Weiterführende Informationen:

Definition ML

- ML als Teilgebiet der KI
- entwickelt Algorithmen und statistische Modelle, die aus Erfahrungen lernen und folglich Muster erkennen, Vorhersagen und Entscheidungen treffen können
- ursprüngliche Idee: Computer lernen durch die Verarbeitung großer Datenmengen durch das Erkennen von Mustern, wodurch sich neue Aufgaben leichter bewältigen lassen
- drei Typen von ML:
    1. *Überwachtes Lernen:* (supervised learning) mittels Datenpaaren zum Training erstellt der Computer eine Zuordnungsfunktion. Diese bildet die passenden Ein- und Ausgabevariablen aufeinander ab. Der Algorithmus wird trainiert, indem er eine vorhergesagte Ausgabe mit der tatsächlichen Ausgabe vergleicht.
    2. *Unüberwachtes Lernen:* (unsupervised learning) der Algorithmus erhält nur noch Eingabe- und keine Ausgabedaten. Er erkennt automatisch Muster und Strukturen in den Daten - ohne jegliche Anleitung. Häufig verwendet für Clusteranalyse, Dimensionalitätsreduktion und Anomalieerkennung.
    3. *Bestärkendes Lernen:* (reinforcement learning) hierbei steht die Interaktion des Algorithmus mit der Umgebung im Fokus. Der Algorithmus erhält Feedback in Form von Belohnung oder Bestrafung, wodurch er lernt eine geeignete Handlungsstrategie zu entwickelt, um die Belohnung zu maximieren.
- Abgrenzung zu kNN's: spezielle Klasse von Algorithmen des ML; basierend auf biologischem Konzept der Neuronen im Gehirn bestehen sie aus mehreren miteinander verbundenen Schichten von künstlichen Neuronen. KNN wird häufig in überwachtem und unüberwachtem Lernen eingesetzt.

Definition kNN

- computergestützte Modelle, die von der Struktur und Funktionsweise der biologischen Neuronen des Gehirns inspiriert sind
- im Bereich ML werden sie dafür eingesetzt, um komplexe Aufgaben wie Mustererkennung, Klassifizierung, Regressionsanalyse und Entscheidungsfindung zu lösen
- Basis ist ein künstliches Neuron, welches Eingaben gewichtet und eine Aktivierungsfunktion ausführt, um die gewichtete Summe zu einem Ausgabewert zu verarbeiten
- Organisation der einzelnen Neuronen in Schichten, jede Schicht hat dabei eine Sammlung von Neuronen; typischer Weise haben kNN's drei Arten von Schichten:
    1. *Eingabeschicht:* (Input Layer) empfängt Rohdaten, die in eingespeist werden
    2. *Verborgene Schicht:* (Hidden Layer) befindet sich zwischen Eingabe- und Ausgabeschicht; verarbeitet Eingabedaten und extrahiert abstrakte Merkmale und Muster
    3. *Ausgabeschicht:* (Output Layer) gibt Ergebnisse/Vorhersagen aus, die interpretiert werden können
- Training des kNN: Optimierung der Gewichtung der Neuronen, um Fehler zwischen vorhergesagten Ausgaben und tatsächlichen Ausgaben zu minimieren - Methode hierzu: Gradientenabstiegsverfahren (Fehler zurückgeben und Gewichtung der Daten anpassen)
- tiefe kNN's = Deep Learning; besteht aus hunderten bis tausenden Hidden Layers (im klassischen kNN sind es nur ein bis zwei), wodurch komplexere Zusammenhänge vom Algorithmus erkannt werden und Probleme gelöst werden können, die weitaus schwieriger

sind; Anwendung: Computer Vision (autonomes Fahren, Gesichtserkennung, Bilderkennung), NLP (Spracherkennung, Übersetzung, Chatbots), Sprachsynthese (Text-to-Speech), Spieltheorien.

Definition white box bzw. black box

WHITE BOX:

- Algorithmus, dessen interne Funktionsweise, Struktur und Entscheidungslogik vollständig transparent ist; die interne Arbeitsweise des Systems ist für den Benutzer/Betrachter nachvollziehbar
- Vorteile: Transparenz und Interpretierbarkeit, da Benutzer alle Entscheidungen und Vorhersagen verstehen können - vor allem bei kritischen Anwendungen (med. Diagnose/Finanzanalyse etc.) von Bedeutung; besseres Vertrauen in die Zuverlässigkeit eines Modells, da Funktionsweise nachvollziehbar
- Bsp.: Entscheidungsbaum - jede Entscheidung basiert auf klarer Sequenz von Regeln, die nachvollziehbar sind

BLACK BOX:

- Algorithmus, bei dem interne Funktionsweise, Struktur und Entscheidungslogik nicht transparent ist; die interne Arbeitsweise des Systems ist für den Benutzer/Betrachter nicht nachvollziehbar
- sehr leistungsstark und effektiv
- im Gegensatz zur white Box in kritischen Anwendungen von Nachteil, da nicht überprüfbar bzw. nachvollziehbar
- Bsp: kNN's und tiefe kNN's sind für den Anwender von außen nicht nachvollziehbar

Wann sollte ML bzw. white box und wann kNN bzw. black box gewählt werden?

- hängt von den spezifischen Anforderungen und Einschränkungen eines bestimmten Problems oder einer Anwendung ab

White Box (ML):

- Transparenz und Interpretierbarkeit entscheidend: inkritischen Anwendungen, wie med. Diagnose oder Rechtsprechung, ist es wichtig, dass Entscheidungen von einem Modell nachvollziehbar sind. Entscheidungsbäume/lineare Regression leichter zu interpretieren + bieten Einblicke in die Entscheidungsgrundlagen
- Daten knapp: White Box-Modelle können effektiver sein, wenn begrenzte Daten zur Verfügung stehen; benötigen i.d.R. weniger Trainingsdaten als Black Box-Modelle, um zu guten Ergebnissen zu gelangen
- schnelle Entscheidungszeiten erforderlich: White Box-Modelle können i.d.R. schneller Vorhersagen treffen als Black Box-Modelle wie tiefe kNN's

Black Box (kNN):

- Modellkomplexität spielt keine Rolle: hohe Genauigkeit erzielen > Entscheidungen des Modells verstehen; tiefe kNN's leistungsfähiger als White Box-Modelle, da sie komplexe Muster in den Daten erkennen können

- große Datenmengen: Black Box-Modelle können mit ausreichend Datenmengen bessere Ergebnisse erzielen
- Fachwissen fehlt: bei komplexen Problemen kann das Fachwissen über die Beziehung zwischen Eingabe und Ausgabe begrenzt sein; Black Box-Modelle hilfreich, da sie eigenständig Muster und Zusammenhänge in den Daten erkennen