

Mögliche Prüfungsfragen und Lösungen zur Vorlesung von Prof. Dr. Adrian Ulges

Frage 1: Overfitting in Neuronalen Netzen

Frage: Was versteht man unter Overfitting in neuronalen Netzen und welche zwei Strategien werden zur Bekämpfung von Overfitting verwendet?

Antwort:

Overfitting tritt auf, wenn ein Modell zu gut auf die Trainingsdaten angepasst wird und dadurch die Fähigkeit verliert, auf neuen, unbekannten Daten gut zu generalisieren. Zwei gängige Strategien zur Bekämpfung von Overfitting sind:

- **Regularisierung:** Hierbei wird ein Bestrafungsterm zum Loss hinzugefügt, der extrem hohe Gewichtswerte im Modell verhindert.
- **Dropout:** In dieser Methode werden während des Trainings zufällig Neuronen deaktiviert, um das Modell zu zwingen, robuster zu lernen und sich nicht auf spezifische Merkmale zu stark zu verlassen.

Frage 2: Kreuzvalidierung im maschinellen Lernen

Frage: Erklären Sie das Konzept der Kreuzvalidierung und ihre Vorteile.

Antwort:

Kreuzvalidierung ist eine Methode zur Evaluierung von Modellen, bei der der Datensatz in mehrere Teilmengen (Folds) aufgeteilt wird. Ein Modell wird dann mehrfach trainiert, wobei jeweils ein Fold als Validierungssatz und die restlichen Folds als Trainingssatz verwendet werden. Dies wird so oft wiederholt, bis jeder Fold einmal als Validierungssatz verwendet wurde. Der Vorteil dieser Methode besteht darin, dass sie eine robustere Schätzung der Modellleistung bietet, da das Modell auf verschiedenen Teilmengen des Datensatzes getestet wird.

Frage 3: Regularisierung in neuronalen Netzen

Frage: Was ist der Zweck der Regularisierung in neuronalen Netzen und wie wird sie mathematisch formuliert?

Antwort:

Der Zweck der Regularisierung in neuronalen Netzen ist es, Overfitting zu verhindern, indem extreme Gewichtswerte vermieden werden. Mathematisch wird dies durch die Hinzufügung eines Regularisierungsterms zur Loss-Funktion erreicht:

$$L_{\text{reg}}(W) = L(W) + \lambda \cdot \|W\|^2$$

wobei λ der Regularisierungsparameter ist, der die Stärke der Regularisierung steuert, und $\|W\|^2$ die quadrierte L2-Norm des Gewichtungsvektors W darstellt.

Frage 4: Informationsgewinn in Entscheidungsbäumen

Frage: Was versteht man unter Informationsgewinn in Entscheidungsbäumen und wie wird er berechnet?

Antwort:

Informationsgewinn misst die Reduktion der Entropie in den Kindknoten nach einem Split im Entscheidungsbaum. Er wird wie folgt berechnet:

$$\text{Gain} = H(D) - \left(\frac{|D_1|}{|D|} \cdot H(D_1) + \frac{|D_2|}{|D|} \cdot H(D_2) \right)$$

wobei $H(D)$ die Entropie des Elternknotens und $H(D_1)$ bzw. $H(D_2)$ die Entropien der Kindknoten sind.

Frage 5: Bedeutung der Lernrate im Training neuronaler Netze

Frage: Warum ist die Wahl der Lernrate im Training neuronaler Netze wichtig und wie beeinflusst sie das Training?

Antwort:

Die Lernrate bestimmt die Größe der Schritte, die das Modell bei jedem Update der Gewichte macht. Eine zu hohe Lernrate kann dazu führen, dass das Modell über das Minimum der Loss-Funktion hinaus schießt und nicht konvergiert, während eine zu niedrige Lernrate das Training extrem langsam macht und in einem lokalen Minimum stecken bleiben kann. Es ist daher wichtig, die Lernrate sorgfältig zu wählen und ggf. während des Trainings anzupassen.

Frage 6: Unterschiede zwischen symbolischer und subsymbolischer KI

Frage: Was sind die Hauptunterschiede zwischen symbolischer und subsymbolischer KI?

Antwort:

Die symbolische KI basiert auf expliziten Regeln und logischen Inferenzverfahren, die Wissen in Form von Symbolen darstellen. Beispiele sind Expertensysteme und regelbasierte Systeme. Die subsymbolische KI hingegen arbeitet

mit numerischen Werten und Signalen, die durch Optimierungsverfahren verarbeitet werden. Typische Beispiele sind neuronale Netze und andere maschinelle Lernverfahren. Die symbolische KI ist interpretierbar, während die subsymbolische KI oft leistungsfähiger, aber weniger transparent ist.

Frage 7: Definition und Anwendung des Bayes'schen Theorems im maschinellen Lernen

Frage: Erklären Sie das Bayes'sche Theorem und seine Anwendung in einem Naive-Bayes-Klassifikator.

Antwort:

Das Bayes'sche Theorem beschreibt die Wahrscheinlichkeit eines Ereignisses, basierend auf vorherigem Wissen über Bedingungen, die mit dem Ereignis zusammenhängen. Es wird in einem Naive-Bayes-Klassifikator verwendet, um die Wahrscheinlichkeit einer Klasse y gegeben die Beobachtungen x zu berechnen:

$$P(y|x) = \frac{P(x|y) \cdot P(y)}{P(x)}$$

Der Naive-Bayes-Klassifikator nimmt an, dass die Merkmale x bedingt unabhängig voneinander sind, was die Berechnung vereinfacht.

Frage 8: Bedeutung des Turing-Tests in der KI

Frage: Was ist der Turing-Test und welche Bedeutung hat er in der künstlichen Intelligenz?

Antwort:

Der Turing-Test, vorgeschlagen von Alan Turing, ist ein Verfahren zur Bestimmung, ob eine Maschine menschenähnliche Intelligenz aufweist. In diesem Test kommuniziert ein Mensch über eine textbasierte Schnittstelle sowohl mit einem anderen Menschen als auch mit einer Maschine. Wenn der Mensch nicht zuverlässig zwischen dem menschlichen Gesprächspartner und der Maschine unterscheiden kann, hat die Maschine den Test bestanden. Der Turing-Test bleibt ein bedeutender, wenn auch umstrittener, Maßstab für die Beurteilung von KI.

Frage 9: Der Fluch der Dimensionalität im maschinellen Lernen

Frage: Was versteht man unter dem Fluch der Dimensionalität“ und wie wirkt er sich auf maschinelles Lernen aus?

Antwort:

Der Fluch der Dimensionalität beschreibt die Probleme, die mit der Analyse und Organisation von Daten in hochdimensionalen Räumen verbunden sind. Mit zunehmender Anzahl der Dimensionen wird der Raum exponentiell größer,

was dazu führt, dass die Daten immer spärlicher werden. Dies erschwert es maschinellen Lernmodellen, Muster zu erkennen und kann zu Overfitting führen, da mehr Daten benötigt werden, um den Merkmalsraum angemessen zu füllen.

Frage 10: Unterschied zwischen supervised und unsupervised learning

Frage: Was sind die Hauptunterschiede zwischen überwachten (supervised) und unüberwachten (unsupervised) Lernmethoden?

Antwort:

Überwachtes Lernen (supervised learning) verwendet gelabelte Daten, um ein Modell zu trainieren, das Vorhersagen oder Klassifikationen auf neue, ungesehene Daten treffen kann. Beispiele sind Klassifikations- und Regressionsprobleme. Unüberwachtes Lernen (unsupervised learning) hingegen arbeitet mit ungelabelten Daten und versucht, Muster oder Strukturen in den Daten zu erkennen, wie z.B. bei Clustering- oder Dimensionsreduktionstechniken. Während das überwachte Lernen klar definierte Ziele hat, ist das unüberwachte Lernen explorativer.

Frage 11: Backpropagation in neuronalen Netzen

Frage: Was ist der Backpropagation-Algorithmus und wie wird er zur Anpassung der Gewichte in neuronalen Netzen verwendet?

Antwort:

Backpropagation ist ein Algorithmus, der verwendet wird, um die Gewichte in neuronalen Netzen anzupassen. Er berechnet den Gradienten des Fehlers der Ausgabeschicht bezüglich der Gewichte durch Anwendung der Kettenregel der Differentiation. Der berechnete Gradient wird dann genutzt, um die Gewichte in Richtung des steilsten Abstiegs zu aktualisieren, um den Fehler zu minimieren. Backpropagation ermöglicht es, tiefere neuronale Netze effizient zu trainieren.

Frage 12: Cross-Validation und ihre Vorteile

Frage: Erklären Sie das Prinzip der Cross-Validation und warum es bei der Modellbewertung im maschinellen Lernen verwendet wird.

Antwort:

Cross-Validation ist eine Technik, bei der der ursprüngliche Datensatz in mehrere Teilmengen (Folds) aufgeteilt wird. Das Modell wird mehrfach trainiert und getestet, wobei in jeder Iteration ein anderer Fold als Testmenge und die restlichen Folds als Trainingsmenge verwendet werden. Der durchschnittliche Fehler über alle Iterationen gibt eine robustere Schätzung der Modelleistung. Diese Methode hilft, das Risiko von Overfitting zu reduzieren und bietet eine genauere Einschätzung der Modellgeneralisation.