# IAML DL - Study Guide - Week 6

Sambit Paul, Pavlos Andreadis, Nigel Goddard

January 2022

## 1 Introduction

Week 6 starts with an examination of ethical issues in machine learning research and practice. For more material, see these class notes by Shannon Vallor.

The next section introduces Nearest Neighbours method for classification. This involves using a distance metric to determine clusters of training points and determine classes based on which cluster the new data point falls in. A practical introduction to both supervised and unsupervised method can be found in this article.

## 2 Ethics and Machine Learning

- **General ethical concerns**

  Data technology is not neutral - it has design choices baked in, and use is a choice too. A key perspective is that ethics in machine learning is a process not a checklist. It is an ongoing conversation amongst stakeholders, including you. Key questions to ask are:

  - Who are the stakeholders?
  - Who benefits and how?
  - Who could be harmed and how?

  We can divide the general ethical concerns into three categories: benefits, harms and challenges.

  - Benefits - want to maximise these
    - Increase human understanding of nature, society and our personal lives
    - Increased social, institutional and economic efficiency
    - Accurate prediction and personalization
  - Harms - want to avoid these

        - Harms to fairness and justice
        - Harms to transparency and accountability
        - Harms to privacy and security
    - Data challenges - want to address these
        - Appropriate collection and use
        - Data stewardship
        - Data cleanliness and relevance
        - Ethically harmful data bias
        - Validation/test of models and analytics
        - Human accountability in data systems
        - Understanding personal, social and business impacts of data practices

- **Fairness**

  People should not be discriminated against or disparately impacted based on their membership of a protected group or class, such as race, gender, sexual orientation, *etc.* There can be bias in data collection, bias in data labelling, and outcome bias. Using protected characteristics in models is generally a bad idea; and simply not doing that is not sufficient, as there may be other features that are correlated with a protected characteristic. One way to assess fairness in classification systems is to compare ROC curves for different groups of interest (e.g., partitioned by protected characteristic) - if these differ significantly there may be a problem. How to address it is usually context dependent, and it can be the case that achieving fairness leads to a decrease in overall performance.

- **Accountability**

  Human decision makers can be asked to account for their decisions, including moral and ethical choices. These do not apply to ML models, but they do apply to the designers, implementors, and users of the models. Some design choices can facilitate accountability - e.g. a Decision Tree is human readable with specific choice points, but a deep neural network may have no accessible human interpretation.

- **Transparency**

  There are several aspects including outcome transparancey - knowing how and why a model produced the outcome it did; and process transparency - being open about the processes and choices that went into building the model.

## 3 Nearest Neighbours

- Nearest Neighbours algorithm works under the principle of "*similar things exist in close proximity*".
- A basic mathematical intuition is given in Hastie et al. [2009] Section 2.3.2 where they have given examples of using $N$-neighbours for creating decision boundaries based on local clustering of data.
- *Voronoi Tessellation*: The smaller the number of neighbours for clustering, the more granularly the decision boundaries are fragmented. For very small numbers like 1-2, this fragmentation is called Voronoi tessellation. You can read more about this in this article.
- To understand few of the issues that Kearest Neighbours method suffers from, you can refer to Barber [2012] Section 14.1 (Pg. 317 - Pg. 318).
- 
  - Larger value of K, means all points may be classified as the class with more data points.
  - Smaller value of K, means the model is not generalisable and can cause large fluctuations for small changes in the data.

  The choice of the number of neighbours to use (**K**) can be identified using validation. This can be considered parameter tuning for a machine learning model.
- One of the key differences between a linear decision boundary built using methods like SVM based on least-squares method and nearest neighbours based decision boundary is the fact that there is an underlying assumption about the data distribution being linearly separable. To read further on this, please refer to Hastie et al. [2009] Section 2.3.3.
- Please refer to this video to understand more about Kernels and Parzen Windows.
- To know more about the ongoing research on kNNs and how they are improving upon the existing method, you can refer to Zhang et al. [2017] and Wu et al. [2008] Section 8.4.
- 
  - *KD trees* can be considered a combination of decision trees and kNN algorithm in which each split in the tree is based on the median value of a specific feature. Each leaf node contains "k" points against which the nearest neighbours calculation can be done.
    To know more about this, please refer to Section 6.3 and Section 6.4 of this article
  - This article provides a clear and succinct explanation of *Locality-Sensitive Hashing*. You can refer to this paper Zhang et al. [2013] to understand how LSH improves the efficiency of kNNs. The abstract

of the paper gives a clear overview of the idea, but to get a deeper understanding, it might be useful to refer to Algorithm 2 in the paper.

  - This video provides a good explanation of inverted list based nearest neighbours search.

- For a probabilistic perspective on nearest neighbours, a very succinct explanation is provided in Barber [2012] Section 14.3.

# References

David Barber. *Bayesian reasoning and machine learning*. Cambridge University Press, 2012.

Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009.

Xindong Wu, Vipin Kumar, J Ross Quinlan, Joydeep Ghosh, Qiang Yang, Hiroshi Motoda, Geoffrey J McLachlan, Angus Ng, Bing Liu, S Yu Philip, et al. Top 10 algorithms in data mining. *Knowledge and information systems*, 14 (1):1–37, 2008.

Shichao Zhang, Xuelong Li, Ming Zong, Xiaofeng Zhu, and Ruili Wang. Efficient knn classification with different numbers of nearest neighbors. *IEEE transactions on neural networks and learning systems*, 29(5):1774–1785, 2017.

Yan-Ming Zhang, Kaizhu Huang, Guanggang Geng, and Cheng-Lin Liu. Fast knn graph construction with locality sensitive hashing. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 660–674. Springer, 2013.