

Umelá inteligencia 2
Hlboké učenie
Vysokoškolská učebnica

Recenzenti: doc. Ing. Peter Schreiber, CSc.
doc. Dr. Ing. Ján Vaščák



Pre Žilinskú univerzitu v Žiline vydal:
CEIT, a.s., Univerzitná 8661/6A, 010 08 Žilina.
© Michal Gregor, Dušan Nemeč, Marián Hrušoš, Juraj Spalek, 2017

ISBN 978-80-89865-03-1

Umelá inteligencia (UI) je veda a tiež súbor výpočtových nástrojov a postupov, inšpirovaných spôsobom, akým ľudia používajú svoj nervový systém a telo na vnímanie, rozhodovanie sa, učenie sa i konanie. Aj keď vieme, že tieto zvyčajne fungujú celkom inak...

Pripomeňme si, že základy UI boli položené pred viac ako 60 rokmi, no miera pokroku bola nepravidelná a často nepredvídateľná. Niekoľko významných období bolo v 50. až 70. rokoch 20. storočia, kedy Newell a Simon propagovali heuristické hľadanie ako účinný nástroj na hľadanie riešení vo veľkých kombinatorických priestoroch. Túto myšlienku použili na vytvorenie dôkazov matematických teórií najprv prostredníctvom programu Logic Theorist a následne prostredníctvom programu General Problem Solver.

V oblasti počítačového videnia vytvorila staršia práca Selfridgeho a kol. [1] o rozpoznávaní znakov základ pre komplexnejšie aplikácie, napr. pre rozpoznávanie tvárí. Koncom 60. rokov sa začali práce aj v oblasti spracovania prirodzeného jazyka. Kolesový robot Shakey, vyvinutý spoločnosťou SRI International v rokoch 1966 – 1972, naštartoval oblasť mobilnej robotiky. Program Arthura Samuela pre doskovú strategickú hru Dáma, ktorý sa sám vylepšoval samotnou hrou, bol jedným z prvých fungujúcich príkladov strojového učenia [2]. Rosenblattov perceptrón (výpočtový model založený na biologických neurónoch) sa stal základom pre oblasť umelých neurónových sietí. Feigenbaum a ďalší vytvárali expertné systémy – znalostné databázy pre špecializované oblasti, pre chémiu a lekársku diagnostiku, atď ...

UI 21. storočia zasahuje do nášho každodenného života množstvom nástrojov, algoritmov a technológií. Napríklad počítačové videnie a plánovanie s podporou UI ovládajú videohry, ktoré dnes predstavujú výnosnejší zábavný priemysel ako Hollywood. Podobne hlboké učenie ako forma strojového učenia, založená na vrstvenej reprezentácii premenných, nazývaných neurónové siete, urobilo rozpoznanie reči využiteľným v našich mobilných telefónoch, ba povedzme aj v našich kuchyniach. Jeho algoritmy sú použité na rad aplikácií, založených na rozpoznávaní vzorov.

Napríklad spoločnosť IBM investovala v r. 2014 miliardu dolárov do vzniku novej divízie pre

vývoj jedného z najdokonalejších superpočítačov na svete, nazvaného Watson. Vďaka rozlíšeniu lingvistických špecifik a ich analýze dokáže nielen imitovať myslenie človeka, ale dokáže spracovať oveľa väčšie množstvo informácií a hlavne to vie urobiť omnoho rýchlejšie. Počítač Watson plánuje IBM využiť vo svojej cloudovej infraštruktúre Softlayer. Tento superpočítač dokonca pred časom (a to ešte pred zavedením hlbokého učenia) porazil humánnych protihráčov v jednej z najúspešnejších amerických televíznych vedomostných súťaží Jeopardy! (po slovensky Riskuj!), kde účastníci prezentujú svoje vedomosti, avšak ich odpovede musia byť vo forme otázok. Watson s prehľadom porazil oboch spoluhráčov a výhru milión dolárov jeho materská spoločnosť venovala na charitu. Podobne, ako spracovanie prirodzeného jazyka i reprezentácia vedomostí a uvažovanie umožnili strojom poraziť šampióna Jeopardy!, čo prináša **UI** nové možnosti aj pri vyhľadávaní na webe. Mimochodom, IBM očakáva, že už v r. 2018 by mal Watson zarobiť miliardu USD ročne a v r. 2023 až 10 miliárd USD ročne.

Význam **UI** však treba vidieť aj v širších súvislostiach. Analýzy ukazujú, že viac ako 50 % povolání, ktoré bude vykonávať nastupujúca mladá generácia Z, doposiaľ ešte neexistuje. A naopak, na každé vytvorené miesto päť iných zanikne. To znamená obrovské preskupovanie profesií na pracovnom trhu. Pripomeňme si, koľko nových zamestnaní sa objavilo s nástupom informačných a komunikačných technológií, ktoré väčšine ľudí dodnes priamo, či sprostredkovane zasahuje do pracovnej náplne. Zavádzanie inteligentných systémov, pokročilej automatizácie a robotizácie prinesie ďalšiu výraznú zmenu do ponuky pracovných miest. Na trhu práce sa budú čoraz častejšie objavovať profesie ako špecialista pre kybernetickú bezpečnosť, vývojár a návrhár vysoko sofistikovaných zariadení a systémov pre digitálne podniky, či smart mestá a domácnosti. Firmy sa pripravujú na éru digitalizácie a robotizácie výroby i procesov, v tímoch už vytvárajú nové pracovné pozície. To sú výzvy, s ktorými treba rátať a včas ponúknuť sortiment vedomostí a zručností, bez ktorých sa dnešní študenti nezaobídu v zajtrajšej praxi.

Je o to zvláštnejšie, že dodnes neexistuje presná a všeobecne akceptovaná definícia **UI**. Výskumní pracovníci i vývojári **UI** sa často riadia intuíciou, či trendovým zmyslom pre smer a imperatív typu „dajme to s ňou“. Preto ostáva potreba definovať **UI** naďalej dôležitá. Jednu praktickú ponúkol Nils J. Nilsson [3]:

Umelá inteligencia je to, čo robí stroje inteligentnými a inteligencia je kvalita, ktorá umožňuje, aby daná entita fungovala správne a vo svojom prostredí predvídavo.

K čitateľovi sa dostáva v poradí tretia zo série vysokoškolských učebníc o umelej inteligencii, ktoré vznikli na Katedre riadiacich a informačných systémov Elektrotechnickej fakulty Žilinskej univerzity v Žiline. Prvá bola z iniciatívy širšieho kolektívu katedry vydaná ešte v roku 2005 pod názvom Rozhodovanie a riadenie s podporou umelej inteligencie. Poskytla veľmi dobrý prehľad o viacerých prístupoch a metódach z oblasti **UI**. Druhú vytvoril Michal Gregor – študent, neskôr denný doktorand a v súčasnosti výskumný pracovník na tomto pracovisku – v roku 2014 s orientáciou na vybrané témy, fuzzy prístupy a umelé neurónové siete. Jej názov Umelá inteligencia 1. už evokoval

pokračovanie diela, ktorým je práve táto učebnica. K jej zostaveniu prizval spolupracovníkov zo svojej materskej katedry, zaoberajúcich sa teóriou a aplikáciami **UI**.

V 1. kapitole sa autori venujú objasneniu základných pojmov z oblasti strojového učenia. Aj keď do **UI** patria okrem tejto aj ďalšie oblasti, napr. problematika reprezentácie poznatkov (explicitných/implicitných, vágnych/ostrých), ako aj problematika inteligentného plánovania, optimalizácie a ďalšie, sú v tejto kapitole zhrnuté predovšetkým základné pojmy, spoločné pre všetky smery strojového učenia a tiež vysokoúrovňový pohľad na niektoré vybrané smery umelej inteligencie.

2. kapitola obsahuje súhrn základných poznatkov z teórie umelých neurónových sietí. Opísaný je koncept umelého neuróna a princíp jeho činnosti, ilustrovaný na probléme lineárnej separovateľnosti. Nasledujú viacvrstvové neurónové siete a rozličné typy architektúr umelých neurónových sietí.

V ďalších kapitolách sa čitateľ dozvie o úlohách kontrolovaného učenia v umelých neurónových sieťach (kapitola 3), o rekurentnom učení (kapitola 4), o hlbokom učení (kapitola 5), o nekontrolovanom učení (kapitola 6) a tiež o modeloch, založených na energii (kapitola 7).

V osobnosti hlavného autora a architekta tejto učebnice čitateľ isto rozpozna nielen mimoriadne talentovaného a vyspelého teoretika a súčasne skvelého aplikátora **UI**, ale aj dobrého pedagóga, ktorý využijúc svoje skúsenosti z vyučovacieho procesu ušil štruktúru učebnice na mieru odborne vyspelého čitateľa. „Hĺbku ponoru“ však servíruje adaptívne s ohľadom na teoretickú, či programátorskú obťažnosť konkrétnej témy. Tej následne podriadil aj ponuku aplikačných a demonštračných príkladov vo forme zdrojových kódov v jazyku Python, ilustrujúcich princípy jednotlivých metód, ale aj možnosti ich uplatnenia v praxi. Z rovnakých dôvodov zaradil do záveru nosných kapitol zoznam odporúčanej literatúry, online kurzov a návodov, softvérových nástrojov a pod.

Autori tejto učebnice vyjadrujú nádej, že jej obsah i forma spracovania čitateľov zaujme, bude pre nich prínosom pri prehlbovaní svojich poznatkov o teórii strojového učenia, no zároveň inšpiráciou pri ich aplikovaní v praktickom živote.

September 2017, Žilina

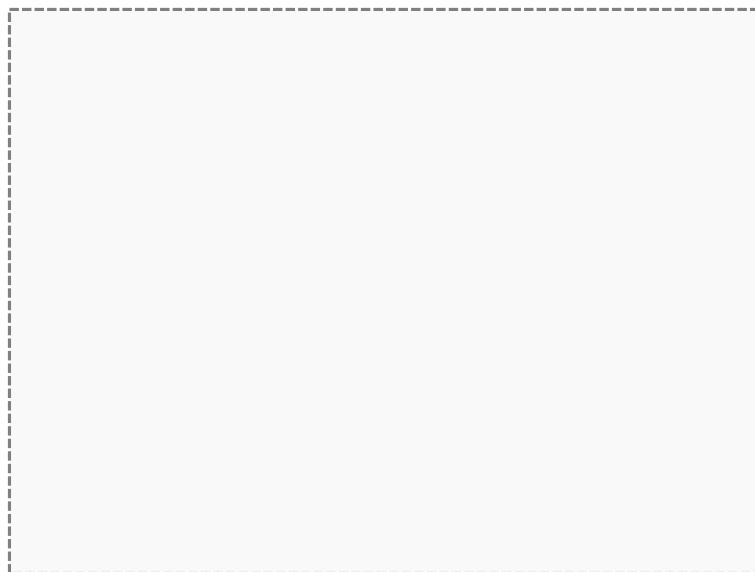
Juraj Spalek

Tento diel učebnice je v poradí už druhým dielom. V prvom dieli sme sa zaoberali fuzzy systémami a klasickou teóriou umelých neurónových sietí. Keďže niektoré základné pojmy z tejto teórie sa používajú aj v modernom hlbokom učení, na začiatku ich v krátkosti zopakujeme. Nasledujúci obsah však už bude prepracovaný novým spôsobom – vrátane nového, grafického a vrstvom orientovaného prístupu ku výkladu princípu spätného šírenia chyby. Pôvodný, klasickejší výklad je stále možné čerpať z prvého dielu. Rovnako sa v tomto dieli nebudeme venovať niektorým typom nekontrolovaného učenia (napr. konkurenčné učenie a samoorganizujúce sa mapy), ktoré je tiež možné nájsť v prvej časti.

Keďže jedným z cieľov učebnice je poskytnúť znalosti o hlbokom učení takým spôsobom, aby ich bolo možné ľahko aplikovať, učebnica kombinuje teoretický výklad s príkladmi pre jazyk Python. Takisto na viacerých miestach prikladáme zdrojové kódy použité na vygenerovanie vizualizácií, aby ich bolo pre čitateľa ľahšie reprodukovať. Takýmto spôsobom chceme jednak lepšie ilustrovať princípy jednotlivých metód a jednak naznačiť, ako ich možno v praxi použiť.

Vzhľadom na to, že drvivá väčšina materiálov týkajúcich sa hlbokého učenia, je prirodzene v anglickom jazyku, je naším zámerom čitateľovi predovšetkým poskytnúť základy dostatočné na to, aby dokázal ďalšiu literatúru študovať samostatne. Aby sme prechod od učebnice ku iným materiálom čo najviac uľahčili, prikladáme v úvode slovensko-anglický slovník použitých pojmov. Keďže v učebnici prezentujeme novú problematiku, a to ešte z oblasti, kde dochádza v súčasnosti ku rapidnému vývoju, mnoho pôvodne anglických výrazov zatiaľ nemá (a možno nikdy nebude mať) žiadne ustálené slovenské ekvivalenty. Snažili sme sa preto – tam, kde to bolo možné – preferovať preklady využívajúce pôvodné výrazové prostriedky, ak sú tieto v slovenskom jazyku k dispozícii ako cudzie slová. V mnohých prípadoch to možné nebolo – práve v tých prípadoch najviac posluží priložený slovník.


V závere hlavných kapitol sme – podobne ako v predchádzajúcom dieli – vyhradili priestor na odporúčanie ďalšej literatúry, online kurzov a návodov, softvérových nástrojov a pod. Veríme, že aj tieto odkazy uľahčia čitateľom prechod od obsahu knihy ku iným zdrojom.



Obr. P.1: Príklad priloženého videa. ▶ ■

POZNÁMKA | Poznámky

Učebnica na niektorých miestach obsahuje poznámky, ktoré ilustrujú a dovysvetľujú zmysel textu.

V elektronickej verzii knihy sú na viacerých miestach k textu priložené aj súbory – tieto sú označené . Extrahovať ich možno typicky buď ľavým alebo pravým kliknutím (v závislosti od konkrétnej aplikácie použitej na čítanie PDF).

Aplikácia Adobe Reader má (vo verziách aktuálnych v čase písania tohto textu) predvolene zakázané akokoľvek manipulovať so súbormi v jazyku Python – či už ich ukladať na disk, alebo ich priamo otvoriť v inej aplikácii. Nastavenie je možné zmeniť jedine prostredníctvom systémových registrov[4] – preto odporúčame na prezeranie dokumentu použiť radšej inú, menej excentricky konfigurovanú aplikáciu.

Na niektorých miestach môžu byť v texte priložené aj videá, animácie a pod. Na ich prezeranie je potrebné, aby mal váš prehrávač videa podporu pre kodek H.264. V niektorých aplikáciách na prezeranie PDF súborov je navyše potrebné prehrávanie videí najprv osobitne povoliť. Príklad videa je na Obr. P.1.

Poznámka ku značeniu

Čo sa týka značenia, budeme sa v texte snažiť, pokiaľ to bude možné, konzistentne dodržiavať značenie matíc veľkými písmenami, napr. W a značenie vektorov hrubými písmenami, napr. \mathbf{x} .

Je tiež dôležité poznamenať, že pri značení intervalov sa budeme pridržovať západnej konvencie, ktorá hranice uzavretých intervalov označuje hranatými zátvorkami, napr. $[a, b]$ a nie špicatými zátvorkami, napr. $\langle a, b \rangle$. Túto konvenciu volíme najmä preto, aby sa značenie intervalov nemýlilo s inými značením, ktoré používame – napríklad s označením n -tíc alebo so značením skalárneho súčinu v kontexte vektorových priestorov.

Licenčné podmienky

Všetky zdrojové kódy – obsiahnuté priamo v texte aj priložené ako súbory – sa spravujú (ak nie je vyslovene povedané inak) *New BSD licenciou*:

Copyright (c) 2017, Michal Gregor
All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- * Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
- * Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
- * Neither the name of the author nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Všetky obrázky a grafické prílohy vytvorené autormi (t.j. kde nie je označený iný zdroj pôvodu) sa spravujú podľa [Creative Commons Attribution](#) licencie. Možno ich teda ľubovoľne šíriť – jedinou podmienkou je uvedenie pôvodného zdroja.

Licenčné práva sa netýkajú samotného textu učebnice – ten nesmie byť šírený bez výslovného povolenia autorov.

Predslov	i
Ku obsahu knihy	iv
Obsah	vii
Zoznam skratiek	x
Slovník pojmov	xii
1 Umelá inteligencia a strojové učenie	1
1.1 Typy učenia: rozdelenie podľa druhu úlohy	3
1.2 Zovšeobecnenie	8
1.3 Globálne a lokálne zovšeobecnenie	16
1.4 Znalosti	21
2 Umelé neurónové siete	24
2.1 Model a základná teória	24
2.2 Teoréma o univerzálnej aproximácii	37
2.3 Oblúbené aktivačné funkcie	37
3 ANN: Kontrolované učenie	41
3.1 Iteratívne učenie	41
3.2 Metóda klesajúceho gradientu	42
3.3 Klesajúci gradient v lineárnej regresii	52
3.4 Klesajúci gradient v neurónových sieťach	56
3.5 Dávkové a inkrementálne učenie	67
3.6 Úlohy kontrolovaného učenia	71
3.7 Nadstavbové aktivačné a chybové funkcie	85

3.8	Nadstavby metódy klesajúceho gradientu	93
3.9	Problém preučenia a ako mu predísť	100
3.10	Kompresia neurónových sietí	104
3.11	Inicializácia váhovej matice	106
4	ANN: Učenie v rekurentných sieťach	109
4.1	Jordanova a Elmanova sieť	109
4.2	Úlohy kontrolovaného rekurentného učenia	111
4.3	Spätné šírenie chyby v čase (BPTT)	111
4.4	Rekurentné učenie v reálnom čase (RTRL)	114
4.5	Echo state siete	117
4.6	Dlhá krátkodobá pamäť	118
4.7	Metóda postupnosť na postupnosť	122
4.8	Metóda dropout v rekurentných sieťach	124
4.9	Predikcia časových radov	125
5	Hlboké učenie v ANN	136
5.1	Intuícia: prečo hlboké učenie?	136
5.2	Hlboké modely konštruujú reprezentáciu	138
5.3	Miznúci/explodujúci gradient	139
5.4	Nekontrolované vrstvové učenie	140
5.5	Metódy využívajúce kombináciu trikov	142
5.6	Optimalizačné metódy vyššieho rádu	143
5.7	Hlboké autoenkodéry	143
5.8	Konvolučné siete	144
5.9	Dávková normalizácia	170
5.10	Reziduálne siete	172
5.11	Diaľnicové siete	181
5.12	Protivnicke príklady	184
5.13	Neuromalby	189
6	ANN: Nekontrolované učenie	198
6.1	Autoenkodéry	198
6.2	Generatívne protivnicke siete	211
7	ANN: Modely založené na energii	221
7.1	Hopfieldova sieť	221
7.2	Boltzmannov stroj	236
7.3	Zjednodušený Boltzmannov stroj	247
	Kam ďalej?	251

8	Vyhodnocovanie úspešnosti učenia	260
8.1	Matica zámen binárneho klasifikátora	260
8.2	Ukazovatele pre binárnu klasifikáciu	261
8.3	ROC analýza	264
8.4	Matica zámen pre viacero tried	271
8.5	Ukazovatele pre klasifikáciu do viacerých tried	273
8.6	ROC analýza pre viacero tried	275
8.7	Učenie pri nerovnomernom zastúpení tried	275
8.8	Chybové miery pre regresiu	283
	Kam ďalej?	290
	Literatúra	292
	Poznámky	301
	Prílohy	302
A	Optimalizačné metódy pre neurónové siete	I
A.1	Klasifikácia optimalizačných metód	I
A.2	Optimalizačné metódy prvého rádu	VI
A.3	Optimalizačné metódy vyššieho rádu	XVII
B	Nguyen-Widrowova inicializačná metóda	XXXIII
B.1	Pre sieť s jedným vstupom	XXXIII
B.2	Zovšeobecnenie: iné vstupy/aktivačné funkcie	XXXVI
B.3	Pre sieť s viacerými vstupmi	XXXVIII
B.4	Zovšeobecnenie pre iný aktívny rozsah	XXXVIII
B.5	Zovšeobecnenie pre iný rozsah vstupov	XXXIX
B.6	Vlastnosti Nguyen-Widrowovej metódy	XLI

- AN** umelý neurón (angl. artificial neuron), s. 25, 27
- ANN** umelá neurónová sieť (angl. artificial neural network), s. 24, 33, 41, 79, 109, 136, 198, 221, 252, 256
- AUC** plocha pod (ROC) krivkou (angl. area under (an ROC) curve), s. 270, 275
- BFGS** Broyden–Fletcher–Goldfarb–Shanno algoritmus, s. XXIV, XXV
- BPTT** spätné šírenie chyby v čase (angl. backpropagation through time), s. 111–113, 117, 120, 125, 127, 132
- CG** združený gradient (angl. conjugate gradient), s. XIII–XVI, XXXII
- DIKW** dáta-informácie-znalosti-múdrosť (angl. data-information-knowledge-wisdom pyramid), s. 20, 21
- ELU** exponenciálna lineárna funkcia (angl. exponential linear unit), s. 91, 92
- FN** falošne negatívny prvok (angl. false negative), s. 261
- FP** falošne pozitívny prvok (angl. false positive), s. 261–263
- GPGPU** univerzálne výpočty na grafických kartách (angl. general-purpose computing on graphics processing units), s. 104
- HF** bezhessiánovská optimalizácia (angl. Hessian-free optimization), s. 143, XXXI, XXXII
- L-BFGS** Broyden–Fletcher–Goldfarb–Shanno algoritmus s obmedzenou pamäťou (angl. limited-memory BFGS), s. 143, XXIV, XXV
- LM** metóda Levenberg-Marquardt, s. XXII–XXIV
- LReLU** jednotka (neurón) s presakujúcou rektifikovanou lineárnou funkciou (angl. leaky rectified linear unit), s. 28, 90, 91
- MAE** stredná absolútna chyba (angl. mean absolute error), s. 284, 288, 291
- MAPE** stredná absolútna percentuálna chyba (angl. mean absolute percentage error), s. 287
- MCMC** (angl. Markov Chain Monte Carlo), s. 244, 248
- MdAE** medián absolútnej chyby (angl. median absolute error), s. 284, 291
- MdAPE** medián absolútnej percentuálnej chyby (angl. median absolute percentage error), s. 287
- MLP** viacvrstvový perceptrón, viacvrstvová neurónová sieť (angl. multi-layer perceptron), s. 36

MSE stredná kvadratická chyba (angl. mean squared error), s. 283, 284

msMAPE modifikovaná symetrická stredná absolútna percentuálna chyba (angl. modified symmetric mean absolute percentage error), s. 288

NAG Nesterovov akcelerovaný gradient (angl. Nesterov accelerated gradient), s. 99

PReLU jednotka (neurón) s parametrická rektifikovanou lineárnou funkciou (angl. parametric rectified linear unit), s. 28, 90, 91

R^2 koeficient determinácie (angl. coefficient of determination), s. 289

ReLU jednotka (neurón) s rektifikovanou lineárnou funkciou (angl. rectified linear unit), s. 28, 76, 90–92, 143, 172, 173, 187, XLVI

RMdSPE odmocnina mediánu kvadratickej percentuálnej chyby (angl. root median square percentage error), s. 287

RMSE odmocnina strednej kvadratickej chyby (angl. root mean squared error), s. 283, 284, 286

RMSLE odmocnina strednej kvadratickej logaritmickej chyby (angl. root mean squared logarithmic error), s. 289

RMSPE odmocnina strednej kvadratickej percentuálnej chyby (angl. root mean square percentage error), s. 287

ROC z angl. receiver operating characteristic; v preklade „operačná charakteristika prijímača“ – väčšinou sa používa len ako skratka ROC, s. 264–267, 269, 270, 274, 275

SGD stochastický klesajúci gradient (angl. stochastic gradient descent), s. 67, 69

sMAPE symetrická stredná absolútna percentuálna chyba (angl. symmetric mean absolute percentage error), s. 288

sMdAPE symetrický medián absolútnej percentuálnej chyby (angl. symmetric median absolute percentage error), s. 288

SMOTE technika syntetického nadvzorkovania minorít (angl. synthetic minority oversampling technique), s. 278–280, 290

TN skutočne negatívny prvok (angl. true negative), s. 261

TP skutočne pozitívny prvok (angl. true positive), s. 261–263

UI umelá inteligencia, s. i–iii

Autoenkodér: autoencoder.

BEGAN: Hranične rovnovážne generatívne protivnícke siete (angl. boundary equilibrium generative adversarial networks).

Bielenie: whitening.

BN: Dávková normalizácia (angl. batch normalization).

Boltzmannov stroj: Boltzmannov machine.

CD: Contrastive divergence (metóda učenia pre zjednodušený Boltzmannov stroj).

CD: Metóda združeného poklesu (angl. conjugate descent method).

Čiastočne kontrolované učenie: semi-supervised learning.

Citlivosť: sensitivity. Tiež: senzitivita; TP miera – TP rate; úplnosť – recall.

Kontraktívny autoenkodér: contractive autoencoder.

CycleGAN: Cyklicky konzistentné generatívne protivnícke siete (angl. cycle-consistent generative adversarial networks).

Dávková normalizácia: batch normalization (BN).

Odšumujúci autoenkodér: denoising autoencoder.

Detektor príznaku: feature detector.

Diaľnicové siete: highway networks.

Dlhá krátkodobá pamäť: long short-term memory.

Učenie pomocou komisií: ensemble learning.

ESN: Echo state sieť (angl. echo state network).

Exponenciálna lineárna funkcia: exponential linear unit.

F miera: F-measure.

Falošne negatívny prvok: false negative.

Falošne pozitívny prvok: false positive.

FP miera: FP rate.

GAN: Generatívne protivnícke siete (angl. generative adversarial networks).

Generatívne protivnícke siete: generative adversarial networks (GAN).

Wassersteinovské GAN

Wasserstein GAN; WGAN.

hranične rovnovážne GAN

boundary equilibrium GAN; BEGAN.

cyklicky konzistentné GAN

cycle-consistent adversarial networks; CycleGAN.

Gibbsovo vzorkovanie: Gibbs sampling.

Gradient: gradient.

Miznúci gradient
vanishing gradient.

Explodujúci gradient
exploding gradient.

Gram matica: Gram matrix.

GRU: Hradlovaný rekurentný neurón (angl. longgated recurrent unit).

Hebbovo pravidlo: Hebb's rule.

Hebbovské učenie: Hebbian learning.

Bez Hessiánovská optimalizácia: Hessian-free optimization (HF).

Hĺbkové učenie: deep learning. My v učebnici preferujeme výraz hlboké učenie¹.

Hlboké siete: deep neural networks.

Hlboké učenie: deep learning (v literatúre sa možno stretnúť aj s pojmom hĺbkové učenie¹).

Hradlovaný rekurentný neurón: gated recurrent unit.

Induktívne preferencie: inductive bias.

Kolaps diverzity: mode collapse.

Kompetitívne učenie: competitive learning.

Konkurenčné učenie: competitive learning.

Kontrastívna divergencia: contrastive divergence.

Kontrolované učenie: supervised learning. Tiež *učenie s predlohou*, al. *učenie s učiteľom*.

Konvolúcia: convolutional.

Dilatovaná konvolúcia
dilated convolution, al. *trous convolution*.

Kroková konvolúcia
strided convolution.

Konvolučná sieť: convolutional network.

Kritérium celkových zmien: total variation loss.

Krížová validácia: cross-validation.

Latentná sémantická analýza: latent semantic analysis.

LSA: Latentná sémantická analýza (angl. latent semantic analysis).

LSTM: Dlhá krátkodobá pamäť (angl. long short-term memory).

Stredná absolútna chyba: mean absolute error.

Makro priemerovanie: macro-averaging.

Mapa príznakov: feature map.

Mapa zásahov: hitmap.

Stredná absolútna percentuálna chyba: mean absolute percentage error.

Matica zámen: confusion matrix.

Medián absolútnej chyby: median absolute error.

Medián absolútnej percentuálnej chyby: median absolute percentage error.

Metóda združeného poklesu: conjugate descent method (CD).

Miera chýb: error rate.

Mikro priemerovanie: micro-averaging.

Stredná kvadratická chyba: mean squared error.

Modifikovaná symetrická stredná absolútna percentuálna chyba: modified symmetric mean absolute percentage error.

Nekontrolované učenie: unsupervised learning. Tiež *učenie bez predlohy*, al. *učenie bez učiteľa*.

Nesterovov akcelerovaný gradient: Nesterov accelerated gradient [learning method] (NAG).

Umelá neurónová sieť: artificial neural network (ANN).

Nevrstvená sieť: non-layered network.

PCA: Analýza hlavných komponentov (angl. principal component analysis).

Analýza hlavných komponentov: principal component analysis.

Plocha pod (ROC) krivkou: area under (an ROC) curve.

Združovacia vrstva: pooling layer. Niekde tiež subsampling layer.

Posun vnútorných kovariát: internal covariate shift.

Pozitívna prediktívna hodnota: positive predictive value. Tiež: presnosť – precision.

Prahovanie: thresholding.

Presnosť: precision. Tiež: pozitívna prediktívna hodnota – positive predictive value.

Profesorské vnucovanie: professor forcing.

Protivnícke príklady: adversarial examples.

Protivnícke učenie: adversarial training.

RBM: Zjednodušený Boltzmannov stroj (angl. restricted Boltzmann machine).

Rekurentné učenie v reálnom čase: real-time recurrent learning.

Rektifikovaná lineárna jednotka (neurón): rectified linear unit; ReLU.

Presakujúca ReLU
leaky ReLU; LReLU.

Parametric ReLU
parametric ReLU; PReLU.

Rekurentné diaľnicové siete: recurrent highway networks (RHN).

Zjednodušený Boltzmannov stroj: restricted Boltzmann machine.

Reziduálne siete: residual networks (res-nets).

RHN: Rekurentná diaľnicová sieť (angl. recurrent highway network).

Odmocnina mediánu kvadratickej percentuálnej chyby: root median square percentage error.

Odmocnina strednej kvadratickej chyby: root mean squared error.

Odmocnina strednej kvadratickej logaritmickéj chyby: root mean squared logarithmic error.

Odmocnina strednej kvadratickej percentuálnej chyby: root mean squared percentage error.

RTRL: Rekurentné učenie v reálnom čase (angl. real-time recurrent learning).

Senzitivita: sensitivity. Tiež: citlivosť; TP miera – TP rate; úplnosť – recall.

Skórovací klasifikátor: scoring classifier.

Skratkové spojenia: shortcut connections.

Skutočne negatívny prvok: true negative.

Skutočne pozitívny prvok: true positive.

Symetrická stredná absolútna percentuálna chyba: symmetric mean absolute percentage error.

Symetrický medián absolútnej percentuálnej chyby: symmetric median absolute percentage error.

Technika syntetického nadvzorkovania minorít: synthetic minority oversampling technique (SMOTE).

Samoorganizujúca sa mapa: self-organising map.

Špecifita: specificity.

Správnosť: accuracy.

Temporálne kontrolované učenie: temporal supervised learning.

Tepelná rovnováha: thermal equilibrium.

Tomekova väzba: Tomek link.

TP miera: TP rate. Tiež: senzitivita, citlivosť – sensitivity; úplnosť – recall.

Transfer učenie: transfer learning.

U-matica: U-matrix.

Učenie bez predlohy: unsupervised learning. Tiež *nekontrolované učenie*, al. *učenie bez učiteľa*.

Učenie bez učiteľa: learning without a teacher. Tiež *nekontrolované učenie*, al. *učenie bez predlohy*.

Učenie s odmenou: reinforcement learning.

Učenie s predlohou: supervised learning. Tiež *kontrolované učenie*, al. *učenie s učiteľom*.

Učenie s učiteľom: learning with a teacher. Tiež *kontrolované učenie*, al. *učenie s predlohou*.

Učiteľské vnucovanie: teacher forcing.

Umelý neurón: artificial neuron (AN).

Úplnosť: recall. Tiež: citlivosť, senzitivita – sensitivity; TP miera – TP rate.

VAE: Variačný autoenkodér (angl. variational autoencoder).

Váhový návrat: weight-backtracking.

Validácia rozdelením: split validation.

Variačný autoenkodér: variational autoencoder (VAE).

Voľný beh: free running.

Vrstvená sieť: layered network.

Výber modelu: model selection.

WGAN: Wassersteinovské generatívne protivnícke siete (angl. Wasserstein generative adversarial networks).

Zákonitosti: regularities.

Združený gradient: conjugate gradient (CG).

Nelineárny združený gradient
nonlinear conjugate gradient (NCG).

Zväčšovanie dátovej množiny: data augmentation.

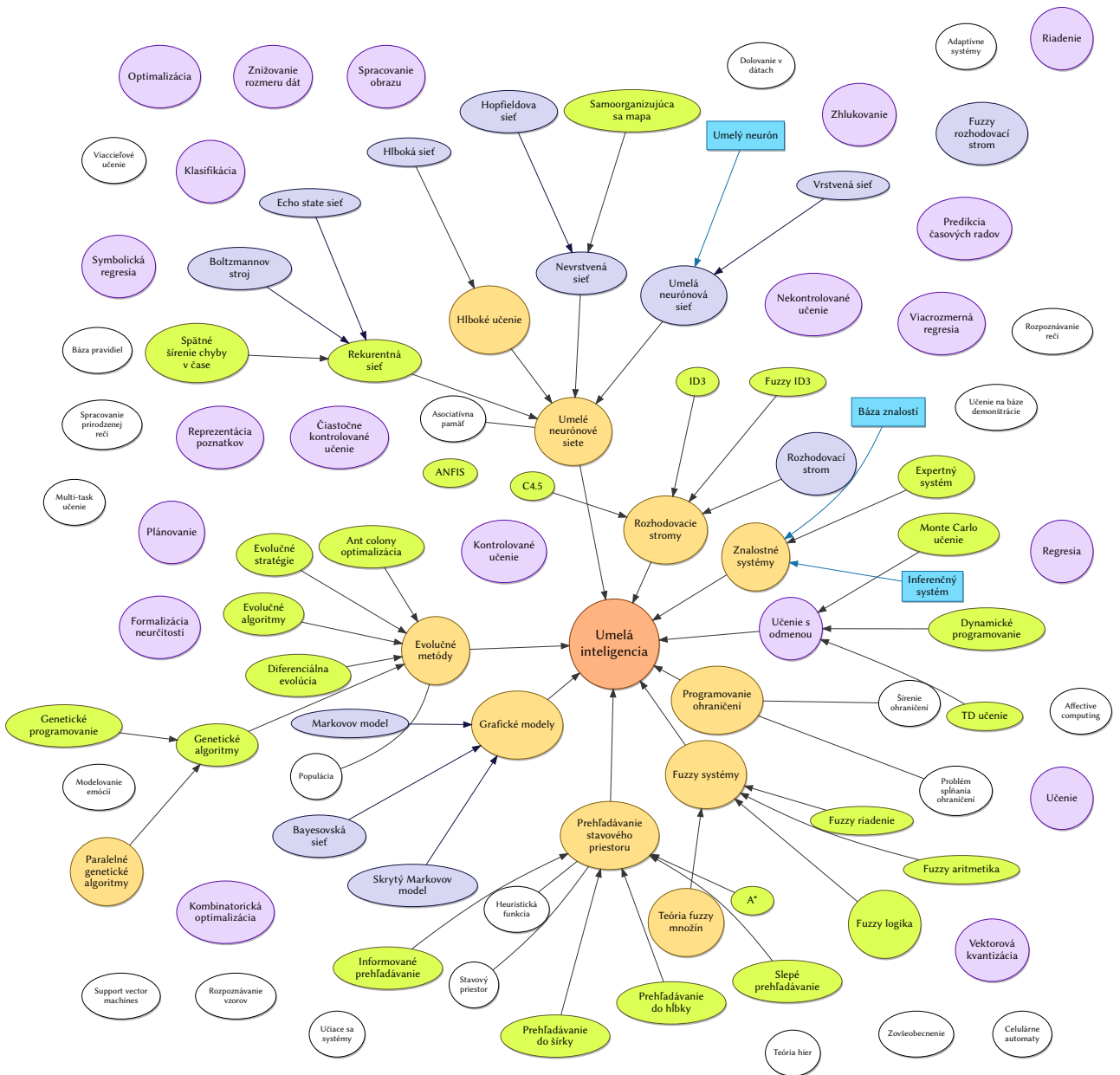
Aby sa čitateľ mohol lepšie zorientovať v oblasti umelej inteligencie a strojového učenia a tiež aby plnšie porozumel kontextu, do ktorého zapadá hlboké učenie, budeme prvú kapitolu venovať základným pojmom z tejto oblasti. Strojovým učením rozumieme súbor prístupov a metód, ktoré umožňujú strojom učiť sa, t.j. získavať určité užitočné správanie bez toho, aby ho bolo potrebné vopred explicitne naprogramovať.

Do oblasti umelej inteligencie patria okrem strojového učenia aj ďalšie oblasti – napríklad problematika reprezentácie poznatkov (explicitných/implicitných, vágnych/ostrých), problematika inteligentného plánovania, optimalizácie a ďalšie. My v tejto kapitole zhrnieme predovšetkým základné pojmy spoločné pre všetky smery strojového učenia a tiež predstavíme vysokoúrovňový pohľad na niektoré vybrané smery umelej inteligencie všeobecne. Pre širší úvod však možno pozrieť aj iné zdroje – napr. [5–7].

Diagram na obrázku Obr. 1.1 ukazuje prehľad vybraných pojmov z oblasti umelej inteligencie a čiastočne prezentuje aj vzťahy medzi nimi. V žiadnom prípade nie je naším cieľom poskytnúť úplný prehľad existujúcich oblastí a metód umelej inteligencie. Pri súčasnom obrovskom a neustále narastajúcom množstve znalostí z tejto oblasti by to bola prakticky nadľudská úloha. Budeme sa preto sústrediť predovšetkým na metódy a prístupy, ktoré sú v ďalšom texte diskutované a na pojmy s nimi súvisiace.

Ako vidno, diagram zobrazuje predovšetkým nasledujúce oblasti:

- Fuzzy systémy (viac informácií možno nájsť napr. v prvom dieli učebnice [8]);
- Grafické modely;
- Evolučné metódy;
- Umelé neurónové siete (klasickú teóriu umelých neurónových sietí možno takisto nájsť v prvom dieli učebnice; modernejší výklad a prístupy v tomto dieli);
- Učenie s odmenou;
- Programovanie ohraňení;
- Expertné a znalostné systémy (základy



Obr. 1.1: Strokové učenie – pojmy.

možno pozrieť napríklad v [6]).

Skôr, než prejdeme priamo ku témam týkajúcim sa hlbokého učenia, vysvetlíme ešte niektoré základné pojmy, ktoré sú spoločné pre celú teóriu strojového učenia.

1.1 | Typy učenia: rozdelenie podľa druhu úlohy

Ak budeme hovoriť už konkrétne o oblasti strojového učenia, ako o jednom z odvetví umelej inteligencie, je potrebné povedať, že strojové učenie môže slúžiť rôznym cieľom. V teórii strojového učenia sa preto – podľa druhu úlohy, ktorú riešia – rozlišujú nasledujúce typy učenia [9, 10]:

- **kontrolované učenie** (resp. učenie s učiteľom, angl. supervised learning, learning with a teacher);
- **učenie s odmenou** (angl. reinforcement learning);
- **nekontrolované učenie** (učenie sa bez učiteľa, angl. unsupervised learning, learning without a teacher);
- **čiastočne kontrolované učenie** (angl. semi-supervised learning).

Uvedme teraz krátku definíciu pre každý z týchto typov.

1.1.1 Kontrolované učenie

Charakteristickou vlastnosťou *kontrolovaného učenia* je, že učiaci sa systém dostáva spätnú väzbu. Systém dostáva na začiatku takzvanú **tréningovú množinu dát**, ktorá sa skladá z dvojíc (vstup, požadovaný výstup). Keďže pre každý vstup je známy aj požadovaný výstup, možno jednoducho vypočítať, nakoľko sa od neho odlišuje skutočný výstup učiaceho sa systému – t.j. možno určiť chybu.

Následne sa parametre systému nastavujú tak, aby chybu minimalizovali, čím dochádza k prenosu poznatkov obsiahnutých v dátach do štruktúry a parametrov systému, teda k učeniu. Hovoríme, že systém sa naučil tréningovú množinu. Ako o tom budeme ešte hovoriť nižšie (časť 1.2), vlastnosti takéhoto učiaceho sa systému sa následne typicky testujú na **testovacej množine dát**, ktorá nie je podmnožinou tréningovej množiny. Teda testovacia množina obsahuje také dáta, z ktorých sa systém neučil. Takto možno otestovať, či systém správne zovšeobecňuje.

Treba poznamenať, že v niektorej literatúre sa aj učenie s odmenou klasifikuje ako špecifický druh kontrolovaného učenia. Takémuto deleniu nechýba vnútorná logika, keďže učenie s odmenou, ako uvidíme, tiež operuje so spätnou väzbou, hoci v inej podobe. Rovnako však má zmysel – keďže