

Übung zur Vorlesung **Maschinelles Lernen**

Wintersemester 2008/2009

Blatt 7

Wiederholung

1. Sei eine beliebige Hyperebene H gegeben als

$$H = \left\{ \vec{x} \mid \langle \vec{\beta}, \vec{x} \rangle + \beta_0 \right\}.$$

Wieso liefert $\frac{1}{\|\vec{\beta}\|} (\langle \vec{\beta}, x \rangle + \beta_0)$ den Abstand eines Punktes \vec{x} zur Ebene?

2. Welche Eigenschaften hat ein Skalarprodukt?

Aufgabe

10 Punkte

In der Einführung zur *Support Vector Machine* wurde ein einfacher Algorithmus zur Bestimmung einer trennenden Hyperebene über die Zentroiden zweier Klassen vorgestellt. Mit den bisherigen Implementierungsaufgaben sollte es Ihnen nicht schwer fallen, diesen Algorithmus ohne Vorgabe von Hilfsklassen zu implementieren.

Die Daten sollen zunächst nur numerische Attribute und ein binominales Label $Y \in \{-1, +1\}$ enthalten.

1. Entwerfen und implementieren Sie eine Modell-Klasse, die von der Klasse

```
com.rapidminer.operator.learner.SimplePredictionModel
```

erbt und die Klassifikation anhand einer trennenden Hyperebene vornimmt.

2. Implementieren Sie einen Lern-Operator `NaiveSVM`, der auf Basis eines `ExampleSet` eine Instanz Ihrer Modell-Klasse zurückliefert. Sie sollten dazu von der Klasse `AbstractLearner` erben, die sich hier befindet:

```
com.rapidminer.operator.learner.AbstractLearner
```

3. Tragen Sie ihren Operator in die Datei `operators.xml` ein und Testen Sie Ihre Implementierung mit Hilfe des `ExampleSet-Generators`. Sie können dazu auch das Experiment `LinReg.xml` aus dem Übungsprojekt anpassen.