

Dimensionality Reduction

*** imp

* this technique reduce no. of input variables in a dataset

* Higher no. of features is harder to visualize the training set and to work in that

* features can be correlated and hence redundant
in such case we need DR (Dimensionality Reduction)

* DR is a process of reducing no. of random variables under correlation by updating a set of principle variables

* DR can be divided into 2 types

1. feature selection

2. feature extraction

1. feature selection

* it is the process of selecting the subset of the relevant feature and leaving out the irrelevant feature present in the dataset to build a model of high accuracy

* it is the way of selecting optimal features from input dataset

Methods of feature selection

1. filter
2. wrapper
3. Embedded

1. Filter method

→ Data set filtered with only relevant features

1. correlation
2. chi square test
3. ANOVA
4. Information gain

2. Wrapper method

→ filtered relevant feature using machine learning technique.

1. forward selection
2. Backward selection
3. Bidirectional elimination

3. Embedded

→ It checks different training iteration of ML model and ~~also~~ evaluate important of each feature

1. Lasso
2. Elastic net
3. Ridge regression

2. Feature Extraction

It is the process of transforming the space containing many dimensions into space with fewer dimensions.

1. PCA

2. LDA (Linear Discriminant Analysis)

3. QDA

Applications

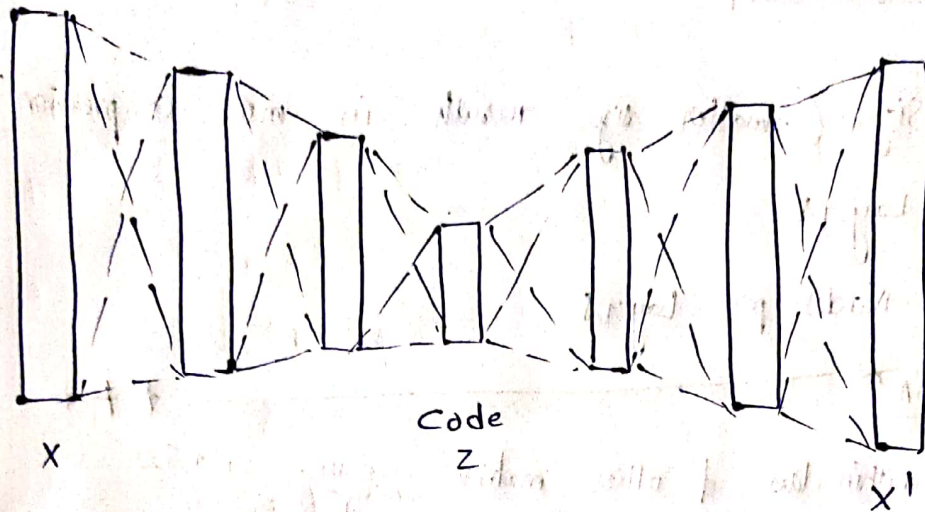
1. Speech Recognition

2. Signal Processing

3. Bio Informatic

Autoencoder

⇒ it is an unsupervised ANN that compresses the data to lower dimension and then reconstructs the input back



x - input \rightarrow encode

x' - output \rightarrow decode

z - code / reconstruct

⇒ it finds the representation of data in a lower dimension by focusing more on the important features getting rid of noise and redundancy

⇒ Based on Encoder and Decoder task, Encoder encodes the high dimensional data to low dimensional data and decoder takes the low dimensional data to reconstruct the high dimensional data

⇒ The mapping of higher to lower dimension can be linear or non-linear depending upon the choice of activation function

Hyper parameter

→ Hyperparameter that we need to set before training the autoencoder

1. Code Size (smaller size results in more compression)
2. No. of Layers
3. No. of Nodes per Layer