eCornell

CHEAT SHEET

Perceptron

Algorithm Name	Perceptron
Description	The Perceptron attempts to converge to a weight vector $\mathbf{w}$ that separates the data from two classes. In essence, it iterates through the whole dataset and when it encounters a misclassified point, it updates using $\mathbf{w} \leftarrow \mathbf{w} + y\mathbf{x}$.
Applicability	Binary classification problems.
Assumptions	Data are linearly separable and the labels have to be ± 1.
Underlying Mathematical Principles	A hyperplane is defined as $\mathbf{w}^T\mathbf{x} = 0$, where $\mathbf{w}$ determines the the orientation of the hyperplane. Predictions are the sign of $h(\mathbf{x}) = sign(\mathbf{w}^T\mathbf{x})$.
Additional Details	• Perceptron is a mistake-driven algorithm in the sense that it will update the parameter $\mathbf{w}$ when it incorrectly classifies any data. • You can append bias b to $\mathbf{w}$ by appending a 1 to all data points. • Depending on how our data points are ordered, the parameter $\mathbf{w}$ will be different. For example, if two points in different classes are misclassified at time t, then updating $\mathbf{w}$ with the first might fix classification for the second as well, in which case perceptron algorithm will stop. In another run, if updating $\mathbf{w}$ with the second point fixes classification for the first and perceptron algorithm stops, we end up with a $\mathbf{w}$ different from the first one.
Example	You want to classify cats and dogs based on number of meows (x axis) and number of woofs (y axis). The data should be linearly separable.