

Štátnicové okruhy mlNF - dátová veda 2025

(v hranatej zátvorke sú uvedené predmety, ktorých sa daná otázka týka:

SU - strojové učenie, NS - neurónové siete, MTSU - moderné techniky strojového učenia, VPDŠ - vybrané partie z dátových štruktúr, VŠA - viacrozmerná štatistická analýza, AZKD - analýza zhlukov a klasifikácia dát, RDD - redukcia dimenzie dát, ČR - časové rady)

1. Regresné modely: lineárna a generalizovaná lineárna regresia, metóda najmenších štvorcov, riešenie pomocou metódy spätného šírenia chyby, testovanie hypotéz o parametroch v regresnom modeli. [SU, VŠA]
2. Teória strojového učenia: štatistický model strojového učenia, výchylka vs. rozptyl, preučenie a podučenie, regularizácia regresných modelov (kombinatorická selekcia premenných, hrebeňová regresia, LASSO). [SU, RDD]
3. Teória strojového učenia: PAC učenie, odhady generalizačných vlastností pomocou VC dimenzie [SU]
4. Neurónové siete: viacvrstvový perceptrón, metóda spätného šírenia chyby, automatická diferenciácia (autograd), chain rule s viacerými cestami. [SU, MTSU, NS]
5. Neurónové siete pre sekvenčné dáta: model Simple Recurrent Network, bránové modely. [NS]
6. Neurónové siete - transformer architektúra, použitie ako klasifikátor pre obraz a pre generovanie textu [SU, MTSU]
7. Neurónové siete - autoasociatívna pamäť: Lineárny model, deterministický Hopfieldov model. [NS]
8. Stochastické modely neurónových sietí: Boltzmannov stroj, Deep Belief Network - základné koncepty. [NS]
9. Reinforcement learning. Policy gradient. [MTSU]
10. Techniky na šetrenie pamäte v neurónových sieťach: Gradient checkpointing, low rank adaptation a kvantizácia váh. [MTSU]
11. Lineárna a kvadratická diskriminačná analýza [AZKD]
12. Bayesov klasifikátor, metóda k najbližších susedov [AZKD]
13. Rozhodovacie stromy, náhodné lesy, bagging a boosting [SU, AZKD]
14. Metóda oporných bodov (angl. support vector machines), kernelová metóda [SU, AZKD]
15. Analýza zhlukov založená na metódach k-means a k-medoids, DBSCAN [SU, AZKD]
16. Zhlukovanie založené na zmesi gaussovských rozdelení [AZKD]
17. Hierarchické metódy analýzy zhlukov [AZKD]
18. Analýza hlavných komponentov (PCA), kernelová PCA [SU, RDD]
19. Mnohorozmerné škálovanie: klasické MŠ, nemetrické MŠ. [RDD]
20. Isomap a t-SNE. [SU, RDD]
21. Dátové štruktúry pre prácu s textom: sufixové stromy a polia, Burrowsova-Wheelerova transformácia a FM index, úsporné dátové štruktúry (rank a select, wavelet tree) [VPDŠ]
22. Hašovanie: univerzálna hašovacia funkcia, perfektné hašovanie, Bloomov filter [VPDŠ]
23. Externé dátové štruktúry: I/O model (triedenie, B-stromy), cache-oblivious model a statický binárny strom s van Emde Boas rozložením [VPDŠ]

24. Dátové štruktúry pre intervaly: range minimum query, lowest common ancestor, segmentové stromy, rozsahové stromy [VPDŠ]
25. Približné hľadanie najbližších susedov (approximate nearest neighbours). Hierarchical Navigable Small World a Locality sensitive hashing. [MTSU, VPDŠ]
26. Náhodný vektor, jeho rozdelenie, distribučná funkcia a kovariančná matica. Náhodný výber a výberové charakteristiky [VŠA]
27. Viacrozmerné normálne rozdelenie. Testy hypotéz o parametroch normálneho rozdelenia [VŠA]
28. Kopula. Ohraničenia kopúl. Sklarova veta. Korelačné koeficienty [VŠA]
29. Analýza rozptylu. Model jednofaktorovej a dvojfaktorovej analýzy rozptylu. Interpretácia a odhadovanie parametrov. Testovanie hypotéz o parametroch. [VŠA]
30. Autoregresné procesy, stacionarita, Woldova reprezentácia [ČR]
31. Autoregresné procesy, autokorelačná a parciálna autokorelačná funkcia [ČR]
32. Moving average procesy, invertovateľnosť [ČR]
33. Zmiešané ARMA procesy [ČR]
34. Diferencovanie časového radu, jednotkový koreň, ADF test [ČR]

Pôvodné podklady:

2-INF-150 Strojové učenie - Vladimír Boža, Marek Šuppa, Tomáš Vinař

- Regresia: lineárna a generalizovaná lineárna regresia, metóda najmenších štvorcov, regularizácia
- Neurónové siete: viacvrstvový perceptrón, metóda spätného šírenia chyby
- Klasifikačné modely: support vector machines, rozhodovacie stromy, náhodné lesy, bagging, boosting
- Strojové učenie bez učiteľa: zhukovanie, analýza hlavných komponentov
- Teória strojového učenia: štatistický model strojového učenia, výchylka vs. rozptyl, preučenie a podučenie, PAC učenie, odhady pomocou VC dimenzie

2-INF-188 Moderné techniky strojového učenia - Vladimír Boža

- Automatická diferenciácia (autograd). Chain rule s viacerými cestami.
- Transformer. Architektúra. Použitie ako klasifikátor pre obraz a pre generovanie textu.
- Reinforcement learning. Policy gradient.
- Približné hľadanie najbližších susedov (approximate nearest neighbours). Hierarchical Navigable Small World a Locality sensitive hashing.
- Techniky na šetrenie pamäte: Gradient checkpointing, low rank adaptation a kvantizácia váh.

2-AIN-132 Neurónové siete - Igor Farkaš

- Viacvrstvový perceptrón, metóda spätného šírenia chyby.
- Vizualizácia vysokorozmerných dát: Samoorganizujúca sa mapa (vektorová kvantizácia, topografické zobrazenie)
- Autoasociatívna pamäť: Lineárny model, deterministický Hopfieldov model.
- Predikcia sekvenčných dát: model Simple Recurrent Network, bránové modely.

- Stochastické modely NS: Boltzmannov stroj, Deep Belief Network - základné koncepty.

2-INF-237 Vybrané partie z dátových štruktúr - Jakub Kováč, Bronislava Brejová

- Dátové štruktúry pre prácu s textom: sufixové stromy a polia, Burrowsova-Wheelerova transformácia a FM index, úsporné dátové štruktúry (rank a select, wavelet tree)
- Hašovanie: univerzálna hašovacia funkcia, perfektné hašovanie, Bloomov filter
- Externé dátové štruktúry: I/O model (triedenie, B-stromy), cache-oblivious model a statický binárny strom s van Emde Boas rozložením
- Dátové štruktúry pre intervaly: range minimum query, lowest common ancestor, segmentové stromy, rozsahové stromy

2-PMS-115 Viacrozmerné štatistické analýzy - Lenka Filová

- Náhodný vektor, jeho rozdelenie, distribučná funkcia a kovariančná matica. Náhodný výber a výberové charakteristiky
- Viacrozmerné normálne rozdelenie. Testy hypotéz o parametroch normálneho rozdelenia
- Kopula. Ohraničenia kopúl. Sklarova veta. Korelačné koeficienty
- Viacrozmerný regresný model. Metóda najmenších štvorcov. Testovanie hypotéz o parametroch v regresnom modeli.
- Analýza rozptylu. Model jednofaktorovej a dvojfaktorovej analýzy rozptylu. Interpretácia a odhadovanie parametrov. Testovanie hypotéz o parametroch.

2-PMS-116 Analýza zhlučiek a klasifikácia dát - Radoslav Harman, Samuel Rosa

- Analýza zhlučiek založená na metódach k-means a k-medoids, DBSCAN
- Zhlučovanie založené na zmesi gaussovských rozdelení
- Hierarchické metódy analýzy zhlučiek
- Bayesov klasifikátor, metóda k najbližších susedov
- Lineárna a kvadratická diskriminačná analýza
- Klasifikačné stromy a lesy, bagging a boosting
- Metóda oporných bodov (angl. support vector machines)

2-PMS-142 Redukcia dimenzie dát - Samuel Rosa

- Analýza hlavných komponentov.
- Mnohorozmerné škálovanie: klasické MŠ, nemetrické MŠ.
- Isomap a t-SNE.
- Bias-variance tradeoff. Kombinatorická selekcia premenných (krokové metódy, informačné kritériá, validácia, krosvalidácia).
- Penalizované regresie: hrebeňová regresia a Lasso.

2-INF-191 Časové rady - Beáta Stehlíková

- Autoregresné procesy, stacionarita, Woldova reprezentácia
- Autoregresné procesy, autokorelačná a parciálna autokorelačná funkcia
- Moving average procesy, invertovateľnosť

- Zmiešané ARMA procesy
- Diferencovanie časového radu, jednotkový koreň, ADF test