ECE324 Lectures

Hanhee Lee

January 9, 2025

Contents

Definition:

Process:

Motivation:

Derivation:

Warning:

Summary:

Algorithm:

Example:

FAQ:

1 Why Systems Software?

1.0.1 Learning algorithm

1.0.2 Linear regression

Example:

$$W \cdot x = y \tag{1}$$

- E: x, y
- P: Error rate (e.g. mean squared error)
- T: Predict y from x (i.e. $\hat{y}$)
- M: Linear model (W)
- Learn:

1.1 Linear transformations

Definition: Transform from one vector space to another.

$$W \cdot x$$

Example: Linear transformations warped to a prediction target.

$$\text{Link}(W \cdot x) = y$$

- x: Input features
- y: Output / target
- W: Linear transformation
- Link: Warping function
- If x has dim 50 and W projects to dimension 100, what is the shape of W? 100×50.
- If W is learnable, how many parameters does W have? $100 \times 50 = 5000$.

1.2 How does a generalized linear model make a prediction?

Definition: By either mapping to a line (regression) or separating data by a line (i.e. hyperplane) (classification).

1.3 Representation

Definition:

- Feature vectors
- Embeddings
- Latent codes
- Intermediate activations, etc.

1.4 Neural networks

Definition: Learnable (optimizable) transformations of data

1.4.1 2-Layer Neural Net (MLP)

Example: By stacking linear transforms with activation functions.

$$\text{Link}(W_2 \cdot \text{relu}(W_1 \cdot x)) = y$$

- E: x and y

- P: Mean squared error
- T: Predict y from x
- M: Neural net (W_1, W_2)
- Learn: Gradient descent

1.4.2 Encoder-Decoder View of Neural Networks

$$x \underset{\text{Encode}}{\mapsto} z \underset{\text{Decode}}{\mapsto} \hat{x}'$$

Language translation via text encoder/decoders.

Example: Using learned representations for new problems

Encoder and decoder view of Supervised Learning.

Example: PCA

1.5 Specifying an ML task

1.6 Landscape of tools and approaches

2 Kernels