

Otázky k magisterské státní závěrečné zkoušce z předmětu Umělá inteligence a strojové učení

Obor: Datová věda

Akademický rok 2023/2024

- 1. Principy strojového učení** (rozdíl mezi klasickým programováním a strojovým učením, kategorie strojového učení (s učitelem, bez učitele, učení posilováním), průběh strojového učení, zdroj dat a nejnámější datasety, trénovací a testovací data, jazyky a knihovny pro strojové učení, aplikační oblasti).
- 2. Základní metody strojového učení** (přehled metod a předpoklady jejich použití (jednorozměrná a vícerozměrná lineární regrese, algoritmus gradient descent, rozhodovací stromy, logistická regrese, shluková analýza, analýza hlavních komponent, metoda podpůrných vektorů).
- 3. Neuronové sítě** (principy a vlastnosti umělých neuronových sítí (inspirace biologií, umělý neuron, váha, práh, aktivační funkce, bias, synapse), perceptron, vícevrstvá dopředná síť, hluboké neuronové sítě, aplikační oblasti).
- 4. Detekce anomálií v datech** (rozdíly mezi detekcí anomálií a klasifikací, druhy anomálií, základní algoritmy (statistické metody, metoda nejbližších sousedů, algoritmus Local Outlier Factor, algoritmus Isolation Forest), aplikační oblasti).
- 5. Doporučující systémy** (účel a principy fungování (rating, metody měření vzdálenosti a podobnosti, kolaborativní filtrování), aplikační oblasti).
- 6. Komplexní systémy a jejich vlastnosti** (systémy jednoduché, chaotické a komplexní (rozdíly mezi nimi, příklady), definice komplexních systémů, emergence, nelinearita, zpětná vazba, růsta a adaptace, intuitivní lidské uvažování a chyby s ním spojené, principy systémového myšlení, redukcionismus a holismus, cíle vědy o komplexních systémech).
- 7. Principy modelování komplexních systémů** (význam výpočetních modelů při zkoumání komplexity, buněčné automaty (buněčný systém, přechodová funkce, typy buněčných automatů, jedno- a dvourozměrné automaty, Conwayův model Life, další příklady a aplikace), agentové modely (agent, prostředí, principy modelování, standardy a software, experimenty s modely, model Segregation, model Sugarscape, dynamika davu a evakuační modely)).
- 8. Komplexní síť a jejich vlastnosti** (uplatnění statistiky a počtu pravděpodobnosti při zkoumání sítí, stupeň uzlu a rozdělení stupňů uzlů v komplexní síti, obří komponenta, Milgramův experiment se zásilkami, Erdősovo číslo, obecné modely sítí (náhodný graf, graf malého světa - Watts-Strogatzův model, bezškálová síť - Barabási-Albertův model), odolnost sítě, aplikace síťových modelů (jednoduchá a komplexní nákaza v síti, epidemiologické modely, modely šíření inovací)).
- 9. Automatizované zpracování textu** (základní oblasti pro automatizované zpracovávání textu (text mining, textová analytika, NLP, NLU, NLG) a odlišnosti mezi nimi, data (typy dat, druhy datových souborů a jejich použití, způsoby získávání dat), text mining (historie, úlohy, vývojové prostředky, aplikace, trendy a výzvy)).
- 10. Zpracování přirozeného jazyka** (jazyk a jeho stavební bloky, úrovně analýzy textu a příklady technik, korpus a jeho využitelnost, fáze vývoje NLP systému, vývojové prostředky, aplikace, trendy a výzvy).
- 11. Reprezentace textu** (úloha vektoru a vektorové transformace při zpracování textů, vektorový prostorový model, základní přístupy pro vektorizaci textu, technika vnoření slov a její základní algoritmy, vývojové prostředky, aplikace).

- 12. Extrakce informací z textů** (typické úlohy extrakce informací z textů, pojmenovaná entita a její typ, fáze vývoje extraktoru, chunking, chunking, obtížnost extrakce, vizualizace výstupu extraktoru, vývojové prostředky, aplikace).
- 13. Modelování témat** (význam modelování, pojem téma, úrovně generování témat, fáze realizace modelování, algoritmy pro generování témat, vyhodnocení modelu, vizualizace, aplikace).