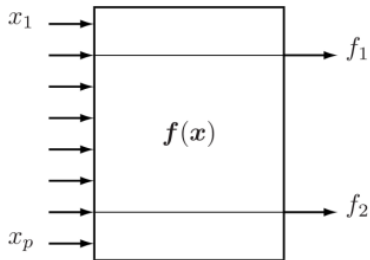


Feature Selection

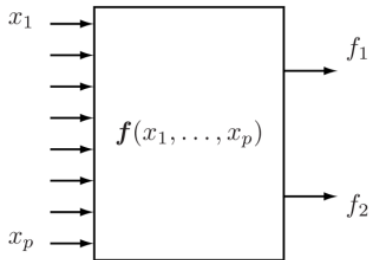
Clayton W. Seitz

February 21, 2022

Feature Selection and Extraction



(a) feature selector



(b) feature extractor

Feature Selection

What is it?

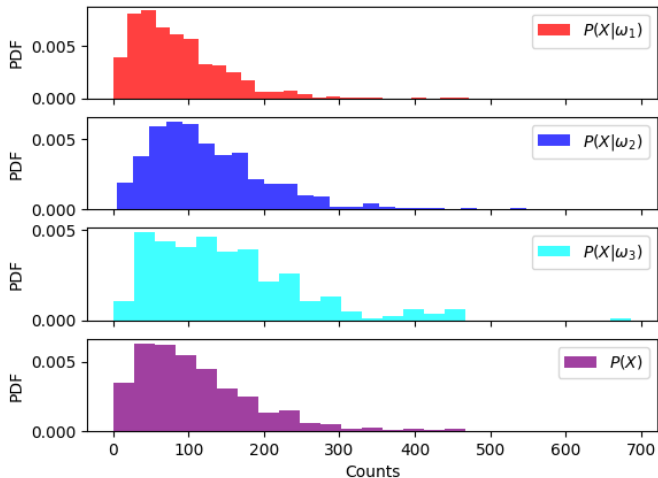
A special type of dimensionality reduction where we select a subset of features, in contrast with feature extraction like in PCA, VAE

Why do we do it?

- ▶ Quality of the input data is just as important as the algorithm you choose
- ▶ The volume of a feature space grows exponentially in the number of dimensions n
- ▶ But we often have a small number of samples $p \ll n$

Probabilistic Distance Measures

How do we define a notion of distance for probability distributions?



Symmetric Kullbeck-Leibler (KL) Divergence

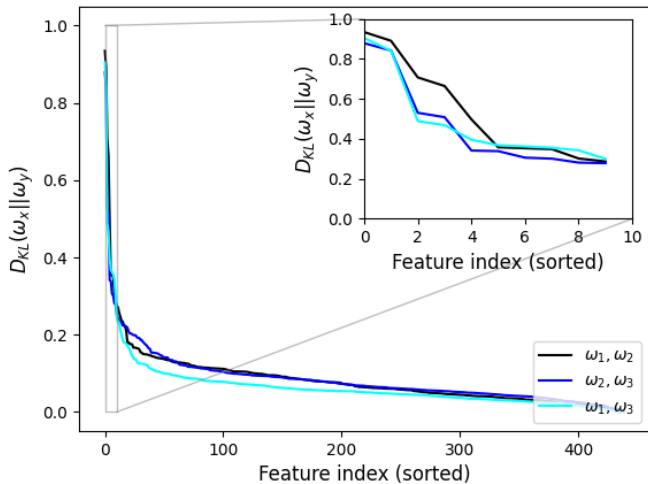
The standard definition of KL-Divergence is

The form that has been used in feature selection for classification tasks reads

$$S = D_{KL}(P||Q) + D_{KL}(Q||P)$$

The major drawback of univariate methods is that we do not consider contextual changes

Result



References I