



# Úvod do deep learning

---

Zdeněk Linhart

Edice učebních textů

---

Odborné nakladatelství Vysoké školy ekonomie a managementu

 **VŠEM**

VYSOKÁ

ŠKOLA

EKONOMIE

A MANAGEMENTU

# ÚVOD DO DEEP LEARNING

Vysoká škola ekonomie a managementu

2021

# Úvod do deep learning

doc. Ing. Zdeněk Linhart, CSc.

Copyright © Vysoká škola ekonomie a managementu, 2021

Vydání první. Všechna práva vyhrazena.

ISBN: 978-80-88330-19-6

Vysoká škola ekonomie a managementu

[www.vsem.cz](http://www.vsem.cz)

Žádná část této publikace nesmí být publikována a šířena žádným způsobem a v žádné podobě bez výslovného svolení vydavatele.

# Obsah

Seznam tabulek a rovnic	5
Zkratky	8
Značky a symboly v učebním textu	10
<b>1. KAPITOLA: ÚVOD</b>	<b>12</b>
<b>2. KAPITOLA: STANDARDY UMĚLÉ INTELIGENCE</b>	<b>18</b>
Pochopení a interpretace výroků AI	18
Učení a testování neuronu	19
Principy deep learning a odlišení od machine learning	20
Zpracování přirozeného jazyka (NLP)	24
Zpracování jazyka a obrazu (CLIP)	24
<b>3. KAPITOLA: DOPADY AI</b>	<b>29</b>
3.1 Dopady AI na společnost	30
Simulace chování	30
Zbystření smyslů pomocí AI	33
3.2 Dopady na ekonomiku	37
Náhrada pracovní síly	37
Ziskovost AI	39
3.3 Psychologie podnikatele v AI	39
Využívání hromadných dat	39
Příklady poslání AI firem	40
<b>4. KAPITOLA: ANALÝZA POMOCÍ AI</b>	<b>46</b>
4.1 Specifika chování v podkladových záznamech	47
Specifika opačných úhlů pohledu konkurentů	47
Kód k zachycení a zobrazení specifík funkcemi z načtených dat	48
Zapojení falešně pozitivních a falešně negativních v absentujících datech	50
Vyvažování dat v sešikmeném rozdělení	51
4.2 Emoce	56
Zdroje rozpoznávání vztahů, výrazu obličeje a doporučení	56
Teorie emocí v AI	58
4.3 Řeč – NLP	61
Sentiment investorů	62
Analýza sentimentu v NLP	63
4.4 Funkce k dekódování podvodů v hloubce učení	65
Záznamy kroků a forem optimalizace nákladové funkce	65
Detekce chyby při ladění modelu DL	71

---

<b>5. KAPITOLA: PŘÍPADOVÁ STUDIE APLIKACÍ AI</b>	<b>76</b>
5.1 Opatřování textů a obrázků (scraping)	77
Ověřování stažených dat	77
Oblasti aplikace AI: obchod, investice, marketing, personalistika	78
5.2 Využití odchylek pro konkurenční účely zcizení a vztahů	80
Funkcionalita a fyzika	81
Úmysl a kauzalita	82
Užitečnost	83
5.3 Mikrokontroléry pro AI	85
Důvody k osvojení	85
Porozumění Raspberry Pi	86
<b>6. KAPITOLA: VYUŽITÍ SOFTWARE PRO AI (PYTHON)</b>	<b>90</b>
6.1 Modifikace kódu	91
Stažení a instalace software	91
N-tice (Tuples)	93
Slovníky (Dictionaries)	94
Sady (Sets)	95
6.2 Využívání knihoven podle návodu převzatého kódu	96
Zrychlení prototypováním algoritmů ML (Python)	97
Zrychlení běhu pomocí transformátorů (transformers)	99
Nástroje (editory) zpřístupňující knihovny	101
6.3 Odstraňování chyb při modifikaci kódu	101
Poznávání příčin chybových hlášení a odstraňování chyb	101
Návody na odstraňování chyb na webu (debugging)	102
<b>7. KAPITOLA: ZÁKLADY PROGRAMOVÁNÍ V PYTHONU</b>	<b>107</b>
7.1 Spuštění editoru Jupyter notebook	108
7.2 Kód z Guithub	109
7.3 Ukázky programování	110
Příklad stažení (scrapingu) dat z webu	110
Úvod k manipulaci jádra (kernel) dat pro funkce a modely	111
7.4 Kontrola porozumění pojmům	112
Literatura	120
Glosář	125
Vzorový test	128

## Seznam tabulek a rovnic

Tabulka 1 Příklad typografického útoku pro testování AI	19
Tabulka 2 Časy interakcí v experimentu	33
Tabulka 3 Nové zacházení s již nakoupenými zařízeními	34
Tabulka 4 Zatřídění nemoci indexem závažnosti CNN	35
Tabulka 5 Zatřídění nemoci indexem závažnosti CNN	35
Tabulka 6 Výsledky zdánlivě nesouvisící regrese	38
Tabulka 7 Výsledky zdánlivě nesouvisící regrese s naivní kalibrací	38
Tabulka 8 Poslání, obraty a počty zaměstnanců dodavatelů AI	41
Tabulka 9 Klastry metody K-Means (% tříd)	59
Tabulka 10 Přesnost klasifikátorů ML	60
Tabulka 11 Funkce gradientu sestupu pro specifické podmínky	68

## Seznam obrázků

Obrázek 1 Podmnožiny ML a DL v rámci AI a oboru datové vědy	13
Obrázek 2 Typy klasifikace, metody a užití ML	13
Obrázek 3 Symbolika vstupů a výstupů regresní rovnice v AI	14
Obrázek 4 Detekce multimodálních neuronů	19
Obrázek 5 Klasifikace rodin neuronů pro machine learning	20
Obrázek 6 Odchytky od standardu nulové aktivity	25
Obrázek 7 Aktivační grafika u neuronu Donald Trump	25
Obrázek 8 AI v týmu	30
Obrázek 9 Korekce postupu výstavby podle posouzení alternativ s a bez AI	31
Obrázek 10 Čtení záznamu výroků AI na výkresech stavby	32
Obrázek 11 Zavádění propojeného textu a obrazu v instituci	34
Obrázek 12 Diagnóza a řešení příznaků s pomocí AI a IoT	35
Obrázek 13 Heatmap	36
Obrázek 14 Vznik hromadných dat instalací kamer do strojů	40
Obrázek 15 Vzdalování psychologických smluv k opačným pólům	48
Obrázek 16 Spuštění statistiky k výběru nejlepší funkce metodou SelectKBest	48
Obrázek 17 Skóre funkcí vybraných metodou SelectKBest	49
Obrázek 18 Výběru nejlepší funkce metodou ExtraTreesClassifier	49
Obrázek 19 Graf s výstupy metody ExtraTreesClassifier	50
Obrázek 20 Úpravy absentujících dat ve zdroji	50
Obrázek 21 Vyvážení dat metodou RandomUnderSampler	51
Obrázek 22 Vyvážení dat metodou RandomOverSampler	51
Obrázek 23 Vyvážení dat metodou SMOTE	51
Obrázek 24 Krabicový graf s odlehlým bodem	52
Obrázek 25 Výpočet 25 % a 75 % kvantilů	52
Obrázek 26 Náhrada odlehlých hodnot	53
Obrázek 27 Odhazování dat do koše při převodu spojitych na diskrétní data	53
Obrázek 28 Převody slovních označení na čísla	54
Obrázek 29 Dvě metody standardizace	55
Obrázek 30 Metody normalizace	55
Obrázek 31 Přesnost měření emocí robotem při prodeji oblečení	56
Obrázek 32 Použití emocí v cloudových databázích ke zkoušení a prodeji robotem	57
Obrázek 33 Odvození chytré otázky a odpovědi AI	58
Obrázek 34 Reprezentace intenzity emocí účastníků pomocí GSOM	59
Obrázek 35 Škála afektů	60
Obrázek 36 Zmatek sebepoškozujících se	61
Obrázek 37 Zlom sentimentu za druhým vrcholem	62
Obrázek 38 Analýza sentimentu	63
Obrázek 39 Mrak slov	64
Obrázek 40 Užívání unikátních slov	64
Obrázek 41 Sestup s hory poslepu jen hmatem	66
Obrázek 42 Geometrický zápis gradientu sestupu	66

---

Obrázek 43 Zátěž učením	67
Obrázek 44 Zápis funkce gradientu sestupu kódem	67
Obrázek 45 3D vizualizace testu účelové funkce	68
Obrázek 46 Zobrazení 2D optima Nesterova Momentu	69
Obrázek 47 Příklad dotazu a odpovědi k hyperparametru nákladové funkce	72
Obrázek 48 Sentiment retweetů	77
Obrázek 49 Střední intervaly spolehlivosti a bodový záznam sentimentu retweetů	78
Obrázek 50 AI prosazovaná jako zařízení managementem	79
Obrázek 51 Integrace scénářů (FPICU)	81
Obrázek 52 Vyrovnávání rozdílných rychlostí v bodu na kritické cestě	82
Obrázek 53 Pravda (t – true) a faleš (f – false) v úmyslech CBN	83
Obrázek 54 Nejlacinější mikrokontrolér	85
Obrázek 55 Porovnání SVD s PCA	98
Obrázek 56 Porovnání SKLEARN a ztrátové funkce PCA	99
Obrázek 57 Transformace výsečí obrazu	100
Obrázek 58 Generalizace v detekci obrazu (DETR)	101
Obrázek 59 Příklad hlášení chyby	102
Obrázek 60 Odvozování štítku z obrazu	102
Obrázek 61 Hierarchie ladění rozhraní a ladícího nástroje	103
Obrázek 62 Otevření prvního notebooku	109
Obrázek 63 Ověření a prohloubení integrity AI	111



## Zkratky

AI (artificial intelligence) je umělá inteligence

ANN (artificial neural network) je zpětnovazební umělá neuronová síť

API (application programming interface) je rozhraní, které definuje interakce mezi více softwarovými aplikacemi

CBN (Causal Bayes Networks) je síť pravděpodobnostní příčinnosti

CLIP je neuronová síť, která se učí vizuální koncepty z dohledu nad NLP

CMA (Causal Markov Assumption) je předpoklad podmíněné nezávislosti

CNN (Convolutional Neural Network) je asi nejznámější algoritmus AI pro rozpoznávání a zatřídění obrázků

CPM (Critical Path Method) je jednou ze základních matematických modelů k nalezení úzkého místa na kritické cestě, jehož zprůchodněním se zvyšuje výkon celku

CPU (central processing unit) je obvod, který provádí pokyny

DAG (directed acyclic graph) směřovaný acyklický graf. To znamená, že je nemožné procházet celým grafem počínaje od jedné hrany. Graf je topologické třídění, kde je každý uzel v určitém pořadí.

DL (deep learning) je označení pro způsob učení se AI, který nevyžaduje klasifikaci vlastností, ale počítá si je sám v hloubce neuronových sítí

FPICU (functionality, physics, intent, causality, and utility) jsou prvky, kterými AI z kamerového záznamu vyhodnocuje interakce mezi funkcionalitou, fyzikou, úmyslem, kauzalitou, a užitečností, třeba pro potřeby průmyslové špionáže bez účasti lidí

GAN (generative adversarial network) je třída rámců strojového učení, kde dvě neuronové sítě navzájem soupeří ve hře s nulovým součtem

GDA (Generalized Discriminant Analysis) je zobecněná diskriminační analýza. Používá se k redukci dimenzionality

GDPR (General Data Protection Regulation) je obecné nařízení o ochraně osobních údajů je EU ke zvýšení ochrany osobních dat občanů, aby se nedal dohledat zločin

GPIO (General Purpose Input Output Pins) je standard pro označování pinů na svorkovnicích mikrokontrolérů

GPU (Graphics Processing Unit) je čip nebo elektronický obvod schopný vykreslit grafiku pro zobrazení na elektronickém zařízení

GSOM (Growing Self-Organizing Maps) jsou metodou zobrazování rostoucích sebeorganizujících se map

GUI (The graphical user interface) je grafické uživatelské rozhraní

HLLAI (Human Level AI) konference k AI v oblasti lidských zdrojů

IoT (Internet of Things) je internet věcí. Od běžného internetu se liší tím, že v něm mezi sebou komunikují zařízení typu vysavač, spínač topení, domovní zvonek, vařič a podobně. IoT nemá servery ani obsluhu

ItemCF (item-based collaborative filtering algorithm) je algoritmus filtrování kolaborativních položek (filmů s příbuznou tematikou pro nalezenou emoci). Z ItemCF se odvozuje podobnost zájmů dvou uživatelů, tedy potencionální kolaborativnost

LDA (Latent Dirichlet Allocation) je latentní Dirichletova alokace generativním statistickým modelem, který umožňuje vysvětlit množiny pozorování nepozorovanými skupinami, které vysvětlují, proč jsou některé části dat podobné

LDA (Linear Discriminant Analysis) je lineární diskriminační analýza k redukci dimenzionality nebo

LSTM (Long short-term memory) je architektura umělé rekurentní neurální sítě (RNN) používaná k učení v hloubce sítě (deep learning)

MAE (mean absolute error) je střední absolutní chyba

- 
- MEU (maximum expected utility) je maximální očekávaný užitek, ke kterému se AI přibližuje nebo od něj odměřuje chybu v podobě rozdílu
- ML (machine learning, strojové učení) je označení pro způsob učení se AI, který vyžaduje klasifikaci vlastností, kterým se říká neurony nebo rodiny neuronů
- MLP (multilayer perceptron) je třída zpracování obrazu v rámci ANN
- MORL (multiobjective RL) je nastavení mnoha cílů algoritmu posíleného učení
- MSE (mean squared error) je střední kvadratická chyba, která se používá k minimalizaci funkce ztráty pomocí Adamovy verze stochastického sestupu gradientu
- NA (not available) jsou absentující data.
- NLP (natural language processing) je zpracování přirozeného jazyka
- PCA (Principal Component Analysis) je analýza základních komponent. Používá se k redukci dimenzionality
- PES je systém protiepidemických opatření
- ReLU (Rectified Linear Unit) je aktivační funkce, která prahuje vstupy na nulu, tj. ReLU nastaví všechny záporné hodnoty na vstupu do neuronu na nulu
- RGB (Red, Green and Blue) je vymezení hranic prostoru barev digitálního obrazu
- RL (reinforcement learning) je posílené učení
- RMSE (root mean squared error) je model regresní predikce
- SA (static analysis) je software pro hledání chyb programového kódu
- SaMD (software as medical device) software AI jako (medicínské) zařízení
- SMOTE (Synthetic Minority Oversampling Technique) je technika umělého převzorkování minority, aby mohla být využita data většiny při zpracování statistickými metodami
- SVD (singular value decomposition) je v lineární algebře rozklad singulární hodnoty faktorizací
- SVM (Support Vector Machine) je metoda strojového učení s učitelem sloužící zejména pro klasifikaci a také pro regresní analýzu.
- TanH (Hyperbolic Tangent Activation Function) je aktivační funkce, která nabývá hodnot v intervalu  $-1, 1$  s vyšší přesností než Sigmoid a nižší než ReLU
- Tensor Processing Unit (TPU) je akcelerátor umělé inteligence integrovaný do specifického aplikačního obvodu vyvinutý společností Google speciálně pro stroj neuronové sítě
- UserCF (user-based collaborative filtering) je algoritmus filtrování kolaborativních uživatelů
- WOM (word-of-mouth) je ústní předávání informace, bez kterého neosobní propagace v médiích není vnímána
- ZSL (Zero-shot learning) je nastavení strojového učení v testovacím čase i na vzorky, které nebyly během tréninku pozorovány, aby pro odvození jejich třídy nemuselo být restartováno učení

---

# Značky a symboly v učebním textu

Struktura distančních učebních textů je rozdílná již na první pohled, a to např. v zařazování grafických symbolů – značek.

*Specifické grafické značky umístěné na okraji stránky upozorňují na definice, cvičení, příklady s postupem řešení, klíčová slova a shrnutí kapitol. Značky by měly studenta intuitivně vést tak, aby se již po krátkém seznámení s distanční učebnicí dokázal v textu rychle a snadno orientovat.*

## Definice



*Upozorňuje na definici nebo poučku pro dané téma.*

---

## Příklad



*Označuje příklad praktické aplikace učiva včetně řešení.*

---

## Otázky k procvičení a úkoly



*Označuje otázky a úkoly s postupem řešení na konci kapitoly.*

---

## Klíčová slova



*Upozorňuje na důležité výrazy či odborné termíny nezbytné pro orientaci v daném tématu.*

---

## Shrnutí kapitoly



*Shrnutí kapitoly se zařazuje na konec dané kapitoly. Přehledně, ve strukturovaných bodech shrnuje to nejpodstatnější z předchozího textu.*

---

# 1

kapitola

---

---

## Úvod

# 1. kapitola

## Úvod

### Cíle kapitoly

- Kapitola k ovlivnění učení uvádí, kde končí metody úvodu, které budou výstupem tohoto předmětu, a kde začíná datová věda a praxe aplikací AI, které se mohou stát předmětem kariéry.
- Pozornost je věnována tomu, jak se počítač učí se software AI a jak se učí jeho uživatel.
- Kapitola přináší souhrn a aplikaci poznatků o množinách a podmnožinách metod a o souvislostech mezi aplikacemi, které jsou používány k učení počítačem.

#### DEFINICE



#### Metody učení se strojem (machine learning), též strojové učení

Regrese (Regression)

Klasifikace (Classification)

Shlukování (Clustering)

Reduce dimenzí (Dimensionality Reduction)

Soubory metod (Ensemble Methods)

Neuronové sítě a hluboké učení (Neural Nets and Deep Learning)

Přenos učení (Transfer Learning)

Posílení učení (Reinforcement Learning)

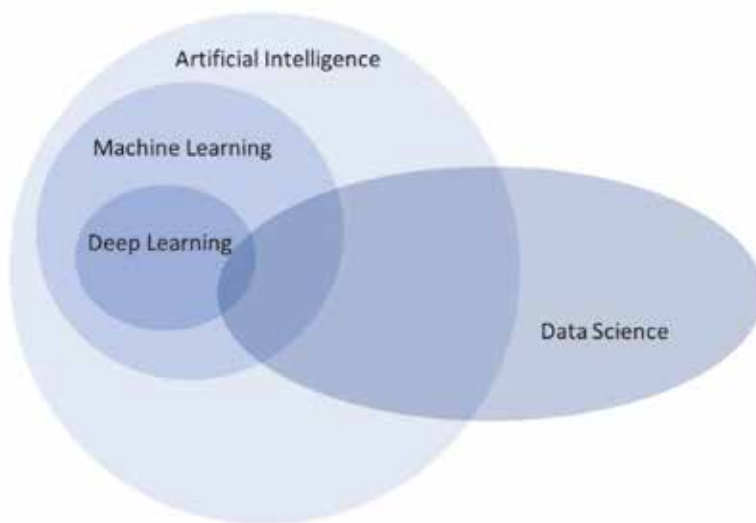
Strojové učení (ML) se učí z dat a hluboké učení (DL) z vrstev „umělé neuronové sítě“, která se může učit. Výstup ML i DL slouží jako pomoc při rozhodování. Vstupním předpokladem pro předmět úvod do deep learning je základní znalost principů práce s počítači a webem. Důvod, proč název předmětu dal přednost DL místo ML je ten, že algoritmy vyvíjí efektivitu DL jak v soukromé, tak veřejné sféře tak rychle, že nahrazují práci s klasifikací i pomaleji se vyvíjející hardware. Autoři předpokládají, že algoritmy DL se budou brzy prodávat nebo se stanou součástí produktů, mezi kterými ekonomové budou muset vybírat, když budou dělat analýzu proveditelnosti pro žádosti o úvěr v bance.

Předmět proto seznamuje studenta s aktuální problematikou směřování umělé inteligence. Student získá schopnost popsat rozdílnými technikami machine learning and deep learning pochopit souvislosti a dopad výstupů umělé inteligence na ekonomiku, společnost a sociální status obyvatel, včetně své podnikatelské schopnosti.

**Cílem předmětu je** seznámit studenty s technikami AI k prozkoumání a překonání rozdílu mezi řečí a chováním pomocí technik ML a DL, protože ten se stává konkurenční výhodou nebo zátěží, není-li poznán a využit AI (Obrázek 1).

## OBRÁZEK 1

### Podmnožiny ML a DL v rámci AI a oboru datové vědy



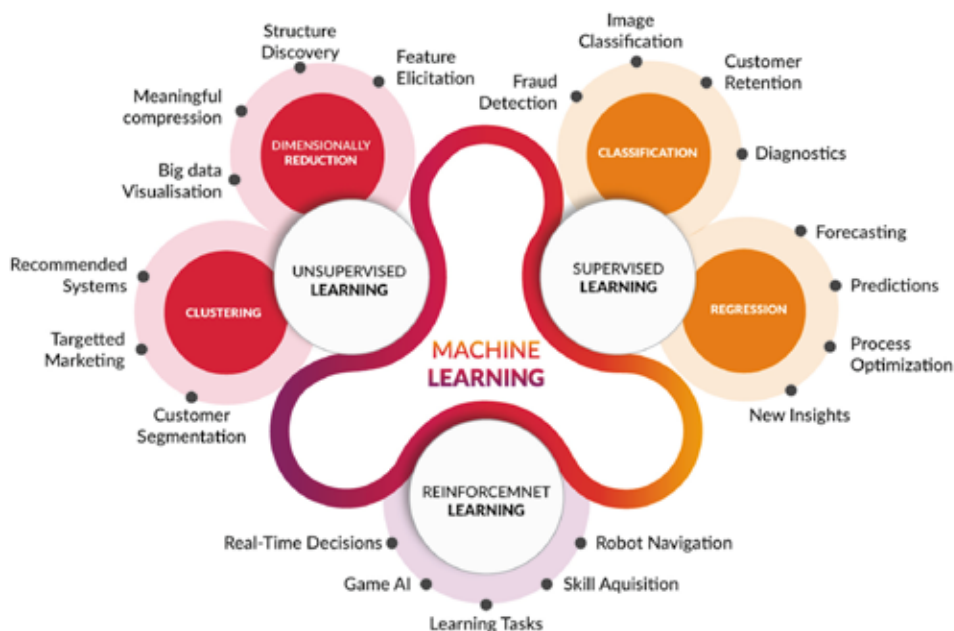
Zdroj: Ghavardan@123, 2021

Metodami je dosahováno cílů nastavením středové osy náměrů, interpretací měření v kartézských souřadnicích sentimentu a technik vizualizace fazet neuronů a výskytu polysémantických neuronů v databázích. Zvládnutí způsobí měření jevů s globálním nebo dosud nevnímaným dopadem na nepojmenované jevy je znakem úspěšného absolvování předmětu i zvýšené podnikatelské schopnosti účastníků kurzu.

Četnost pojmů a jejich vazby s kontrolovatelným ML (Obrázek 2) demonstrují, že ti, pro které je předmět novinkou budou potřebovat bližší seznámení se s citovanými zdroji, včetně dostudování novinek, které se objevují denně.

## OBRÁZEK 2

### Typy klasifikace, metody a užití ML



Zdroj: Ghavardan@123, 2021

Tento předmět s názvem „Úvod do deep learning“ splnil své poslání, když výsledky učení odpovídají výstupním cílům předmětu:

- student umí vysvětlit koncept umělé inteligence,
- student se umí orientovat v pojmech používaných umělou inteligencí,
- student dokáže propojit příčiny v sentimentu a spojit je s důsledky v chování, včetně svého sentimentu a chování,
- student dokáže formulovat požadavky na zavádění umělé inteligence,
- student umí analyzovat a využít výstupy AI pro zdokonalování sebe i obchodu,
- student umí aplikovat analýzu pomocí vybraných nástrojů AI.

## CVIČENÍ 1



Jaký vliv má znalost AI na kariéru jednotlivce a udržitelnost průmyslu?

Na uplatnění výsledků učení a schopností čekají firmy, které nabízí například pozice marketér nebo data scientist, aby jim pomohly transformovat inovace vědeckým základem projektů při přechodu na průmysl 4.0 nebo automatizaci služeb.

## CVIČENÍ 2



Jaký vliv má znalost AI na profesi ekonoma?

Data Science není obor, kde teoretické porozumění pomůže začít kariéru. AI připočítává falešně pozitivní a falešně negativní hodnoty k výstupům, aby odvodila vliv držitelů zájmů, což teorie nedělají. Ke konstrukci funkcí pro projekt (Obrázek 3) pomáhá délka praxe.

## ➔ OBRÁZEK 3

**Symbolika vstupů a výstupů regresní rovnice v AI**



Zdroj: Kumar, 2021

## CVIČENÍ 3



Jaká je průměrná vypovídací schopnost modelu AI bez vyčištění dat?

**Pracovní funkce pro absolventy kurzu**

- |                              |  |
|------------------------------|--|
| A) AI Engineer               | F) ML Architect                          |
| B) Machine Learning Engineer | G) ML Specialist                         |
| C) Data Scientist            | H) Researcher                            |
| D) Data Analyst              | I) Applied ML Scientist a mnoho dalších. |
| E) Data Engineer             |  |

Pomoc s odlišením dovedností v průběhu vývoje projektů s odlišnou tematikou si můžete vyžádat i na Všem nebo od Coursera, Udemy, Edx atd. Certifikace přidávají hodnotu vašim dovednostem při žádosti o zaměstnání. Plnohodnotný programátor a vedoucí projektů, kromě praktických znalostí, potřebuje mít také dobré teoretické znalosti.

## Klíčová slova



- Deep learning (DL), též hluboké učení
- Machine learning (ML), též strojové učení
- Data science, též datová věda

## Řešení ke cvičením



### Cvičení 1

AI se učí a zobrazuje zajímavé výstupy, které laik nedokáže interpretovat. Proto budou oceňováni zaměstnanci, kteří jak přeloží výstupy AI, tak dokážou objednat aplikaci AI k napojení projektu k úspoře pracovníků průmyslu a služeb.

### Cvičení 2

Profese ekonoma kvantifikuje a do vzorců podnikání přeloží každou inovaci. Ekonomem se uchazeč o profesi stává, když si inovaci osvojí, aby prodával se ziskem její produkty. Pokud není schopen kvantifikovat a se ziskem prodat inovace na světovém trhu, je to v nejlepším případě účetní. Zde se věnujeme AI, ale u CRISPR, ale i jiné exoticky znějící inovace obdobně po ekonomech vyžadují orientaci v prvních nálezech ve vědeckých článcích a rychlé zavedení inovace na trh. Pokud ekonom stále vidí svou budoucnost v rozúčtovávání covidové nebo jiné podpory na zaměstnance, tak si může být jist, že brzy nebude z čeho rozesílat podpory či nárokové mzdy.

### Cvičení 3

Přesnost modelu AI bez vyčištění dat je cca 70%. Po vyčištění dat lze ve strojovém učení zvýšit přesnost klasifikací a zaučením modelu (Feature Engineering), v hlubokém učení porovnáním kontrastnějších zdrojů.

## Otázky s možností výběru odpovědi:



1. V průběhu školní docházky žák a student zvyšuje efektivitu učení tříbením
  - a) stylu
  - b) metody
  - c) nástrojů
  - d) učení se počítačem
  - e) učení se od počítače
2. AI zvyšuje efektivitu učení tříbením
  - a) stylu
  - b) metody
  - c) nástrojů
  - d) učení se počítačem
  - e) učení se od počítače



- 
3. Vědec v oboru velkých dat (Data scientist) zvyšuje efektivitu učení tříbením
- a) stylu
  - b) metody
  - c) nástrojů
  - d) učení se počítačem
  - e) učením se od počítače

## Řešení k otázkám s možností výběru odpovědí:



1 a, d; 2 b, c, d; 3 b, c, d, e;

# 2

---

kapitola

---

## Standardy umělé inteligence

# 2. kapitola

## Standardy umělé inteligence

### Úvod

Standardy AI se stále vyvíjí, ale měřeno stářím Pythonu, který se standardem stal, mají za sebou relativně dlouhou historii. Relativně malý počet lidí, kteří AI vyvíjí a používají, není důsledkem toho, že by neuměli programovat, ale spíš toho, že výsledky AI nedokážou interpretovat a implementovat, aby dosáhli zisku ať sami nebo přes prodej služeb nějaké firmě nebo trhu. Z toho důvodu je standard AI v této kapitole pojat tak, aby pomohl s interpretací výroků AI. Tato dovednost pomůže při nákupu zboží, které AI obsahuje, ale i při objevování netušeného kolem sebe, zejména v jednání a při řízení lidí. Teprve po dotažení do standardu se v dalších kapitolách budeme věnovat nadstandardu, a dovednostem pro excelenci.

### Cíle kapitoly

- Kapitola uvádí metody, které usnadňují pochopení a interpretaci výroků AI.
- Pozornost je věnována poznávání vlivu funkcí a polysémantiky neuronů na aktivaci.
- Kapitola přináší souhrn a aplikaci poznatků o deep learning a machine learning.

#### DEFINICE



##### Učení se z funkcí vizualizace na štitcích neuronů

AI na výstupu vizualizuje detaily (funkce), které nemusí dávat smysl. Proto je funkcím po pozorování stovek stimulů, které aktivují neuron, přiřazován štítek s označením konceptu, který dává neuronu smysl. Koncepty mohou být emoce, regiony, a podobně (Obrázek 2). Z názvů funkcí na štitcích neuronu se v průběhu učení odvozují polysémantické neurony.

##### Aktivace

Aktivace je proces, kterým se funkce (detail) neuronu z jednoho případu, ať slova nebo skupiny pixelů na obrázku, spojuje s větším nebo menším počtem jiných slov nebo obrázků. Aktivace tak prohlubuje jak učení, tak uplatnění poznatků AI.

### Pochopení a interpretace výroků AI

Chápání a interpretace výroků AI jsou odvozovány z četnosti výskytu sledovaných prvků, ať pixelů nebo číselné hodnoty slov a podobně. Dle četnosti výskytu RGB (Red, Green and Blue) pixelů jsou klasifikovány neurony (Obrázek 4). Obraz každého neuronu, tedy třídy s velmi podobnými vlastnostmi, interpretuje nejčastěji se vyskytující motivy velikostí zobrazení. Pokud AI místo RGB pixelů zobrazuje četnost výskytu a interakce slov, je nejzákladnější interpretací výroku AI mrak slov (word cloud).

Rozdíly mezi lidskou a počítačovou inteligencí jsou patrné po srovnání toho, jak pracují bio-neurony, kterými rozlišujeme osoby v rolích, například herce s postupy neuronů v síti AI. Bio-neurony z pozorování odvozují identitu přiřazováním ke známým obrazům, ke kterým se pozorované fragmenty transformovat. Transformace spojují výsledky proměrování tvarů až do koncepční informace (Quiroga et al., 2005). Důkaz CLIP RN50\_4x ukazuje identickou vypovídací schopnost jako elektrody v mozku. To znamená, že umí na fotografii identifikovat osobu, ztotožnit ji s kresbou a s textem při daném rozlišení. Detektory osob umí jen ztotožnit osobu bez symboliky kresby a bez popisku.

## OBRÁZEK 4

### Detekce multimodálních neuronů

Důkaz elektrodami v mozku



Fotografie Halle Berry v kostýmu



Fotografie Berry na fotografii byla spojena s kresbou



Obrazy Halle Berry byly spojeny s jejím psaným jménem



Důkaz CLIP RN50 4x



Fotografie Spidermana v kostýmu testovala asociace s pavouky



Kresba Spidermana s charakteristikou pavoučí sítě



K fotografií Spidermana byl přiřazen tento obrázek z databáze s textem i sítí



Důkaz generickým detektorem osob Inception v1



Předložená fotografie



Vygenerovaná kresba neodpovídá



Vygenerovaný text neodpovídá



Zdroj: Goh et al., 2021

### Učení a testování neuronu

Některé neurony připomínají témata osnov mateřské školky: počasí, roční období, písmenka, počítání a základní barvy. Ač triviální, jsou tato témata multimodální. Například neuron pro žlutou barvu nestačí, aby rozlišil slova „žlutá“, „banán“, a „citron“.

## TABULKA 1

### Příklad typografického útoku pro testování AI

	iPod	95 %
	Lasička	0,5 %
	Ovladač	0,4 %
	Křeček	0,4 %
	Surikata	0,2 %
	schéma pro modelování dat aplikace	0,1 %

Zdroj: Goh et al., 2021

Modely AI jsou typografickými útoky záměrně testovány ke zvyšování vypovídací schopnosti rozlišování reálných objektů, například

- vizualizací a porozuměním opakujícím se sítím (Karpathy, Johnson, Fei-Fei, 2015),
- detektory objektů se objevují v hluboké scéně CNN (Zhou et al., 2014),
- při kvantifikaci interpretovatelnosti hlubokých vizuálních reprezentací (Bau, 2017),
- při zaostřování na obvyklé (Olah et al., 2020).

Cílem typografických útoků je nalézt a posunout hranice špatného ve prospěch přesného zatřídění objektů.

Goh, et al. (2021) pozorovali efektivitu chápání aktivace neuronů, tedy učení, jak na straně počítače, tak na straně uživatele AI komparací vizualizovaných funkcí s jejich výskytem v databázi.

- Vizualizace funkcí byla pozorována podle aktivace počtu připojených neuronů v průběhu optimalizace vstupů (scraping). Zde se čekalo na výstup počítače s AI.
- Databáze umožňovala v rozdělení neuronů vybrat ze souboru dat ty, které aktivovaly svými odlišnostmi ostatní neurony nejvíce a spolehlivě. Zde šlo o lidské pozorování.