

aufgabe2

December 6, 2018

1 Aufgabe 20

1.1 a)

Die Lossfunktion ist ein Maß für die Ungenauigkeit einer Vorhersage zum Beispiel innerhalb eines Klassifikationsproblems. Man kann beispielsweise den quadratischen Abstand zwischen wahrem und vorhergesagten Output verwenden. Bei der Klassifikation wird mit einer Auswertungsfunktion entschieden zu welcher Klasse ein zu klassifizierender Datenpunkt (zum Beispiel ein Messpunkt). Dies geschieht auf Basis gelabelter Daten, dem Testdatensatz. Ziel des machine learning ist es die Lossfunktion für alle Inputs eines Testdatensatzes zu minimieren, um die beste Auswertungsfunktion zu finden

1.2 b)

Für unterschiedliche Klassifikationen gibt es unterschiedliche Lossfunktionen. In vielen Fällen kann die Minimierung durch Differenzieren der Lossfunktion bezüglich der Gewichte der Auswertungsfunktion erfolgen. In jedem Schritt wird das Neuronale Netz bzw. das Modell trainiert, indem entlang der Gradienten nach unten gegangen wird, die Gewichtsmatrix und der Biasvektor werden dadurch geupdated.

Als Beispiel wird die residual sum of squares (RSS) minimiert.

$$RSS(\beta) = (y - X\beta)^T(y - X\beta) \quad (1)$$

Mit Hilfe von Matrix Calculus wird die Ableitung berechnet und für die notwendige Bedingung gleich Null gesetzt.

$$2 \cdot X^T(y - X\beta) = 0 \quad (2)$$

Daraus folgt für die Gewichte.

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad (3)$$

1.3 c)

Die Aktivierungsfunktion simuliert die Aktivierung des Zellkerns einer Nervenzelle. Die Ausgabe des künstlichen Neurons wird also durch die Aktivierungsfunktion bestimmt. Die Aktivierung wird beeinflusst durch die Netzeingabe aus der Übertragungsfunktion. Mithilfe von nicht-linearen Aktivierungsfunktion ist es möglich, den Raum zu verzerren und damit nicht-linear trennbare Datensätze zu trennen, ohne weitere Layer einzuführen. Auch die Verbindungen

mehrerer Layer mit nichtlinearen Aktivierungsfunktion kann sehr mächtig sein. Beispiele für Aktivierungsfunktionen sind Sigmoidfunktionen, Tangens Hyperbolicus und ReLu (Rectified Linear Unit).

1.4 d)

Ein Neuron ist ein informationstechnisches Modell für eine reale Nervenzelle. An das Neuron übertragene Eingaben x_i werden Gewichte w_{ij} zugeordnet. Eine Übertragungsfunktion $f(x_i, w_{ij})$ berechnet aus der Wichtung der Eingaben die Netzeingabe net_j .

$$net_j = \sum_{i=0}^n x_i w_{ij} \quad (4)$$

Durch eine Aktivierungsfunktion Φ wird aus der Netzeingabe die Aktivierung o_j berechnet.

$$o_j = \Phi(net_j) \quad (5)$$

1.5 e)

Für Probleme, bei denen kein oder kaum explizites (systematisches) Wissen vorliegt und viele Daten in eine geringe Anzahl an Ergebnissen überführt werden müssen, eignen sich neuronale Netzwerke. Beispiele sind Gesichts-/Spracherkennung, Prognose von Wetter, Aktien usw. und Robotik/KI.