Distance Measures

Distance Measure	Formula	Description
Euclidean (L_2) norm	$\|u,v\|_2 = \sqrt{\sum_i (u_i - v_i)^2}$	Standard distance in Euclidean space. Most common geometric distance.
Manhattan (L_1) norm	$\|u,v\|_1 = \sum_i \lvert u_i - v_i \rvert$	Sum of absolute differences. Useful for sparse data and robust models.
Maximum (L_∞) norm	$\|u,v\|_\infty = \max_i \lvert u_i - v_i \rvert$	Only considers the largest component-wise difference.
L_p-norm	$\|u,v\|_p = \left(\sum_i \lvert u_i - v_i \rvert^p\right)^{1/p}$	Generalized distance measure. $p = 1, 2, \infty$ are special cases.
Mahalanobis Distance	$D(u,v) = \sqrt{(u-v)^T S^{-1} (u-v)}$	Takes correlations and scaling into account via covariance matrix S. Useful in classification.
Cosine Distance	$D(u,v) = 1 - \frac{u \cdot v}{\|u\|\|v\|}$	Measures angle between vectors. Often used in NLP, recommender systems.
Hamming Distance	$D(u,v) = \sum_i [u_i \neq v_i]$	Counts differing positions in binary vectors. Common in genetics, information theory.
Jaccard Distance	$D(A,B) = 1 - \frac{\lvert A \cap B \rvert}{\lvert A \cup B \rvert}$	Measures dissimilarity between sets.
Canberra Distance	$D(u,v) = \sum_i \frac{\lvert u_i - v_i \rvert}{\lvert u_i \rvert + \lvert v_i \rvert}$	Sensitive to small differences when components are near zero.
Bray-Curtis Distance	$D(u,v) = \frac{\sum_i \lvert u_i - v_i \rvert}{\sum_i \lvert u_i + v_i \rvert}$	Used in ecology for compositional dissimilarity.
Earth Mover's Distance (Wasserstein)	Informal: minimal "work" to transform one distribution into another	Popular in image comparison, optimal transport, GANs. Metric over distributions.
Bregman Divergence	$D_F(u,v) = F(u) - F(v) - \nabla F(v)^T (u - v)$	General class of distances (e.g., squared Euclidean, KL divergence).
KL Divergence (Kullback–Leibler)	$D_{\mathrm{KL}}(P\|Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}$	Measures how one probability distribution diverges from another. Not symmetric.