Modell

Autor und Herausgeber

Alexander Jung*, Department of Computer Science, Aalto University, Espoo, Finland (ORCID 0000-0001-7538-0990)
Editors: Konstantina Olioumtsevits (ORCID 0000-0003-2697-8659), Ekkehard Schnoor (ORCID 0000-0001-8252-1499), Aalto University, Espoo, Finland
*Corresponding author: alex.jung@aalto.fi

Zusammenfassung

In der machine learning (ML) wird das Wort *Modell* auf verschiedene Weisen verwendet. Formal ist ein Modell ein Hypothesenraum $\mathcal{H}$: die Menge der Kandidaten-Hypothese-Abbildungen, aus denen eine ML-Methode eine auswählt (oder lernt). In der angewandten ML-Literatur und in Softwarebibliotheken bezieht sich das Wort *Modell* auch auf den trainierten predictor, d.h. auf die gelernte Hypothese, die von einem ML-Algorithmus zurückgegeben wird; der EU AI Act verwendet diese Bedeutung im zusammengesetzten Begriff general-purpose AI model (GPAI model). Eine andere Art von mathematischem Modell ist ein Wahrscheinlichkeitsmodell, das eine Familie von Wahrscheinlichkeitsverteilungen spezifiziert, die einen Daten-Generierungsprozess charakterisieren.

Definition

Man betrachte das Problem der Vorhersage der morgigen maximum Tagestemperatur anhand der heutigen Morgentemperatur. Machine learning (ML)-Methoden lernen eine Hypothese-Abbildung $\widehat{h}$, die die heutige Morgentemperatur, das Merkmal, auf eine genaue Vorhersage der morgigen maximum Temperatur, das Label, abbildet. Diese gelernte Hypothese wird aus einer (typischerweise großen) Menge von Kandidaten-Hypothese-Abbildungen, d.h. einem Hypothesenraum $\mathcal{H}$, ausgewählt.

Der einem ML-Verfahren zugrunde liegende Hypothesenraum ist sein Modell: die Menge der im Voraus gewählten Kandidaten-Hypothese-Abbildungen, aus denen das Verfahren eine auswählt (Shalev-Shwartz and Ben-David, 2014, Sect. 2.3; Vapnik, 2000, Sect. 1.1). Zum Beispiel verwenden lineare Regression und logistische Regression das lineares Modell als ihren Hypothesenraum. Eine Entscheidungsbaum-Methode verwendet einen Hypothesenraum, der aus Hypothese-Abbildungen besteht, die durch ein Flussdiagramm erzeugt werden. Großes Sprachmodell (LLM)-Systeme verwenden einen Hypothesenraum, der aus nicht-lineare Abbildungen besteht, die durch ein künstliches neuronales Netz (KNN) repräsentiert werden.

Das Modell-Training bezieht sich auf den Prozess des Findens einer Hypothese im Hypothesenraum, die genaue Vorhersagen liefert. Ein trainiertes (oder angepasstes) Modell ist eine gelernte Hypothese $\widehat{h} = \mathcal{A}(\mathcal{D}^{(\text{train})})$, die Ausgabe der Abbildung $\mathcal{A}$, durch die die ML-Methode einen Trainingsdatensatz

$\mathcal{D}^{(\text{train})}$ in eine Hypothese umwandelt. In der angewandten ML-Literatur und in Softwarebibliotheken bezieht sich das Wort *Modell* typischerweise auf ein trainiertes Modell, d.h. auf eine spezifische Hypothese-Abbildung $\widehat{h}$ (Goodfellow et al., 2016, Sect. 5.1; Murphy, 2012, Sect. 3.5.2; Prince, 2023). Der EU AI Act verwendet das Wort *Modell* im selben Sinne (d.h. eine gelernte Hypothese) im zusammengesetzten Begriff general-purpose AI model (GPAI model) (European Parliament and Council of the European Union, 2024, Art. 3(63)).

Eine andere Art von mathematischem Modell ist ein Wahrscheinlichkeitsmodell, das eine Familie von Wahrscheinlichkeitsverteilungen spezifiziert. Jede solche Wahrscheinlichkeitsverteilung definiert einen Daten-Generator: Datenpunkte sind Realisierungen von unabhängig und identisch verteilt (u.i.v.) Zufallsvariablen (ZV) mit der gegebenen Wahrscheinlichkeitsverteilung.

Abb. 1 veranschaulicht diese drei Verwendungen des Wortes *Modell*: den Hypothesenraum, das trainierte Modell und das Wahrscheinlichkeitsmodell.

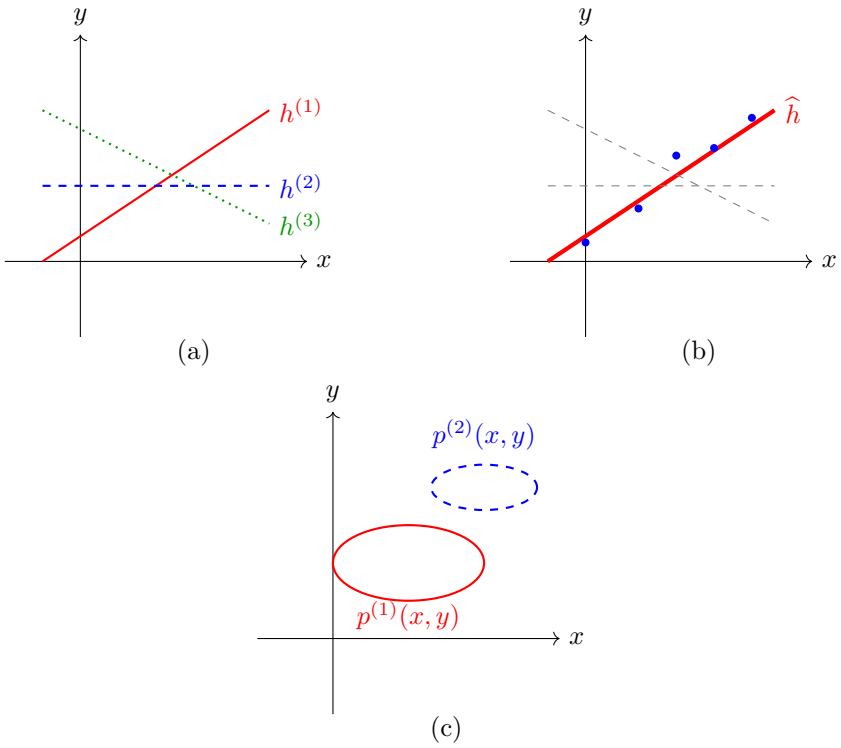


Abbildung 1: Drei Verwendungen des Wortes *Modell* in der ML-Praxis. (a) Hypothesenraum, bestehend aus drei lineare Abbildungen (durchgezogen, gestrichelt, gepunktet). (b) Trainiertes Modell: die gelernte Hypothese $\widehat{h} = \mathcal{A}(\mathcal{D}^{(\text{train})})$, ausgewählt aus dem Hypothesenraum in (a) durch eine ML-Abbildung $\mathcal{A}$, die auf einem Trainingsdatensatz $\mathcal{D}^{(\text{train})}$ von Datenpunkte (blaue Kreise) operiert. (c) Wahrscheinlichkeitsmodell, bestehend aus zwei Wahrscheinlichkeitsverteilungen (durchgezogene und gestrichelte Umrisse) zur Generierung von Datenpunkte mit Merkmal x und Label y.

Synonyme: Hypothesenraum, Hypothesenklasse, Modellklasse.

Siehe auch: Hypothesenraum, Hypothese, Training, Wahrscheinlichkeitsmodell, parametrisches Modell, Modellparameter, GPAI model.

Querverweise

Literatur

1. European Parliament and Council of the European Union (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU)

No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance). Official Journal of the European Union, L series, Jul. 12. URL: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng.

2. Goodfellow et al. (2016). Deep Learning. MIT Press, Cambridge, MA, USA.

3. Murphy (2012). Machine Learning: A Probabilistic Perspective. MIT Press, Cambridge, MA, USA.

4. Prince (2023). Understanding Deep Learning. MIT Press, Cambridge, MA, USA.

5. Shalev-Shwartz and Ben-David (2014). Understanding Machine Learning: From Theory to Algorithms. Cambridge Univ. Press, New York, NY, USA.

6. Vapnik (2000). The Nature of Statistical Learning Theory. 2nd ed., Springer, New York, NY, USA.