ML in Fundamental Physics – Summary / Overview lecture

→ Basics of Machine Learning:

- × how to approach problem :– ML (split into training / test set)
- × Bias vs variance tradeoff

→ Optimisers:

- × gradient descent
- × Newton's method vs. gradient descent
- × Learning rate : choices
- × SGD
- × Modifications to gradient descent
  - × What's meant with momentum?
  - × ~~Hessian~~ 2nd moment methods (RMS-Prop, Adam)
  - × What's the idea behind modifications?

→ Simple regression:

- × What's a regression problem
- × Least square regression
- × Ridge & Lasso regularisation
- × Bayesian formulation

Logistic / multi-class classification

- Logistic binary classification with logistic sigmoid
- MLE
- Multi-class classification (softmax)
- Perceptron as a binary classifier & perceptron capabilities
- Shannon information content

Other classifiers / methods

- SVM (kernel trick)
- Decision trees (How do they work)
- Bagging
- Boosting
- Random Forests

Neural Networks

- × Activation Functions
- × Dense Layers
- × Cost functions
- × Universal Approximation Theorem
- × Backpropagation – aka how to optimise a neural network
  - ↳ What can happen with gradients
- × Regularisation methods: l_1, l_2, dropout, batch-normalisation, training with mini-batches
- × Practicalities: How do you set up NN
- × Convolutional layer
  - ↝ Motivation
  - ↝ ~~Realisation~~ Implementation
  - ↝ Connection to dense layers
- × Pooling layers
- × Standard datasets: MNIST, ImageNet
- × Training strategies, hyperparameter tuning
- × RNNs (idea, problems, LSTM)

Unsupervised learning

→ What's that?

- PCA (what's interpretation of first component) [important]
- Multi-dim scaling
- K-means clustering
- Variations of clustering (agglomerative clustering, density based clustering
- Autoencoders
- t-SNE

Advanced topics

a) Variational methods

× KL-divergence

× variational free energy minimisation (def, why)

× Example: var free energy for spin systems (mean field equations)

× Hopfield networks (def, training, capacity)

× Boltzmann-machines (defn.) / energy based models

→ Negative log-likelihood

- Example: Restricted Boltzmann Machines
  (how to train them?)
- Defn. Deep Boltzmann machine
- Variational autoencoder (reparametrization trick)
  implementation

b) Generative Adversarial Networks

- Idea
- Minimisation objective
  → limit of JS-divergence
- Variation: Wasserstein-GAN (Idea?)
- Implementation (→ potential problems)

c) Reinforcement learning

- Vocabulary (agent/worker, environment, actions, policy, reward, return)
- General procedure
- Policies: ε-greedy, greedy
  on-policy, off-policy (defn.)
- SARSA, Q-learning
- How do you deal with large state spaces?
  ↳ (e.g. Atari games)

Applications of ML in fundamental Physics

- ML in string landscape, model building, string geometries
- ML & holography
- ML & differential equations
- ML and phase classification (supervised & unsupervised)