$ and $\dfrac{\partial\lambda}{\partial\mathbf{b}^l}$

Use the chain rule to find how the loss function changes with respect to the weights and biases.

How do neural networks work?

Loss function, λ

Calculate difference between prediction and label.

Backpropagate to find $\dfrac{\partial\lambda}{\partial\mathbf{w}^l}$ and $\dfrac{\partial\lambda}{\partial\mathbf{b}^l}$

Use the chain rule to find how the loss function changes with respect to the weights and biases.

Find the error, $\dfrac{\partial\lambda}{\partial\phi(v^L)}$

Calculate gradient of the loss function with respect to the output.

How do neural networks work?

Loss function, λ

Calculate difference between prediction and label.

Backpropagate to find $\dfrac{\partial \lambda}{\partial \mathbf{w}^l}$ and $\dfrac{\partial \lambda}{\partial \mathbf{b}^l}$

Use the chain rule to find how the loss function changes with respect to the weights and biases.

Find the error, $\dfrac{\partial \lambda}{\partial \phi(v^L)}$

Calculate gradient of the loss function with respect to the output.

Update weights and biases

$$\mathbf{w}^l \rightarrow \mathbf{w}^l - \eta \frac{\partial \lambda}{\partial \mathbf{w}^l} \quad \text{and} \quad \mathbf{b}^\mathbf{l} \rightarrow \mathbf{b}^\mathbf{l} - \eta \frac{\partial \lambda}{\partial \mathbf{b}^l}$$

Types of neural networks

Convolutional neural networks

Weights are contained in a convolutional filter and shared over the input.

Useful for work with images and high dimensional, spatially invariant problems.

Types of neural networks

Convolutional neural networks

Weights are contained in a convolutional filter and shared over the input.

Useful for work with images and high dimensional, spatially invariant problems.

Recurrent neual networks

Output of previous input in a sequence of data is used to influence the training of the next input.

Useful for time-series data or for semantic processing.

Types of neural networks

Convolutional neural networks

Weights are contained in a convolutional filter and shared over the input.

Useful for work with images and high dimensional, spatially invariant problems.

Recurrent neual networks

Output of previous input in a sequence of data is used to influence the training of the next input.

Useful for time-series data or for semantic processing.

Generative neural networks

Generates outputs by learning how to map from latent space to true distributions.

Useful for making realisitic looking, but non-real images (data).

How do I use machine learning at the IAP?

Current projects and collaborations

GRAND project

Anne Zilles & Florian Führer

BORG bias

Guilhem Lavaux (and the BLSS group)

Learned-inference

Alessandro Manzotti

Painting halos

Doogesh Kodi-Ramanah

DEEPVIS

DEEPVIS team

Information maximising neural networks

Guilhem Lavaux & Ben Wandelt

Various statistical projects

Ben Wandelt (and members of the Flatiron institute)

Cosmic strings

Ed Copeland, Tasos Avgoustidis & Adam Moss

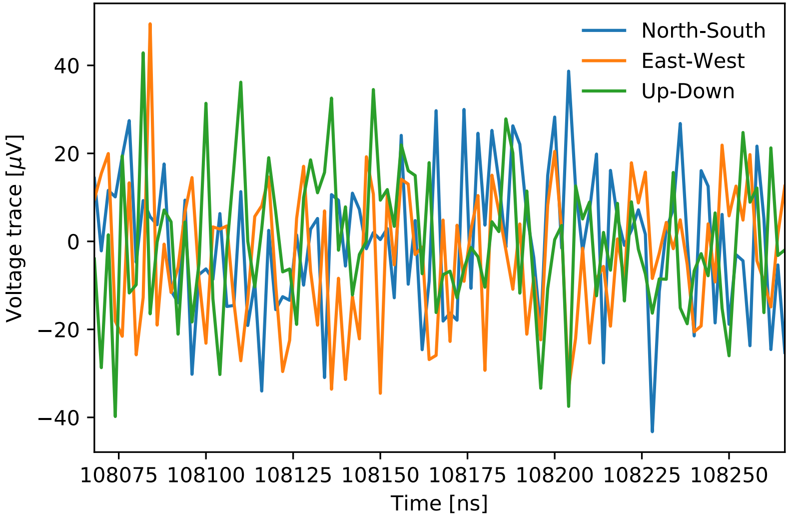
GRAND project - Anne Zilles & Florian Führer

GRAND

GRAND project - Anne Zilles & Florian Führer

Radio shower trigger

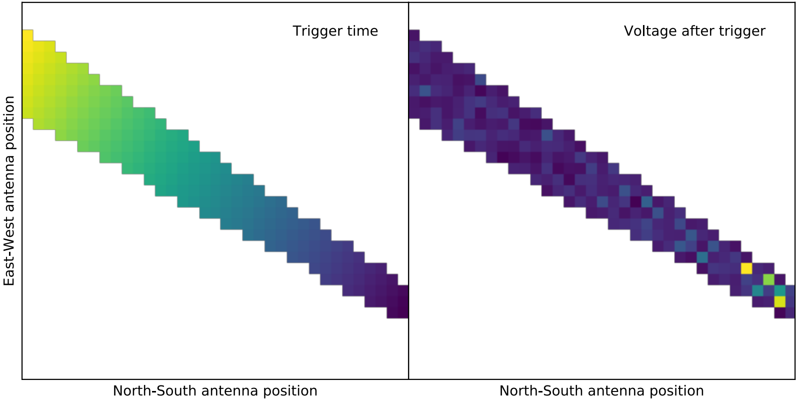
Using recurrent neural networks to learn when antenna becomes active.



GRAND project - Anne Zilles & Florian Führer

Near instant classification

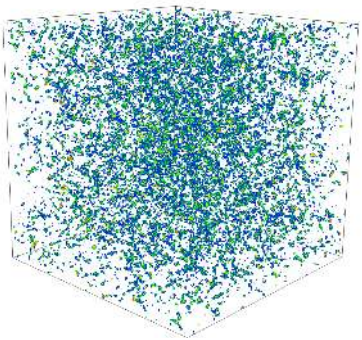
Use recurrent convolutional neural network to detect particle shower direction, particle type, event energy, etc.

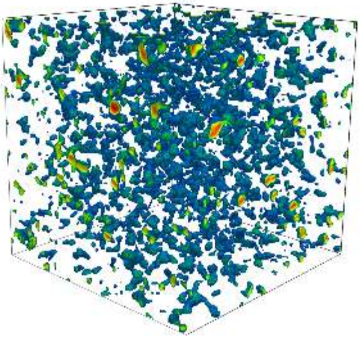


Painting halos - Doogesh Kodi-Ramanah

Paint halo fields from dark matter density

Use WGAN and variational techniques to find the map from δ field to halo field using only simulations.

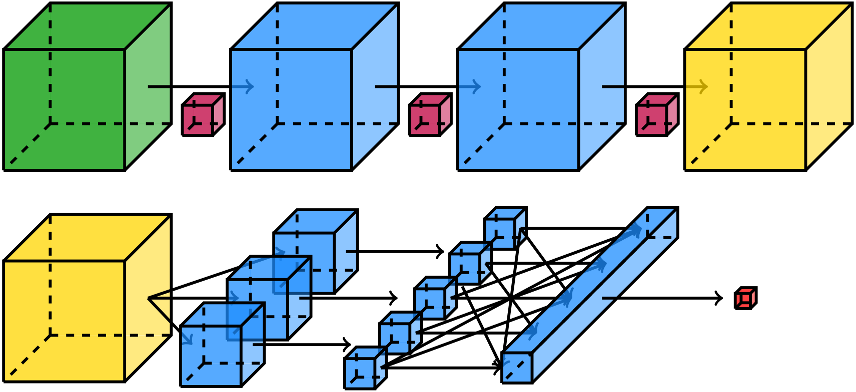




Painting halos - Doogesh Kodi-Ramanah

Paint halo fields from dark matter density

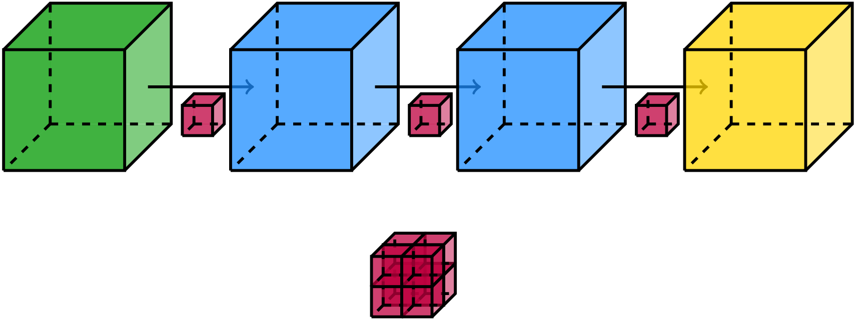
Use WGAN and variational techniques to find the map from δ field to halo field using only simulations.



BORG - BLSS group

Painting galaxy counts from dark matter density

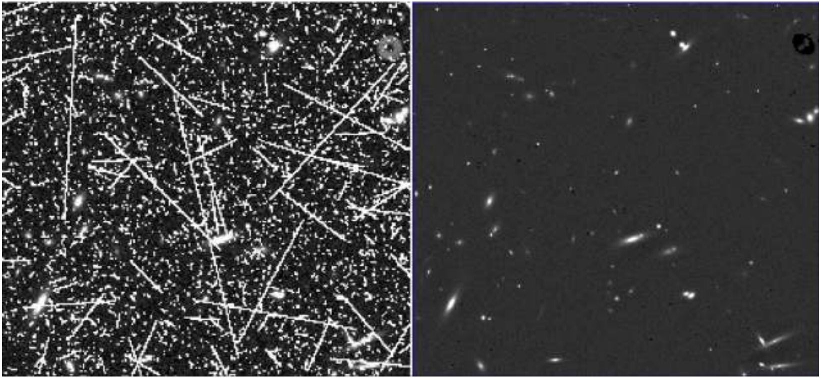
Incorporate learned bias map into the BORG framework.



DEEPVIS - DEEPVIS team

Anomaly detection in EUCLID imaging pipeline

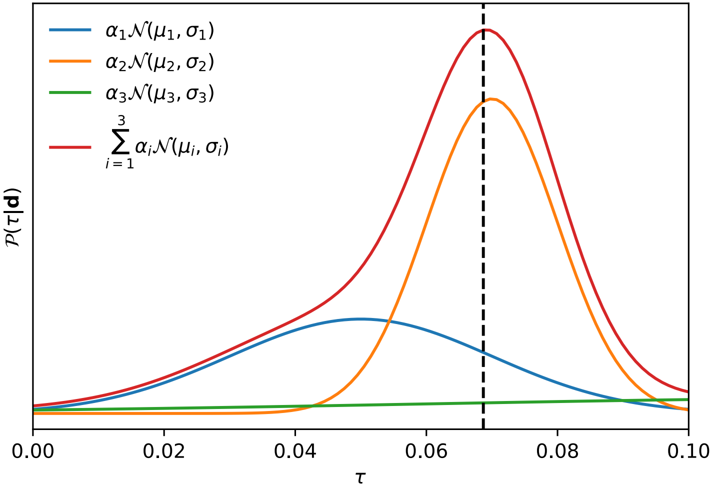
Learn EUCLID imaging pipeline, detect if network maps incorrectly.



Learned-inference - Alessandro Manzotti

Conditional probability distributions of parameters

Learn shape of the likelihood distribution function using mixture density networks.



Current projects and collaborations

GRAND project

Anne Zilles & Florian Führer

BORG bias

Guilhem Lavaux (and the BLSS group)

Learned-inference

Alessandro Manzotti

Painting halos

Doogesh Kodi-Ramanah

DEEPVIS

DEEPVIS team

Information maximising neural networks

Guilhem Lavaux & Ben Wandelt

Various statistical projects

Ben Wandelt (and members of the Flatiron institute)

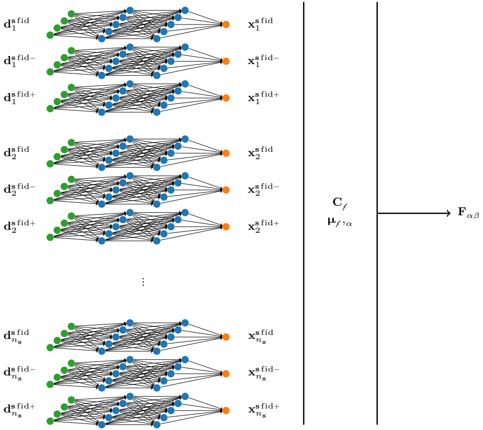
Cosmic strings

Ed Copeland, Tasos Avgoustidis & Adam Moss

Information maximising neural networks

Learn to compress data without a known likelihood

Train a neural network to find the best function to compress data, non-linearly, by maximising the Fisher information.



Information maximising neural networks

Automatic physical inference

Can use the compressed summaries to do ABC to find posterior distributions of parameters by comparing compressed summary of simulations to compressed summary of real data.