

# Využití umělé inteligence pro data mining v oblasti ekonomie a managementu

Martin Kubíček

---

Bakalářská práce  
2022

 Univerzita Tomáše Bati ve Zlíně  
Fakulta managementu a ekonomiky

---

Univerzita Tomáše Bati ve Zlíně  
Fakulta managementu a ekonomiky  
Ústav ekonomie

Akademický rok: 2021/2022

# ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Martin Kubíček**  
Osobní číslo: **M190591**  
Studijní program: **B0413A050024 Ekonomika a management**  
Specializace: **Ekonomika a management podniku**  
Forma studia: **Prezenční**  
Téma práce: **Využití umělé inteligence pro datamining v oblasti ekonomie a managementu**

## Zásady pro vypracování

Úvod

Definujte cíle práce a použité metody zpracování práce.

I. Teoretická část

- Zpracujte literární rešerši z oblasti dataminingu a systémů umělé inteligence.

II. Praktická část

- Proveďte analýzu možností využití umělé inteligence v kontextu potřeb současné ekonomiky a moderního managementu.
- Vyhodnoťte možné nástroje z hlediska jejich efektivnosti a účelnosti.

Závěr

Rozsah bakalářské práce: **cca 40 stran**  
Forma zpracování bakalářské práce: **tištěná/elektronická**

Seznam doporučené literatury:

- DOSTÁL, Petr. *Soft computing v podnikatelství a veřejné správě*. Brno: Akademické nakladatelství CERM, 2015, 130 s. ISBN 978-80-7204-896-0.
- HAN, Jiawei, Micheline KAMBER a Jian PEI. *Data mining: concepts and techniques*. 3rd ed. Waltham: Elsevier, 2012, 703 s. ISBN 978-0-1238-1479-1.
- HOLOŠKA, Jiří. *Artificial intelligence applied on cryptanalysis aimed on revealing weaknesses of modern cryptology and computer security*. Zlín: Tomas Bata University in Zlín, 2012, 30 s. ISBN 978-80-7454-144-5.
- HOLUBOVÁ, Irena, Jiří KOSEK, Karel MINAŘÍK a David NOVÁK. *Big Data a NoSQL databáze*. Praha: Grada, 2015, 281 s. ISBN 978-80-2475-466-6.
- LABERGE, Robert. *Datové sklady: agilní metody a business intelligence*. Brno: Computer Press, 2012, 350 s. ISBN 978-80-2513-729-1.

Vedoucí bakalářské práce: **Ing. Milan Damborský, Ph.D.**  
Ústav ekonomie

Datum zadání bakalářské práce: **11. února 2022**  
Termín odevzdání bakalářské práce: **20. května 2022**

L.S.

---

**prof. Ing. David Tuček, Ph.D.**  
děkan

---

**doc. Ing. Petr Novák, Ph.D.**  
garant studijního programu

Ve Zlíně dne 11. února 2022

## **PROHLÁŠENÍ AUTORA**

### **BAKALÁŘSKÉ/DIPLOMOVÉ PRÁCE**

#### **Prohlašuji, že**

- beru na vědomí, že odevzdáním diplomové/bakalářské práce souhlasím se zveřejněním své práce podle zákona č. 111/1998 Sb. o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších právních předpisů, bez ohledu na výsledek obhajoby;
- beru na vědomí, že diplomová/bakalářská práce bude uložena v elektronické podobě v univerzitním informačním systému dostupná k prezenčnímu nahlédnutí, že jeden výtisk diplomové/bakalářské práce bude uložen na elektronickém nosiči v příruční knihovně Fakulty managementu a ekonomiky Univerzity Tomáše Bati ve Zlíně;
- byl/a jsem seznámen/a s tím, že na moji diplomovou/bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, zejm. § 35 odst. 3;
- beru na vědomí, že podle § 60 odst. 1 autorského zákona má UTB ve Zlíně právo na uzavření licenční smlouvy o užití školního díla v rozsahu § 12 odst. 4 autorského zákona;
- beru na vědomí, že podle § 60 odst. 2 a 3 autorského zákona mohu užít své dílo – diplomovou/bakalářskou práci nebo poskytnout licenci k jejímu využití jen s tím, že vyrovnání případného přiměřeného příspěvku na úhradu nákladů, které byly Univerzitou Tomáše Bati ve Zlíně na vytvoření díla vynaloženy (až do jejich skutečné výše) bude rovněž předmětem této licenční smlouvy;
- beru na vědomí, že pokud bylo k vypracování diplomové/bakalářské práce využito softwaru poskytnutého Univerzitou Tomáše Bati ve Zlíně nebo jinými subjekty pouze ke studijním a výzkumným účelům (tedy pouze k nekomerčnímu využití), nelze výsledky diplomové/bakalářské práce využít ke komerčním účelům;
- beru na vědomí, že pokud je výstupem diplomové/bakalářské práce jakýkoliv softwarový produkt, považují se za součást práce rovněž i zdrojové kódy, popř. soubory, ze kterých se projekt skládá. Neodevzdání této součásti může být důvodem k neobhájení práce.

#### **Prohlašuji,**

1. že jsem na diplomové/bakalářské práci pracoval samostatně a použitou literaturu jsem citoval. V případě publikace výsledků budu uveden jako spoluautor.
2. že odevzdaná verze diplomové/bakalářské práce a verze elektronická nahraná do IS/STAG jsou totožné.

Ve Zlíně

Jméno a příjmení: .....

.....  
podpis diplomanta

## **ABSTRAKT**

Práce se zabývá umělou inteligencí, použitím umělé inteligence v podnicích, a především využití data miningu. Jsou zde popsány základní pojmy, které jsou potřeba pro porozumění praktické části. V praktické části se zhodnocují výsledky na základě veřejně dostupných dat na internetu. Data odhalí oblasti využívání data miningu, trend umělé inteligence a také míru digitalizace. Dále je zde uveden výčet moderních trendů v data miningu a interpretace výsledků.

Klíčová slova: data mining, Big data, umělá inteligence, sociální sítě, software, strojové učení

## **ABSTRACT**

The work deals with artificial intelligence, the use of artificial intelligence in companies, and especially the use of data mining. It describes the basic concepts that are needed to understand the practical part. The practical part evaluates the results based on data publicly available on the internet. The data will reveal the areas of data mining use, the trend of artificial intelligence and also the degree of digitization. There is a list of modern trends in data mining and interpretation of results as well.

Keywords: Data Mining, Big Data, Artificial Intelligence, Social Medias, Software, Machine Learning

## Poděkování

Děkuji rodině za podporu a vedoucímu práce Ing. Milan Damborský, Ph. D. za odborné vedení a rady, které mi poskytnul při zpracování bakalářské práce.

Prohlašuji, že odevzdaná verze bakalářské/diplomové práce a verze elektronická nahraná do IS/STAG jsou totožné.

# OBSAH

<b>ÚVOD</b> .....	<b>9</b>
<b>CÍLE A METODY ZPRACOVÁNÍ</b> .....	<b>10</b>
<b>I TEORETICKÁ ČÁST</b> .....	<b>11</b>
<b>1 VYMEZENÍ ZÁKLADNÍCH POJMŮ</b> .....	<b>12</b>
1.1 SOFTWARE, HARDWARE A BIG DATA .....	12
1.1.1 Velikost Big dat.....	12
1.2 ALGORITMUS A DATA MINING.....	13
1.3 HISTORIE DATA MININGU .....	14
1.4 VYUŽITÍ DATA MININGU .....	14
1.5 STROJOVÉ UČENÍ (ML) A DIGITALIZACE.....	15
<b>2 UMĚLÁ INTELIGENCE (AI)</b> .....	<b>17</b>
2.1 ALAN TURING A TURINGŮV TEST .....	17
2.2 BUSINESS INTELLIGENCE (BI) .....	17
2.3 UMĚLÁ INTELIGENCE PŘI VZDĚLÁVÁNÍ .....	17
2.4 UMĚLÁ INTELIGENCE VE FINANČNICTVÍ.....	18
2.5 UMĚLÁ INTELIGENCE V MEDICÍNĚ.....	18
2.6 UMĚLÁ INTELIGENCE V ZEMĚDĚLSTVÍ.....	18
2.7 VÝHODY UMĚLÉ INTELIGENCE .....	19
2.7.1 Efektivita .....	19
2.7.2 Snížení lidského faktoru.....	19
2.7.3 Předpovědi.....	20
<b>3 ÚVOD DO NEURONOVÉ SÍTĚ</b> .....	<b>21</b>
3.1 NEURON .....	21
3.2 NEURONOVÁ SÍŤ V POČÍTAČOVÝCH VĚDÁCH.....	21
3.3 DŮLEŽITOST DAT.....	22
3.4 SOUBOR DAT .....	22
3.5 ZÁKLADNÍ DRUHY NEURONOVÝCH SÍTÍ .....	22
<b>4 NÁSTROJE DATA MININGU</b> .....	<b>24</b>
4.1 MONKEYLEARN .....	24
4.2 INTELLIGENT MINER .....	24
4.3 RAPIDMINER.....	24
4.4 ORACLE DATA MINING .....	24
4.5 DARWIN .....	25
4.6 WEKA.....	25

4.7	H2O .....	25
4.8	IBM SPSS MODELER.....	25
4.9	PYTHON.....	26
4.10	ENTERPRISE MINER.....	26
<b>5</b>	<b>SOCIÁLNÍ SÍŤE .....</b>	<b>28</b>
5.1.1	Facebook .....	29
5.1.2	YouTube.....	29
5.2	MARKETING, PODNIKÁNÍ A SOCIÁLNÍ SÍŤE.....	29
5.3	VÝHODY A NEVÝHODY SOCIÁLNÍCH SÍTÍ .....	30
<b>II</b>	<b>PRAKTICKÁ ČÁST.....</b>	<b>31</b>
<b>6</b>	<b>ANALÝZA MOŽNOSTÍ VYUŽITÍ UMĚLÉ INTELIGENCE A DATA MININGU V KONTEXTU POTŘEB SOUČASNÉ EKONOMIKY A MODERNÍHO MANAGEMENTU .....</b>	<b>32</b>
6.1	MÍRA VYUŽÍVÁNÍ DIGITALIZACE VE SVĚTĚ.....	32
6.2	VYUŽÍVÁNÍ UMĚLÉ INTELIGENCE VE SVĚTĚ.....	33
6.2.1	Grafické porovnání velikosti trhu s umělou inteligencí a četnosti vyhledávání pojmu umělá inteligence na webu. ....	36
6.2.2	Big data .....	36
6.2.3	Trendy vývojem blízke k trhu s Big daty.....	37
6.3	VYUŽÍVÁNÍ DATA MININGU .....	38
6.3.1	Využití data miningu společnostmi.....	39
6.3.2	Data mining v jednotlivých oblastech.....	41
6.3.3	Nástroje data miningu z hlediska efektivnosti a účelnosti .....	45
6.3.4	Činnosti, které jsou předmětem data miningu.....	48
<b>7</b>	<b>ZHODNOCENÍ VÝSLEDKŮ, NÁSTROJŮ Z HLEDISKA JEJICH EFEKTIVNOSTI A ÚČELNOSTI A IDENTIFIKACE TRENDŮ V OBLASTI DATA MININGU A UMĚLÉ INTELIGENCE.....</b>	<b>51</b>
	<b>ZÁVĚR .....</b>	<b>55</b>
	<b>SEZNAM POUŽITÉ LITERATURY.....</b>	<b>57</b>
	<b>SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK.....</b>	<b>63</b>
	<b>SEZNAM OBRÁZKŮ .....</b>	<b>64</b>
	<b>SEZNAM TABULEK.....</b>	<b>65</b>



## ÚVOD

Práce se bude zabývat především digitalizací, umělou inteligencí a data miningem. Veškeré tyto pojmy jsou dnes již poměrně známé a využitelné napříč všemi odvětvími.

Digitalizace je nyní velmi rozšířená především kvůli pandemii COVID-19, která mnoho podnikům znemožnila fungovat bez elektronické podoby. S touto digitalizací se momentálně rozmohl i trend využívání umělé inteligence. Zde se budeme především zabývat data miningem a jeho působením v ekonomické sféře a managementu. Ukážeme si využití data miningu v oborech, v odvětvích, jeho důležitost a výhody. Tento trend se dá dobře propojit s marketingem a finančním odvětvím celkově.

Trend data mining se samozřejmě užívá tam, kde je velké množství dat. Dnes v internetovém světě se dá právě tato těžba dat využít téměř ke každodenním problémům, kdy musíme vytáhnout nějaká data z komplexního balíčku, který pro nás nemá jako celek žádnou hodnotu. Zde se budeme také bavit o fungování a popularitě sociálních sítích, které jsou pro data mining jako dělané, je zde velké množství potenciálně užitečných dat pro společnosti, které něco nabízí. Tato popularita s sebou nese i nevýhody, a to poměrně významné. Může se jednat například o podvody nebo hacknuté účty. Na sociálních sítích i mimo ně ve většině segmentech na internetu probíhá lehký monitoring, který do sebe nabírá stále větší množství dat, která chtějí společnosti miningovat a použít. Tato data, která pochází z internetu totiž společnosti mohou využít na nabízení a „vnucování“ svých produktů přímo ke spotřebiteli.

Toto je dnes naprosto běžné a my si zde ukážeme, jak časté jsou tyto činnosti a co se k tomu využívá za softwary. Také si uvedeme momentální trendy v data miningu. Ohledně umělé inteligence si ukážeme posuny i potenciál, který tento fenomén nabízí.

Využijeme dostupných dat k určení a identifikaci jednotlivých problematik data miningu, uvedeme si společnosti, která data mining zavedli a také důvody společností, které se tomuto trendu vyhýbají. Tato data poté zhodnotíme, vyvodíme výsledky a závěrečná doporučení společností a managementu pro případné užití data miningu a oblastí, ve kterých se bez data miningu dnes již neobejdete.

## **CÍLE A METODY ZPRACOVÁNÍ**

Hlavním cílem práce je identifikovat trendy v data miningu a návazně vyhodnotit možné nástroje z hlediska efektivnosti a účelnosti. Cíl je splněn skrze analýzu a konzultací s expertem na danou problematiku.

### **METODY**

#### **Literární rešerše**

Literární rešerši jsme použili v teoretické části. Na základě odborné literatury je nastíněna problematika, jenž bude sloužit jako podklad pro tvorbu praktického celku.

#### **Sběr a analýza dat**

Teoretický výzkum vedený v této práci je tvořen díky sběru a analýze dat. Informace, tvořící teoretickou část jsou především z odborných článků, dostupných na science direct. Dále bylo použito několik zdrojů odborné literatury na dané téma. Podklady pro praktický celek jsou brány z odborné literatury, článků a internetových zdrojů, jenž jsou veřejně dostupné. Analýza probíhá na základě těchto dat, které se vyhodnocují a vedou ke splnění cíle práce.

#### **Metoda indukce**

Shromažďují se empirické poznatky a z nich se vyvozují obecné i konkrétní závěry o stavu dané reality a formulují se zákonitosti jejího pohybu. Tato metoda byla použita v teoretické i praktické části, kde jsme na základě přijímání informací z teorie, byli schopni provést daný výzkum a návazné vyhodnocení výsledků.

#### **Expertní metody**

Rozhovory proběhly s expertem v oblasti data miningu a umělé inteligence. Expertní metoda byla vedena napříč celou prací. Nejvíce jsme se zaměřili na praktickou část, kde jsme konzultovali především vztah práce s realitou. Zmíněná metoda je dle autora užitečná a velmi přínosná pro náš výzkum.

## **I. TEORETICKÁ ČÁST**

# 1 VYMEZENÍ ZÁKLADNÍCH POJMŮ

## 1.1 Software, Hardware a Big data

Software můžeme definovat jako systémové vybavení počítače, tedy to, co nelze nahmatat. Hmatatelné vybavení počítače je naopak hardware a jedná se především o grafické karty, procesory, základové desky, skříně, zdroje, disky a paměti RAM. Software jsou programy, které jsou již díky hardwaru spustitelné. Jedná se o systém například v České republice velmi známý Microsoft Windows a dále také programy, které na tomto systému můžeme používat. Může se jednat například o hry nebo profesionální programové vybavení pro střihačskou práci nebo grafickou činnost (Braden, 2021).

Big data, nebo česky Velká data byla popsána jako revoluce. Mayer-Schönberger a Cukier definovali Big Data jako nový možný zdroj pro objevy a inovace. Big Data jsou jedním z nejžhavějších a nejrychleji rostoucích trendů v technologii a předpokládá se, že budou mít dramatický dopad na ekonomii, vědu a společnost jako celek. Big Data se dají charakterizovat jako obrovské množství dat, které je dnešními běžnými softwary v reálném čase nezpracovatelné (Mayer- Schönberger, Cukier, 2014).

Big data slouží k rychlejšímu a efektivnějšímu dosažení výsledků. Na světě je stále více a více firem, které tato data využívají. Dají se využít v téměř každém odvětví a různých oborech, přičemž například v ekonomii, kterou se zde hodláme ve velkém zabývat, jsou už společnosti poměrně pokročilé, a to především na trhu s nebankovními půjčkami, kde se využívají například pro rozpoznání potenciálních dlužníků (Mayer- Schönberger, Cukier, 2014).

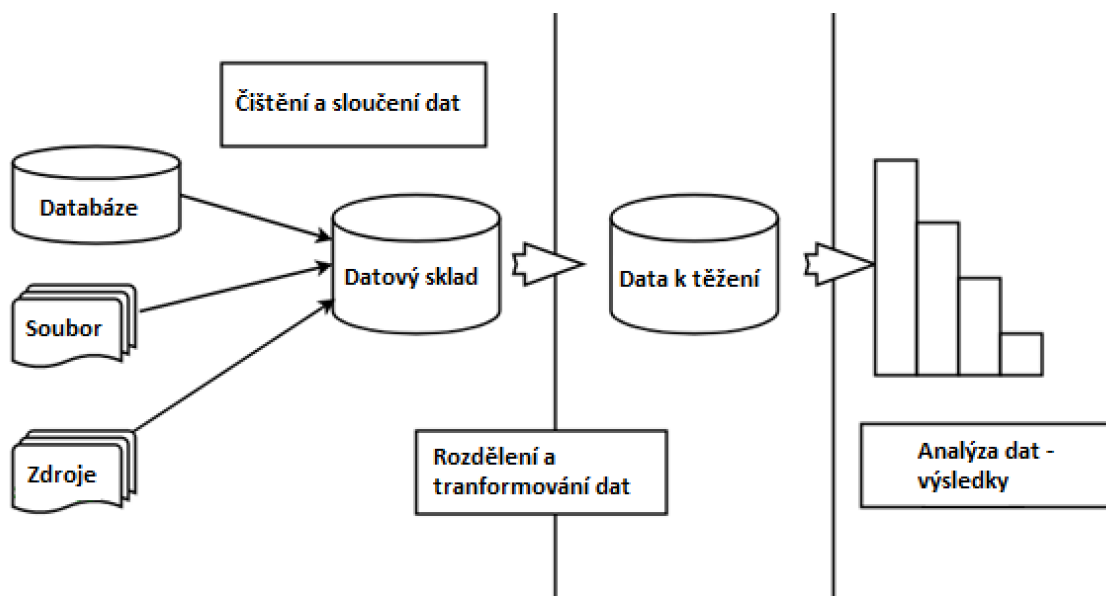
### 1.1.1 Velikost Big dat

Společnost IBM odhadovala, že v roce 2020 bude každý den vznikat 2,8 kvintilionů bajtů nových dat a na discích by mohlo být uloženo kolem 40 zettabajtů dat (pro přiblížení, 1 zettabajt =  $10^{21}$  bajtů, tedy obrovské množství dat). Tento odhad se časem potvrdil. Nejčastějším zdrojem Big dat jsou v dnešním světě sociální sítě například YouTube nebo Twitter (Holubová, Kosek, 2015).

## 1.2 Algoritmus a Data mining

V matematice a informatice je algoritmus konečná posloupnost dobře a kvalitně definovaných instrukcí, které se často používají k řešení třídy specifických problémů nebo k provedení výpočtu (Brinton, 2018).

Data mining je definován jako proces objevování vzorců v datech. Proces bývá automatický nebo (obvykleji) poloautomatický. Vzory, které jsou objevené musí být smysluplné v tom, že vedou k nějaké výhodě – např ekonomická výhoda, z těchto výhod se následně dají tvořit modely. Data jsou vždy dostupná ve značném množství. Užitečné vzory nám umožňují dělat netriviální a potenciálně užitečné předpovědi nových dat. Existují dva extrémy pro vyjádření vzoru: jako černá skříňka, jejíž vnitřnosti jsou nepochopitelné a dále jako průhledná krabice, jejíž konstrukce odhaluje strukturu vzoru. Obě tyto varianty tvoří pro uživatele data mining softwaru užitečné předpovědi. Rozdíl je v tom, zda vzory, jenž jsou těženy, jsou reprezentovány z hlediska struktury, kterou lze zkoumat, odůvodňovat, a která slouží k informování o budoucích rozhodnutích. Takové vzory nazýváme strukturální, protože zachycují rozhodovací strukturu explicitním způsobem. Jinými slovy, pomáhají vysvětlit něco o datech. Data mining se využívá nejen k děláni předpovědí, ale i kvalitnímu získávání informací. Techniky a metody, které se používají při data miningu a obecně pro zdokonalování a učení se v této oblasti se vyvinuli v takzvané „Strojové učení“ (Witten, 2017).



Obrázek 1 – Data mining a výsledek (vlastní zpracování dle Data Mining – Time-Series, Symbolic and Biological Sequences Data, 2022)

Na obrázku číslo 1 můžeme vidět, že základ pro data mining je velký počet dat, která jsou pro tuto činnost připravena. Tato data jsou shromážděna z různých zdrojů a databází, které úzce souvisí s tím, co chceme mít jako výstup. Vlastně vytěžíme data a zanalyzujeme je tak, abychom měli pohromadě ty opravdu důležité a ty, které pro nás mohou být pouze okrajové.

Data mining je rozdělen do jednotlivých činností, které si zde uvedeme (HAN, 2012):

1. Čištění dat
2. Integrace dat (lze použít více zdrojů dat)
3. Selektce dat
4. Datová transformace
5. Data mining neboli těžba dat
6. Vyhodnocení výsledků
7. Prezentace výsledků

### **1.3 Historie Data miningu**

V roce 1991 byla na svět uvedena první všeobecně uznávaná definice data miningu, kterou napsal William J. Frawley: „Data mining je netriviální získávání předtím neznámé a potenciálně užitečné informace ukryté v datech.“ Data mining byl potřebný díky rychle se rozšiřujícím informačním a technickým zařízením a službám, které měly za následek obrovský nárůst dat. Tato data byla potřeba zpracovat na získání užitečných informací, které mají za následek zvýšení efektivity a zároveň cílenost činností, přímo na danou problematiku. Z tohoto důvodu vznikl data mining, který data analyzoval a roztrídil (Úvod do data miningu, 2013).

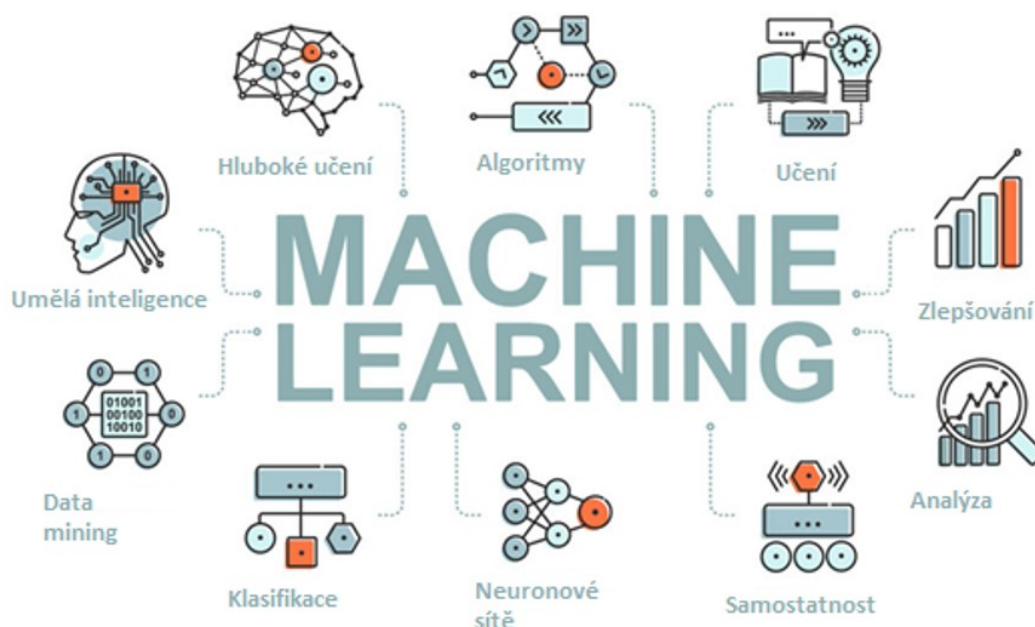
### **1.4 Využití data miningu**

Těžba dat se využívá ve všech možných odvětvích. Velmi důležitou roli hraje například v marketingu, kde si můžeme všimnout, že například reklamy na internetu či sociálních sítích jsou pro nás jako stvořené a pasují pro naši poptávku a potřebu. Zde už data mining proběhl, a to je ten důvod, proč jsou tyto nabídky „šité na míru.“ Jednoduše se sbírají data z toho, co daný potenciální zákazník hledá a shání a na základě toho se mu doporučí daný produkt či služba. Další velmi rozšířené využívání těžby dat je v oblasti návštěvnosti internetových stránek a webů, kde majitelé mohou například zjistit, jaký segment lidí je

sdílený s jinými stránkami. Dále se výrazně používá ve finančnictví, momentálně i v moderní medicíně a třeba i v telekomunikačních podnicích (Twin, 2021).

## 1.5 Strojové učení (ML) a digitalizace

Strojové učení neboli Machine learning je věda o tom, jak přimět počítače, aby fungovaly, aniž by byly explicitně naprogramovány. V moderní době nám strojové učení (ML) dalo například samořídící auta, praktické rozpoznávání řeči nebo efektivní vyhledávání na webu. Strojové učení (ML) je dnes už tak všudypřítomné, že jej pravděpodobně používáme i my několikrát denně, aniž bychom o tom věděli. Mnoho výzkumníků si myslí, že se jedná o nejlepší způsob, jakým může společnost dosáhnout pokroku směrem k umělé inteligenci na lidské úrovni. V této lekci se dozvíme o nejučinnějších technikách strojového učení a získáme praxi při jejich implementaci a přimějeme je, aby fungovaly pro nás. Ještě důležitější je, že se dozvíme nejen o teoretických základech učení, ale také získáme praktické know-how potřebné k rychlému a účinnému použití těchto technik na nové problémy. Nakonec se dozvíme o některých osvědčených postupech Silicon Valley v oblasti inovací, které se týkají strojového učení a umělé inteligence (Ng, 2022).



Obrázek 2 – Machine Learning (vlastní zpracování dle EURIX develops Machine Learning models, 2022)

V obrázku číslo 2 můžeme vidět počítačové vědy a segmenty, které dnes spadají do pojmu ML tedy strojového učení. Je zde mnoho pojmů, kterými se budeme zabývat. Především se

jedná o data mining, umělou inteligenci, zlepšování a analýzu. Definujeme si neuronové sítě a popíšeme si, jak fungují. Nicméně naše hlavní oblast ze strojového učení bude opravdu již zmíněný a definovaný data mining.

Digitalizace spadá pod disciplínu strojového učení. Digitalizace obecně je proces, při kterém se tvůrce snaží převést dostupná data z analogové formy na digitální formu, kterou poté dokáže přečíst počítač (Oxford University, 2022).



## 2 UMĚLÁ INTELIGENCE (AI)

Umělá inteligence (AI), také známá jako strojová inteligence, je odvětví počítačové vědy, které se zaměřuje na budování a řízení technologií, které se mohou naučit samostatně rozhodovat a provádět akce podobné, jako lidské bytosti (Hološka, 2012).

Zájem o digitalizaci a umělou inteligenci neustále roste (Brejčák, 2022).

### 2.1 Alan Turing a Turingův test

Alan Turing byl britský matematik, logik a informatik. Zabýval se algoritmy a je zakladatelem dnešní moderní informatiky jako vědy. Jako profesor se zasloužil o vynalezení šifrovacího stroje Enigma, který se používal během 2. světové války (Hodges, 2018).

Po zanechání obecné filozofické otázky, co je to vlastně umělá inteligence nebo entita se Alan Turing rozhodl vytvořit test pro umělou inteligenci. Tento test pojmenoval Turingův test, anglicky Turing test. Tento test spočíval v poznání umělé inteligence. Probíhal tak, že byly 2 místnosti. V jedné místnosti byl testovací objekt, tudíž nějaký člověk, který měl poznat komunikaci. Ve druhé místnosti byl člověk a počítač se softwarem, který měl simulovat lidské chování. Počítač se střídal s člověkem v různé komunikaci s testovacím objektem, jednalo se o psanou formu, protože počítač neuměl simulovat řeč. Testovací objekt měl za úkol poznat, jestli se jedná o počítač nebo člověka. Pokud počítač nepoznal a myslel si, že komunikuje s opravdovým člověkem, tento test byl brán jako úspěšný a tento software byl považován za umělou inteligenci (Neapolitan, Jiang, 2018).

### 2.2 Business intelligence (BI)

Termín business intelligence obsahuje znalosti týkající se procesů, technologií, aplikací a postupů, které zjednodušují rozhodování managementu podniku. Tato inteligence je založena na práci s již dostupnými daty, které se zpracují a následně fungují jako podklad pro rozhodování společnosti. Výstupy se dají použít pro SWOT a PEST analýzu. Obecně se BI používá především pro interní informace (Laberge, 2012).

### 2.3 Umělá inteligence při vzdělávání

Vzhledem k reakci na koronavirus je zcela na místě přidat i toto téma. Při učení dnes máme obrovské množství nových možností. Většina programů, které nás mohou vyučovat už samozřejmě umělou inteligenci nabízí a není to pouze CD, které někdo namluvil. Dnes můžeme komunikovat s interaktivním krokodýlem nebo s panem Donaldem v různých

programech a procvičovat si například svoji angličtinu. Tento segment zaznamenal velké množství zlepšení za poslední 2 roky. Dnes už je svět připraven na virtuální učení samozřejmě v rámci možností a s tím spojenou výukou (Kisk, 2020).

## **2.4 Umělá inteligence ve finančnictví**

Očekává se, že nasazení umělé inteligence ve financích bude stále více podporovat konkurenční výhody finanční firmy, a to dvěma hlavními způsoby: a) zlepšením efektivity firem snížením nákladů a zvýšením produktivity, což vede k vyšší ziskovosti (např. lepší rozhodování, zisky ze zlepšení řízení rizik, back-office a další optimalizace procesů); a b) zlepšením kvality služeb a produktů nabízené spotřebitelům (např. nabídka nových produktů, vysoké přizpůsobení produktů a služby). Taková konkurenční výhoda může být také přínosem i pro spotřebitele, a to buď zvýšením kvality produktů, rozmanitost možností a personalizací nebo snížením jejich nákladů (Artificial Intelligence, Machine Learning and Big Data in Finance, 2021).

Techniky umělé inteligence (AI) jsou stále častěji nasazovány ve finančních oblastech, jako je správa aktiv, algoritmické obchodování, upisování úvěrů nebo financování založené na blockchainu, což umožňuje množství dostupných dat a dostupná výpočetní kapacita. Modely strojového učení využívají velká data ke zlepšování, předvídatelnosti a výkonu automaticky prostřednictvím zkušeností a dat, aniž by k tomu byly naprogramovány lidmi (Artificial Intelligence, Machine Learning and Big Data in Finance, 2021).

## **2.5 Umělá inteligence v medicíně**

AI prokázala významný potenciální přínos pro lékaře a pacienty. Tyto přínosy transformují terapeutický vztah z tradiční dyády lékař–pacient na triadický vztah lékař–pacient–stroj. Nové technologie umělé inteligence však vyžadují pečlivé prověření, právní normy, záruky pacientů a vzdělávání poskytovatelů. Lékaři by měli rozpoznat limity a rizika AI a také její potenciální přínosy (Vearrier, 2022).

## **2.6 Umělá inteligence v zemědělství**

Umělá inteligence v kombinaci s autonomními traktory dokáže vyřešit jeden z nejčastějších problémů v zemědělství: nedostatek pracovních sil. Tyto technologie jsou také nákladově méně zatěžující, protože jsou přesnější, a tím snižují chyby. Dohromady jsou AI a autonomní

traktory klíčem k preciznímu zemědělství (Artificial Intelligence in Agriculture: Rooting Out the Seed of Doubt, 2022).

Další méně běžnou, ale rychle rostoucí technologií je robotika. Zemědělství roboti se již používají pro manuální práce, jako je sběr ovoce a zeleniny nebo třeba prořezávání salátu. Výhody robotů oproti farmářům jsou značné. Mohou pracovat déle, jsou přesnější a jsou méně náchylné k chybám (Artificial Intelligence in Agriculture: Rooting Out the Seed of Doubt, 2022).

Umělá inteligence v zemědělství opravdu velmi rychle mění směr, kterým se toto odvětví udává. Tyto technologie jsou nejen v zemědělství ale i v přepravě. Traktory a nákladní automobily, které jsou moderní, obrovským způsobem ulehčují řidiči práci. Pokud nebereme přímo autonomní traktory, tak i ty běžné jsou téměř totožné, a i když v nich můžeme zahlédnout řidiče, tak by traktor byl schopný jezdit sám. Tato schopnost je pro hospodářství jako celek poměrně revoluční a pro zemědělství možná nejvýznamnější za poslední dobu (Artificial Intelligence in Agriculture: Rooting Out the Seed of Doubt, 2022).

## **2.7 Výhody umělé inteligence**

### **2.7.1 Efektivita**

Jednou z největších výhod umělé inteligence je, že umožňuje lidem být efektivnější. Umělou inteligenci lze využít k rychlejšímu provádění malých, opakujících se úkolů, nebo ji lze použít k dokončení mnohem větších a složitějších úkolů. Bez ohledu na její použití jsou systémy umělé inteligence nespoutané lidskými omezeními a nikdy se neunaví (Callahan, 2020).

### **2.7.2 Snížení lidského faktoru**

Při vysoce automatizovaných činnostech, je zde snížena lidská chybovost. Umělá inteligence je dnes precizní a velmi přesná, dále je také neúnavná, tudíž dokáže pracovat 24 hodin 7 dní v týdnu. AI je potřeba kvalitně naprogramovat, aby vše fungovalo přesně tak jak má, tudíž lidský faktor je při plně automatizovaném procesu pouze již zmíněné naprogramování a určitá kontrola (Duggal, 2022).

### 2.7.3 Předpovědi

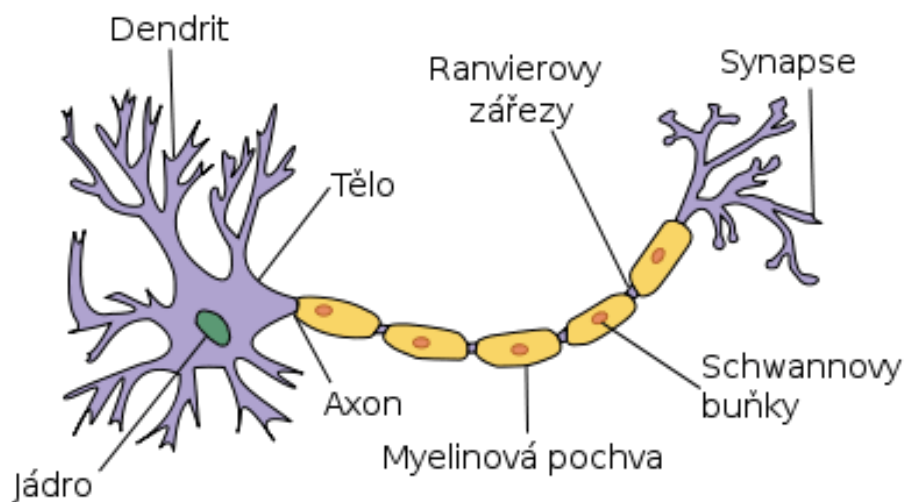
Velmi výrazná pomoc umělé inteligence je také predikování. Předpověď počasí umí umělá inteligence výborně. Dnes již je tato předpověď velice přesná i na několik dní dopředu, díky podobnostem v počasí a vyvozování následků (Vyas, 2021).

### 3 ÚVOD DO NEURONOVÉ SÍTĚ

#### 3.1 Neuron

V průměrném lidském mozku je odhadem osmdesát šest miliard neuronů. Tyto neurony jsou navzájem propojený počet strukturních prvků, které do nich vstupují. Tyto prvky nazýváme dendrity a axony, dále je zde centrální jednotka zvaná soma, kde se informace zpracovávají. Typický neuron má náhodný počet dendritů a axonů v závislosti na roli, kterou hraje v konkrétní funkci (Leterrier, 2016).

Neurony posílají příchozí signály, které do nich vstupují, přes dendrity do soma a to vyvolá výstupní signál na axonu. Tyto vstupní signály jsou považovány za aktivace. Axon jednoho neuronu se větví ven a je spojen s dendrity mnoha dalších neuronů. Místo, kde se vyskytuje dendrit a větev axonu, se kterou se setkává, se nazývá synapse, kterou neuron přenáší signály z jednoho do druhého. Odhaduje se, že v jednom je sto miliard synapsí průměrný lidský mozek a nové synapse se vytvářejí a lámou každou sekundu (Leterrier, 2016).



Obrázek 3 – Neuron (*Anatomie neuronu – odkud přichází a kam směřuje nervový vzruch, 2015*)

V obrázku číslo 4 můžeme vidět neuron, kde se nachází axon i dendrit a následně je zde jádro, které charakterizuje centrální jednotku soma.

#### 3.2 Neuronová síť v počítačových vědách

Neuronová síť obecně je propojení více neuronů v jeden celek. Síť se v čase vyvíjí. Zde se budeme zajímat především o umělé typy těchto sítí, protože fungují jako metoda pro data

mining. Neuronové sítě, zastávají postavení univerzálního matematicko-informatického přístupu k učení a modelování procesů. Jedná se o adaptace umělých systémů, které jsou založené na fungování lidského mozku. Strukturu těchto neuronových sítí není možné projít detailně, proto je často nazývána jako „černá skříňka“ (Dostál, 2015).

### 3.3 Důležitost dat

Data jsou souborem kvalitativních nebo kvantitativních proměnných. Je to entita, která nese informace týkající se konkrétní události. Je to skutečná reprezentace skutečného světa události v zakódovaném formátu pro univerzální reprodukovatelnost. Důležitost dat závisí na informacích, které data obsahují. Může to být sbírka čísel, abecedy, obrázků, symbolů a zvuků. Elektronická reprezentace dat je prováděna binárním kódem, což je sekvence 0 a 1 v překódovaném formátu, se kterým lze manipulovat technikami elektronického zpracování. Surová data jsou nejdostupnějším formátem, který je směsí různých hodnot a faktorů. Nezpracovaná data jsou také známá jako univerzální sada. Cenné informace se získávají z nezpracovaných dat použitím technik předběžného zpracování/filtrování dat. Neuronové sítě zcela spoléhají na data během trénovací fáze, kde sítě extrahují vlastnosti, formulují vztahy a přizpůsobují se určitým úkolům. Použitím dat poskytovaných do sítě se upravují hodnoty, vychýlení a související parametry. Poskytování dat ve správném formátu a správném množství je pro nás důležitým úkolem trénování neuronových sítí (DO, 2021).

### 3.4 Soubor dat

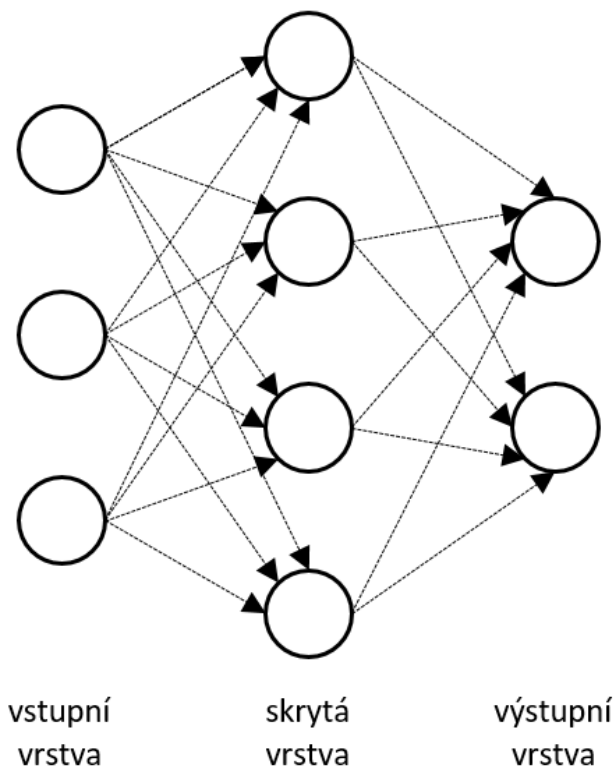
Soubor dat je organizovaná a uspořádaná struktura. Tato data jsou řazena podle určité podobnosti. Tato podobnost je zastřešuje a dává právě do jednoho souboru, který nese daný informační název hlavního tématu, jenž mají data společný. Obecně soubor vyjadřuje nějakou relevantnost k danému tématu – například lékařské spisy, soudní spisy nebo podklady pro učení v daném tématu. Soubory dat jsou data již vyčištěná, a tudíž mohou sloužit jako podklad pro nějakou analýzu nebo modelování (What Is a Data Set? (With Definition, Components and Types), 2022).

### 3.5 Základní druhy neuronových sítí

Neuronové sítě jsou rozděleny do druhů podle jejich složení a uspořádání. Autor rozděluje neuronové sítě do těchto 5 základních druhů.

Mezi základní druhy patří (DO, 2021):

- Feedforward neuronová síť
- Radial Basis Function neuronová síť
- Kohonen Self Organizing neuronová síť
- Recurrent neuronová síť
- Convolutional neuronová síť



Obrázek 4 – Příklad neuronové sítě (Durčák, 2017)

Zde můžeme vidět příklad neuronové sítě, kde vstupní vrstva pouze zopakuje a převede hodnoty vstupů do neuronů, tudíž do skryté vrstvy, kterou vidíme uprostřed. Výstupy neuronů jsou zároveň vstupem pro dva neurony do vstupní vrstvy. Tyto hodnoty výstupů z neuronů poté kódují výstup celé neuronové sítě (Durčák, 2017).

## 4 NÁSTROJE DATA MININGU

V této kapitole si ukážeme a popíšeme deset momentálně nejčastěji využívaných nástrojů, tudíž softwarů pro data mining.

### 4.1 MonkeyLearn

Jedná se o software pro data mining. V uživatelsky přívětivém prostředí můžete snadno integrovat MonkeyLearn s vašimi stávajícími nástroji a provádět data mining v reálném čase (Wolff, 2020).

Nástroje pro těžbu dat MonkeyLearn se již používají k automatizaci vyvození výsledků, nasměrování zákazníků přímo k podpoře, automatickému zjišťování negativní zpětné vazby na sociálních sítích a poskytování podrobných informací, které vedou k lepšímu rozhodování (Wolff, 2020).

### 4.2 Intelligent Miner

Intelligent Miner je integrovaný software řešící sofistikovanou analýzu dat. Umožňuje nám získávat cenné informace z databází a prezentovat je obchodním uživatelům k analýze a rozhodování. Tento software je určen především pro ekonomické, obchodní a marketingové činnosti. Pomůže k vytvoření ekonomických nebo prodejních modelů (Intelligent Miner Visualizer, 2021).

### 4.3 RapidMiner

Rapidminer je nástroj pro data mining s vizualizací pracovních postupů a plnou automatizací. To znamená, že nemusíme dělat ruční kódování pro jednotlivé úlohy data miningu. Rapidminer je jedním z nejpopulárnějších nástrojů dnešní doby (Arnaldo, 2021).

Jedná se tedy o grafické uživatelské rozhraní procesů v Rapidmineru. Má úložiště, které obsahuje naše datové soubory. Také můžeme importovat vlastní soubory. Dokonce nabízí i mnoho veřejných souborů, které můžeme vyzkoušet. Můžeme pracovat i s databázovým připojením (Arnaldo, 2021).

### 4.4 Oracle data mining

Díky Oracle data mining softwaru můžete vytvářet a používat trendové a prediktivní modely, které spadají do Oracle databáze. Tyto modely nám pomohou predikovat chování zákazníků



na trhu, cílit na naše nejlepší zákazníky, vytvářet profily zákazníků, identifikovat příležitosti křížového prodeje a odhalovat anomálie a potenciální podvody. Jedná se o silně ekonomicky a marketingově orientovaný data miningový program (Oracle Data Miner, 2022).

#### **4.5 Darwin**

Darwin je software pro těžbu dat, který je nyní součástí Oracle. Využívá různé techniky dolování dat, jako je klasifikace, neuronové sítě, genetické algoritmy a regrese. Vyplývá to z architektury klient-server, přičemž klient může pracovat na platformě Windows i Unix. Je schopen získávat data z široké škály zdrojů (Palus, 2014).

#### **4.6 WEKA**

Jedná se spíše o základní open source software pro data mining. Využívá se pro výuku spojenou s data mining a také zaučením, tedy seznámením s tím, jak programy na těžbu obecně fungují. Weka je velmi jednoduchý na ovládání s velmi jednoduchým grafickým rozhraním, jenž uživateli nenabízí miliony funkcí, ale pouze základní. Ovšem kromě tohoto také obsahuje poměrně mnoho implementovaných algoritmů a také nabízí supervised i unsupervised ML, což jsou základní techniky strojového učení (Chudán, 2020).

#### **4.7 H2O**

Je open source software, jehož cílem je zpřístupnit technologie vznikající společně s umělou inteligencí. Podporuje nejběžnější algoritmy strojového učení a nabízí funkci Auto ML, která uživatelům pomáhá vytvářet a využívat modely strojového učení rychlým a jednoduchým způsobem, i když nejsou odborníky. Jedná se tedy o jednodušší software (Wolff, 2020)

H2O lze propojit prostřednictvím API, které je k dispozici ve všech hlavních programovacích jazycích. Také využívá distribuované výpočty v paměti, díky čemuž je ideální při analýze velkých datových souborů (Wolff, 2020).

#### **4.8 IBM SPSS Modeler**

Velmi často využívaný program ve větších, menších a středních podnicích. Dobře se kombinuje s předními technologiemi pro implementaci robustního celopodnikového řešení. IBM SPSS Modeler je užitečný i pro vědu díky vizualizaci dat a strojového učení, které šetří čas potřebný k dosažení výsledků (Sarangam, 2020).

Tento software se používá pro přípravu dat, zjišťování, prediktivní analýzu, trendovou analýzu a vytváření modelů. Tento nástroj značně pomůže organizacím využívat jejich data a aplikace. Jednou z výhod tohoto softwaru je jeho schopnost splnit hlavní požadavky na správu a zabezpečení organizace na podnikové úrovni (Sarangam, 2020).

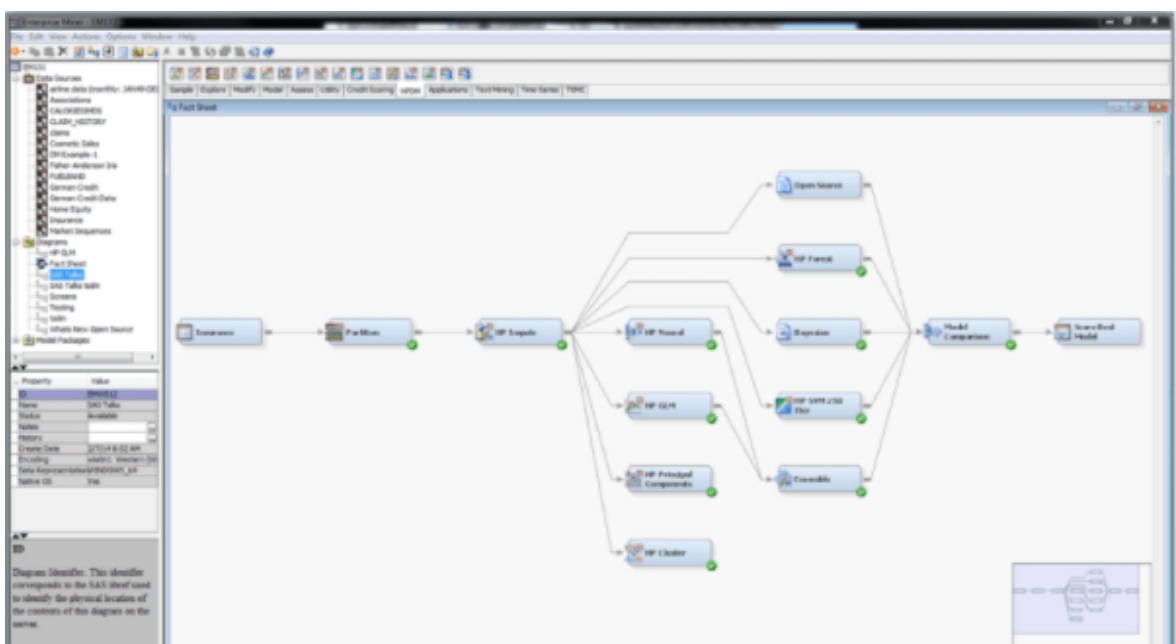
## 4.9 Python

Python je volně dostupný program využívající open-source jazyk. V kombinaci se schopností univerzálního jazyka a široké škály dostupných balíčků pro tento program, tvoří Python velmi flexibilní nástroj pro data mining. Firma si ho může takto upravit dle jejich potřeb přímo na „míru.“ Také je zde možná kombinace s modelováním (Sarangam, 2020).

S Pythonem nezískáme to nejlepší, co dnes trh s data mining softwarem nabízí, ale funkce je tu pro každého a můžeme si vytvořit vlastní prostředí dle libosti. Co také podporuje Python, je velká online dostupnost vývojářských balíčků, kteří zajišťují, že nabízené balíčky jsou robustní a bezpečné. Jednou z funkcí, kterými je Python v této oblasti známý, jsou výkonné funkce pro vizualizaci za běhu, které nabízí (Sarangam, 2020).

## 4.10 Enterprise Miner

SAS Enterprise Miner je pokročilý analytický nástroj pro data mining, který má uživatelům pomoci rychle vytvořit popisné a prediktivní modely prostřednictvím zjednodušeného procesu data miningu. Grafické rozhraní Enterprise Miner umožňuje uživatelům logicky



Obrázek 5 – Enterprise Miner (Enterprise Miner, 2022)

procházet pětikrokovým procesem SAS SEMMA: vzorkování, průzkum, modifikace, modelování a hodnocení. Uživatelé si vyberou danou kartu, jak můžeme vidět na obrázku, kde jsou jednotlivé karty s modrým pozadím, k té připevní určitou činnost a propojí jednotlivé karty ve finální proces. Jedná se tedy o mapu v data miningovém programu (Reifer, 2017).

## 5 SOCIÁLNÍ SÍŤE

Název sociální média označuje moderní počítačovou a internetovou technologii, která usnadňuje sdílení nápadů, pocitů, myšlenek a informací prostřednictvím virtuálních sítí a komunit. Sociální média jsou součástí internetu. Právě internet umožňuje uživatelům rychlou elektronickou komunikaci obsahu. Tímto obsahem mohou být osobní informace, dokumenty, videa anebo fotografie. Uživatelé sociálních sítí musí být připojeni na internetu, tudíž se mohou připojit pomocí počítače, tabletu nebo chytrého telefonu. Zatímco sociální média jsou všudypřítomná v Americe a Evropě, asijské země, jako je Indonésie a Čína, vedou seznam používání a zakázaných sociálních médií. K říjnu 2021 sociální média používalo 4,5 miliard lidí (Social Media: What Countries Use It Most & What Are They Using?, 2021).

Uveďme si 6 momentálně nejužívanějších a nejpopulárnějších sociálních sítí (Lua, 2022):

- Facebook
- YouTube
- WhatsApp
- Instagram
- Facebook Messenger
- TikTok

### 5.1.1 Facebook

Facebook založil Mark Zuckerberg jako student Harvardu v roce 2004. K roku 2021 bylo měsíčně aktivních 2,7 miliardy lidí, což jen poukazuje na to, kolik informací a dat se musí na facebooku během malého časového úseku nakupit. Toto obrovské množství dat je možné společnostmi těžit za účelem budoucí prosperity (Facebook - statistics & facts, 2022).

### 5.1.2 YouTube

YouTube je spíše než facebook mediální platforma, kde se jedná především o videa a takzvaná „Shorts,“ což jsou velmi krátká videa, která vyjadřují nějakou pointu. Je dobré říci, že pokud vlastníme účet Google, což je také obrovská platforma, kterou každý z nás denně využívá, tak se můžeme přes tento účet přihlásit i na YouTube. YouTube totiž patří pod Google a tato společnost se snaží všechno co nejvíce propojit. K roku 2021 má YouTube měsíčně dvě miliardy aktivních uživatelů (Za deset let se stal YouTube druhou nejnavštěvovanější stránkou, 2015, Ceci, 2022).

## 5.2 Marketing, podnikání a sociální sítě

Můžeme si dát příklad, že si vyrobíme nějakou stránku. Například by se mohlo jednat stránku pro mobilní telefony – poradna a popřípadě i e-shop. Pokud si takto založíme skupinu například na Facebooku, můžeme zde pozvat všechny lidi, které známe během pár sekund. Tito lidé to mezi sebou mohou sdílet a můžeme mít potenciálně velké množství zákazníků během relativně krátké doby. Nyní jsme si ukázali, jak bychom mohli sehnat lidi. Nyní ta zajímavější část, což je komunikace. Pokud tyto lidé tedy jsou členem naší skupiny, budou dostávat naše příspěvky, kde můžeme dávat nějaké slevy nebo akce a oni si toho rozhodně všimnou. Takovéto krátké a zajímavé reklamy mohou mít velmi kladný důsledek na využívání našich služeb a produktů. Další velmi klíčovou věcí je to, že je zde velmi lehký a rychlý přístup zákazníka k prodejci a naopak. Jelikož je zde chat a dá se psát, můžeme využít přímý kontakt zákazníka, který nám může sdělit, jaké jsou jeho požadavky, co by si tak představoval a jestli je možné to vůbec dostat. Tato komunikace může být velmi důležitá i později, pokud se něco s produktem stane anebo pro zpětnou vazbu, kde se můžeme přímo zákazníka zeptat, jestli byl spokojen nebo ne (Haye, 2021).

### 5.3 Výhody a nevýhody sociálních sítí

V dnešním světě je obrovský rozmach sociálních médií a je opravdu hodně možností, kam jít a co využívat. Tato média spolu ale nesou značná rizika. Nejčastější riziko v běžném světě je samozřejmě únik informací, které mohou být špatně použity a mohou se dostat do nesprávných rukou anebo klidně k celé veřejnosti. Je obrovské množství situací, kdy nějaká osoba například sdílela cizí fotky, které mohou například zesměšnit či jinak ublížit osobě, která je pořídila. Proto je velmi důležité se na sociálních sítích chovat dle pravidel. Poté co jsme zmínili stinné stránky sociálních sítí, zaměříme se i na dobré, kterých je mnohem více. Dříve jsme mohli pouze psát SMS nebo posílat MMS, ovšem pokud jsme je chtěli posílat mimo území České republiky, bylo to komplikované a finančně náročné. Skrz sociální sítě můžeme snadno sdílet média a informace během okamžiku, a to napříč celým světem, což je opravdu velká výhoda. Také si můžeme zavolat, a to pouze hlasově nebo i videohovorem.

## **II. PRAKTICKÁ ČÁST**

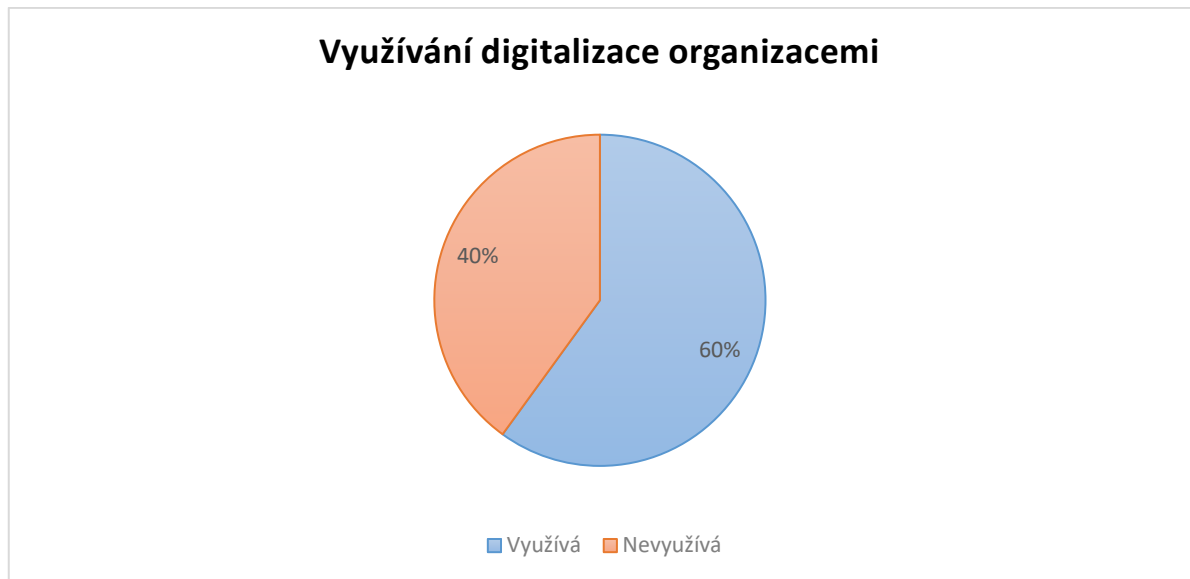
## **6 ANALÝZA MOŽNOSTÍ VYUŽITÍ UMĚLÉ INTELIGENCE A DATA MININGU V KONTEXTU POTŘEB SOUČASNÉ EKONOMIKY A MODERNÍHO MANAGEMENTU**

V této části se budeme zabývat působením digitalizace a umělé inteligence v oblasti spotřebitelů, trhu, a především v marketingu. Dále budeme sledovat společnosti, kde se umělá inteligence i data mining využívají hojně. Graficky si znázorníme procentuální podíly společností, které využívají digitalizaci a umělou inteligenci. Také nesmíme zapomenout na prezentaci vývoje popularity pojmu umělá inteligence na a růst Big data trhu a trhu s umělou inteligencí. Tyto trhy zaznamenaly rapidní nárůst a zaznamenávají jej stále každým rokem. Jedná se tedy o velmi významnou část trhu a pravděpodobně o budoucnost poskytování služeb a výroby produktů. Později v kapitole narazíme na data mining. Tento trend si zde rozvedeme a zhodnotíme jeho nástroje, jenž využívají již zmíněnou umělou inteligenci.

### **6.1 Míra využívání digitalizace ve světě**

Digitalizace je velmi rozsáhlé téma, především v dnešní době, kdy začíná převládat průmysl 4.0, který tento trend vysoce prosazuje a ukazuje její důležitost. Ukážeme si tedy, jaká část společností trend přijala. Data uvedená v grafu pochází od Blake Morgana, který je sesbíral a uvedl v roce 2019. Jsou podložena interními daty společností, které Morgan Blake získal a na základě nich udělal statistiky o digitalizační přeměně společností, včetně využití digitalizace společnostmi. Tato data pochází výhradně ze Spojeného království napříč obory, aby jejich hodnota byla co nejvíce vypovídající. Dle něj se tento trend zvyšuje a v následujících letech můžeme očekávat ještě vyšší podíl digitalizace mezi společnostmi (Blake, 2019).





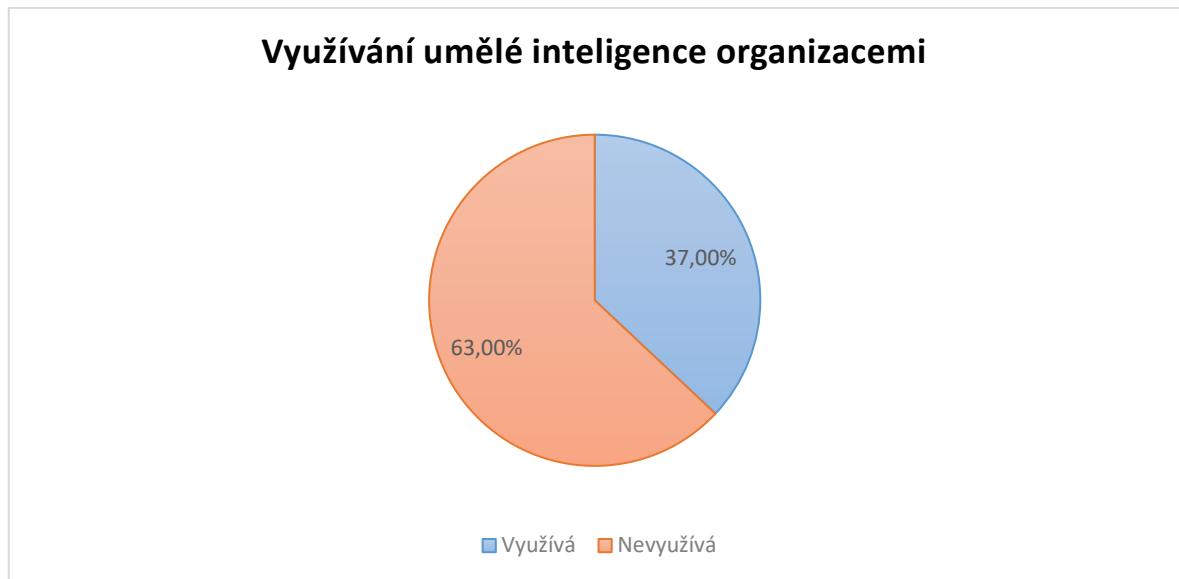
Obrázek 6 – Využívání digitalizace organizacemi (vlastní zpracování dle Blake, 2019)

Z obrázku číslo 6 je patrné, že většina firem zahrnutá v článku využívá větší či menší míru digitalizace. Digitalizace velmi usnadňuje fungování společnosti a často zvyšuje spokojenost zákazníka.

Průzkumy zmíněné Blakem také ukázaly, že spokojenost zákazníka při digitalizaci vzroste průměrně o 25 % a ekonomická situace podniku po digitalizaci se zvýší od 20 % do 50 %. Na základě těchto dat můžeme konstatovat, že digitalizace bude s největší pravděpodobností stále růst. Autor se domnívá, že by se mohlo v intervalu následujících deseti let jednat i o devadesáti procentní podíl digitalizovaných firem.

## 6.2 Využívání umělé inteligence ve světě

Obrázek číslo 7 zobrazuje graf využívání umělé inteligence světovými organizacemi. Data, která jsou použita v grafu, pocházejí z portálu dataprot.net, který je publikoval v roce 2022. Jedná se o globální průzkum na základě jejich statistik, ohledně využití umělé inteligence společnostmi. Tyto společnosti odpovídaly na dotazník spojený s umělou inteligencí včetně jejím užívání. Dále zde budou uvedena data, která zobrazují trend, tedy popularitu a známost studijních oborů, spojených s umělou inteligencí, které vytvořil Google v roce 2022. Jedná se o jejich průzkum, který monitorují díky data miningu přes internet. Tento graf bude zobrazovat rozpětí pěti let.

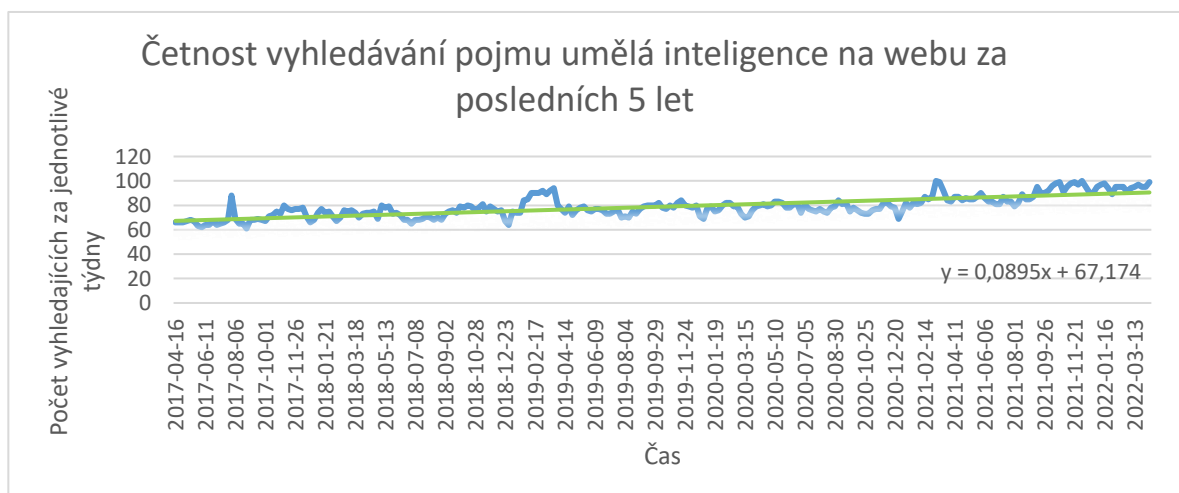


Obrázek 7 – Využívání umělé inteligence organizacemi (vlastní zpracování dle Jovanovic, 2022)

Vidíme, že celkově se ve světě umělá inteligence využívá v 37 % organizací, 63 % ji prozatím nevyužívá (z organizací zúčastněných průzkumů). Z těchto 37 % můžeme vyzdvihnout především pět gigantů, které ročně investují miliardy dolarů do umělé inteligence. Jsou to Amazon, Google, IBM, Microsoft a Aliba.

Umělá inteligence je součástí digitalizace, tudíž můžeme předpokládat, že 23 % digitalizovaných firem umělou inteligenci zatím nevyužívá a spadá pod těchto celkových 63 %. Data mining i umělá inteligence spadají do strojového učení. Oněch 37 % firem tudíž využívá strojové učení, které přispívá k vysoké míře automatizace.

Zobrazme si graf, ukazující četnost vyhledávání pojmu umělá inteligence na webu. Tento graf je vytvořen na základě dat z Google trends (Umělá inteligence, 2022).

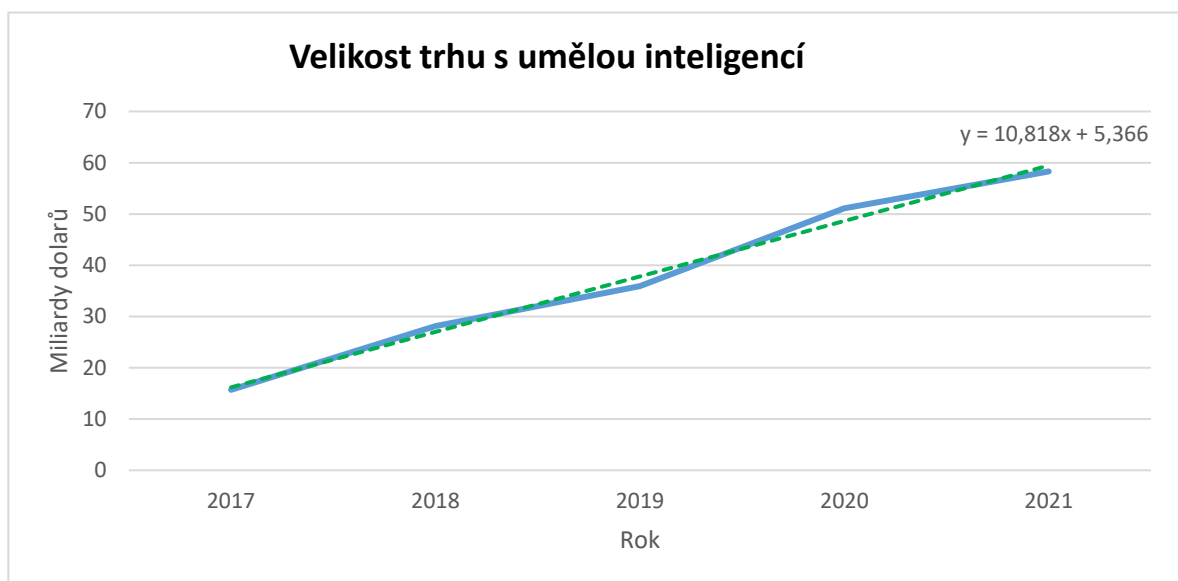


Obrázek 8 – Četnost vyhledávání pojmu umělá inteligence na webu za posledních 5 let (vlastní zpracování dle Umělá inteligence, 2022)

Trend za posledních pět let je stále mírně rostoucí. Umělá inteligence by měla mít velkou budoucnost, pokud se bude zájem o tento výraz dále zvětšovat. Graf ukazuje vývoj trendu vždy za jednotlivý týden. Uvedli jsme si i lineární rovnici trendu. Ta má následující tvar:

$Y = 0,0895x + 67,174$ . Jedná se tedy o mírnou vzestupnou tendenci

Nyní si ukážeme trend velikosti trhu s umělou inteligencí a poté porovnáme zmíněný graf s grafem předchozím, tedy s četností vyhledávání pojmu umělá inteligence. Dozvíme se, zda je v růstu určitá spojitost. Data, která jsou použita pro graf, vycházejí z dat společnosti IBM, Yahoo a Statista Research Department. Tato data představují velikost trhu každý rok, což je následně předěláno autorem do grafu. Hodnoty jsou v miliardách dolarů.

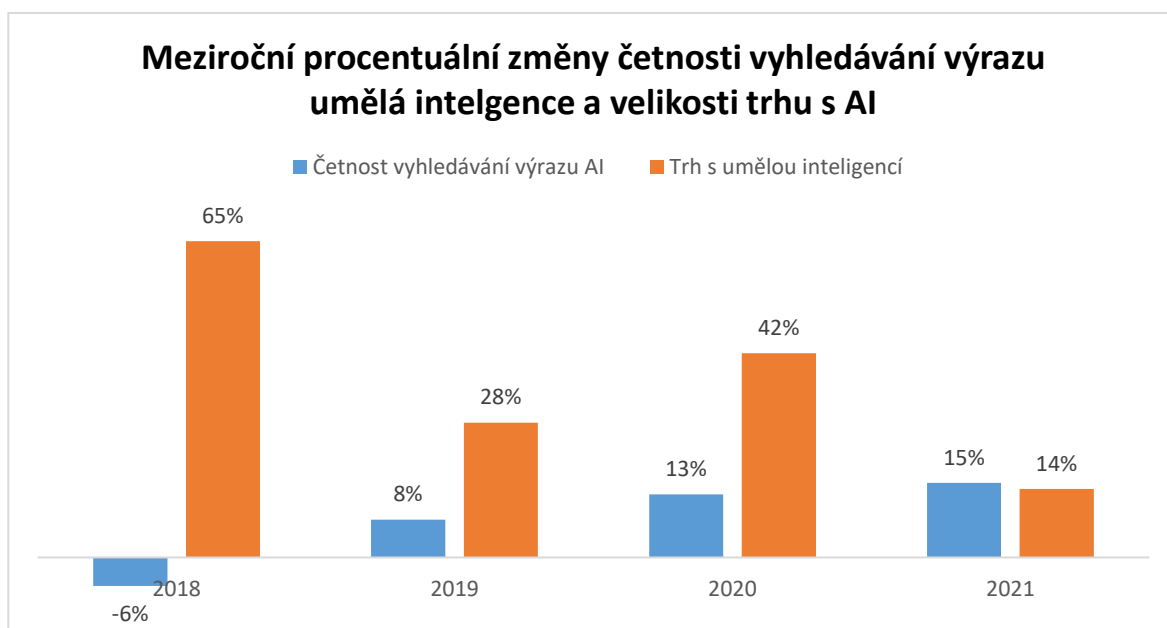


Obrázek 9 – Velikost trhu s umělou inteligencí (vlastní zpracování dle Revenues from the artificial intelligence (AI) software market worldwide from 2018 to 2025, 2022, Jyoti, 2018)

Na základě údajů si můžeme všimnout, že z hlediska trhu se jedná o obrovský nárůst a velmi rostoucí tendenci. Celková hodnota trhu vzrostla více než třikrát z původní hodnoty. Můžeme také vidět, že lineární rovnice grafu  $Y = 10,818x + 5,366$  má poměrně rychle rostoucí tendenci. Trh tak dle modelu roste o cca 10,8 mld. USD ročně.

### 6.2.1 Grafické porovnání velikosti trhu s umělou inteligencí a četnosti vyhledávání pojmu umělá inteligence na webu.

V této podkapitole si rozebereme vývoj trendů. Přesně se bude jednat o četnost vyhledávání pojmu umělá inteligence na webu, tedy popularitu výrazu v oblasti vyhledávání (oranžová barva v grafu) a velikosti trhu s umělou inteligencí (modrá barva v grafu).



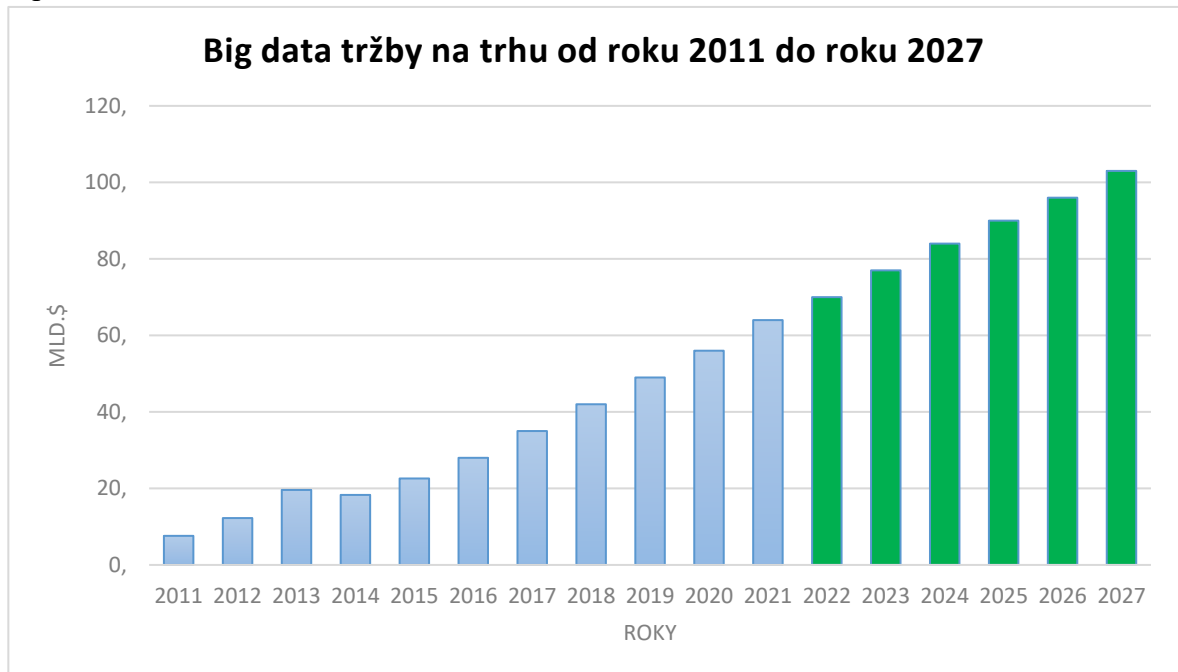
Obrázek 10 – Meziroční procentuální změny četnosti vyhledávání výrazu umělá inteligence a velikosti trhu s AI (vlastní zpracování dle *Revenues from the artificial intelligence (AI) software market worldwide from 2018 to 2025*. Jyoti, 2018 a *Umělá inteligence, 2022*)

Lze vyvodit, že se jedná v obou případech o rostoucí tendenci. V podání četnosti vyhledávání výrazu umělá inteligence se jedná o vcelku rostoucí trend. Z hodnoty -6 % se dostáváme na růst 15 % meziročně. Pokud se bavíme o velikosti trhu, zaznamenáváme zde rapidní nárůst především v prvním roce, kdy se jedná o růst 65 %. Nicméně z grafu lze také vyvodit, že rychlost růstu velikosti trhu s AI polevuje a blíží se k menším hodnotám.

### 6.2.2 Big data

V souvislosti s data miningem a umělou inteligencí, přes kterou funguje, je zde další velmi důležitý faktor. Big data, jak již víme, jsou data, která bývají velmi často těžena pro různé účely. Nyní si ukážeme vývoj a predikci tržeb na big data trhu pomocí sloupcového grafu. Data použitá na grafu byla publikována v roce 2021 společností Statista Research Department. Zahrnuje celosvětové tržby na trhu s big daty již od roku 2011. Od roku 2021

se jedná o pouhou predikci, kdy na základě funkce, kterou data vykazují, dochází k predikovanému rostoucímu efektu.



Obrázek 11 – Big data tržby na trhu od roku 2011 do roku 2027 (vlastní zpracování dle Big data market size revenue forecast worldwide from 2011 to 2027, 2022)

Jak můžeme vidět, jedná se opravdu o rapidní a každoroční nárůst s téměř stejnou rostoucí funkcí. Modře vybarvená data jsou do roku 2021 a tato data jsou již potvrzená. Zeleně zbarvená data jsou pouhou predikcí. Pokud budeme brát v potaz tento růst a chtěli bychom zjistit funkci, nejbližší je tomu lineární funkce, která by v tomto grafu vypadala následovně:

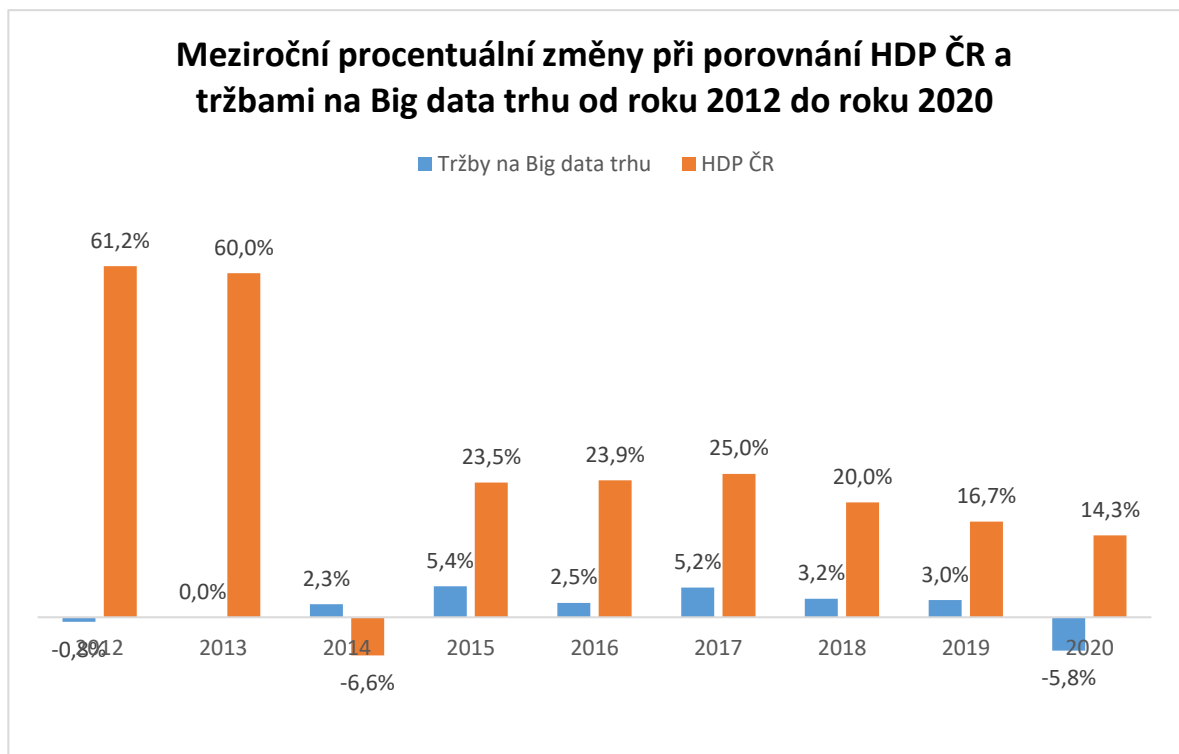
$$Y = 6,1665x - 4,0665$$

Nicméně z grafu je čitelné, že trh s big daty rapidně roste, a to se odráží i v popularitě tohoto segmentu. Společnosti využívající data mining skupují balíčky s Big daty ve značném množství.

### 6.2.3 Trendy vývojem blízké k trhu s Big daty

Jak jsme již zjistili, trh s Big daty má skutečně velký potenciál a průzkumy do roku 2027 se ve většině situacích potkávají s velmi podobným výsledkem. Abychom viděli, jak tento trend rapidně roste, srovnáme jej s vývojem HDP tady u nás v ČR. HDP v ČR samozřejmě zaznamenal menší pokles meziročně kvůli koronaviru v letech 2019 a 2020. Nicméně odráží se nám zde hodnota a mezi lety 2015 až 2018 se jedná o prudký nárůst HPD, který je právě srovnatelný s vývojovým trendem tržeb na Big data trhu. Pojdme si tedy ukázat graf a zhodnotit výsledky. Data ohledně HDP jsou pořizena z portálu kurzy.cz, kde je vedena

kompletní statistika meziročních i kvartálních hodnot hrubého domácího produktu v České republice. Data pochází z roku 2022, protože jsou neustále aktualizovaná.



Obrázek 12 – Meziroční procentuální změny při porovnání HDP ČR a tržbami na Big data trhu od roku 2012 do roku 2020 (vlastní zpracování dle Big data market size revenue forecast worldwide from 2011 to 2027, 2022 a HDP 2022, vývoj hdp v ČR, 2022)

Vycházejme ze stejného bodu, abychom zviditelnili čistý trend. Jak můžeme vidět, HDP z roku 2011 do roku 2019 vzrůstal, kromě menšího propadu z roku 2011 na rok 2012. Ovšem z roku 2019 na rok 2020 se jedná o propad v hodnotě 5,80 %. Toto je samozřejmě velkou částí způsobeno koronavirovou krizí, která postihla celý svět. Kromě těchto dvou propadů si ale můžeme všimnout mírného trendu, který je celou řadu let rostoucí. HDP tedy ze začátku lehce narůstá a poté na konci upadá. Pokud se podíváme na Big data, můžeme od roku 2014 vidět téměř totožný trend. Vycházíme z meziročního poklesu a poté lehce rosteme do roku 2017. Poté opět lehký úpadek v tempu růstu. Jistá spojitost tam tedy může nastat i když se jedná o naprosto odlišný segment. Graf slouží spíše pro zajímavost než pro důležitost našeho výzkumu.

### 6.3 Využívání data miningu

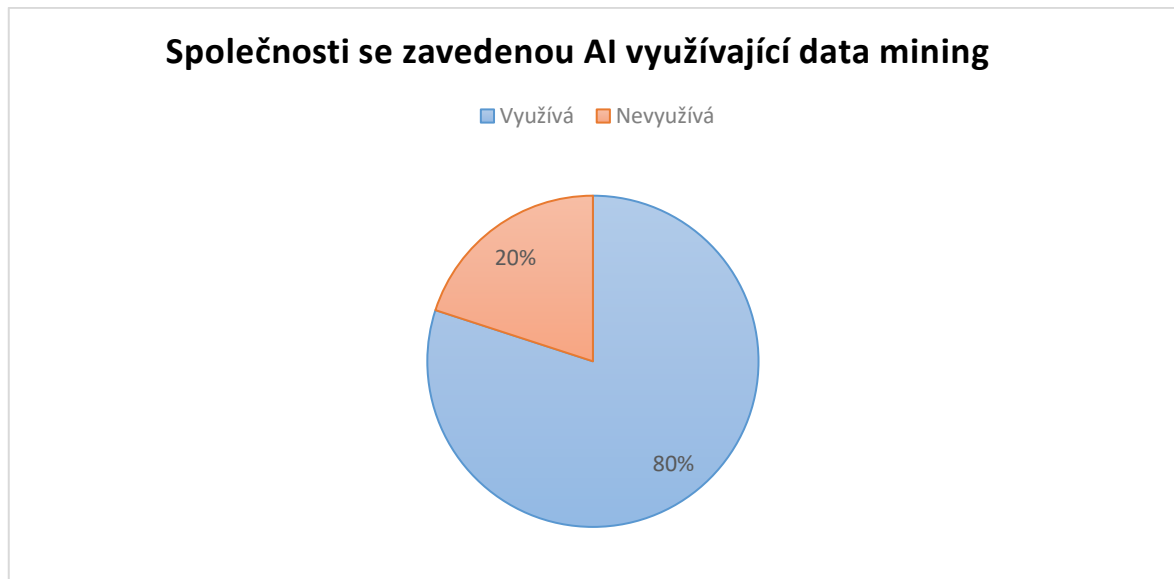
Nyní se dostáváme k tomu zřejmě nejdůležitějšímu tématu této práce. Data mining, neboli těžení obrovského množství dat pro nějaký účel, je dle statistiky v momentálním

digitalizovaném světě velmi důležitou součástí funkčností a prosperity firem, především v daných segmentech. Pro tyto statistiky se budeme řídit již vzniklými daty a dotazníky, které jsou veřejně dostupné na internetu a nabízí možnost statistického porovnání v těchto oblastech. Pro první podtéma využijeme data z internetového webu, který nám nabídl % zastoupení společností, které data mining využívají a naopak které ne. Tyto společnosti jsou už v rámci digitalizace pokročilé a využívají i umělou inteligenci. Bude se tedy jednat o společnosti se zavedenou AI. Článek zde bude citován (How Is Data Mining Used in Marketing?, 2022). Další data, která budeme používat v této kapitole, pocházejí z odborných článků, které jsou dostupné na sciencedirect.com. Tato data jsou vytvořena na základě dotazníků a sbírání jednotlivých odpovědí. Pro účel této práce jsou velmi vhodná a splňují potřebné parametry na použití. V první řadě se bude jednat o práci, která obsahuje dotazník. Jde o článek s názvem Data mining techniques in social media: A survey, který zhotovili Mohammadnoor Ahmad Mohammad Injadat, Ali Bou Nassif a Fadi Salo. Všichni tři zde uvedení pochází z The University of Western Ontario, tedy z univerzity v západním Ontariu. Zmínění jsou studenti nebo absolventi této kanadské univerzity. Data uvedená v jejich práci byla publikována roku 2016, tudíž budou pro naši práci relevantní a přínosná. Opět je zde uvedena citace (Injadat, Salo a Nassif, 2016). Jako poslední článek, který je v této kapitole použit, je opět odborný text, jenž je dostupný na stránce researchgate.net. Z tohoto článku byl použit graf, zachycující vývoj využití data miningu v oblasti výzkumu, tedy kvartérního sektoru. Citace opět zde (Lausch, Schmidt a Tischendorf, 2015).

Dále je potřeba zdůraznit, že výzkum, který se v této práci uskuteční, není vytažen z těchto zdrojů. Pouze jsou použita data, která jsou díky těmto článkům dostupná, a tudíž je autor využil pro vyhotovení vlastního výzkumu a vyhodnocení. Všechny komentáře, které zde autor uvádí, jsou psané jeho rukou a vymyšlené pouze ním spolu s potřebnými znalostmi z teoretické části, které nabyl před a v průběhu vyhotovení práce.

### 6.3.1 Využití data miningu společnostmi

V této podkapitole budeme zjišťovat, kolik společností data mining využívá. Budeme zde hovořit o společnostech, které data mining využívají anebo hodlají využívat, a pak o společnostech, které těžbu dat považují za méně důležitou nebo zbytečnou. Berme v potaz, že společnosti využívající data mining už podle následovného grafu umělou inteligenci pro svůj business používají. Data jsou pořízena z článku, který na tuto procenta odkazuje. Jedná se tedy o veřejná data.



Obrázek 13 – Využívání data miningu společnostmi (vlastní zpracování dle *How Is Data Mining Used in Marketing?*, 2022)

Z obrázku číslo 13 můžeme vidět, že 80 % společností data mining používají. Tyto společnosti budou zřejmě spadat do digitalizované sféry s velkou konkurencí a velmi silným využíváním umělé inteligence. Tato metoda je v moderních společnostech, které chtějí být konkurenceschopné, poměrně vyžadovaná. Převážná většina digitalizovaných společností data mining využívá anebo chce využívat. Data mining je dnes pro téměř všechny společnosti vítaný trend, který je součástí strojového učení. Usnadní podnikům spoustu práce především v oblasti marketingu a zákazníků obecně. Také je nezbytný pro kvalitní charakteristiku podnikových chyb a výhod. To vše odkazuje na dnešní velmi hojnou využívanost tohoto moderního trendu.

20 % společností zatím nebere data mining jako důležitou součást jejich businessu. Tento podíl je podle autora poměrně velký k dnešnímu digitalizovanému světu, kde převládá umělá inteligence napříč obory a odvětvími. Je tedy na místě přemýšlet, z jakých důvodů tomu tak je, a proč se společnosti bojí nebo dokonce vůbec nechtějí těžbu dat použít. Pojďme si tedy možné důvody nevyužívání data miningu přiblížit.

Může se například jednat o průmyslové továrny, které se zabývají výrobou běžných výrobků nebo výrobou zaměřenou na jeden hlavní produkt, což může být například ocelárna a podobně. Samozřejmě toto bude tvořit pouze část. Nicméně by to poměrně odpovídalo, protože umělou inteligenci tyto společnosti zajisté využívají, ale data mining pro ně nemusí být tak moc přínosný.



Dále je také nutné se zabývat otázkou odborného personálu. Je jasné, že potřebujeme někoho, kdo se o data mining a obecně o získávání a tvorbu analýz bude starat. Datový analytik může být pro určité firmy finančně náročným faktorem a například pro odvětví, kde data mining vyloženě potřebný není, se tato varianta nemusí zdát jako výhodná.

Také to mohou být společnosti mající konzervativní vedení, které odmítá další formu využití umělé inteligence. Z grafu vidíme, že tyto společnosti již mají zavedenou umělou inteligenci, nicméně i tato AI mohla být pro vedení dost velký problém a více ji podporovat už nechtějí. Autor se domnívá, že v Českém prostředí by se jednalo poměrně o běžnou věc, zde je konzervativních jedinců velké množství.

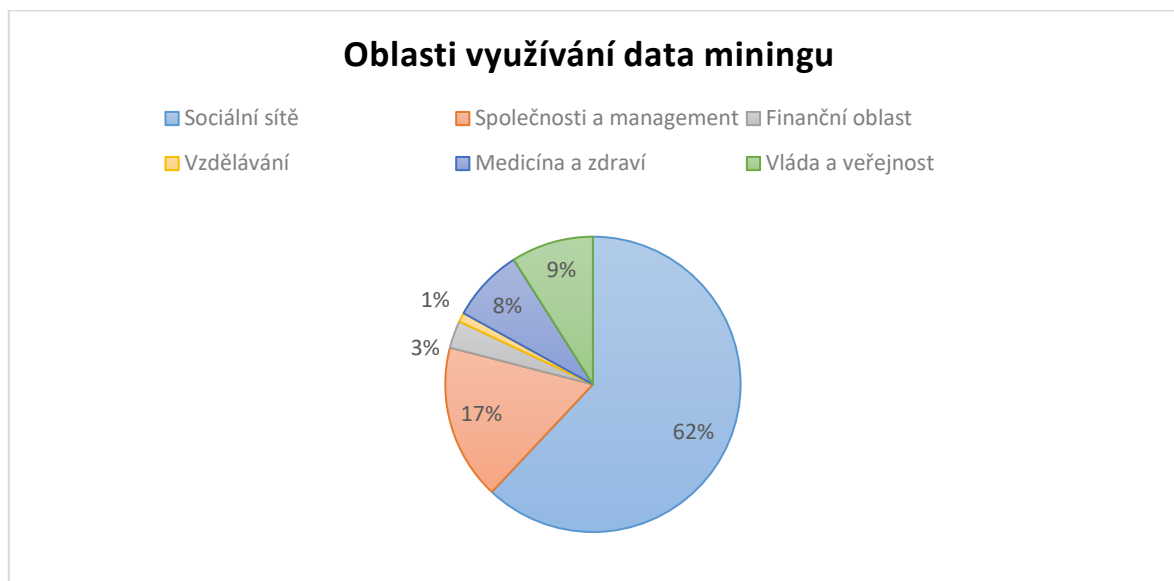
Jak jsme již avizovali, mohlo by jít o zdánlivě finanční nevýhodnost pro firmy. Je třeba brát v potaz skutečnost, že některé firmy a mladé společnosti by mohly data mining velmi dobře využít, ale chybí jim dostatečné množství finančních prostředků, informací a vědomostí potřebných k jeho efektivnímu zavedení. To je možnost, kterou je třeba brát v potaz. Toto se může běžně stát společností, především mladým a nově vznikajícím organizacím. Nemusí mít finance a vědomosti na to, aby data mining využívaly v plném rozsahu. Tudíž by tento důvod patřil do kolonky „Nevyužívá,“ protože by v jejich momentální podnikové strategii tento problém neřešili a data mining by odmítali.

Jako poslední důvod, který autora napadá, lze uvést zbytečnost a nepotřebnost ve vztahu k odvětví a konkurenci. My jsme si již uvedli, že některé podniky nemusejí vidět potenciál v data miningu. Je zde ale možnost, že ho nevidí nejen jednotlivý podnik, ale celé odvětví. Zde by se dalo hovořit pravděpodobně o těžším strojírenství, kde je potřebná umělá inteligence, ale data mining rozhodně ne.

### **6.3.2 Data mining v jednotlivých oblastech**

V minulé podkapitole jsme si ukázali míru využití data miningu společnostmi, které mají již zavedenou AI technologii. Dále jsme si také rozvedli jednotlivé a možné důvody, které brzdí společnosti v zavedení data miningu. Nyní je na čase ukázat si oblasti, kde se data mining využívá nejvíce a ve kterých segmentech dnes opravdu dominuje. V teoretické části jsme si uvedli, že by se mohlo jednat o finanční oblast, dále o zdravotnictví, medicínu a zdraví, marketing a co s ním dnes bezprostředně souvisí. V současnosti je pro marketing velmi důležité monitorování a vystupování na sociálních sítích. V data miningu tomu nebude jinak, pokud budeme potřebovat podklady pro marketing, musíme přímo ke zdroji dat a informací, což jsou dnes sociální sítě. Celkový počet tázaných společností pro tuto kapitolu

je šedesát šest v rámci nového dotazníku, který pochází z roku 2016. Tento dotazník je k dispozici na Science Direct pod názvem Data mining techniques in social media: A survey, který zhotovili Mohammadnoor Ahmad Mohammad Injadat, Ali Bou Nassif a Fadi Salo. Společnosti byly vybírány napříč obory, aby vznikl nějaký obecný přehled a názor o využívání data miningu v rámci oborů. Pojďme si tedy ukázat graf a popsat jeho výsledky (Injadat, Salo a Nassif, 2016). Průzkum zahrnuje 66 náhodných společností, odpovídajících na dotazník.



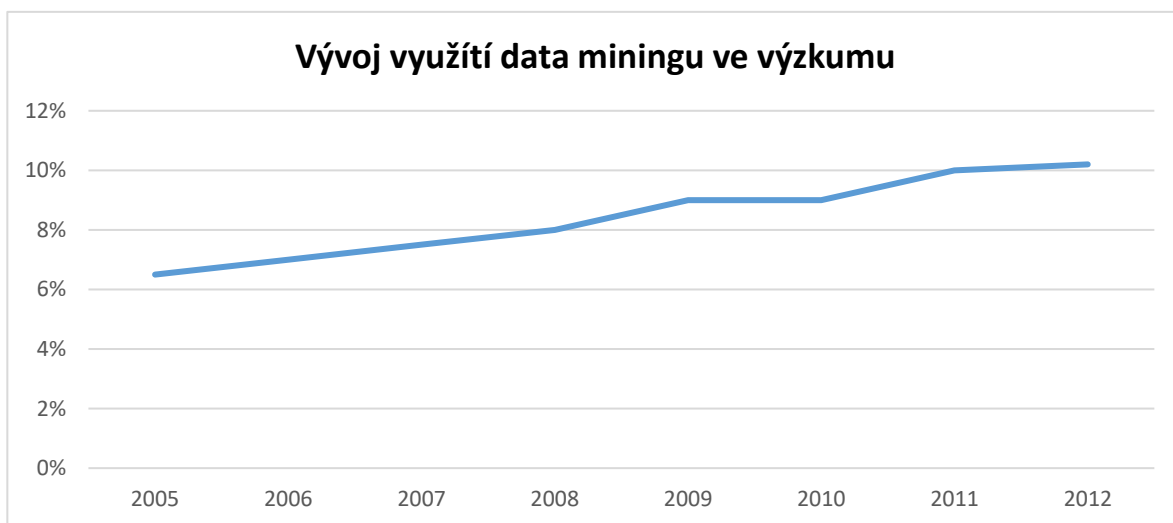
Obrázek 14 – Využívání data miningu v daných oborech (vlastní zpracování dle Injadat, Salo a Nassif, 2016)

Zde v komentování začneme od nejnižších procent, tudíž vzdělávání. Ve vzdělávání se dle grafu (obrázek číslo 14) data mining moc často nevyužívá. Důvodem může být například předepsaná osnova přímo tady v České republice. V osnově máme přímo dané věci, které se budou učit a vyučovat pro danou cílovou skupinu. Tudíž zde není taková potřeba těžby dat. Ovšem pokud se studuje nové téma, pak pro data mining ve vzdělávání rozhodně prostor je. To je důvod, proč je zde tato oblast zastoupena.

Dále se přesuneme k finanční sféře. Zde jsou pouze 3 %. Je třeba zdůraznit, že díky náhodnému výběru firem se toto číslo může značně lišit průzkum od průzkumu. Jak totiž víme i na základě teoretické části, v komerční a finanční sféře se data mining využívá často. Do této sféry patří bankovní sektor, o kterém si můžeme říct pár důležitých věcí především v dnešní době. Banky jsou dnes velmi digitalizované, díky COVID-19 se digitalizace bank posunula obrovským směrem dopředu. S tím souvisí i umělá inteligence a data mining. Jak

víme, tak data mining je vhodný pro získání informací, které jsou velmi důležité pro nabídku produktů zákazníkům. Banka může tedy cílit na potenciální zákazníky díky novým úrokovým mírám nebo zajímavým produktům pro nové klienty. Tyto produkty může přímo nabízet konkrétním skupinám, díky data miningu. Stejně tak pojišťovnictví. Tyto dva sektory zde budou tvořit většinu. Jednoduše budou díky těžbě dat cílit na jednotlivé cílové skupiny u kterých bude úspěch poměrně vysoký. V dnešní době je data mining opravdu velmi důležitý v tomto sektoru, a právě zisk potenciálních zákazníků a udržení stávajících klientů je jeho hlavním úkolem v této oblasti. Finančnictví také zažívá digitální rozmach a veškerá dostupná data, jak o klientech, tak i o obecném stavu a vývoji různých aktiv a komodit se dají analyzovat právě díky data miningu. Data mining zde má i spekulaci úlohu v podobě investic a různých příležitostí, které se díky němu dokážou velmi dobře vyfiltrovat na podstatné a velmi užitečné informace.

Medicína a zdraví má 8 % zastoupení. Zde se data mining může využívat pro predikování. Dále se v medicíně data mining využívá na různé analýzy. Na základě vysokého počtu vytěžených dat můžeme předpovídat vývoj zdravotního stavu u nemocí, kde západní medicína ještě nemá moc zkušeností. Také můžeme uvést extrakci dat z DNA, kde se data mining využívá velmi často. Tato oblast obecně roste s výzkumem, který za poslední roky v oblasti data miningu také značně roste. Názorně si ukážeme právě již zmíněný vývoj data miningu v oblasti výzkumu, kdy použijeme data z webu researchgate.net z roku 2012. Zde máme křivku vývoje neboli trend, který autor zkrátil a využijeme pouze 7 let, tedy od roku 2005 do roku 2012. Díky starším datům se zde jedná pouze o zajímavost doplňující zdravotnickou oblast.



Obrázek 15 – Vývoj data miningu ve výzkumu (vlastní zpracování dle Lausch, Schmidt, Tischedorf, 2015)

Jak můžeme vidět, trend má poměrně rychle rostoucí funkci. Za sedm let vzrostl téměř dvojnásobně. Tento trend samozřejmě do jisté míry ovlivňuje právě zdravotnictví, kterým jsme se zde zabývali.

Nyní přejdeme na vládu a veřejnost, veřejnost je myšlena jako domácnosti. Jak můžeme vidět, tak tento podíl se rovná 9 %, což není úplně málo. Domácnosti v posledních letech využívají data mining čím dál více pro své vlastní účely. Také vláda využívá data mining a to především pro vládní zakázky anebo různé analýzy. Může se jednat o trendové analýzy nebo analýzy trhu, které samozřejmě vládní sektor provádí také. Zde by se mohlo jednat opět i o predikce, můžou to být predikce finanční nebo třeba sociální. Například monitorování trhů práce a na základě tohoto následné vymodelování předpovědi. Pro vládu jakožto orgán, který musí mít v každé oblasti přehled, je data mining opravdu důležitý a autor je přesvědčen, že bez pomoci umělé technologie by se dnes vláda neobešla. Stejně tak by se bez umělé inteligence neobešla většina domácností.

Společnosti a management využívají data miningu v 17 %. Společnosti a management mají za úkol být ziskové a uživit se. Tudíž je pro ně velmi nezbytné a vlastně i nutné přilákat, získat a udržet zákazníky pro své produkty nebo služby. Díky data miningu má společnost ponětí o tom, co komu prodat a v jaké míře. Zná cílové skupiny pro své produkty nebo služby a také může díky tomuto trendu zjistit, co potřebuje pro další zákazníky. Využití data miningu je tedy v tomto případě marketing a kvalitní práce se zákazníkem. Zde to samozřejmě v žádném případě nekončí, neboť tu máme i management, tedy vrchní odpovědné osoby, které zodpovídají za směr a chod společnosti. V návaznosti na již zmíněné je potřeba zapřemýšlet co taková osoba potřebuje pro svoji činnost znát. Potřebuje samozřejmě různé analýzy, ať už je to finanční analýza nebo například SWOT analýza. Tyto analýzy management potřebuje pro svoji činnost a zlepšení své konkurenceschopnosti. Také díky těžbě dat dokáže management zjistit různé problémy a odhalit hrozby, které by společnost mohly poškodit. Může se jednat například o různé podvody nebo problémy se zabezpečením. Všechny tyto problémy musí management řešit. Tento obrovský podíl je tedy celkem pochopitelný a data mining je zde naprosto na svém místě.

Nyní popíšeme nejvíce atraktivní část z grafu, tedy sociální síť. Není náhoda, že se zabýváme sociálními sítěmi v teoretické části. Již od začátku této práce se dalo předpokládat, že většina činností data miningu je dnes soustředěna na sociální média. Sociální média dnes zastřešují různé činnosti lidí, jenž jsou předmětem data miningu. Někteří jedinci tráví na sociálních sítích velkou část svého volného času. Představme si, že máme miliony lidí

v jednom malém bytu a kteroukoliv informací, kterou bychom potřebovali, bychom okamžitě dostali. Takto to funguje na sociálních sítích. Jsou zde tak obrovská kvanta dat, že se jednoduše nejvíce vyplatí všem společnostem využívat data mining právě zde na sociálních sítích. Velmi to souvisí s předchozím tématem, což byl management a společnosti. Nejvíce se na sociálních médiích využívá data mining pro marketing. Zde můžeme mluvit o přímém marketingu, protože zde je vazba zákazník-produkt/služba. Jak můžeme vidět, jedná se o neuvěřitelných 62 %. To je vskutku velké množství procent, které jdou za touto oblastí. Je pravda, že je to poměrně pochopitelné, protože zde dostaneme informace o zákaznících, o potenciálních zákaznících a můžeme je později snáze oslovit díky komunikaci s nimi. Díky data miningu budeme vědět, kdo by mohl být pravidelný klient, a tak se ho pokusíme získat. Dále je zde obrovské množství informací na analýzu, na zlepšení produktů nebo služeb. Také je zde prostor pro inovace. Můžeme získat velké množství myšlenek, které na základě různých analýz a modelů můžeme využít ve svůj prospěch.

V teorii jsme si přesně tento druh marketingu uváděli. Produkty „šité na míru“ jsou na sociálních sítích a obecně na internetu všude. Můžeme se o tom sami přesvědčit, pokud budeme hledat nějaký produkt nebo jen něco, co s ním souvisí, zobrazí se nám na mnoha webových stránkách a sociálních sítích.

### **6.3.3 Nástroje data miningu z hlediska efektivnosti a účelnosti**

V této podkapitole si ukážeme programy a jejich použití při těžbě dat společnostmi i domácnostmi, které v této části rozebíráme. Zjistili jsme, že data mining se využívá především pro zajištění potřeb zákazníků a analyzování. Tudiž zjistíme, co je dnes moderní a co se používá na těžbu dat. Dále to porovnáme s teoretickou částí, kde máme některé z těchto programů vyjmenované a posoudíme, zda jsou skutečně užitečné nebo zda jsou naopak spíše teoreticky přeceňované nebo podceňované. Na tuto část jsou použity hodnoty a data nasbírané Delenem Dursunem, jenž jsou dostupné v jeho knize Predictive Analytics: Data Mining, Machine Learning and Data Science for Practitioners (Dursun, 2020). Autor z této databáze vybral pouze část programů. Programy jsou uvedené v tabulce, kde porovnáme jejich benefity, a naopak i nevýhody. Následně programy vyhodnotíme a interpretujeme výsledky.

Tabulka 1 – Nejpoužívanější nástroje data miningu využívající umělou inteligenci (vlastní zpracování dle Dursun, 2020)

název	počet uživatelů v průzkumu	Druh programu	výhody	nevýhody
Python	1078	Programovací jazyk	Zdarma, snadný, vhodný pro začátečníky, velmi rozšířený a lehce dostupný, modifikovatelný	méně pokročilých funkcí, nutno pořídit balíčky
RapidMiner	839	Nástroj zdarma	Zdarma, možnost prediktivních modelů, zhodnocení výsledků	Složitější
R Language	764	Programovací jazyk	Zdarma, snadno dostupný, velmi rozšířený, vynikající na statistické analýzy	Složitější, Úzší způsob využití
Excel	571	Komerční nástroj	Levný, Snadné používání, známost	méně pokročilých funkcí
Anaconda	556	Nástroj zdarma	Zdarma, napsána v jazyku Python	Složitější, Úzší způsob využití
Ostatní	5877	/	/	/
<b>Celkem uživatelů v průzkumu</b>	<b>9685</b>	/	/	/

Programů v knize dle Delena Dursena bylo okolo padesáti. Z tohoto důvodu autor vybral 5 nejčastěji využívaných programů, které si zhodnotíme.

Python jsme si také uváděli v teoretické části. Program je velmi úspěšný pravděpodobně díky své flexibilitě. Již víme, že tento program je veřejně dostupný a každá společnost si ho může vlastnoručně modifikovat dle své potřeby na základě jednotlivých balíčků. Tyto balíčky dokážou společnosti vytvořit data miningový program přímo na míru a jak můžeme vidět dle dat, funguje to. Program si v průzkumu odnesl první místo v počtu uživatelů a to přesně 1078.

Druhý program s největším zastoupením je RapidMiner, jenž používá dle průzkumu 839 uživatelů z celkového počtu 9685 respondentů. Jak jsme již uvedli v teoretické části, RapidMiner je graficky i technologicky skvěle vyvinutý software. Je dobře optimalizovaný a spolehlivý. Má více funkcí, tudíž je složitější, nicméně pro zkušeného programátora nebo uživatele se bude jednat o poměrně běžný program dle složitosti. Dále se snadno pojí s programem na modelování a trendovou analýzu, kterou můžeme hojně využívat v naší ekonomické oblasti. Také bude kvalitní i pro IT firmy.

764 respondentů uvedlo, že používá programovací jazyk R. Tento software slouží především pro analýzy. Proč má tento program tak velikou popularitu je z velké části prostý – je totiž zdarma. Samozřejmě se dá zlepšovat licence, ale základní je zcela bezplatná. Jedná se i o kvalitní program, tudíž není důvod ho nepoužívat. Výborně zvládá statistické úkony a je zde skvěle využitelný. Těžbu dat umí kvalitně a jak jsme již uvedli následnou analýzu a vyhodnocení také.

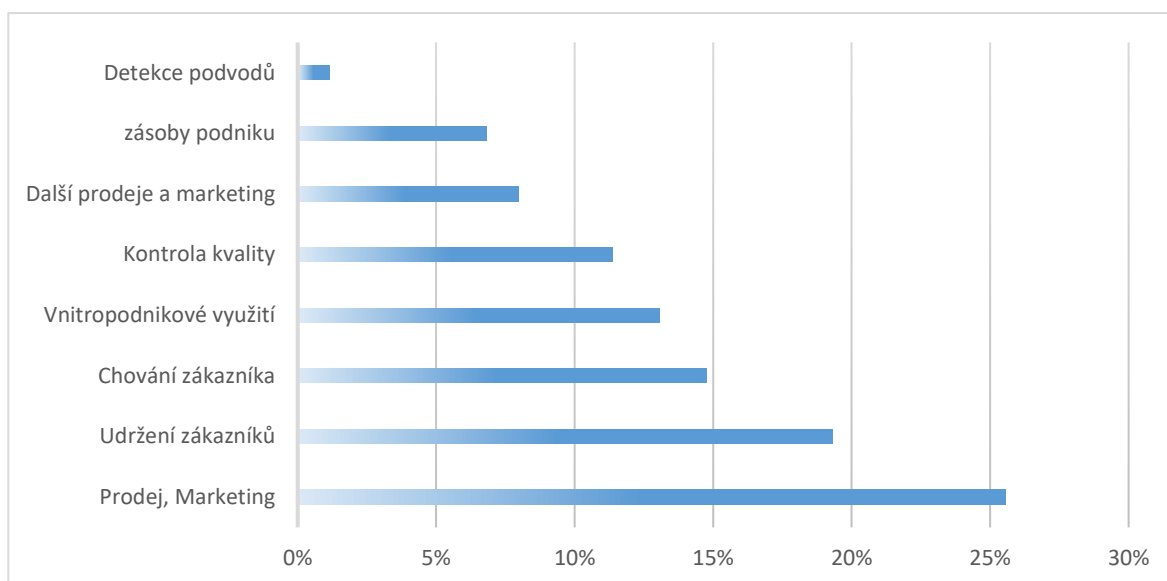
Zajímavým nástrojem pro nás může být MS Excel. Jedná se o námi dobře známý Excel, který používáme při vytváření nejrůznějších tabulek anebo řešení statistických záležitostí. Právě tento produkt od Microsoftu má velké množství doplňků, které lze do excelu doinstalovat. Jeden z nich je právě i nástroj pro těžbu dat. Zde se tedy jedná především o výhodu známého prostředí, ve kterém dnes umí většina z nás skvěle pracovat. Samozřejmě nejmodernějším programům, jenž jsou vytvořeny především na komerční data mining se Excel nevyrovná. Nicméně pro běžné i poměrně náročnější úkony lze využít.

Poslední uvedený program je Anaconda. Tento program je napsán v Python a R jazyce, tudíž nám předává určité výhody s tímto spojené. Anaconda je opět open source program tedy software dostupný zdarma. Uvedený nástroj pro data mining je především statisticky založen a orientován na tvoření analýz.

Můžeme vidět, že většina respondentů, tedy přesně 5877 používá jiné programy, než jsme si zde uvedli. Autor uváděl 5 nejvíce používaných programů a zbytek dal právě do této kolonky. Nicméně se v průzkumu objevily i programy námi známé z teoretické části. Programy jako například WEKA nebo SAS se v průzkumu také objevily. Zmíněná WEKA rozhodně stojí alespoň za zmínku, jelikož je vhodná přímo pro začátečníky na zaučení v oblasti data miningu.

### 6.3.4 Činnosti, které jsou předmětem data miningu

V posledních podkapitolách jsme si ukázali oblasti, ve kterých se data mining a obecně umělá inteligence a digitalizace používá nejvíce. Následně jsme si popsali a vysvětlili, proč jsou jednotlivé oblasti atraktivní pro data mining a nyní tyto oblasti převedeme přímo na jednotlivé potřeby. V této podkapitole si zobrazíme procentuální zastoupení činností, které jsou předmětem data miningu. Graf a hodnoty jsou převzaty z odborného článku, který napsal Asiri v roce 2017.



Obrázek 16 – Nejčastější využívání data miningu (vlastní zpracování dle staršího dotazníku modernizováno Asiri, 2017)

Největší část grafu zde patří marketingu a prodeji. V teoretické části jsme se bavili o sociálních médiích, které je potřeba v této podkapitole řádně připomenout. Sociální média se v dnešní době stávají hlavní platformou, která umožňuje tento druh marketingu. Také jsme si již ukázali, že hlavní oblastí, kde se data mining využívá, jsou právě sociální média. Pokud se půjdeme podívat na Facebook, okamžitě se objeví reklamy zobrazující produkty nebo služby, které jsou pro nás nějakým způsobem atraktivní. To se děje právě z důvodu využívání data miningu. Dnes už funguje běžně i odposlech například v mobilu, kdy se bavíme o nějakém tématu a následně se nám to přímo objeví jako nějaká forma reklamy či odkazu na internetu. Všechny tyto formy direct marketingu se ve velkém měřítku dělají přes data mining. Na základě podkapitoly, která nám zobrazila data miningové oblasti, vidíme zde jistou spojitost. Oblast sociálních médií tvořila 62 %, zde prodej a marketing tvoří cca 26 %. Marketing je tedy na základě našich dat hlavním předmětem data miningu. Nyní něco o prodeji samotném, samozřejmě se pojí s marketingem, nicméně zde také účinkuje jako



samostatný objekt. Prodej se bude do jisté míry pojit i s chováním zákazníků. Můžeme tedy využít dané nástroje, například analýzu tržeb na příští rok, kde na základě již dostupných informací analýzu vytvoříme. Plánované tržby jsou pro podnik poměrně důležité, především pro stanovení směru a strategie podniku.

Pokud se posuneme dál, zjistíme, že s téměř třetinovým zastoupením se zde objevuje chování zákazníka. Toto chování je z velké části založeno na trendové analýze. Trendové analýzy velmi hojně vytváří přímo společnost Google, která na to samozřejmě data mining využívá. V této práci je použit přímo trendový graf, který je vytvořen právě již zmíněnou firmou. Jedná se o trend popularity umělé inteligence na webu. Google má trendový graf za poměrně dlouhý časový úsek téměř v každém segmentu. Samozřejmě analýza trendu nemusí být vhodná pouze pro Google. Řeč zde byla především o chování zákazníka, kde na základě data miningu porozumíme tomu, jak se zákazník může chovat v určitých situacích, a především jak reaguje na trendy. Také investice jsou z velké části trendovou záležitostí, například Bitcoin a jiné komodity. Z investičního hlediska také vznikl slogan „Trend is your friend,“ tedy v překladu „Trend je tvůj přítel“.

Udržení zákazníka a obecně vztah s ním je v dnešní době plné konkurence zcela na místě. Data mining se zde využívá v plném rozsahu. Na základě dat již získaných o zákaznících se dá předpovídat potřeba zákazníka a prohlubovat to, co u společnosti daného člověka drží, tedy upevňovat vztah. Zde patří i různé analýzy spokojenosti a toho, co by zákazník chtěl zlepšit. Jsou to kompletně všechna data dohromady, jenž má společnost k dispozici. Díky data miningu můžeme filtrovat data tak, abychom si zákazníka udrželi nebo zlepšili vztah s dotyčnou osobou či přímo celou cílovou skupinou. S udržením a vztahem se může pojit i zisk zákazníků. Zde je také vhodné použít data mining, díky kterému můžeme opět vyfiltrovat data z komplexní databáze. Tímto zjistíme, co by od nás mohl potenciální zákazník požadovat. Data mining v oblasti podnikání obecně je nesmírně důležitou součástí. Na základě dat a trendové analýzy se dá samozřejmě udělat i vlastní investiční portfolio. Jak jsme již avizovali, investiční tedy trendová analýza se také dělá na základě data miningu. Tudíž ji využijeme k tomu abychom mohli vytvořit vlastní podíl investic v našem celkovém objemu. Investice veškerého majetku pouze do jednoho projektu je velice riskantní, takzvané sázíme na „jednu kartu“. Tudíž je zde velmi vhodné a investory i často doporučované využít diverzifikaci, aby se riziko snížilo a případný propad aktiva pokrylo aktivum druhé. Takže opět se zde výborně využívá data mining.

Vnitropodnikové využití máme na čtvrtém místě v grafu. Není se čemu divit, že se data mining hojně využívá právě na ekonomické činnosti. Zde můžeme opět uvést využití právě v oblasti managementu a společností, kde je finanční analýza velmi potřebná. Finanční analýza je opět proces založen na datech, které je potřeba vyfiltrovat a jak můžeme vidět z grafů, děje se to i v praxi. Informací pro vnější prostředí i vnitřní prostředí firmy je takové množství, že bez data miningu by se jednalo o naprosto zbytečně zdlouhavý, ale zároveň potřebný proces, který by nemusel vést k následovnému přesnému výsledku. Zde do finanční analýzy patří i vyhodnocování rizik pro podnikání, samozřejmě i analýza majetku.

Kontrola kvality a požadavky na ni jsou také častým předmětem data miningu. Společnosti vyhodnocují a analyzují celý produkt. Produkt tedy musí odpovídat kvalitě, kterou podnik prezentuje a zároveň zákazník požaduje. Data mining nám tedy vyhodnotí potřebné informace, na základě kterých můžeme kvalitu kontrolovat a popřípadě zjišťovat důvody nedodržení kvality.

Předposlední nejmenší zastoupení mají zásoby podniku. Zde se může jednat především o analýzy, na základě kterých bude podnik pořizovat určité množství zásob v daném čase. Také zde můžeme propojit právě již zmíněnou práci se zákazníkem, na základě které může podnik vymodelovat analýzy prodeje a zjistit orientační poptávku.

Jako poslední zde máme detekci podvodů. Na základě dobře vyfiltrovaných dat lze poměrně snadno detekovat podvod, například na bázi podobností již proběhnutých podvodů. Tato detekce se také ve velké hojnosti využívá na již zmíněných sociálních sítích. Internet je ohledně podvodů poměrně pestré místo a strůjci těchto ilegálních činností si toho jsou samozřejmě vědomi.

## **7 ZHODNOCENÍ VÝSLEDKŮ, NÁSTROJŮ Z HLEDISKA JEJICH EFEKTIVNOSTI A ÚČELNOSTI A IDENTIFIKACE TRENDŮ V OBLASTI DATA MININGU A UMĚLÉ INTELIGENCE**

Grafy jsme si vypsalí a uvedli možné důvody, které data mohou ovlivňovat. Nyní je na řadě interpretace výsledků jako celek. Z provedených výzkumů můžeme říct, že většina společností, které se zúčastnily dotazníků a jiných průzkumů, na základě kterých je tato práce psána, jsou digitalizované. Zřejmě na tom má velký podíl krize ohledně nemoci COVID-19, při které velká část zemí zaváděla lockdowny a zavírala obchody nebo jiné služby. Pokud se něco podobného stane, je zcela evidentní potřeba zavedení alespoň základní digitalizace ve společnosti, jako jsou například e-shopy a další internetové služby. Zjistili jsme, že počet společností zúčastněných v průzkumech, které minimálně tuto základní formu digitalizace již zavedli, je 60 %.

S digitalizací se pojí i pojem umělá inteligence. V teoretické části jsme měli výrok z roku 2021, kde se hovoří o neustálém růstu zájmu o digitalizaci a umělou inteligenci. Tento výrok můžeme potvrdit i v praktické části, kde jsme mohli vidět vzestupný trend popularity výrazu umělá inteligence. Tento trend podle Google Trends roste již více než 5 let a je velmi pravděpodobné, že se zájem ještě o něco zvýší.

Dále jsme se detailně zaměřili na Big data, především na trh s nimi. Tato data, jak jsme již zjistili v teoretické části, jsou podkladem pro data mining. Po filtraci se tedy dají využít například k identifikaci trendů, modelování různých analýz, detekování podvodů anebo také pro marketingové a ekonomické potřeby. Nicméně ohledně Big dat jsme zde především zjistili, jakým směrem by se mohl vyvíjet jejich trh. Ukázali jsme si, že by se mohlo jednat o růstový trend, kde se tržby na Big data trhu vyvíjí prakticky lineárně a stále bez výkyvů s predikcí až do roku 2027 na základě nyní dostupných dat, jenž zde byly použity. Pro zajímavost jsme zde využili podobnosti rostoucí tendence k porovnání s českým HDP za poslední roky. Ovšem dle předpokladu Big data trh roste daleko rychleji než HDP, což se samozřejmě ve vztahu ke zvyšující se popularitě daného segmentu dalo očekávat.

Nyní už se dostáváme k samotné problematice data miningu, kde jsme udělali poměrně velké množství grafů a komentářů. Je tedy na čase je všechny shrnout a okomentovat finální výsledek práce. Jak již bylo v kapitole o data miningu napsáno, byly zde využity dle autora kvalitní zdroje a podklady, které nám nabídly data pro tuto práci. Jako první jsme se zabývali použitím data miningu z obecného hlediska, kde jsme zjistili že 80 % společností

(účastnících se průzkumů) se zavedenou AI data mining využívá. Zbýlých 20 % data mining nevyužívá například z finančního důvodu nebo zbytečnosti v jejich podnikání.

Pokud se podíváme na důvody nevyužívání data miningu, zjistíme, že by se společnosti mohly obávat finanční a nátlakové stránky při neznalosti tohoto trendu. Další velmi dobrý důvod je právě již zmíněná zbytečnost nebo nepotřebnost v daném odvětví.

Nyní už se dostáváme k více ekonomické části, a to je využití data miningu v konkrétních oblastech. Zde dominuje oblast sociálních médií a také společností a managementu. Dále zde máme finančnictví, které data mining ve značném množství využívá. Z grafu můžeme vidět výsledek 3 %, je třeba si uvědomit, že toto číslo je poměrně zanedbatelné ve vztahu k náhodně vybraným společnostem. Pokud bychom vybrali jiný vzorek a získali tak větší objem dat, oblast finančnictví by měla pravděpodobně větší procentuální zastoupení. Měli bychom zde vyzdvihnout především bankovníctví a pojišťovnictví. Bankovní sektor využívá digitalizaci, umělou inteligenci i data mining poměrně často a je pro ně důležitý z hlediska vyhodnocování rizik, potenciálních klientů i doporučování produktů a služeb daným cílovým skupinám. Investiční sféra zde hraje také důležitou roli, především opět ve vyhodnocování rizik a investičních příležitostech. Rovněž se zde může jednat o tvorbu investičního portfolia. Investiční část se do velké míry řídí právě trendovou křivkou, která vyplývá z trendové analýzy. Trendová analýza je právě jedna z hlavních částí, kterou má data mining na starost. Jednoduše zde sbírá data a poté vymodeluje trendovou křivku. Z výše uvedených výsledků tedy vyplývá důležitost data miningu v tomto segmentu.

Pokud se budeme bavit přímo o oblastech, jednoznačně zde převládají sociální sítě, které především díky marketingu zabírají více než 60 % grafu vztahujícího se k využití data miningu dle oblastí. Ovšem nesmíme zapomenout na poměrně velký úspěch i v oblasti podniků a managementu. Důvody jsou zde poměrně jasné, opět se bude jednat o různé analýzy, jak jsme si ukázali v posledním grafu praktické části, kde byly popsány jednotlivé činnosti, kterými se data mining nejvíce zabývá. Opět se zde může jednat o určité pokusy získat zákazníky, udržet si je, vymodelovat tržby nebo cílit na marketing. Také v této oblasti bude velmi důležitá finanční analýza.

Zajímavý výsledek, který tedy s ekonomickou sférou spojen není, je strojírenství. V teoretické části jsme si uvedli, že umělá inteligence se využívá ve strojírenství a zemědělství poměrně často. To je sice pravda, ovšem zde se jedná pouze o data mining, jenž sice využívá pro svoji funkčnost umělou inteligenci, nicméně pro strojírenství a obecně tento typ průmyslu je pravděpodobně zbytečný, minimálně podle našich výsledků. Jak jsme si

mohli všimnout, tak strojírenství a těžký průmysl dohromady nenasbírali ani 1 %, tudíž je nemáme v našem grafu uvedené.

Velmi důležitým oddílem praktické části je také využívání programů, kde se střetáváme s teorií pouze částečně. Můžeme si povšimnout programů, které nejsou tolik známy, a přitom se využívají nejvíce. Uvedme si například RapidMiner, který zaznamenal největší četnost, a dokonce patří dle teoretických trendů na 1. místo víceméně ve všech žebříčcích uvedených na internetových webech. Je tedy hodnocen jako nejlepší, a takto se jeví i podle dat. Zde jsme si mohli ověřit, že programy jsou stavěné pro budoucí aktualizace (lze vidět u společnosti SAS), aby klienti neměli potřebu měnit své dosavadní softwary. Pak zde máme například software od Python nebo IBM, které jsou jak v trendech, tak i v našich datech. Tudíž z toho můžeme vyvodit, že tyto trendové softwary disponují opravdovou kvalitou. Pokud je společnost mladá a začíná s data miningem, rozhodně by si mohla tyto novějších softwary pořídit.

Autor ve své práci věnoval velkou pozornost právě marketingu na sociálních sítích, s čímž souvisí i cíl práce. Z teoretické části se můžeme dočíst, že na sociálních sítích se provozuje přímý marketing na základě data miningu. Jsou to běžné reklamy zacílené přímo na zákazníka na základě monitoringu a následného zpracování dat. Toto zpracování dat přes data mining, je důvod, proč jsou reklamy přesné a zobrazují produkty či služby, které nás opravdu zajímají. Jak jsme se mohli přesvědčit, tak se data mining pro marketing využívá z poměrně velké části. I společnosti, jenž se marketingem přímo zabývají, tento druh využívají perfektně. Tudíž můžeme na základě výsledků říci, že právě marketing na sociálních sítích je momentálně velký trend v oblasti data miningu. Samozřejmě není to pouze marketing. Předmětem data miningu na sociálních médiích jsou i běžné nicméně velmi důležité činnosti jako je porozumění chování zákazníků a již několikrát zmíněná trendová analýza, pojící se se zájmem a ziskem zákazníka. Výše uvedená fakta nám potvrdil i graf s oblastmi, kde jsou sociální sítě hlavním segmentem, který je využíván technikami těžby dat.

Trendy, které můžeme uvést z výsledků praktické části mohou být následující:

- A) Trh s nabídkou služeb s AI a daty poroste.
- B) AI a data mining bude o součástí základních podnikových činností stejně jako trend zásobování
- C) AI a data mining bude stále významnější nákladová položka.

D) V návaznosti na předchozí bod se bude více řešit efektivnost vynaložených zdrojů.

E) Vznikne více firem nabízejících nástroje data miningu.

Tyto trendy by společnosti měly jednoznačně zvážit a snažit se jim přizpůsobit. Digitalizace roste a s tím i AI technologie, tudíž se zvyšuje i využívání data miningu.

Zhodnotili jsme výsledky a uvedli trendy v oblasti data miningu. Tyto trendy jsou využitelné především v ekonomické sféře, tím je myšlen hlavně marketing a také podnikové vedení, které využívá různé analýzy a vyhodnocování rizik pro své podnikání. Společnosti chtějí být leaderem ve zmíněných oblastech a odvětvích, budou muset zavést silný a kvalitní data mining. Data mining je, jak jsme se mohli přesvědčit, opravdu dobrým nástrojem na získání zákazníků a jejich udržení. Dále je využitelný pro veškeré analýzy, jenž společnosti víceméně ve většině odvětví využívat chtějí, jako například analýza tržeb. Předběžné tržby chce vědět každá společnost z důvodu plánování a volby strategie a vůbec volby budoucího směru firmy. Také jsme zjistili, že je dobré řídit se trendovými softwary, samozřejmě všechny obsažené v grafu vzhledem k teorii nebyly, nicméně většina ano. Z hlediska managementu, který promýšlí zavedení data miningu do vlastní společnosti je na místě, aby zvážily výhody a nevýhody. Pokud dokážou přes data mining využívat kvalitní marketing a zařídit dobrou komunikaci se zákazníky, potom je těžba dat skvělou volbou. Samozřejmě jsou i odvětví, kde se tento moderní trend nehodí a nebude mít právě ten přínosný efekt, který se očekává. Doporučené oblasti pro využití umělé inteligence pro data mining jsou rozhodně ekonomické a podnikové oblasti.

## ZÁVĚR

Hlavním cílem této práce bylo identifikovat trendy v data miningu a návazně vyhodnotit jeho nástroje z hlediska jejich efektivnosti a účelnosti. Tato identifikace se nám povedla a vyhodnotili jsme trendy, které by měli v budoucnu nastat a nástroje, jenž se s velkou mírou pravděpodobnosti budou využívat. Dále jsme zjistili, že starší a nynější trendy jsou především využívání data miningu pro přímý marketing především v oblasti sociálních sítí. Jedná se i o trendovou analýzu a analýzu sledující chování spotřebitele. Využívání data miningu v oblasti ekonomie a finančnictví obecně a používání trendových data miningových softwarů. Lze také vyzdvihnout modelování výsledků založených na data miningu a lákání zákazníků, které souvisí s tímto internetovým marketingem, který se běžně na sociálních médiích využívá.

V teoretické části jsme si vysvětlili základní pojmy a popsali jednotlivé problematiky. Zaměřili jsme se na teoretické definice data miningu nebo například umělé inteligence, které pro náš výzkum mají významnější hodnotu. Dále jsme narazili i na samotné sociální sítě, kde jsme si vyjmenovali ty hlavní, pro které je data mining nejvíce určený. Již víme, co je data mining a jakou má sílu v dnešním digitalizovaném světě.

Po teoretickém celku jsme se přesunuli k části praktické. V praktické oblasti jsme si ukazovali jednotlivá data na základě dostupných informací o dané problematice. Zjistili jsme, že digitalizace firem je momentálně převažujícím a klíčovým faktorem. Také jsme se zaměřili na popularitu umělé inteligence, přesněji na četnost vyhledávání pojmu umělá inteligence. Po ukázce vzestupného trendu četnosti vyhledávání pojmu umělá inteligence jsme si zobrazili trendovou křivku růstu trhu s Big daty, která zaznamenává rapidní růst každým rokem. Big data jsou samozřejmě podklad pro data mining, proto bylo potřeba se o tématu dozvědět více. Data mining jsme rozebrali co možná nejprecizněji. Zjistili jsme kolik společností ho používá, kolik společností o něj má zájem do budoucna a vidí ho jako užitečný nástroj a také jak a na co ho použít. Jaký je důvod společností data mining nevyužít. Zda jsou tyto důvody dostatečné a oprávněné. Dále zde bylo využití tohoto trendu v oblastech, kde jsme zjistili, že sociální média dominují, a to především díky marketingu. Velmi důležité také bylo ukázat si nejvyužívanější programy pro data mining a modeling. V poslední grafu jsme uvedli využívání miningu pro jednotlivé činnosti. Tato podkapitola byla zřejmě nejdůležitější v našem výzkumu, protože nám prakticky shrnula a potvrdila názory, jenž jsme do té doby, díky shrnutí praktické části předpokládaly.

Na úplný závěr jsme si uvedli kompletní shrnutí výsledků a jejich interpretaci, která byla zaměřena především na ekonomickou a komerční sféru. Poté jsme uvedli moderní trendy a doporučení pro podniky, které data mining používají anebo se o něj zatím pouze zajímají.

Také je nutné uvést, že by bylo potřeba mít k dispozici více dostupných zdrojů. Bohužel data s touto tematikou víceméně dostupná nejsou a bylo velmi obtížné dělat tento výzkum. Pro bližší a více hloubkové rozebrání této problematiky je za potřebí mít k dispozici několikoroční mezifiremní dotazník, kde bychom zaznamenali daleko více dat, a především by tato data přímo souvisela s našim výzkumem.

Práce byla projednána s expertem z firmy TIC. Tato firma se zabývá umělou inteligencí i data miningem. Konzultace potvrdila výzkum zde probíraný.



**SEZNAM POUŽITÉ LITERATURY**

DO, Quang Hung, ed. Artificial Neural Network Applications in Business and Engineering [online]. IGI Global, 2021 [cit. 2022-05-04]. Advances in Computational Intelligence and Robotics. ISBN 9781799832386. Dostupné z: doi:10.4018/978-1-7998-3238-6

DOSTÁL, Petr. Soft computing v podnikatelství a veřejné správě. Brno: Akademické nakladatelství CERM, 2015, 130 s. ISBN 9788072048960.

DURSUN, Delen, 2020. Predictive Analytics: Data Mining, Machine Learning and Data Science for Practitioners. 2nd Edition. London: Pearson Education. ISBN 9780136738510.

HAN, Jiawei. Micheline KAMBER a Jian PEI. Data mining: concepts and techniques. 3rd ed. Waltham: Elsevier, 2012, 703 s. ISBN 9780123814791.

HOLOŠKA, Jiří. Artificial intelligence applied on cryptoanalysis aimed on cryptoanalysis aimed [sic] on revealing weaknesses of modern cryptology and computer security. Zlín: Tomas Bata University in Zlín, 2012, 30 s. ISBN 9788074541445.

HOLUBOVÁ, Irena, Jiří KOSEK, Karel MINAŘÍK a David NOVÁK. Big Data a NoSQL databáze. Praha: Grada, 2015, 281 s. ISBN 9788024754666.

INJADAT, MohammadNoor, Fadi SALO a Ali Bou NASSIF. Data mining techniques in social media: A survey. Neurocomputing [online]. 2016, 214, 654-670 [cit. 2022-05-05]. ISSN 09252312. Dostupné z: doi:10.1016/j.neucom.2016.06.045

Intelligent Miner Visualizer, 2021. Ibm.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.ibm.com/docs/en/db2/9.7?topic=overview-intelligent-miner-visualizer>

JOVANOVIĆ, Bojan, 2022. 55 Fascinating AI Statistics and Trends for 2022. Dataprot.net [online]. [cit. 2022-05-05]. Dostupné z: <https://dataprot.net/statistics/ai-statistics/>

LABERGE, Robert. Datové sklady: agilní metody a business intelligence. Brno: Computer Press, 2012, 350 s. ISBN 9788025137291.

LAUSCH, Angela, Andreas SCHMIDT a Lutz TISCHENDORF. Data mining and linked open data – New perspectives for data analysis in environmental research. Ecological Modelling [online]. 2015, 295, 5-17 [cit. 2022-05-05]. ISSN 03043800. Dostupné z: doi:10.1016/j.ecolmodel.2014.09.018

LETERRIER, Christophe. The Axon Initial Segment, 50Years Later. Dynamic Plasma Membranes - Portals Between Cells and Physiology [online]. Elsevier, 2016, 2016, s. 185-

233 [cit. 2022-05-05]. Current Topics in Membranes. ISBN 9780128054048. Dostupné z: doi:10.1016/bs.ctm.2015.10.005

NEAPOLITAN, Richard E. a Xia JIANG. Artificial intelligence: with an introduction to machine learning. Second edition. Boca Raton, FL: CRC Press, Taylor & Francis Group, [2018], 1 online zdroj. Chapman & Hall/CRC artificial intelligence and robotics series. Dostupné z: doi:9781351384384

WITTEN, I. H., Eibe FRANK, Mark A. HALL a Christopher J. PAL. Data mining: practical machine learning tools and techniques. Fourth edition. Amsterdam: Elsevier, 2017, 621 s. ISBN 9780128042915.

## INTERNETOVÉ ZDROJE

Anatomie neuronu - odkud přichází a kam směřuje nervový vzruch, 2015. Mentem.cz [online]. [cit. 2022-05-05]. Dostupné z: <https://www.mentem.cz/blog/neuron/>

ARNALDO, Muhammad, 2021. Intro to Rapidminer: A No-Code Development Platform for Data Mining (with Case Study). Analyticsvidhya.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.analyticsvidhya.com/blog/2021/10/intro-to-rapidminer-a-no-code-development-platform-for-data-mining-with-case-study/>

Artificial Intelligence in Agriculture: Rooting Out the Seed of Doubt, 2022. Intellias.com [online]. [cit. 2022-05-05]. Dostupné z: <https://intellias.com/artificial-intelligence-in-agriculture/>

Artificial Intelligence, Machine Learning and Big Data in Finance, 2021. Medium.com [online]. [cit. 2022-05-05]. Dostupné z: [www.oecd.org/finance/financial-markets/Artificial-intelligence-machine-learning-big-data-in-finance.pdf](http://www.oecd.org/finance/financial-markets/Artificial-intelligence-machine-learning-big-data-in-finance.pdf)

ASIRI, Sidath, 2017. Data Mining in Brief. Towardsdatascience.com [online]. [cit. 2022-05-05]. Dostupné z: <https://towardsdatascience.com/data-mining-in-brief-26483437f178>

BLAKE, Morgan, 2019. 100 Stats On Digital Transformation And Customer Experience. Forbes.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.forbes.com/sites/blakemorgan/2019/12/16/100-stats-on-digital-transformation-and-customer-experience/?sh=3277edaa3bf3>

BRADEN, Abby, 2021. What is Software?. Webopedia.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.webopedia.com/definitions/software/>

BREJČÁK, Peter, 2022. Zájem o digitalizaci roste a s ním i česká Etnetera. Nově řeší i elektrovozy a umělou inteligenci. Cc.cz [online]. [cit. 2022-05-05]. Dostupné z: <https://cc.cz/kvuli-valce-bude-pracovni-trh-v-it-jeste-vice-pod-tlakem-pomoc-muze-vyvoj-na-dalku-rika-sef-cn-group/>

CALLAHAN, Guv, 2020. What Are the Advantages of Artificial Intelligence?. Intellias.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.rev.com/blog/what-are-the-advantages-of-artificial-intelligence>

CECI, Laura, 2022. YouTube - Statistics & Facts. Statista.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.statista.com/topics/2019/youtube/>

Data Mining – Time-Series, Symbolic and Biological Sequences Data, 2022. Geeksforgeeks.org [online]. [cit. 2022-05-05]. Dostupné z: <https://www.geeksforgeeks.org/data-mining-time-series-symbolic-and-biological-sequences-data/>

Digitize: definition. Oxford Dictionaries [online]. Oxford: Oxford University Press, 2017 [cit. 2017-01-11]. Dostupné z: <https://en.oxforddictionaries.com/definition/digitize>

DUGGAL, Nikita, 2022. Advantages and Disadvantages of Artificial Intelligence. Simplilearn.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.simplilearn.com/advantages-and-disadvantages-of-artificial-intelligence-article>

DURČÁK, Pavel, 2017. Neuronové sítě a princip jejich fungování. Napocitaci.cz [online]. [cit. 2022-05-05]. Dostupné z: <https://www.napocitaci.cz/33/neuronove-site-a-princip-jejich-fungovani-uniqueidgOkE4NvrWuNY54vrLeM670eFNQh552VdDDulZX7UDBY/>

Enterprise Miner, 2022. Sas.com [online]. [cit. 2022-05-05]. Dostupné z: [https://www.sas.com/en\\_us/software/enterprise-miner.html](https://www.sas.com/en_us/software/enterprise-miner.html)

EURIX develops Machine Learning models, 2022. Eurixgroup.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.eurixgroup.com/eurix-machine-learning-2/>

Facebook - statistics & facts, 2022. Statista.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.statista.com/topics/751/facebook/#dossierKeyfigures>

HAYE, Adam, 2021. Social Media Marketing (SMM). Investopedia.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.investopedia.com/terms/s/social-media-marketing-smm.asp>

HDP 2022, vývoj hdp v ČR, 2022. Kurzy.cz [online]. [cit. 2022-05-05]. Dostupné z: <https://www.kurzy.cz/makroekonomika/hdp/>

How Is Data Mining Used in Marketing?, 2022. Comptia.org [online]. [cit. 2022-05-05]. Dostupné z: <https://www.comptia.org/content/articles/how-is-data-mining-used-in-marketing>

CHUDÁN, David, 2020. Weka, jednoduchý nástroj pro data mining. Davidchudan.cz [online]. [cit. 2022-05-05]. Dostupné z: <https://www.davidchudan.cz/post/weka-jednoduchy-nastroj-pro-data-mining>

JYOTI, Ritu, 2018. Worldwide Artificial Intelligence Market Shares, 2018: Steady Growth — POCs Poised to Enter Full-Blown Production. Ibm.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.ibm.com/downloads/cas/MK85Y8V3>

KISK, Edtech, 2020. Trendy ve výuce: Umělá inteligence ve vzdělávání. Medium.com [online]. [cit. 2022-05-05]. Dostupné z: <https://medium.com/edtech-kisk/trendy-ve-vyuce-umela-inteligence-ve-vzdelavani-3de531a059d9>

LUA, Alfred, 2022. 20 Top Social Media Sites to Consider for Your Brand in 2022. Buffer.com [online]. [cit. 2022-05-05]. Dostupné z: <https://buffer.com/library/social-media-sites/>

MARR, Bernard, 2021. Future Developments of Artificial Intelligence. Bernardmarr.com [online]. [cit. 2022-05-05]. Dostupné z: <https://bernardmarr.com/future-developments-of-artificial-intelligence>

Nejoblíbenější sociální sítě, 1999–2022. Agionet.cz [online]. [cit. 2022-05-05]. Dostupné z: <https://www.agionet.cz/sluzby/internet-servis/online-podnikani/nejoblibenejsi-socialni-site>

NG, Andrew, 2022. Machine Learning. Investopedia.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.coursera.org/learn/machine-learning>

Oracle Data Miner, 2022. Oracle.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.oracle.com/database/technologies/datawarehouse-bigdata/dataminer.html>

PALUS, Shannon, 2014. Data Mining Proves Darwin's Finches Weren't Really His. Discovermagazine.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.discovermagazine.com/technology/data-mining-proves-darwins-finches-werent-really-his>

REIFER, Abie, 2017. How SAS Enterprise Miner simplifies the data mining process. Techtarget.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.techtarget.com/searchbusinessanalytics/feature/How-SAS-Enterprise-Miner-simplifies-the-data-mining-process>

SARANGAM, Ajay, 2020. Top 14 Data Mining Tools. Jigsawacademy.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.jigsawacademy.com/blogs/data-science/data-mining-tools/>

Social Media: What Countries Use It Most & What Are They Using?, 2021. Digitalmarketinginstitute.com [online]. [cit. 2022-05-05]. Dostupné z: <https://digitalmarketinginstitute.com/blog/social-media-what-countries-use-it-most-and-what-are-they-using>

TWIN, Alexandra, 2021. Data Mining Definition. Investopedia.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.investopedia.com/terms/d/datamining.asp>

Umělá inteligence, 2022. Trends.google.com [online]. [cit. 2022-05-05]. Dostupné z: <https://trends.google.com/trends/explore?date=today%205-y&geo=CZ&q=%2Fm%2F0mkz>

Úvod do data miningu, 2013. Statsoft.cz [online]. [cit. 2022-05-05]. Dostupné z: [http://www.statsoft.cz/file1/PDF/newsletter/2014\\_02\\_26\\_StatSoft\\_Uvod\\_do\\_data\\_miningu.pdf](http://www.statsoft.cz/file1/PDF/newsletter/2014_02_26_StatSoft_Uvod_do_data_miningu.pdf)

VEARRIER, Laura, 2022. Artificial Intelligence in Emergency Medicine: Benefits, Risks, and Recommendations. Sciencedirect.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.sciencedirect.com/science/article/abs/pii/S0736467922000506>

VYAS, Kashyap, 2021. 7 Benefits of AI That Will Help Humanity, Not Harm It. <https://interestingengineering.com> [online]. [cit. 2022-05-05]. Dostupné z: <https://interestingengineering.com/7-ways-ai-will-help-humanity-not-harm-it>

What Is a Data Set? (With Definition, Components and Types), 2022. Indeed.com [online]. [cit. 2022-05-05]. Dostupné z: <https://www.indeed.com/career-advice/career-development/what-is-data-set>

WOLFF, Rachel, 2020. 10 Best Data Mining Tools in 2022. Monkeylearn.com [online]. [cit. 2022-05-05]. Dostupné z: <https://monkeylearn.com/blog/data-mining-tools/>

Za deset let se stal YouTube druhou nejnavštěvovanější stránkou, 2015. Ct24.ceskatelevize.cz [online]. [cit. 2022-05-05]. Dostupné z: <https://ct24.ceskatelevize.cz/media/1527110-za-deset-let-se-stal-youtube-druhou-nejnavsteovanejsi-strankou>

## SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK

ML Machine Learning – strojové učení

AI Artificial Intelligence – umělá inteligence

% Procento

## SEZNAM OBRÁZKŮ

<i>Obrázek 1 – Data mining a výsledek (vlastní zpracování dle Data Mining – Time-Series, Symbolic and Biological Sequences Data, 2022) .....</i>	13
<i>Obrázek 2 – Machine Learning (vlastní zpracování dle EURIX develops Machine Learning models, 2022).....</i>	15
<i>Obrázek 3 – Neuron (Anatomie neuronu – odkud přichází a kam směřuje nervový vzruch, 2015) .....</i>	21
<i>Obrázek 4 – Příklad neuronové sítě (Durčák, 2017).....</i>	23
<i>Obrázek 5 – Enterprise Miner (Enterprise Miner, 2022).....</i>	26
<i>Obrázek 6 – Využívání digitalizace organizacemi (vlastní zpracování dle Blake, 2019)....</i>	33
<i>Obrázek 7 – Využívání umělé inteligence organizacemi (vlastní zpracování dle Jovanovic, 2022) .....</i>	34
<i>Obrázek 8 – Četnost vyhledávání pojmu umělá inteligence na webu za posledních 5 let (vlastní zpracování dle Umělá inteligence, 2022) .....</i>	35
<i>Obrázek 9 – Velikost trhu s umělou inteligencí (vlastní zpracování dle Revenues from the artificial intelligence (AI) software market worldwide from 2018 to 2025, 2022, Jyoti, 2018) .....</i>	35
<i>Obrázek 10 – Meziroční procentuální změny četnosti vyhledávání výrazu umělá inteligence a velikosti trhu s AI (vlastní zpracování dle Revenues from the artificial intelligence (AI) software market worldwide from 2018 to 2025. Jyoti, 2018 a Umělá inteligence, 2022)..</i>	36
<i>Obrázek 11 – Big data tržby na trhu od roku 2011 do roku 2027 (vlastní zpracování dle Big data market size revenue forecast worldwide from 2011 to 2027, 2022).....</i>	37
<i>Obrázek 12 – Meziroční procentuální změny při porovnání HDP ČR a tržbami na Big data trhu od roku 2012 do roku 2020 (vlastní zpracování dle Big data market size revenue forecast worldwide from 2011 to 2027, 2022 a HDP 2022, vývoj hdp v ČR, 2022) .....</i>	38
<i>Obrázek 13 – Využívání data miningu společnostmi (vlastní zpracování dle How Is Data Mining Used in Marketing?, 2022).....</i>	40
<i>Obrázek 14 – Využívání data miningu v daných oborech (vlastní zpracování dle Injadat, Salo a Nassif, 2016) .....</i>	42
<i>Obrázek 15 – Vývoj data miningu ve výzkumu (vlastní zpracování dle Lausch, Schmidt, Tischedorf, 2015) .....</i>	43
<i>Obrázek 16 – Nejčastější využívání data miningu (vlastní zpracování dle staršího dotazníku modernizováno Asiri, 2017).....</i>	48



**SEZNAM TABULEK**

<i>Tabulka 1 – Nejpoužívanější nástroje data miningu využívající umělou inteligenci (vlastní zpracování dle Dursun, 2020) .....</i>	<i>46</i>
---	-----------