Representation Learning
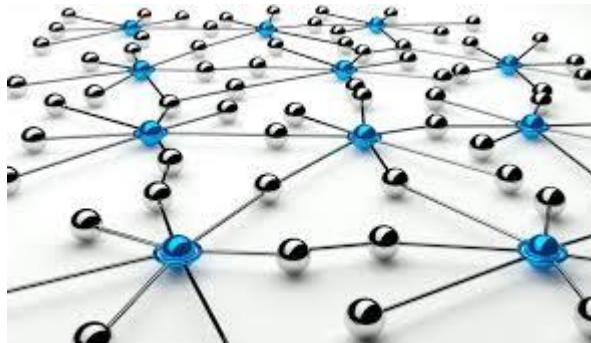# (Embedding)

Representation

In any machine learning task, representation is of the samples in the dataset is the first and foremost task.

Representation

In any machine learning task, representation is of the samples in the dataset is the first and foremost task.

	#Wheel	Height	Weight	Color
	4	6	500	Red
	4	5.5	600	Blue
	4	5	550	Yellow
	2	3	200	Red
	2	3.5	150	blue
	2	4	250	Yellow

Representation

Issues

- Need of domain knowledge
- Expert Bias

Can we automatically learn the representation from the dataset?

Embedding

Embedding is a process of automatically learning representation vector of the samples in the low dimensional space from the data.

Embedding

Embedding is a process of automatically learning representation of the samples in the low dimensional space from the data.

It generally transform a high dimensional data to a lower dimensional data in the form of a vector, in such as way that the lower dimensional vector representation still maintains the semantic meaning of the data in the higher dimensional space.

Why Embedding?

- Challenges of feature engineering to determine appropriate features

- Course of dimensionality

Types of embeddings

Depending on the types of the data, different types of embedding are proposed.

- Text : Word Embedding

- Image: Feature mapping

- Network: Network Embedding