Title: Similarity learning

URL: https://en.wikipedia.org/wiki/Similarity_learning

PageID: 38059657

Categories: Category:Machine learning, Category:Semantic relations

---

Similarity learning is an area of supervised machine learning in artificial intelligence . It is closely related to regression and classification , but the goal is to learn a similarity function that measures how similar or related two objects are. It has applications in ranking , in recommendation systems , visual identity tracking, face verification, and speaker verification.

Learning setup

There are four common setups for similarity and metric distance learning.

A common approach for learning similarity is to model the similarity function as a bilinear form . For example, in the case of ranking similarity learning, one aims to learn a matrix W that parametrizes the similarity function f W (x , z) = x T W z {\displaystyle f_{W}(x,z)=x^{T}Wz} . When data is abundant, a common approach is to learn a siamese network – a deep network model with parameter sharing.

Metric learning

Similarity learning is closely related to distance metric learning . Metric learning is the task of learning a distance function over objects. A metric or distance function has to obey four axioms: non-negativity , identity of indiscernibles , symmetry and subadditivity (or the triangle inequality). In practice, metric learning algorithms ignore the condition of identity of indiscernibles and learn a pseudo-metric.

When the objects x i {\displaystyle x_{i}} are vectors in R d {\displaystyle R^{d}} , then any matrix W {\displaystyle W} in the symmetric positive semi-definite cone S + d {\displaystyle S_{+}^{d}} defines a distance pseudo-metric of the space of x through the form D W (x 1 , x 2) 2 = (x 1 − x 2) ■ W (x 1 − x 2) {\displaystyle D_{W}(x_{1},x_{2})^{2}=(x_{1}-x_{2})^{\top }W(x_{1}-x_{2})} . When W {\displaystyle W} is a symmetric positive definite matrix, D W {\displaystyle D_{W}} is a metric. Moreover, as any symmetric positive semi-definite matrix W ∈ S + d {\displaystyle W\in S_{+}^{d}} can be decomposed as W = L ■ L {\displaystyle W=L^{\top }L} where L ∈ R e × d {\displaystyle L\in R^{e\times d}} and e ≥ r a n k (W) {\displaystyle e\geq rank(W)} , the distance function D W {\displaystyle D_{W}} can be rewritten equivalently D W (x 1 , x 2) 2 = (x 1 − x 2) ■ L ■ L (x 1 − x 2) = ■ L (x 1 − x 2) ■ 2 2 {\displaystyle D_{W}(x_{1},x_{2})^{2}=(x_{1}-x_{2})^{\top }L^{\top }L(x_{1}-x_{2})=\|L(x_{1}-x_{2})\|_{2}^{2}} . The distance D W (x 1 , x 2) 2 = ■ x 1 ′ − x 2 ′ ■ 2 2 {\displaystyle D_{W}(x_{1},x_{2})^{2}=\|x_{1}'-x_{2}'\|_{2}^{2}} corresponds to the Euclidean distance between the transformed feature vectors x 1 ′ = L x 1 {\displaystyle x_{1}'=Lx_{1}} and x 2 ′ = L x 2 {\displaystyle x_{2}'=Lx_{2}} .

Many formulations for metric learning have been proposed. [4] [5] Some well-known approaches for metric learning include learning from relative comparisons, [6] which is based on the triplet loss , large margin nearest neighbor , [7] and information theoretic metric learning (ITML). [8]

In statistics , the covariance matrix of the data is sometimes used to define a distance metric called Mahalanobis distance .

Applications

Similarity learning is used in information retrieval for learning to rank , in face verification or face identification, [9] [10] and in recommendation systems . Also, many machine learning approaches rely on some metric. This includes unsupervised learning such as clustering , which groups together close or similar objects. It also includes supervised approaches like K-nearest neighbor algorithm which rely on labels of nearby objects to decide on the label of a new object. Metric

learning has been proposed as a preprocessing step for many of these approaches. [11]

Scalability

Metric and similarity learning scale quadratically with the dimension of the input space, as can easily see when the learned metric has a bilinear form f W (x , z) = x T W z {\displaystyle f_{W}(x,z)=x^{T}Wz} . Scaling to higher dimensions can be achieved by enforcing a sparseness structure over the matrix model, as done with HDSL, [12] and with COMET. [13]

Software

metric-learn [14] is a free software Python library which offers efficient implementations of several supervised and weakly-supervised similarity and metric learning algorithms. The API of metric-learn is compatible with scikit-learn . [15]

OpenMetricLearning [16] is a Python framework to train and validate the models producing high-quality embeddings.

Further information

For further information on this topic, see the surveys on metric and similarity learning by Bellet et al. [4] and Kulis. [5]

See also

Kernel method

Latent semantic analysis

Learning to rank

References