

# **Big Data koncept a aplikace algoritmů strojového učení v podnikových procesech**

Jakub Dostál

---

Bakalářská práce  
2023



Univerzita Tomáše Bati ve Zlíně  
Fakulta managementu a ekonomiky

---

Univerzita Tomáše Bati ve Zlíně  
Fakulta managementu a ekonomiky  
Ústav statistiky a kvantitativních metod

Akademický rok: 2022/2023

# ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Jakub Dostál**  
Osobní číslo: **M210574**  
Studijní program: **B0413A050024 Ekonomika a management**  
Specializace: **Ekonomika a management podniku**  
Forma studia: **Prezenční**  
Téma práce: **Big Data koncept a aplikace algoritmů strojového učení v podnikových procesech**

## Zásady pro vypracování

### Úvod

Definujte cíle práce a použité metody zpracování práce.

#### I. Teoretická část

- Zpracujte literární rešerši zaměřenou na oblast velkých objemů dat.

#### II. Praktická část

- Pomocí metod strojového učení predikujte vývoj prodeje vybrané společnosti.
- Na základě výsledků predikce navrhněte vhodná doporučení vedoucí k optimalizaci skladových zásob dané společnosti.

### Závěr

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

**Seznam doporučené literatury:**

- HENDL, Jan. *Big data: věda o datech – základy a aplikace*. Praha: Grada Publishing, 2021, 224 s. ISBN 978-80-271-3031-3.  
HOLUBOVÁ, Irena, Jiří KOSEK, Karel MINAŘÍK a David NOVÁK. *Big Data a NoSQL databáze*. Praha: Grada, 2015, 281 s. ISBN 978-80-247-5466-6.  
MARR, Bernard. *Big data in practice: how 45 successful companies used big data analytics to deliver extraordinary results*. Chichester: Wiley, 2016, 308 s. ISBN 978-1119231387.  
MARR, Bernard. *Big data: using smart big data, analytics and metrics to make better decisions and improve performance*. Chichester: Wiley, 2015, 246 s. ISBN 978-1118965832.  
MAYER-SCHÖNBERGER, Viktor a Kenneth CUKIER. *Big data: revoluce, která změní způsob, jak žijeme, pracujeme a myslíme*. Brno: Computer Press, 2014, 256 s. ISBN 978-8-0251-4119-9.

Vedoucí bakalářské práce: **Ing. Tomáš Urbánek, Ph.D.**  
Ústav statistiky a kvantitativních metod

Datum zadání bakalářské práce: **10. února 2023**  
Termín odevzdání bakalářské práce: **19. května 2023**

L.S.

---

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

---

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

Ve Zlíně dne 10. února 2023

**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 připouští-li tak licenční smlouva uzavřená mezi mnou a Univerzitou Tomáše Bati ve Zlíně 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 dne 17.5.2023

Jméno a příjmení: Jakub Dostál

.....  
podpis diplomanta

## **ABSTRAKT**

Bakalářská práce je rozdělena na teoretickou a praktickou část. V teoretické části se nachází vypracovaná obecná literární rešerše na téma big data, strojové učení, podnikové procesy a řízení zásob. Dále se teoretická část věnuje významu strojového učení v kontextu big data a podnikových procesů. V praktické části se nachází popis vybrané společnosti a rozbor konsolidovaných účetních výkazů. Následuje aplikace strojového učení za účelem predikce prodejů nářadí této společnosti. Z této predikce je odvozeno doporučení na optimalizaci skladových zásob.

Klíčová slova: big data, strojové učení, podnikový proces, řízení zásob, podnikové systémy

## **ABSTRACT**

The bachelor thesis is divided into theoretical and practical sections. The theoretical section includes a comprehensive literature review on big data, machine learning, business processes and inventory management. Additionally, the importance of machine learning within the context of big data and business processes is discussed. The practical section contains a description of the selected company and an analysis of its consolidated financial statements. This is followed by the application of machine learning in order to predict sales of this company's tools. Recommendations for inventory optimization are derived from these predictions.

Keywords: big data, machine learning, business process, inventory management, enterprise systems

Na začátek bakalářské práce bych rád poděkoval mému vedoucímu práce, panu Ing. Tomášovi Urbánkovi, Ph.D., za vedení bakalářské práce, jeho cenné rady a přátelský přístup. Rovněž bych chtěl poděkovat tátovi, který mi pomohl se získáním potřebných dat.

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 BAKALÁŘSKÉ PRÁCE .....</b>	<b>10</b>
<b>I TEORETICKÁ ČÁST .....</b>	<b>11</b>
<b>1 BIG DATA.....</b>	<b>12</b>
1.1 HISTORIE A BUDOUCNOST BIG DATA .....	12
1.2 VLASTNOSTI BIG DATA.....	14
1.3 VZNIK BIG DAT.....	14
1.4 UPLATNĚNÍ BIG DAT .....	15
1.4.1 SQL a NoSQL databáze.....	16
1.4.2 Etická problematika big dat .....	17
1.5 PREDIKTIVNÍ ANALYTIKA .....	17
1.6 ZPRACOVÁNÍ BIG DAT.....	18
1.7 STROJOVÉ UČENÍ JAKO NÁSTROJ ZPRACOVÁNÍ DAT .....	19
1.7.1 Vymezení pojmu strojového učení.....	19
1.7.2 Obecný postup při zpracování dat pomocí strojového učení .....	19
1.7.3 Jednoduchá lineární regrese .....	21
<b>2 PODNIKOVÉ PROCESY.....</b>	<b>22</b>
2.1 OPTIMALIZACE PODNIKOVÝCH PROCESŮ .....	22
2.2 APLIKACE STROJOVÉHO UČENÍ V PODNIKOVÝCH PROCESECH .....	23
2.2.1 Strojové učení v podnikových systémech.....	24
2.3 PROCES ŘÍZENÍ ZÁSOB .....	25
2.3.1 Hladiny zásob .....	26
2.4 METODY ŘÍZENÍ ZÁSOB.....	27
2.4.1 Výzvy při řízení zásob .....	28
<b>SHRNUTÍ TEORETICKÉ ČÁSTI.....</b>	<b>29</b>
<b>II PRAKTICKÁ ČÁST.....</b>	<b>30</b>
<b>3 PŘEDSTAVENÍ SPOLEČNOSTI.....</b>	<b>31</b>
3.1.1 Popis společnosti .....	31
3.1.2 Popis účetních výkazů .....	32
3.1.3 Zhodnocení účetních výkazů.....	33
<b>4 APLIKACE STROJOVÉHO UČENÍ V PROCESU ŘÍZENÍ ZÁSOB.....</b>	<b>34</b>
4.1.1 Obdržžený datový set .....	35
4.2 PREDIKCE PRODEJŮ POMOCÍ STROJOVÉHO UČENÍ .....	36
4.2.1 Volba nástroje .....	36
4.2.2 Sběr dat .....	36
4.2.3 Popis dat.....	36

4.2.4	Úprava dat .....	36
4.2.5	Rozdělení dat .....	38
4.2.6	Volba algoritmu a učení .....	39
4.2.7	Testování algoritmu .....	39
4.2.8	Aplikace algoritmu .....	40
4.2.9	Výpočet intervalu spolehlivosti a popis grafů .....	41
4.3	INTERPRETACE VÝSLEDKŮ .....	43
4.3.1	Možná omezení predikce .....	44
<b>5</b>	<b>OPTIMALIZACE SKLADOVÝCH ZÁSOB .....</b>	<b>45</b>
5.1	ANALÝZA SOUČASNÉHO STAVU ZÁSOB SORTIMENTU NÁŘADÍ .....	45
5.2	APLIKACE ABC ANALÝZY .....	46
5.2.1	Závěr z ABC analýzy .....	47
5.2.2	Rychlost obratu zásob .....	48
5.2.3	Závěry z výpočtů .....	48
5.3	NÁVRH DOPORUČENÍ .....	49
	<b>ZÁVĚR .....</b>	<b>50</b>
	<b>SEZNAM POUŽITÉ LITERATURY .....</b>	<b>51</b>
	<b>SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK .....</b>	<b>54</b>
	<b>SEZNAM OBRÁZKŮ .....</b>	<b>55</b>
	<b>SEZNAM TABULEK .....</b>	<b>56</b>
	<b>SEZNAM PŘÍLOH .....</b>	<b>57</b>



## ÚVOD

Obstát v dnešním silně konkurenčním prostředí je pro společnosti čím dál těžším úkolem. Proto se organizace snaží zvyšovat svou konkurenceschopnost i pomocí nově dostupných technologií. Jednou z nich jsou takzvaná big data, díky kterým společnosti mohou nalézat nové cesty jak navýšit výnosy, snížit náklady nebo celkově zefektivnit podnikové procesy. Big data představují revoluční proměnu ve způsobu, jakým společnosti zacházejí s informacemi za účelem získávání výhod v dynamicky měnícím se podnikatelském prostředí.

První kapitola teoretické části práce se věnuje vymezení pojmu Big data, charakteristice, jak vznikají, v jakých oblastech se uplatňují a jak se zpracovávají. V této části se také čtenář dozví více o strojovém učení, neboli jedním z nástrojů, jak s těmito daty pracovat. Také se čtenář seznámí s příklady aplikací big data a strojového učení. Druhá kapitola se věnuje podnikovým procesům, především řízení zásob v podniku, což je proces ve společnosti, který jsem si vybral pro účely praktické části. V této kapitole je také uvedeno několik příkladů aplikace strojového učení pro řízení zásob a metod řízení zásob. Kladen důraz je na pojmy vztahující se k tématu a názvu práce. To je význam informačních technologií, strojového učení a big dat v procesu moderního řízení zásob.

Druhá část práce se zaměřuje na popis společnosti Ptáček – velkoobchod, a.s., jejímu postavení na trhu a popisu volně dostupných účetních výkazů. Znalost těchto dokumentů má smysl z hlediska komplexního pochopení finanční situace firmy. Hlavní kapitola praktické části se zabývá predikcí prodejů vybraného sortimentu náradí na následující období pomocí metod strojového učení. Z toho je poté odvozeno doporučení pro společnost. Skladové zásoby jsou také zhodnoceny analýzou ABC. Výpočet je založen na získaných reálných datech a je zde popsáno, jak se při práci s daty postupovalo. Praktická část vychází z poznatků založených na teoretické části.

## CÍLE A METODY BAKALÁŘSKÉ PRÁCE

Hlavním cílem celé bakalářské práce je smysluplně aplikovat algoritmy strojového učení na podnikový proces tak, aby přinesl ekonomický užitek vybrané společnosti.

Cílem teoretické části je vytvoření dostatečného teoretického podkladu pro praktickou část práce. S použitím odborné literatury objasnit big data, strojové učení, definovat podnikové procesy a vyjmenovat současné aplikace strojového učení v podnikových procesech. Dále seznámit čtenáře s pojmy, které s problematikou souvisí.

Cílem praktické části je aplikovat vhodnou metodu strojového učení na podnikový proces. Z výpočtu poté vyvodit a navrhnout úsporná opatření, vedoucí ke snížení nákladů společnosti Ptáček – velkoobchod, a.s. a přinést jí tak ekonomický prospěch.

Metodou, jaká je zpracována teoretická část, je tvorba literární rešerše s použitím více odborných literárních pramenů na pojmy a témata vycházející ze zadání bakalářské práce.

Metodou zpracování praktické části byly znalosti vyplývající z vypracování teoretické části, analýza získaných interních dat a jejich zpracování metodou strojového učení.

Primárním zdrojem informací pro tuto práci bylo množství tištěné literatury zapůjčené v univerzitní knihovně. Protože se technologie big dat a strojového učení velmi rychle vyvíjí, je obtížné nacházet aktualizovanou tištěnou literaturu, které v češtině není mnoho. Proto bylo nutné pracovat i s literaturou zahraniční a jako sekundárním zdrojem bylo zvoleno také množství internetových zdrojů, týkajících se témat big data a strojového učení.

## **I. TEORETICKÁ ČÁST**

## 1 BIG DATA

Slovní spojení Big data naznačuje, že se jedná o pojem týkající se velkých objemů dat. Tento anglický výraz je i v českém prostředí ustálený, a proto budu tento termín v bakalářské práci používat. Můžeme se v literatuře také setkat s doslovným překladem, velká data.

Existuje více definic tohoto pojmu. Přesně vymezená definice, která by byla všeobecně uznávaná, však neexistuje. Všeobecně se velkými daty myslí shromážděné datové sady, které jsou tak velké a komplexní, že k jejich zpracování vyžadují speciální technologie, jako například strojové učení a umělou inteligenci. Hendl (2021, s. 14) uvádí jednu z přijímaných definic pojmu:

*„Big data jsou data s různorodým formátem, o velkém objemu a rychle se měnící, což v souhrnu způsobuje, že je nelze spravovat pomocí konvenčních databázových prostředků.“*

Technologicko-výzkumná společnost Gartner, která působí v oblasti informačních technologií v USA, uznává další, často uváděnou definici:

*„Big Data jsou data, jejichž velikost, rychlost nárůstu a různorodost neumožňují zpracování pomocí doposud známých a ověřených technologií v rozumném čase.“* (Holubová, et al., 2015, s. 19)

### 1.1 Historie a budoucnost Big data

Z historie můžeme vidět, jak se tento koncept vyvíjel od prvotních počítačových systémů až po současný ekosystém založený na cloudových technologiích a umělé inteligenci. Big data byla v minulosti omezena na zpracování velkých objemů strukturovaných dat, ale dnes zahrnuje i množství dat nestrukturovaných. Postupně vznikly metody a techniky pro uchovávání, zpracování a analýzu různých typů dat, jako jsou Hadoop, NoSQL databáze a MapReduce. Ty zásadně zvětšily praktická využití těchto dat. (Holubová, et al., 2015, s. 64)

První článek věnující se detailněji analýze dat se objevuje v 60. letech minulého století. *„V roce 1962 John Tukey popsal oblast, kterou nazval analýza dat. Jeho popis se podobal moderní vědě o datech.“* (Hendl, 2021, s. 19)

Termín "big data" začal být častěji používán v prostředí informačních technologií od 90. let, doby vzniku internetu. Michael Cox a David Ellsworth publikovali v roce 1997 článek

s názvem *Application-controlled demand paging for out-of-core visualization*. To bylo asi poprvé, kdy byl pojem velká data použit v kontextu, který je v souladu s dnešním pojetím:

*„Vizualizace představuje zajímavou výzvu pro počítačové systémy: datové sady jsou obecně poměrně velké a zdaňují kapacity hlavní paměti, místního disku a dokonce i vzdáleného disku. Tomu říkáme problém velkých dat. Pokud se datové sady nevejdou do hlavní paměti (v jádru) nebo když se nevejdou ani na lokální disk, nejběžnějším řešením je získat více prostředků.“* (vlastní překlad, Cox, Ellsworth, 1997, s. 1)

Od přelomu tisíciletí se stal pojem v technologickém průmyslu všudypřítomným. Nejdříve se big dat chopily společnosti jako Google, Amazon, Yahoo a Ebay, které shromažďovaly obrovské množství dat od svých uživatelů v reálném čase. Tyto údaje používaly ke zlepšení svých vyhledávacích algoritmů, doporučování produktů a poskytování personalizované reklamy. Brzy je následovala další odvětví.

Jednou z klíčových událostí, které urychlily růst velkých dat, byl vznik systému Hadoop v roce 2006. Hadoop je software, který umožňuje zpracování velkých datových souborů na větším množství počítačů najednou. Výhodou je, že soubory nemusí být uloženy na jednom velkém uložišti. Hadoop umožnil společností ukládat a analyzovat obrovské objemy dat za zlomek nákladů tradičních metod.

Dalším důležitým vývojem v historii velkých dat byl vzestup cloud computingu. Poskytovatel cloudových služeb Amazon Web Services umožnil od roku 2006 firmám všech velikostí ukládat a zpracovávat velká data, aniž by musely investovat do drahého hardwaru a softwaru. V tomto odvětví brzy následovala další významná společnost Microsoft Azure.

Znalost big data vedla nezdědka k velkým osobním ziskům. Oren Etzioni vyvinul aplikaci, která předpovídala, jestli budou ceny letenek stoupat, nebo klesat. Podle toho doporučovala uživatelům vhodné načasování koupě letenky a tím jim šetřila peníze. Tento startup s názvem Forecast prodal Microsoftu v roce 2008 za 115 milionů dolarů. Ten ho poté implementoval do svého vyhledávače Bing. (Mayerschönberger, 2014, s. 5)

V budoucnu se bude význam velkých dat neustále zvyšovat, protože objem informací generovaných v digitálním světě se zdá mít velký růst. Kontinuálně se vyvíjejí nové technologie a nástroje pro zpracování a analýzu dat. Je zřejmé, že se význam strojového učení a umělé inteligence v analýze velkých dat stane zásadní. Tyto technologie se postupně stávají důležitým nástrojem pro manažery a vedoucí pracovníky pro získávání důležitých podkladů k přesnému rozhodování.

## 1.2 Vlastnosti Big data

Aby se dataset mohl považovat za big data, musí splňovat určité charakteristiky. Autorem těchto charakteristik byl datový specialista Doug Laney, který definoval několik nových pojmů. V roce 2001 publikoval článek, kde poprvé zmiňuje takzvané 3V: volume (objem), velocity (rychlost), variety (rozmanitost). Postupem času se k těmto 3 charakteristikám přidaly ještě termíny veracity (věrohodnost) a value (hodnota). V překladu se tedy big data vyznačují objemem, rychlostí, rozmanitostí, věrohodností a hodnotou.

Objem dat je nejdůležitějším parametrem. V kontextu big dat se jedná minimálně o TB (terabajty) shromážděných dat, avšak častěji spíše o PB (petabajty). Rychlost se týká rychlosti, s jakou jsou data generována, zpracovávána a analyzována. Většinou je požadováno, aby byla data zpracovávána nepřetržitě a téměř okamžitě. Rozmanitost se týká různých typů a formátů dat, která jsou k dispozici, včetně strukturovaných, nestrukturovaných a polostrukturovaných dat. Věrohodnost se týká kvality a přesnosti dat, která mohou být ovlivněna chybami, zkresleními a nekonzistencemi. A nakonec hodnota se týká potenciálních poznatků a přínosů, které lze získat analýzou těchto dat. Získané poznatky poté vedou například k přesnějšímu rozhodování vedoucích pracovníků, lepší zkušenosti zákazníků a vyšší efektivitě a produktivitě. (Erl, Khattak, Buhler, 2016, s. 13)

## 1.3 Vznik big dat

Velká data pocházejí z nejrůznějších zdrojů v různých formátech. Dnešní technologie nám již umožňují, aby takto velká množství dat mohla být shromažďována a ukládána téměř v reálném čase. Nyní budou představeny hlavní zdroje těchto dat.

Zařízení internetu věcí (IoT), jako jsou výrobní stroje a chytré spotřebiče, jsou často vybavené senzory, které generují velká množství dat o svém používání a výkonu. Ty lze využívat pro prediktivní údržbu, optimalizaci a automatizaci. Data z GPS lokátorů mobilních telefonů a chytrých automobilů se používají ke zmírnění dopravní zácpy a v neposlední řadě satelity shromažďující údaje o atmosféře pomáhají tvořit předpovědi počasí. (Kod'ousková, 2023) Podniky a banky generují velká množství dat prostřednictvím transakčních činností a zpracováním plateb. Webová data zahrnují informace shromážděné z webových stránek, vyhledávačů a dalších online zdrojů, které lze využít pro analýzu konkurence, průzkum trhu a identifikaci trendů. Vlady a další veřejné organizace (například Český statistický úřad, World Bank, nebo WHO) poskytují prostřednictvím portálů velká množství dat, která lze volně využívat. V neposlední řadě sociální média generují obrovská množství dat,

pocházející z příspěvků, komentářů, sdílení a chování uživatelů. Díky tomu mohou sledovat informace o preferencích, chování a náladách uživatelů.

## 1.4 Uplatnění big dat

S velkými daty se pojí celá řada příležitostí v různých oblastech ekonomiky. Dnes jsou důležitou součástí většiny průmyslových odvětví, od zdravotnictví přes finančnictví až po maloobchod. Jejich využití zahrnuje předpovídání chování spotřebitelů, vývoj nových léků, ale také optimalizaci řízení dodavatelského řetězce. Velká data zkrátka mají potenciál transformovat jak jednotlivé společnosti, tak celá odvětví. Zde je uvedené množství příkladů konkrétních využití podle různých sektorů: (Česká spořitelna, 2018)

### **Průmysl**

Big data nabízejí společnostem možnost inovací. Prostřednictvím lepšího porozumění potřebám a požadavkům lidí tyto znalosti pomáhají v zavádění zcela nových produktů. Efektivnější využití průmyslových dat vede k inovacím ve výrobě. Analýza velkých dat může pomoci nalézat řešení na problémy ve výrobě a zvyšovat tak produktivitu. Zároveň snižovat náklady, například prostřednictvím předpovědi prodeje nebo optimalizace údržby v továrnách vybavených inteligentními systémy. (Europa, 2021)

### **Obchod a finance**

Na základě předpovědí prodeje mohou společnosti přesněji upravovat své plány, jako je nákup dalšího zboží. Zdravotní pojišťovny mohou přesněji předvídat trendy a stanovovat pojistné, což vede k dosahování většího zisku a jsou tak na trhu konkurenceschopnější.

### **Životní prostředí**

Data ze satelitů mají potenciál vylepšovat výzkum v tomto odvětví a přispívat ke snížení emisí skleníkových plynů. Současně mohou usnadňovat prevenci přírodních katastrof, jako jsou požáry, a lépe se na ně připravovat. Efektivnější výrobní a jiné procesy by mohly vést ke snížení emisí, vyšší produkci a menšímu množství odpadu.

### **Zdravotní péče**

Prozkoumání klinických datových souborů, jako jsou zdravotní záznamy, informace získané dotazníky a data z digitálních metod vyšetření může přispět k vylepšení diagnostiky, přesnější zamíření léčby a pokroku v medicíně.

## **Zemědělství**

S využitím dat ze satelitů a senzorů mohou zemědělci efektivněji hospodařit se zdroji a snížit tak náklady na jednotku vypěstované produkce. Taková data zemědělcům umožňují přesněji nastavovat dávkování hnojiv, herbicidů a lépe hospodařit s vodou.

## **Veřejný sektor**

Využití dat ve veřejném sektoru může přispět ke zvyšování účinnosti veřejných služeb, poskytování lépe přizpůsobených služeb občanům a také občany lépe chránit. V oblasti zabezpečení se nejedná pouze o několik hackerů, ale státy čelí celým týmům odborníků. Bezpečnostní prostředí kyberprostoru a požadavky na regulaci se neustále mění. Velká data umožňují rozpoznávat vzory, které naznačují podvodná jednání a poté na ně varovat ještě dříve, než se stanou. Je tak možné rychleji připravovat hlášení pro regulační a bezpečnostní instituce.

## **Doprava**

S využitím dat získaných z GPS a sociálních médií je možné zmírňovat dopravní zácpy. Zlepšená regulace dopravního provozu přispívá k úspoře času, snižování spotřeby paliva a emisí. Tato data v dnešní době nevyužívají pouze řidiči aut, ale celý logistický sektor, například k přesnější lokalizaci vhodných míst pro logistická překladiště.

## **Vývoj produktů**

Firmy jako Netflix a Procter & Gamble využívají data k předpovídání zákaznické poptávky po jejich produktech. Pro vývoj nových produktů a služeb vytvářejí tyto společnosti prediktivní modely. Z nich se snaží zjistit, jaké mají mít jejich produkty vlastnosti a vzhled, aby se jich prodalo co nejvíce. Dané modely hledají vztah mezi atributy jejich produktů a obchodním úspěchem. Společnosti jako P&G používají data a analýzy získané ze zájmových skupin, sociálních médií a testovacích trhů pro vyhledání vhodného umístění produktů v obchodech, správného načasování a určení optimálního objemu výroby. (Oracle, 2023)

### **1.4.1 SQL a NoSQL databáze**

SQL a NoSQL databáze jsou základní typy databázových systémů a jejich znalost je důležitá pro pochopení správy a analýzy velkých objemů dat.



Databáze jsou optimalizované systémy pro ukládání a vyvolávání dat. Relační databáze je nejrozšířenějším typem databáze. Můžeme ji přirovnat k organizované množině tabulek v programu Microsoft Excel. Je navržena tak, aby umožnila pomocí příkazů dotazovat určité množiny hodnot. K tomuto účelu bylo vyvinuto mnoho dotazovacích jazyků a nejznámějším z nich je právě strukturovaný dotazovací jazyk SQL (z angl. Structured Query Language). Pomocí tohoto jazyka můžeme ovládat a manipulovat relační databáze. Z názvu tohoto jazyka tedy plyne název SQL databáze. (Hendl, 2021, s. 59)

Nedostatky při práci s relačními databázemi vedly ke vzniku NoSQL databází. Jsou vhodnější pro zpracování velkého množství nestrukturovaných dat. Jsou použitelné pro aplikace, které vyžadují vysokorychlostní přístup k velkému množství dat s nízkou latencí, jako jsou platformy sociálních médií nebo webové stránky elektronického obchodu. Další výhodou NoSQL nad klasickou SQL databází je, že je rychlejší ve spojování a úpravách velkého množství tabulek, obsahující velká množství dat. V případě použití klasické relační databáze dochází dříve k zahlcení výpočetního systému. (Holubová, et al., 2015, s. 88)

#### **1.4.2 Etická problematika big dat**

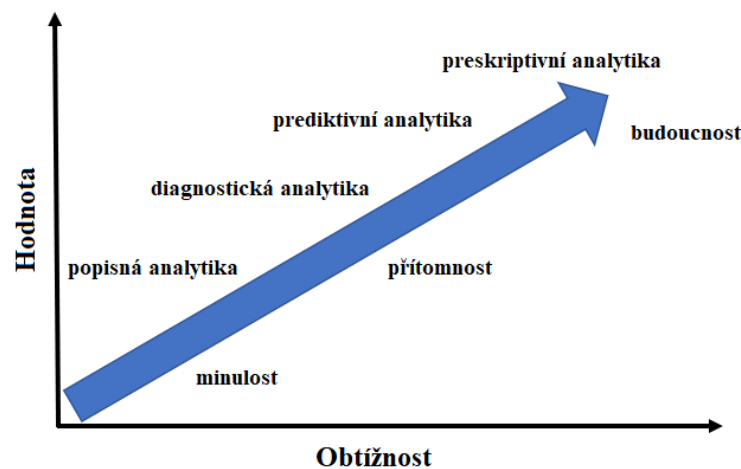
S tématem uplatnění velkých objemů dat se pojí etická problematika, která se s nárůstem cloud computingu začala častěji vyskytovat ve zpravodajství a mediálním prostoru. V současné době se velké společnosti a státní orgány musí se zaváděním těchto technologií zabývat také otázkami týkající se ochrany soukromí a bezpečnosti, aby bylo zajištěno odpovědné nakládání s těmito daty. Evropská Unie tuto problematiku na českém území řeší pomocí právního rámce chránícího osobní údaje, všeobecně známého pod zkratkou GDPR (Zákony pro lidi, 2019). Uplatnění těchto technologií v Číně začíná nést i náznaky zneužívání. Automatické sledování a identifikace lidí, sociální kredit nebo párování nejružnějších typů informací mohou být začátkem technologické diktatury. (Pietraš, 2020) Dalším příkladem zneužití bylo odhalení v roce 2013, že agentura NSA v USA shromažďovala ve svých datacentrech asi 1,7 miliardy emailů a další komunikace, ale také rozsáhlé záznamy platebních transakcí občanů. (Mayerschönberger, 2014, s. 156)

#### **1.5 Prediktivní analytika**

Zmíněná využití dat jsou založena na takzvané prediktivní analytice. Prediktivní modelování je technika zaměřená na předpovídání budoucího chování a událostí. Tato metoda využívá historická a aktuální data k vytváření modelů, které pomáhají předpovídat budoucí výsledky.

Při prediktivním modelování se shromažďují data, vytváří se statistický model, provádějí se předpovědi a model se ověřuje či reviduje, jakmile jsou k dispozici další data. Prediktivní modely analyzují dění v minulosti, aby vyhodnotily pravděpodobnost určitého budoucího chování. Prediktivní analytika například předpovídá poptávku, což umožňuje organizacím využít historická data k odhadu budoucího vývoje, například příštího měsíce. (Hendl, 2021, s. 33 a Gartner Information Technology Glossary)

Pro zařazení prediktivní analytiky do širšího kontextu pomůže často citovaný graf společnosti Gartner, který identifikuje čtyři různé druhy analytik z pohledu obtížnosti a hodnoty. K těmto analytikám také uvádí otázky, čemu se věnují. Popisná analytika (co se stalo?), diagnostická analytika (proč se to stalo?), prediktivní analytika (co se pravděpodobně stane?) a preskriptivní analytika (jak můžeme zajistit, aby se to stalo?). Z grafu na obrázku 1 lze vyčíst, že čím hodnotnější data chceme získat z analýzy, tím je to náročnější. Většinou jsou totiž nejhodnotnější ty informace, které ještě nemůžeme znát, protože se stanou až v budoucnu.



Obrázek 1 Model analytik (vlastní zpracování podle Maoz 2013)

## 1.6 Zpracování big dat

Analytické metody jako statistické, matematické a grafické analýzy jsou pro účely dnešních podniků často pomalé. Společnosti jsou v době informačních technologií čím dál více zahlcené daty v množstvích, která jsou již někdy těžko zpracovatelná i na dostupných počítačích. Proto je také potřeba nalézat rozumné postupy a řešení, jak s nimi pracovat. K analýze big dat se používá strojové učení a k prezentaci big dat grafický software.

Obecně platí při zpracování dat tento postup: definování problému, sběr dat, čištění dat, analýza dat (kvantitativní analýza, kvalitativní analýza, data mining, statistická analýza, strojové učení), prezentace dat a využití dat (vyvození doporučení).

Business intelligence (BI) a datové inženýrství jsou dva nově vzniklé a rychle se rozvíjející obory zabývající se zpracováním dat. Specialisté z oblasti datového inženýrství jsou zodpovědní za návrh, přípravu, udržení a rozšiřování potřebného toku dat. BI analytici ve společnostech provádí kvantitativní analýzy, data vizualizují a nakonec prezentují návrhy efektivních strategií pro řízení firmy. (Hendl, 2021, s. 27)

## 1.7 Strojové učení jako nástroj zpracování dat

Pro úspěšné splnění cíle bakalářské práce je z hlediska přiměřeného rozsahu práce nadbytečné objasňovat technické detaily takto složitého nástroje. Zároveň je tato kapitola nezbytná, jelikož se popsané postupy budou používat v praktické části.

### 1.7.1 Vymezení pojmu strojového učení

Čím dál aktuálnější technologií je v tomto oboru strojové učení. Poskytnout jednoduchou a srozumitelnou definici strojového učení je obtížné. Strojové učení je podmnožinou umělé inteligence. Díky algoritmům a technikám umožňuje počítači učit se na základě dat, aniž by k tomu byl přímo naprogramován. Jednou z klíčových aplikací je analýza velkých objemů dat, u nichž je vzhledem k jejich objemu a složitosti nepraktické používat tradiční metody analýzy. Pomocí algoritmů, které se dokáží automaticky učit z dat, může strojové učení identifikovat vzorce, provádět předpovědi a poskytovat poznatky, které by pro člověka bylo nemožné odhalit. Díky této schopnosti rychle a přesně zpracovávat velké množství dat je strojové učení základním nástrojem pro datové a business analytiky. (Hendl, 2021, s. 26)

### 1.7.2 Obecný postup při zpracování dat pomocí strojového učení

Celý proces strojového učení probíhá s použitím počítačového softwaru. Jupyter Notebook je příkladem aplikace, která umožňuje upravovat a spouštět dokumenty prostřednictvím webového prohlížeče. Tato aplikace bude použita v praktické části.

**1) Sběr dat:** Z různých zdrojů a v různých formátech se shromáždí data.

**2) Úprava dat:** Tento krok spočívá v čištění a úpravě dat. Pro následující práci se soubory je nutné odstranit nesrovnalosti v datech (např. chybějící údaje a extrémní hodnoty). Poté data převedeme do společného formátu, který je vhodný pro aplikaci strojového učení. Často

se v takovém souboru nachází data přebytetěčná. Zjistíme, která data budeme skutečně potřebovat a zbavíme se dat přebytetěčných.

**3) Volba správného modelu:** Zvolíme vhodný model strojového učení. K tomu se často používá knihovna Scikit-learn. Na této webové stránce můžeme nalézt již ověřené modely rozdělené podle funkce, připravené na použití. Vhodný model strojového učení vybíráme podle našich potřeb. Žádný model totiž neřeší všechny možné scénáře a každý má své jedinečné vlastnosti a uplatnění. Modelů strojového učení je celá řada. Nejprve je nutné zvážit, zda se jedná o problém klasifikace, nebo regrese. Mezi základní typy ještě patří algoritmy řešící problémy doporučovací a shlukovací. Pokud jde o klasifikaci, lze použít klasifikační algoritmy, jako jsou náhodné lesy nebo metodu podpůrných vektorů. Pokud jde o regresi, mohou být použity algoritmy, jako je lineární regrese nebo neuronové sítě. Jednoduchá lineární regrese je především vhodná, pokud jsou data lineárně rozdělena a obsahují pouze jednu nezávislou proměnnou. (Zubair, 2023) Dalším faktorem je velikost a rozmanitost datové sady. Je třeba zvážit, jaké jsou požadavky na přesnost a rychlost modelu a najít tu správnou rovnováhu. Některé jsou přesné, ale pomalé, zatímco jiné jsou rychlé, ale méně přesné.

**4) Rozdělení dat:** Data jsou dále rozdělena na dvě části: trénovací a testovací. Trénovací množina je sada dat, na které algoritmus aplikuje učení pro nalezení určitých vztahů a obvykle tvoří přibližně 70 % až 80 % celkových dat. Testovací množina je sada dat, která se používá pro ověření přesnosti naučeného modelu. Měla by být odlišná od trénovací sady a obsahovat zbylých 20 % až 30 % dat. Existují i jiné přístupy k tomuto rozdělení. Jako příklad lze uvést cross-validaci, což je metoda rozdělení dat s využitím různých technik. Mezi tyto techniky patří například K-Fold a Leave-One-Out.

**5) Učení algoritmu:** Vybraný model se „trénuje“ pomocí první sady připravených dat. Hledá v nich vztahy a spojitosti. Tento krok můžeme ovlivnit vhodným nastavením parametrů modelu strojového učení.

**6) Testování algoritmu:** Pomocí druhé sady připravených dat se model testuje. Při tomto kroku se vypočte jeho přesnost a rychlost.

**7) Aplikace:** Když ověřený model na množině dat odpovídá stanoveným požadavkům, lze ho využít na řešení daného problému. Pokud současný model nespĺňuje požadované vlastnosti, je nutné vrátit se ke kroku 5 a lépe nastavit parametry daného modelu. Popřípadě lze vybrat úplně jiný, vhodnější model strojového učení.

### 1.7.3 Jednoduchá lineární regrese

Tato metoda bude použita v praktické části práce, a proto je zde popsána. Lineární regrese je tradiční statistická metoda, která si našla široké uplatnění také v oblasti strojového učení. Vhodná je pro předpovídání prodejů, predikci vývoje na akciovém trhu, nebo například odhadu tržní ceny nemovitosti na základě znalosti více než dvou parametrů, což může být stáří, lokalita a počet pokojů.

Problémy regrese se liší od klasifikačních problémů tím, že se snaží predikovat spojité proměnné. Na získaná data se aplikuje vhodný regresní model. Tento model hledá matematickou funkci, která by co nejlépe popisovala křivku procházející množinou bodů, jejichž hodnoty jsou získány ze shromážděných dat.

Lineární regrese slouží k tomu, abychom získali předpis, pomocí kterého budeme schopni předpovědět hodnotu jedné proměnné ze znalosti hodnoty jiné proměnné. U lineární regrese hledáme neznámé konstanty  $\beta_1$  a  $\beta_0$ . Jejich hodnoty nalézáme nejčastěji pomocí metody nejmenších čtverců. (Berka, 2003, s. 78 a Otyepka, Banáš a Otyepková, 2013, s. 59)

$$\mathbf{Y} = \beta_1 \mathbf{X} + \beta_0$$

Obecný tvar lineární regresní funkce

## 2 PODNIKOVÉ PROCESY

V odborné literatuře se nachází více definic podnikového procesu. Podnikový proces je podle Svozilové (2011, s. 14): „*Série logicky souvisejících činností nebo úkolů, jejichž prostřednictvím – jsou-li postupně vykonány – má být vytvořen předem definovaný soubor výsledků.*“ Řepa (2007, s. 15) podnikový proces popisuje následovně: „*Podnikový proces je souhrnem činností, transformujících souhrn vstupu do souhrnu výstupu (zboží nebo služeb) pro jiné lidi nebo procesy, používající k tomu lidi a nástroje.*“ Nakonec Vyskočil a Štrup (2003, s. 84) definují proces jako: „*způsob, jak se abstraktní cíl, postavený zákazníky na první místo, změni na praktické postupy.*“

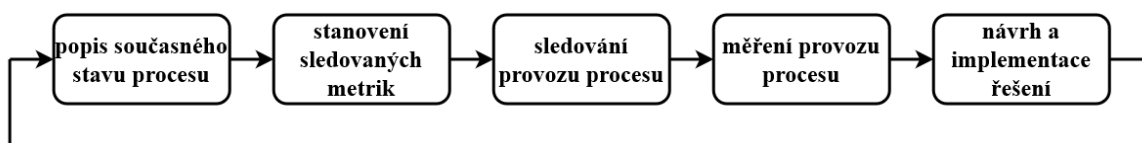
Z těchto definic vyplývá, že proces je soustava souvisejících činností, které společně transformují vstupy na výstupy, čímž vytvářejí hodnotu pro zákazníky a další procesy. Jinými slovy, proces je uspořádaný sled kroků nebo činností, které na sebe navazují s cílem dosáhnout žádoucího výsledku a zároveň vytvářet hodnotu. Procesy obvykle zahrnují lidi a nástroje, které umožňují efektivní a hladké provádění těchto činností.

### 2.1 Optimalizace podnikových procesů

V dnešním rychle se měnícím prostředí může optimalizace procesů přispět k zvýšení efektivity a konkurenceschopnosti podniků. Cílem je analýza a zlepšení stávajících postupů a činností, aby podniky lépe plnily potřeby zákazníků a maximalizovaly hodnotu pro všechny zúčastněné strany. Důležitým měřítkem míry optimalizace je přidaná hodnota pro zákazníka.

Zlepšování podnikových procesů definuje Svozilová (2011, s. 19): „*Činnost zaměřenou na postupné zvyšování kvality, produktivity nebo doby zpracování podnikového procesu prostřednictvím eliminace neproduktivních činností a nákladů.*“

Průběžné zlepšování procesu ve společnosti znázorňuje Řepa (2007, s. 16) obrázkem 2:



Obrázek 2 Průběžné zlepšování procesu (vlastní zpracování podle Řepa, 2007)

Optimalizace procesů zahrnuje identifikaci slabých míst a neefektivit, zavedení kontrolních mechanismů, průběžné sledování a aktualizaci postupů. Zásadní roli v dnešní době hrají informační systémy, neboť umožňují efektivněji plánovat, organizovat a řídit činnosti. Integrace a propojení různých částí dodavatelských řetězců prostřednictvím informačních technologií usnadňuje koordinaci a zvyšuje efektivitu. Informační systémy umožňují flexibilnější postupy, přispívají k úspěšné optimalizaci procesů a udržení konkurenceschopnosti podniků. Technologie strojového učení hraje v podnikových informačních systémech a v optimalizaci podnikových procesů zásadní roli.

Hlavními přístupy k optimalizaci systémů jsou stručně metody: Kaizen, Six Sigma, Lean Manufacturing a Reinženýring. Kaizen se snaží o vytrvalé a postupné zlepšování procesů. Six Sigma se snaží odstraňovat příčiny chyb. Lean Manufacturing eliminuje plýtvání zdroji. Reinženýring představuje kompletní předělání podnikových procesů.

Snahou každé organizace by mělo být neustálé optimalizování podnikových procesů. Motivací pro procesní optimalizaci je, že společnosti potřebují snižovat náklady, vylepšovat kvalitu výrobků a služeb, adaptovat se legislativním požadavkům nebo provádět organizační změny. Optimalizace může vyřešit problémy s kvalitou, když si například zákazníci stěžují na vysokou poruchovost a často produkty reklamují. Pokud dojde k akvizici firmy, je třeba sjednotit podnikové procesy v zakoupené společnosti s mateřskou. V takovém případě je potřeba procesy obou firem sjednotit a synchronizovat. Největším důvodem optimalizace procesů je pro firmy potřeba snižovat náklady, čehož se v praxi často docílí snížením počtu pracovníků, či zvýšením produktivity.

## **2.2 Aplikace strojového učení v podnikových procesech**

Aplikace jsou podobné těm, které již byly zmíněny, jelikož se na zpracování velkých dat strojové učení standardně používá. Další příklady praktického využití strojového učení ve výrobních, či obchodních společnostech jsou následující:

### **Rozpoznávání obrázků**

Rozpoznávání obrázků vyfocených kamerou může být využito k identifikaci vadných nebo poškozených produktů na výrobní lince, což zlepšuje úroveň kvality, snižuje náklady na lidskou pracovní sílu a reklamace. Rozpoznávání obličejů může být použito pro zlepšení zabezpečení a sledování činností zaměstnanců nebo zákazníků v obchodech či kancelářích.

### **Detekce poruch**

Strojové učení umožňuje automatickou detekci anomálií a chyb ve výrobních procesech. Tyto algoritmy mohou odhalit poruchu výrobního stroje ještě před produkčním cyklem, ve kterém k poruše dojde, a ve výrobním softwaru pomáhají proaktivně řešit problémy včas.

### **Doporučovací algoritmy**

Pro oddělení marketingu jsou tyto moderní technologie velmi užitečné, protože zajišťují zpracovávání velkého množství informací o zákaznících, jako je jejich nákupní chování, demografické údaje, věk, pohlaví atd. Díky těmto údajům je poté možné zákazníkům nabízet přesně to, co potřebují. Algoritmy strojového učení také umí analyzovat historii interakcí se zákazníkem a jeho předchozích nákupů, což také umožňuje poskytovat nabídky na míru.

### **Prediktivní modelování**

Učením se na základě vzorců v historických datech mohou algoritmy strojového učení předpovídat budoucí trendy v poptávce, ideální úroveň zásob a další klíčové ukazatele podniků s větší přesností než tradiční metody prognózování. Předpovídat množství prodeje je jednou z možných aplikací strojového učení. Této funkci se bude více věnovat praktická část. Algoritmy strojového učení mohou také například identifikovat vzorce plýtvání nebo neefektivity ve skladování, což manažerům umožní přijmout nápravná opatření. Strojové učení tak pomáhá manažerům zvyšovat rychlost a přesnost jejich rozhodování. Tyto algoritmy mohou také v těchto systémech odhalit opakující se činnosti pracovníků a díky tomu je automatizovat. (Kubina, Varmus a Kubinová, 2015, s. 3)

### **Zákaznická podpora**

Virtuální asistenti založení na technologii strojového učení umožňují zákazníkům poskytnout rychlé a přesné odpovědi na jejich dotazy. To snižuje čekací dobu klientů na odpověď a ulehčuje zatížení lidských operátorů.

#### **2.2.1 Strojové učení v podnikových systémech**

Pro řízení zásob a správné rozhodování o jeho stavu je nezbytný přístup k jeho aktuálnímu stavu. To umožňují Informační a komunikační technologie ICT (z angl. Information and Communication Technology). Zajišťují sběr, analýzu a vyhodnocování dat v každém okamžiku. Pomáhají v plánování, organizaci, provozu a administrativě, spolu se všemi souvisejícími řídicími procesy. Všechny části dodavatelských řetězců a požadavky, které se týkají plánování, organizování, provozu a administrativy jsou napojeny na tyto systémy.



Integrace ICT do dodavatelských řetězců umožňuje efektivnější a rychlejší řízení, což vede ke zlepšení celkové výkonnosti a konkurenceschopnosti podniků. (Emmett, 2008, s. 125) Informační technologie, která poskytuje řešení na zmíněné požadavky je systém ERP.

Součástí podnikových informačních systémů ERP (z angl. Enterprise Resource Planning) je v dnešní době již často implementováno množství funkcí využívajících strojového učení. Implementace strojového učení do systémů ERP v nedávné době přinesla významné pokroky a vylepšení do celkového fungování těchto systémů. Systémy ERP se stávají silným nástrojem pro většinu oddělení společností.

Příkladem aplikace může být detekce poruchy výrobního stroje a na první pohled nepatrně menší produkce. Když se množství vyrobených kusů ve výrobě začne lišit od standardu, systém to ihned pozná. Přesně identifikuje, kde nastala chyba a vyšle varovné oznámení. Na to navazuje část systému spravující finance podniku, která odhadne, že dojde k opožděné fakturaci. Systém na to upozorní příslušné oddělení a navrhne objem potřebných peněžních rezerv, aby nedošlo k nedostatku cash flow. Gatema je příkladem společnosti, která implementuje metody strojového učení do systémů ERP a tím optimalizuje výrobní a podnikové procesy. (GatemaIT, 2023)

Konkrétně ERP systém Helios od společnosti Gatema IT a.s. využívá strojového učení tak, že na základě vyrobených kusů v minulosti předpovídá ve výrobě budoucí počty vyrobených kusů. Díky tomu může snadno vypočítat požadavek na nákup surovin, určit potřebný dostatečný počet zaměstnanců a spočítat vhodné vytížení výrobních strojů. To vše pomůže firmě v tvorbě inventury, jelikož systém disponuje aktuálními informacemi o počtu kusů na skladě a jejich umístění v každém okamžiku. (GatemaIT, 2023)

### 2.3 Proces řízení zásob

Tato kapitola bakalářské práce se již zaměřuje na konkrétní podnikový proces. Strategie a metody řízení zásob a jejich naskladňování se liší mezi výrobními podniky, maloobchodním a velkoobchodním prodejem. Takové firmy běžně generují big data a mají svá analytická oddělení na zpracování těchto dat. Jedni z nejvýznamnějších lídrů v sektoru velkoobchodních společností využívajících těchto dat jsou Amazon.com, Inc. a Wal-Mart Stores, Inc. pocházející z USA. V Česku je zástupcem obchodní společnosti Alza.cz a.s. a výrobní společnosti Škoda Auto a.s. Proces řízení zásob má významný dopad na konkurenceschopnost podniku.

Existuje několik pohledů na to, jak zásoby definovat. „*Pojem zásob může být chápán jako neoddělitelná součást výrobních, obchodních či distribučních subjektů, které takto označují materiál, suroviny, paliva, nářadí, obaly, náhradní díly, polotovary a hotové výrobky.*“ (Jurová, 2016, s. 223) Účetní pohled je dán zákonem o účetnictví č. 563/1991 Sb., který vymezuje zásoby jako materiál, nedokončená výroba a polotovary, výrobky, zvířata, zboží a poskytnuté zálohy na zásoby. V knize Podnikové řízení se dělí zásoba dle funkce na obratovou, maximální, minimální, pojistnou, technickou, sezónní a spekulativní. Je zde uvedena jiná definice: „*Zásoby mají formu vstupů (materiál), výstupů (hotové výrobky) a polotovarů (nedokončená výroba), přičemž přístup ke každé z nich musí být odlišný.*“ (Váchal a Vochozka, 2013, s. 153)

### 2.3.1 Hladiny zásob

Pro řízení zásob jsou často také využívány takzvané řídicí hladiny zásob. Jejich konkrétní stanovení je důležité. Pokud jsou příliš nízké, může to vést k nedostatku zásob a firma nebude moct dodávat zboží, když je potřeba, nebo může docházet k výrobním zpožděním. Pokud jsou příliš vysoké, tak společnost váže příliš mnoho prostředků v zásobách. To znamená pro společnost pomalejší obrátku zásob a s tím související vyšší náklady ušlých příležitostí. Tedy ušlý zisk, jestliže by prostředky vložené do zásob použila jiným způsobem.

**Obratová (běžná) zásoba** se používá pro výpočet průměrné hladiny zásob během daného období. Výše obratové zásoby se pohybuje od maxima k minimu. Průměrná zásoba je ukazatel pro řízení zásob v podniku a je vhodný pro plánování nákupů. Pro výpočet průměrné zásoby se používá uvedený vzorec. (Kalouda, 2015, str. 163)

$$\text{Průměrná zásoba} = \frac{\text{minimální zásoba} + \text{maximální zásoba}}{2}$$

Minimální zásoba je nejnižší úroveň zásob, kterou potřebuje firma udržovat, aby se vyhnula jejich vyprodání. Její výše je v souladu se stavem zásob před příchodem další dodávky.

Maximální zásoba je nejvyšší úroveň zásob, kterou firma má zájem udržovat. Je to výše zásob v době přijetí nové dodávky.

Pojistná zásoba je množství zásob, které firma udržuje navíc k běžné zásobě pro případ neočekávaných událostí. Může to být náhlé zvýšení poptávky, zpoždění dodávek, nebo jiné nepředvídatelné okolnosti, které by mohly ohrozit fungování firmy. Její správně nastavená výše pomáhá snižovat riziko výpadků zásob, aniž by se zbytečně zvyšovaly náklady na skladování. Velikost pojistné zásoby je obvykle stanovena na základě rizikového profilu

firmy a její ochoty akceptovat riziko vyprodání zásob. Firma s vyšší tolerancí rizika může udržovat nižší pojistnou zásobu a naopak. (Váchal a Vochozka, 2013, s. 154)

## 2.4 Metody řízení zásob

Cílem řízení zásob je minimalizovat náklady spojené se skladováním, ale zároveň uspokojovat poptávku zákazníků a zajišťovat hladký chod podniku. Pro dosahování cílů v této oblasti se využívají různé metody. Klíčovou aktivitou v řízení zásob je stanovit optimální množství zásob. Žádná společnost nechce mít příliš velké, nebo naopak příliš malé zásoby. Optimalizace zásob může firmě přinést významný ekonomický přínos. „*Optimalizace zásob má za cíl minimalizaci nákladů na pořízení a skladování při zachování plynulosti výrobního procesu.*“ (Váchal a Vochozka, 2013, s. 153)

### Metoda ABC

Metoda ABC je technika řízení zásob, která kategorizuje položky zásob na základě jejich hodnoty a důležitosti. Metoda vychází z Paretova pravidla, které říká, že 80% zisku tvoří 20% produktů. "A" představuje nejvyšší kumulativní hodnotu a jsou to položky pro společnost nejvýznamnější a nejprodávanější. Těchto položek by mělo být na skladu co nejvíce. "B" položky jsou střední hodnoty a "C" položky jsou nevýznamné, nejnižší hodnoty. Zvoleným parametrem může být obrat podniku a položky kategorizuje podle toho, kolik tržeb podniku přináší. Tato metoda umožňuje efektivnější správu zásob, protože se většina úsilí soustředí na nejdůležitější položky (A), zatímco položky s nižší hodnotou (B a C) vyžadují méně intenzivní dohled. Pro manažery je důležité koncentrovat se především na nejdůležitější skupinu položek. Podle této metody se dají segmentovat například i zákazníci. (Váchal a Vochozka, 2013, s. 156)

### Just-in-Time (JIT)

Tato metoda staví na myšlence doručování zásob do výroby přesně v potřebný moment, což minimalizuje nutnost zboží skladovat. Podstatou je, že dva blízké články dodavatelského řetězce velmi úzce spolupracují a odběratel dostává materiál přesně v čas, kdy jej potřebuje. Pokud je dodavatel přesný, není potřeba vytvářet vlastní zásobu. Je tedy až paradoxní mluvit o JIT jako o metodě řízení zásob, vzhledem k tomu, že ideální aplikace této metody vede k nulovému stavu zásob. JIT vede k nižším nákladům na skladování, ale vyžaduje těsnou spolupráci s dodavateli a efektivní řízení logistiky. (Váchal a Vochozka, 2013, s. 158)

### **Řízení zásob dodavatelem (VMI)**

VMI (z angl. Vendor Managed Inventory) je metoda, při které dodavatel spravuje zásoby zákazníka a zodpovídá za jejich doplňování. Tento způsob řízení zásob byl vyvinut společností Walmart koncem 80. let 20. století. Princip VMI spočívá ve zřízení spolupráce založené na důvěře a koordinaci, která je zajištěna prostřednictvím sdílení dat v podnikovém informačním systému o aktuálním stavu zásob. Dodavatel má neustále přístup k aktuálním informacím o zásobách u zákazníka a předpokládané poptávce, což využívá při rozhodování o množství a čase dodávky. Implementace VMI vyžaduje elektronickou komunikaci mezi dodavatelem a zákazníkem. Tento přístup vede ke zlepšení efektivity dodavatelského řetězce a udržení nižších hladin zásob, což přináší úspory oboum stranám. (Jurová, 2016, s. 231)

#### **2.4.1 Výzvy při řízení zásob**

Řízení zásob s sebou přináší mnoho výzev. Efektivní řízení zásob může vést ke snížení nákladů, zlepšení výkonnosti a zvýšení spokojenosti zákazníků. Naopak neefektivní řízení zásob může způsobit vyšší náklady a poškození pověsti podniku.

Výrobky s omezenou trvanlivostí nebo rychle měnící se technologiemi mohou zastarat a ztrácet hodnotu. Podniky musí efektivně řídit své zásoby, aby minimalizovaly tato rizika.

Sezónnost a fluktuace poptávky ovlivňuje potřebu zásob v průběhu roku. Podniky se musí přizpůsobovat měnícím se podmínkám trhu. Efektivní řízení zásob zohledňuje sezónní výkyvy a fluktuaci poptávky při plánování nákupu zboží.

Spolupráce s důvěryhodnými dodavateli a správa dodacích lhůt je nezbytná pro bezproblémové naskladňování zásob. Podniky si musí zajistit, aby dodavatelé dodržovali smlouvené termíny a měli dostatečnou kapacitu k dodání požadovaného množství zboží.

Efektivní skladování vyžaduje správné využití skladového prostoru a jeho uspořádání. Podniky by měly optimalizovat skladovací kapacitu a uspořádání plochy tak, aby minimalizovaly manipulační náklady a zajistily snadný přístup k zásobám.

## SHRUTÍ TEORETICKÉ ČÁSTI

V teoretické části byly shrnuty základní pojmy. Jednalo se především o klíčová slova big data, strojové učení a podnikové procesy. Důraz byl kladen také na řízení zásob a podnikové informační systémy. Každé z těchto částí byla věnována samostatná kapitola, rozvedená do několika podkapitol. Celý text čerpal z mnoha knižních a internetových zdrojů a byl doplněn o pár obrázků, týkajících se tématu.

První kapitola začínala definicí big dat a následoval souhrn historie této technologie. Poté byl uveden nutný výtazek z literatury týkající se základních pojmů a charakteristik big dat. Byly objasněny jejich vlastnosti a pojmy SQL a NoSQL datábáze. Další kapitola řešila, kde big data vznikají. Byly zde vyjmenovány hlavní zdroje těchto dat. Z hlediska aktuálního vsudypřítomného sběru big dat byla nejzajímavější kapitolou jejich aktuální uplatnění.

Druhá kapitola začínala stručným úvodem a definicí podnikových procesů. Poté bylo blíže rozvedeno téma optimalizace podnikových procesů. Z této teoretické podkapitoly se postupně přenášela pozornost na konkrétní aplikace strojového učení v podnikových procesech. Bylo zde rozepsáno, jak využívají podniky strojového učení pro zvyšování svého zisku díky úsporám nákladů, nebo zvyšování efektivity výroby a prodeje. Aplikace algoritmů strojového učení nacházíme v dnešní době téměř ve všech průmyslových odvětvích. Bakalářská práce se nejvíce věnuje využití této technologie v obchodní a výrobní oblasti, čemuž napomáhají podnikové systémy ERP, jak zmínila následující kapitola. Technologie strojového učení je totiž velmi často implementována do těchto systémů. Kapitola například zodpověděla na otázku, jak pomáhají systémy ERP v řízení zásob. Řízení zásob je problematikou, kterou se zabývala poslední podkapitola, ale především praktická část. Bylo zde zmíněno několik dalších metod, konkrétně ABC, JIT a VMI, které se k řízení zásob používají. Metodou řízení zásob v praktické části je aplikace strojového učení. Nejlepší volbou pro komplexní podnikové řízení je vycházet ze všech těchto metod, nebo s nimi být alespoň obeznámen.

## **II. PRAKTICKÁ ČÁST**

### 3 PŘEDSTAVENÍ SPOLEČNOSTI

Název: Ptáček – velkoobchod, a.s.

Sídlo: U Velké ceny 413/4, Kohoutovice, 623 00 Brno

IČO: 25501143

Vznik společnosti: 1992

Právní forma: akciová společnost



Obrázek 3 Logo společnosti Ptáček – velkoobchod, a.s. (Ptáček, 2023)

#### 3.1.1 Popis společnosti

Společnost Ptáček - velkoobchod, a.s., byla založena Zdeňkem Ptáčkem v roce 1992 a během své historie prošla významným růstem a expanzí. Za firmou stojí vedle zakladatele Zdeňka Ptáčka, který ji buduje už téměř 30 let, i jeho dva synové Jiří a Zdeněk, mezi něž je firma také podílově rozdělena. Ti figurují ve firmě jako členové správní rady a předsedou správní rady je zakladatel Ing. Zdeněk Ptáček.

Ptáček – velkoobchod, a.s. je největším prodejcem kompletního sortimentu topení, plyn, voda, vybavení koupelen, inženýrských sítí, krbů a kamen v České republice a na Slovensku. Ve svém sortimentu má také klimatizace, řízené větrání, solární kolektory a příslušenství. Dále zákazníkům nabízí velké množství služeb, od zpracování aktuální nabídky zboží, technického poradenství v oboru, servisu, školení o novinkách, rozvozu zboží dle požadavků zákazníka, až po poradenství v oblasti dotací alternativních zdrojů. Zákazníky jsou odborné topenářské a instalátérské firmy, developerské společnosti, stavební firmy, ale i koncoví spotřebitelé. Součástí rozsáhlé prodejní sítě jsou také specializovaná studia koupelen známé pod názvem Koupelny Ptáček. Všechny tyto služby by společnost nemohla poskytovat bez velkého počtu zaměstnanců, který v roce 2021 činil zhruba 1900 osob.

Hlavní sídlo společnosti se nachází v brněnských Modřicích, ale neustále se rozšiřuje síť prodejen po celém území České republiky a Slovenska, s hlavními logistickými centry v Brně, Praze, Olomouci, Jihlavě, Bratislavě a Žilině. Společnost je tedy lidem snadno

dostupná díky velmi husté síti 120 poboček a garantuje zákazníkům stejné nákupní podmínky při nákupu v kterékoli z nich, viz obrázek 4.



Obrázek 4 Rozložení poboček v roce 2023 v České Republice (Ptáček, 2023)

Společnost je akcionářem společnosti EDT GmbH - Evropského sdružení profesionálních velkoobchodů - rodinných firem v oblasti sanitárních zařízení a topení. Díky tomuto kroku může nabízet zboží od exkluzivních evropských dodavatelů, které jinde zákazník nenajde a vyjednávat nižší nákupní ceny. Dále je díky tomuto spojení na trhy Česka a Slovenska zavedena celoevropsky známá nabídka produktů pod označením CONCEPT. (Ptáček, 2023)

### 3.1.2 Popis účetních výkazů

Popis konsolidovaných výkazů, konkrétně rozvahy a výkazu zisků a ztrát, představuje souhrnný pohled na finanční situaci společnosti. V rámci představení společnosti stručně popíši, jak si firma finančně vede z hlediska těchto účetních dokumentů. Takové informace mohou pomoci v utvoření komplexního návrhu na úsporná opatření ve společnosti.

Z nahlédnutí do majetkové struktury společnosti lze vidět, že celková bilanční suma byla 5 654 503 tis. Kč. Z toho vlastní kapitál společnosti v roce 2021 činil 3 408 591 tis. Kč. V nejnovějším dostupném výkazu zisků a ztrát se výrazně projevil zvýšený zájem Čechů o rekonstrukce domácností a investice do nemovitostí, který výrazně posílil během nedávné pandemie Covidu. Obrat společnosti se zvýšil v roce 2021 ve srovnání s rokem 2020 zhruba o dvě miliardy Kč. Výsledek hospodaření meziročně výrazně narostl a dosáhl 618 263 tis. Kč, to je zhruba o 256 milionů čistého zisku více, než v předchozím roce. Z údajů výsledku hospodaření a vlastního kapitálu můžeme dále spočítat rentabilitu vlastního kapitálu ROE, která činila v roce 2021 přibližně 18 %.



Průměrná marže na zboží v procentech, kterou společnost měla v letech 2020 a 2021, představuje výnos z prodeje. Výsledek bude využit v analytické části. V roce 2020 byl objem tržeb za zboží 8 679 461 tis. Kč a náklady vynaložené na prodané zboží činily 6 469 008 tis. Kč. Tržby za prodej zboží dosahovaly v roce 2021 10 720 051 tis. Kč a náklady vynaložené na prodané zboží činily 8 086 157 tis. Kč. Dosazením těchto čísel do vzorce, kde  $m$  je marže,  $p$  je prodejní cena a  $n$  je nákupní cena, získáme průměrnou marži na zboží v %, se kterou společnost pracovala. V obou letech se jednalo přibližně o 25 %. Tento postup má svá omezení, protože nemůžeme spočítat marži jednotlivého zboží.

$$m = \left( \frac{p - n}{p} \right) \times 100$$

Výpočet pro roky 2020 a 2021:

$$\frac{8\,679\,461 - 6\,469\,008}{8\,679\,461} \times 100 \doteq 24,57 \%$$

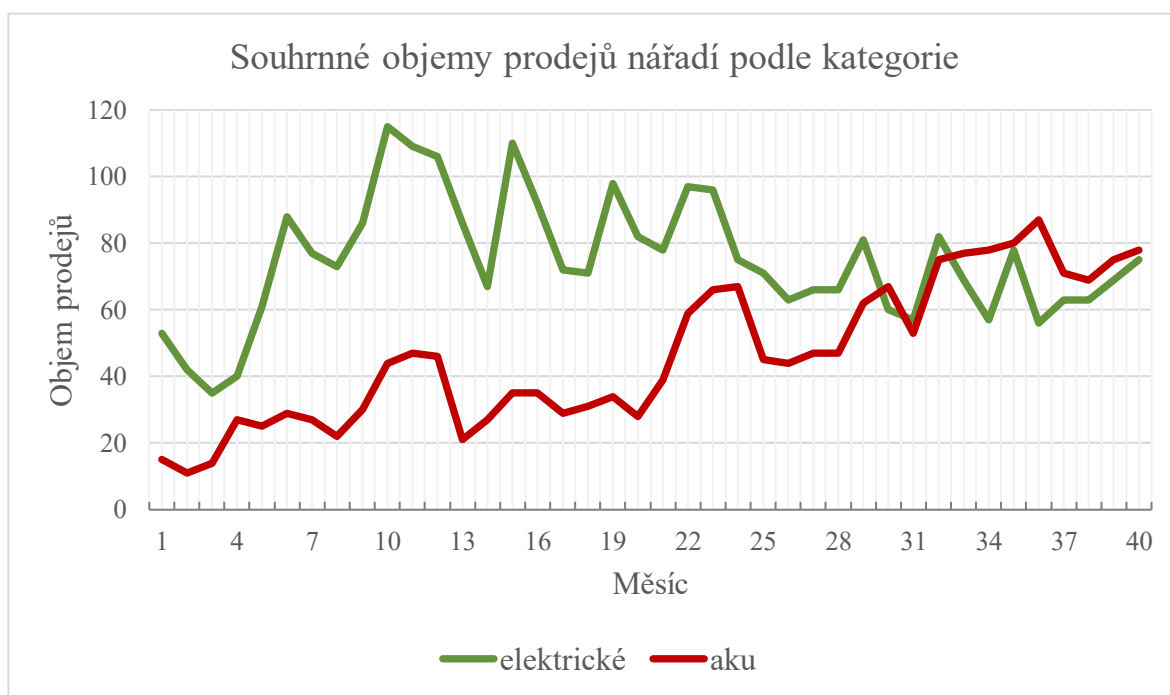
$$\frac{10\,720\,051 - 8\,086\,157}{10\,720\,051} \times 100 \doteq 24,57 \%$$

### 3.1.3 Zhodnocení účetních výkazů

Na základě uvedených údajů se zdá, že finanční zdraví této společnosti je poměrně silné. Pozitivní je, že vlastní kapitál společnosti je významný a činí více než polovinu celkových pasiv. Vyšší podíl vlastního kapitálu také naznačuje, že společnost není příliš závislá na dluhovém financování, což je obecně považováno za zdravé. Rovněž obrat společnosti výrazně vzrostl v roce 2021 oproti roku 2020, což je další pozitivní signál. Výsledek hospodaření rovněž významně vzrostl, což ukazuje zvýšení provozní efektivity a ziskovosti společnosti. Vyšší hodnota ROE obecně naznačuje, že společnost efektivně využívá svůj vlastní kapitál ke generování zisku. Bez konkrétních údajů o průměrném ROE v tomto odvětví je těžké posoudit, zda je dobré nebo špatné. Obecně lze říci, že pokud je ROE výrazně vyšší než průměrná úroková sazba, jedná se o efektivní využití vlastního kapitálu. Celkově, vzhledem k významnému růstu obratu a zisku, silnému vlastnímu kapitálu a efektivnímu využití vlastního kapitálu, lze konstatovat, že si společnost vede velmi dobře. Optimální čisté ziskové rozpětí je do značné míry závislé na specifikách konkrétního odvětví. Pro společnost je důležité sledovat konkurenční prostředí a porovnávat každoroční vlastní čisté ziskové marže s těmi, které vykazují konkurenti.

## 4 APLIKACE STROJOVÉHO UČENÍ V PROCESU ŘÍZENÍ ZÁSOB

Na úvod této kapitoly bude nejprve shrnuto, jak se při práci postupovalo. Je to důležité pro lepší přehlednost. Po obdržení dat od společnosti Ptáček – velkoobchod, a.s. byl vytvořen graf, který spočíval v počtu objemu prodejů v závislosti na kategorii zboží. Bylo zjištěno, že prodejnost nářadí napájeného baterií stoupá, zatímco objem prodejů kabelového nářadí v průběhu let spíše stagnuje. Tohoto zajímavého trendu bylo povšimnuto už dříve v samotné společnosti. Zobrazeno grafem na obrázku 5. V grafu je zobrazeno období od ledna 2020 do dubna 2023, jednotlivé zlomy představují data za každý měsíc. Z tohoto zjištění vyplývá potenciál přesněji stanovit objem zásob tohoto sortimentu tak, aby to bylo pro společnost co nejvýhodnější. K tomu byl použit algoritmus strojového učení a znalosti prediktivní analytiky. Je u toho vycházeno z předpokladu, že by bylo vhodné do budoucna snížit poměr zásob u produktů, jejichž prodejnost bude klesat, nebo stagnovat a zvýšit poměr zásob u produktů, jejichž prodejnost bude stoupat. Smyslem tedy je ušetřit společnosti náklady spojené s vázaností prostředků v zásobách. Na základě výsledků byla společnosti navržena úsporná opatření na budoucí období. Podnikovým procesem, ve kterém se tato úsporná opatření hledají, je proces řízení zásob a s tím související proces nákupu zboží. Takový výstup může být použitelný jako pomocný bod pro nákupčího firmy, když se rozhoduje o množstvích, kolik by měl zboží objednat.



Obrázek 5 Graf souhrnných objemů prodejů nářadí podle kategorie (vlastní zpracování)

#### 4.1.1 Obdržený datový set

Na obrázku 6 můžeme vidět drobnou ukázkou dat, která jsem obdržel od analytického útvaru společnosti. Tabulka obsahuje všech 243 různých druhů kabelového a 229 druhů akumulátorového nářadí, která firma nabízí svým zákazníkům. Elektrické aku v práci znamená, že se jedná o nářadí napájené baterií a elektrické znamená, že se jedná o nářadí napájené kabelem z elektrické rozvodné sítě. Pro konkrétnější představu, nářadí na baterii, které firma nabízí, jsou například vrtačky, pily a kladiva. Kabelovým nářadím jsou v tabulce zastoupené taktéž vrtačky, pily, rádia, vysavače, nebo třeba brusky. V záhlaví tabulky z levé strany vidíme: ID produktu, Zboží – popis, Vlastnost (elektrické / elektrické aku.), Jednotková cena, Zásoba – množství a nakonec je v tabulce seznam měsíců. Pod seznamem jednotlivých měsíců se v tabulce nachází čísla označující počty prodaných kusů výrobků. Čísla označují, kolik se jich za každý kalendářní měsíc od ledna 2020 do dubna 2023 prodalo. ID produktu je pouze číslo, pravděpodobně používané pro identifikaci zboží pomocí čárového kódu. Zboží – popis obsahuje stručný popis a základní charakteristiku zboží. Vlastnost výrobku říká, zda se jednalo o nářadí elektrické, nebo elektrické aku. V datech se dále nachází doporučené ceny jednotlivých výrobků, za kolik je firma přibližně nabízí. Další informací je počet kusů, kolik jednotlivých nářadí měla firma na skladě k datu 1.5.2023. Přesné nákupní ceny nemohly být zveřejněny, jelikož se jedná o citlivé interní informace společnosti a poškodilo by to budoucí vyjednávací pozici s dodavateli. Proto se tuto hodnotnou informaci, se kterou by šlo dobře pracovat, nepodařilo získat.

V příloze práce se nachází tento popisovaný dataset, ze kterého jsem vycházel, s názvem *Export\_dat.csv*. Ne všechny informace z tohoto souboru jsou pro účely analýzy v praktické části důležité.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	ID	Zboží - popis	Vlastnost	Jednotková cena	Zásoba - množství	2020 - Jan	2020 - Feb	2020 - Mar	2020 - Apr	2020 - May	2020 - Jun	2020 - Jul	2020 - Aug	2020 - Sep	2020 - Oct	2020 - Nov	2020 - Dec
2	20800900	Nářadí elektrické Bosch	elektrické	7929	1			1	1	1	1						
3	20800900	Nářadí elektrické Bosch	elektrické	3019	0	1											
4	20800900	Nářadí elektrické Bosch	elektrické	3019	0				1								
5	20800900	Nářadí elektrické Bosch	elektrické	5286,97	1												
6	20800900	Nářadí elektrické Bosch	elektrické	5863,73	0												1
7	20800900	Nářadí elektrické Bosch	elektrické	4554,72	0									1			
8	20800900	Nářadí elektrické Bosch	elektrické	4236,57	8							1	1				
9	20800900	Nářadí elektrické Bosch	elektrické	10165,54	0						1						1
10	20800900	Nářadí elektrické Bosch	elektrické	10350	0												
11	20800900	Nářadí elektrické Bosch	elektrické	2401,48	0										1		
12	20800900	Nářadí elektrické Bosch	elektrické	2231	0						1						
13	20800900	Nářadí elektrické Bosch	elektrické	4908,9	0						1						
14	20800900	Nářadí elektrické Bosch	elektrické	5890,94	0									1			
15	20800900	Nářadí elektrické Bosch	elektrické	1835,87	0												

Obrázek 6 Obdržená data (vlastní zpracování)

## 4.2 Predikce prodejů pomocí strojového učení

Na základě teoretického podkladu získaného z teoretické části, z kapitoly „Obecný postup při zpracování dat pomocí strojového učení“, je následovně postaven postup pro prediktivní analýzu v praktické části.

### 4.2.1 Volba nástroje

V případě většího množství dat, které tabulka obsahuje, by bylo komplikované vypočítávat výsledná čísla na kalkulačce. Další možností je použití funkce v Excelu, ale v případě větších objemů dat se Excel stává pomalým ve srovnání s metodami strojového učení. Excel je dále omezen limitem řádků, což je problém u velkých souborů typu big dat. Využití nástroje strojového učení je univerzálnější a především výhodnější, pokud chceme model automatizovat. Při změně vstupních dat je poté rychlejší získat nový výsledek, pokud jsou vstupní data komplikovaná.

### 4.2.2 Sběr dat

Data byla exportována z interního systému ERP ve formátu .xlsx souboru, který lze otevřít v programu Microsoft Excel. Nemám přístup k informacím, jak konkrétně data ve společnosti systém ERP zaznamenává a zpracovává. Obdržená data jsou pravá a byla vygenerována přímo analytickým oddělením společnosti.

### 4.2.3 Popis dat

V obdržených datech je možné nalézt všechny důležité informace, které na predikci budou potřeba. Data, se kterými se bude pracovat při predikci jsou: Vlastnost (elektrické, nebo elektrické aku.) a objemy prodejů náradí za jednotlivé měsíce. Zbylá data nebudou potřeba vůbec, nebo budou využita až v dalším kroku, kdy se výsledky budou analyzovat. Prozatím budou data odložena, aby se k nim mohlo vrátit později.

### 4.2.4 Úprava dat

Nejprve se musí data připravit do vhodného tvaru tabulky. Dalším krokem je tedy úprava dat do potřebné formy vhodné pro aplikaci strojového učení. V datech nebyly nalezeny žádné extrémní hodnoty, které by mohly být chybou a měly by být smazány. Byl obdržen pouze jeden soubor v jednom formátu, tudíž se nevyskytl problém převodu dat do jednoho společného formátu.

Pro tento krok je potřeba nejdříve otevřít Jupyter Notebook a provést import a instalaci knihoven, které budou potřeba. Jedná se o několik nástrojů, avšak nejpodstatnější je knihovna od scikit-learn s názvem LinearRegression, v níž se skrývá samotný model strojového učení. Pomocí knihovny scipy.stats bude nakonec určen 95 procentní interval spolehlivosti této predikce a díky knihovně matplotlib.pyplot bude tento graf zobrazen. Data jsou nahrána do tohoto prostředí a nářadí rozděleno do dvou kategorií podle vlastností, tedy „elektrické“ a „elektrické aku“. Zároveň příkazem .sum sečteny objemy prodejů za každý měsíc, podle vlastnosti produktu.

```
In [1]: # Import knihoven
import pandas as pd
import numpy as np
import openpyxl
import datetime
import statsmodels.api as sm
import scipy.stats as stats
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error, r2_score

In [2]: # Import dat a pre-processing → rozdělení objemu prodejů do dvou kategorií po měsících
df = pd.read_excel('Export_dat.xlsx', sheet_name='Sheet1')
df = df.groupby('Vlastnost').sum()
```

Obrázek 7 Import knihoven a dat (vlastní zpracování)

V rámci úpravy dat je potřeba zbavit se dat, která nebudou potřeba. Příkazem se zbaví nepotřebných sloupců tabulky, konkrétně 'ID', 'Zboží - popis', 'Jednotková cena', 'Zásoba – množství' a zůstanou pouze ty sloupce, které jsou potřeba pro tuto predikci. Dále jsou osy tabulky převráceny pro přehlednost a je zvolen vhodnější formát datumu. Po těchto úpravách se původní Excelová tabulka velmi zmenšila. Na obrázku číslo 8 je vidět prvních 10 řádků nové tabulky, která obsahuje sečtené objemy prodejů za každý měsíc, rozdělené do kategorií.

```
In [4]: # Zbavení se dat, která nebudou potřeba
df = df.drop(['ID', 'Zboží - popis', 'Jednotková cena', 'Zásoba - množství'], axis=1)
# Transponování tabulky a lepší formát času
df = df.transpose()
df.index = pd.to_datetime(df.index, format='%Y - %b')
```

```
In [5]: # Takto nyní vypadá tabulka ve tvaru, kterém potřebujeme
df.head(10)
```

```
Out[5]:
```

Vlastnost	elektrické	elektrické aku.
2020-01-01	53.0	15.0
2020-02-01	42.0	11.0
2020-03-01	35.0	14.0
2020-04-01	40.0	27.0
2020-05-01	61.0	25.0
2020-06-01	88.0	29.0
2020-07-01	77.0	27.0
2020-08-01	73.0	22.0
2020-09-01	86.0	30.0
2020-10-01	115.0	44.0

Obrázek 8 Úprava dat (vlastní zpracování)

Nyní je třeba zvolit jednu z kategorií, v níž je zájem predikovat vývoj prodejů. To je zobrazeno obrázkem 9. Jelikož se pracovalo pouze se dvěma kategoriemi, je jednodušší poté celý proces zopakovat pro druhou kategorii, než vytvářet speciální script, který by tento proces automatizoval. V reálném použití v praxi je to však většinou potřeba. Proto druhou kategorii následující postup bude tedy zopakován stejně, což je snadný a rychlý proces v prostředí Jupyter Notebook.

```
In [6]: # Nejprve si zvolíme pouze jednu kategorii produktů, například "elektrické aku." (nářadí na baterii)
elektrické_df = df.drop(['elektrické'], axis=1)

In [7]: # Tabulka vypadá nyní takto
elektrické_df.head(10)

Out[7]:
```

Vlastnost	elektrické aku.
2020-01-01	15.0
2020-02-01	11.0
2020-03-01	14.0
2020-04-01	27.0
2020-05-01	25.0
2020-06-01	29.0
2020-07-01	27.0
2020-08-01	22.0
2020-09-01	30.0
2020-10-01	44.0

Obrázek 9 Volba kategorie produktu (vlastní zpracování)

Dále je datum potřeba převést do indexu, protože model lineární regrese umí pracovat pouze s číselným datovým typem, obrázek 10. Nyní proměnná X obsahuje pořadové číslice od 0 do 39, protože má data za 40 měsíců. Proměnná Y obsahuje sečtené objemy prodejů za každý měsíc vybrané kategorie.

```
In [8]: # Datum si převedeme do indexu, protože lineární regrese umí pracovat pouze s čísly a ne daty
X = np.array(range(len(elektrické_df)))[:, np.newaxis]
y = elektrické_df['elektrické aku.'].values
```

Obrázek 10 Volba kategorie nářadí, převedení datumu do indexu (vlastní zpracování)

#### 4.2.5 Rozdělení dat

Nyní data již lze rozdělit do dvou sad. První sada je sada trénovací. Obsahuje 80% položek. Druhá sada je sada testovací a obsahuje zbylých 20% položek. To je zajištěno vynásobením koeficientem 0,2 v příkazu „test\_size=0,2“, jak je vidět na obrázku 11. Parametr random\_state ovlivňuje proces náhodného rozdělení dat na tréninkovou a testovací množinu.

```
In [9]: # Rozdělení dat do trénovací a testovací sady, zvolil jsem poměr 80% a 20%
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=41)
```

Obrázek 11 Rozdělení dat (vlastní zpracování)

#### 4.2.6 Volba algoritmu a učení

Přichází fáze volby modelu. Na tento typ problému lze použít celá řada modelů, jako například LSTM, Xgboost, nebo Decision tree. V případě většího množství dat by mohly být přesnější, ale zároveň by bylo nutné u těchto modelů upravovat jejich takzvané hyperparametry. I v případě, že by byla neuronová síť přesnější, velmi špatně by se poté výsledky vysvětlovaly. V práci bude použit model strojového učení s názvem lineární regrese. Díky své snadné interpretovatelnosti a univerzálnosti je pro tento problém vhodný. Další podstatnou vlastností, která hrála roli při výběru byla obrovská jednoduchost této metody, díky absenci hyperparametrů a jejímu snadnému použití. To všechno jsou důvody, proč byl zvolen tento algoritmus, jak je vidět na obrázku 12. Dále je model aplikován na trénovací sadu dat a nechán se „naučit“ naše data.

```
In [10]: # Aplikace modelu strojového učení, Lineární regrese
model = LinearRegression()
model.fit(X_train, y_train)

Out[10]: LinearRegression
LinearRegression()
```

Obrázek 12 Volba modelu strojového učení (vlastní zpracování)

#### 4.2.7 Testování algoritmu

Nyní je otestován a vyhodnocen vybraný model. Na obrázku 13 je vidět, jak se algoritmus testuje na druhé sadě dat. Ve spodní části jsou vidět statistické veličiny, určující přesnost zvoleného modelu. Čím nižší jsou tyto veličiny, tím je model přesnější. R-squared, neboli koeficient determinace v hodnotě 0,89 je dobrý výsledek a model můžeme vyhodnotit jako vhodný k použití.

```
In [11]: # Vytvoření predikce na testovací soubor
y_pred = model.predict(X_test)

In [12]: # Vyhodnocení přesnosti modelu
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
print("Mean Squared Error: {:.2f}".format(mse))
print("R-squared: {:.2f}".format(r2))

Mean Squared Error: 51.03
R-squared: 0.89
```

Obrázek 13 Testování modelu a vyhodnocení (vlastní zpracování)

#### 4.2.8 Aplikace algoritmu

Aplikací modelu se dostane výstup v podobě 12 čísel, znamenajících predikci objemů prodeje na následujících 12 měsících. To je posledním krokem pro tuto kategorii náradí, protože je predikce hotová. Nyní celý sled těchto kroků je třeba zopakovat stejně pro druhou kategorii náradí. V případě potřeby je možné v tomto kroku přepsat číslo „+ 12“ a získat tak predikci na libovolný počet měsíců. Většinou se však doporučuje predikovat pouze objemy na nejbližší následující období, protože poté výrazně stoupá nepřesnost, jak bude znázorněno grafem intervalu spolehlivosti.

```
In [13]: # Budoucí předpověď
future_X = np.array(range(len(elektrické_df), len(elektrické_df)+12))[:, np.newaxis]
future_y = model.predict(future_X)
print("Future Predictions:", future_y)

Future Predictions: [79.08814726 80.68759517 82.28704308 83.88649099 85.48593889 87.0853868
88.68483471 90.28428262 91.88373052 93.48317843 95.08262634 96.68207425]
```

Obrázek 14 Aplikace modelu a výsledek (vlastní zpracování)

Nakonec je potřeba vytvořit Excelový soubor, kam se nechají získané predikce vepsat. Tento krok je libovolný, protože je predikce již hotová. Obecně je to považováno za správný, ne-li nutný postup, protože je výstup rovnou zobrazen v Excelovém souboru a neudělá se chyba ručním přepisováním. V praxi je to vyžadováno, aby data byla automaticky připravena pro interpretaci.

```
In [14]: # Vytvoření souboru Excel, kam se data uloží
workbook = openpyxl.Workbook()

# Výběr aktivního pracovního listu
worksheet = workbook.active

# Nastavení záhlaví sloupců
worksheet.cell(row=1, column=1, value="Produkt")
worksheet.cell(row=1, column=2, value="Měsíc")
worksheet.cell(row=1, column=3, value="Predikovaná hodnota")
# Cyklus pro vepsání předpokládaných hodnot pro každý měsíc
now = datetime.datetime.now()
for i in range(len(future_y)):
    next_month = now + datetime.timedelta(days=30*i)
    value = future_y[i]
    worksheet.cell(row=i+2, column=1, value='elektrické aku.')
    worksheet.cell(row=i+2, column=2, value=next_month.strftime("%Y-%m"))
    worksheet.cell(row=i+2, column=3, value=value)

In [15]: # Uložit .xlsx soubor do počítače
workbook.save("Předpověď_pro_elektrické_aku.xlsx")
```

Obrázek 15 Tvorba Excelového souboru (vlastní zpracování)



	A	B	C	D
1	Produkt	Měsíc	Predikovaná hodnota	
2	elektrické aku.	2023-05	79,08814726	
3	elektrické aku.	2023-06	80,68759517	
4	elektrické aku.	2023-07	82,28704308	
5	elektrické aku.	2023-08	83,88649099	
6	elektrické aku.	2023-09	85,48593889	
7	elektrické aku.	2023-10	87,0853868	
8	elektrické aku.	2023-11	88,68483471	
9	elektrické aku.	2023-12	90,28428262	
10	elektrické aku.	2024-01	91,88373052	
11	elektrické aku.	2024-02	93,48317843	
12	elektrické aku.	2024-03	95,08262634	
13	elektrické aku.	2024-04	96,68207425	
14				

Obrázek 16 Výsledná tabulka v Excelu (vlastní zpracování)

Výsledná tabulka, která byla vygenerována, je zobrazena na obrázku 16. Zobrazuje, kolik kusů náradí na baterii se bude prodávat následujících 12 měsících. Nakonec se výsledky zaokrouhlí na celá čísla. Tab. 1 vznikne spojením predikovaného objemu kusů elektrického, neboli kabelového náradí a náradí aku napájeného baterií.

Tabulka 1 Zaokrouhlené výsledky predikce (vlastní zpracování)

Měsíc, rok	V.23	VI.23	VII.23	VIII.23	IX.23	X.23	XI.23	XII.23	I.24	II.24	III.24	IV.24
Elektrické (ks)	74	74	74	74	74	74	74	73	73	73	73	73
Aku (ks)	79	81	82	84	85	87	89	90	92	93	95	97

#### 4.2.9 Výpočet intervalu spolehlivosti a popis grafů

V této fázi byl vypočítán 95 % interval spolehlivosti predikce obou kategorií pomocí kódu zobrazeného na obrázku číslo 17. Tímto postupem se získají grafy na obrázcích 18 a 19.

```
In [20]: model = LinearRegression()
model.fit(X_train, y_train)

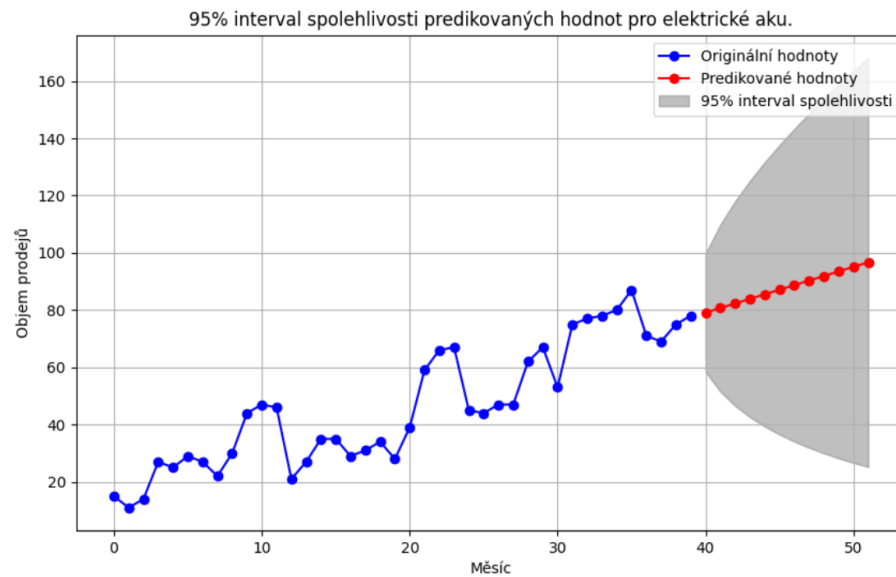
y_train_pred = model.predict(X_train)
residuals = y_train - y_train_pred
residual_std = np.std(residuals)

future_y = np.array([79.08814726, 80.68759517, 82.28704308, 83.88649099, 85.48593889, 87.0853868,
88.68483471, 90.28428262, 91.88373052, 93.48317843, 95.08262634, 96.68207425])
future_X_range = range(len(y), len(y) + len(future_y))
confidence_interval = [residual_std * stats.t.ppf
((1 + 0.95) / 2, len(y_train) - 1) * np.sqrt(i+1) for i in range(len(future_y))]

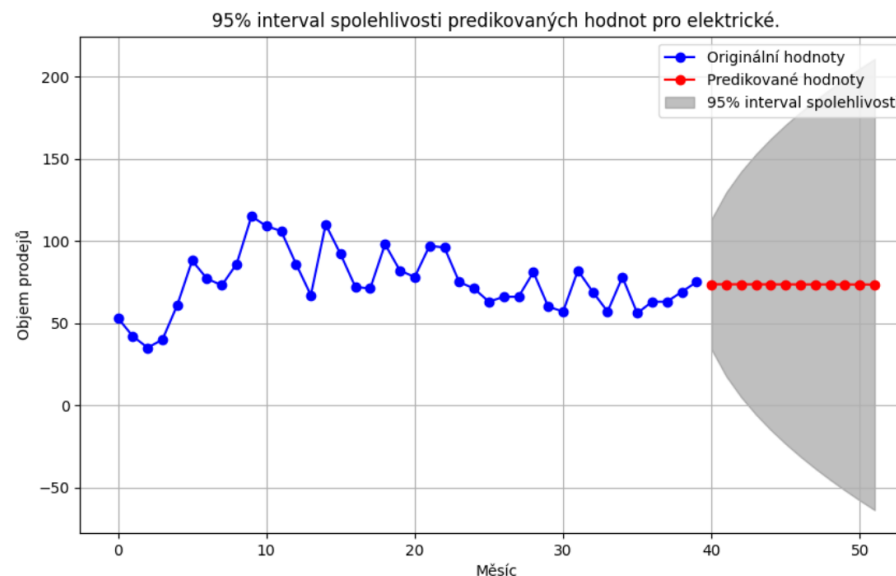
plt.figure(figsize=(10,6))
plt.plot(range(len(y)), y, 'o-', color='blue', label='Originální hodnoty')
plt.plot(future_X_range, future_y, 'o-', color='red', label='Predikované hodnoty')
plt.fill_between(future_X_range, future_y - confidence_interval, future_y + confidence_interval,
color='gray', alpha=0.5, label='95% interval spolehlivosti')

plt.grid(True)
plt.legend()
plt.title("95% interval spolehlivosti predikovaných hodnot pro elektrické aku.")
plt.xlabel("Měsíc")
plt.ylabel("Objem prodeju")
plt.show()
```

Obrázek 17 Tvorba grafu intervalu spolehlivosti (vlastní zpracování)



Obrázek 18 Interval spolehlivosti pro nářadí elektrické aku. (vlastní zpracování)

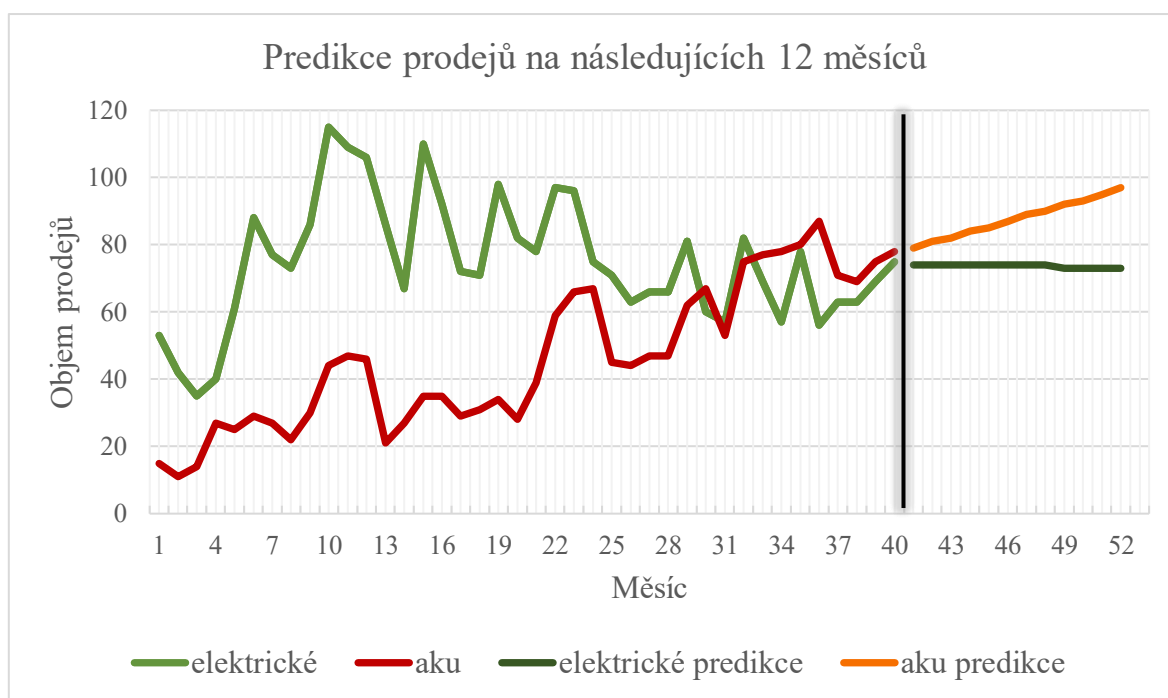


Obrázek 19 Interval spolehlivosti pro elektrické nářadí (vlastní zpracování)

Na obrázcích 18 a 19 je vidět predikce z Tab. 1 zobrazená grafy. Červené body reprezentují predikované hodnoty a modré body představují skutečné původní hodnoty. Grafy dále zobrazují 95 % interval spolehlivosti predikovaných hodnot. Že se skutečný objem prodeje bude pohybovat v tomto rozmezí lze očekávat s 95 % spolehlivostí. Jedná se o statistický termín, který v tomto případě dává rozsah hodnot znázorněný šedivou plochou, tedy míra této jistoty je zobrazena šedivým polem trychtýřovitého tvaru. S časově vzdálenější predikcí je vidět, že nepřesnost velmi rychle narůstá a vypovídající hodnota se snižuje. Proto se obecně doporučuje predikovat prodeje pouze na další nejbližší nutné časové období.

### 4.3 Interpretace výsledků

Na grafu 2 je vizualizovaný vývoj objemů prodaného zboží podle kategorie za každý měsíc začínající v minulosti a končící v budoucnu. Na grafu jsou vidět předchozí uskutečněné prodeje, začínající lednem 2020, až po prodeje za měsíc duben 2023. Na tato data navazují křivky z předikce, které jsou vizuálně odděleny svislou čarou, začínající na hodnotách 74 kusů elektrického zboží a 79 kusů zboží na baterii v květnu 2023. Křivky končí v dubnu roku 2024 na hodnotách 73 kusů elektrického zboží a 97 kusů zboží na baterii. Celkový objem prodejů kategorie nářadí na aku baterii viditelně rostl a z výsledků predikce je zřejmé, že v tomto trendu bude nejpravděpodobněji i nadále pokračovat. Celkové prodeje kategorie elektrického nářadí se zdály, že spíše klesaly a z predikce vyšlo, že budou stagnovat, nebo velmi pomalu klesat.



Obrázek 20 Predikce prodejů na následujících 12 měsících (vlastní zpracování)

Z grafu vyplývá, že v únoru roku 2023 byl poprvé objem prodejů nářadí na baterii vyšší, než kabelového a z přijatých firemních informací by se tento celosvětový trend růstu prodejů nářadí na baterii neměl změnit. Na tuto změnu v poptávce po různých typech zboží se společnost musí včas připravit a zavést vhodná opatření, aby tak zamezila propadu tržeb. Obchodní a jiné společnosti, které se dokážou na základě této znalosti přizpůsobit trhu, budou mít konkurenční výhodu nad ostatními.

Na tomto grafu si je dobré povšimnout cykličnosti prodejů aku náradí, které mělo každé předvánoční období nárůst a každý leden poté strmý pokles o desítky procent. Je to pravděpodobně způsobeno tím, že největšími zákazníky společnosti jsou především firmy, které si mohou tímto nákupem před koncem účetního období, tedy před účetní závěrkou snížit základ pro výpočet daně a tím využívají možnosti daň optimalizovat.

#### 4.3.1 Možná omezení predikce

Omezením výpočtu predikce bylo, že je založena na datech začínajících až rokem 2020. Je tomu tak z důvodu, že společnost dříve artikl náradí vedla jen málo. Teprve v roce 2020 se společnost rozhodla tento typ zboží více rozšířit ve svých prodejnách. Ani v roce 2023 ale není náradí prioritou a společnost se na něj příliš nezaměřuje v porovnání s jiným sortimentem. Z malých objemů prodejů poté plyne nižší přesnost prognózy.

Nedostatkem také je, že časově delší prognózy nemají příliš smysl, jelikož omezená vstupní data byla dostačující pouze na nejbližší budoucí časové období. S každým dalším měsícem se poté rychle zvyšuje nepřesnost predikce, jak bylo znázorněno obrázky 18 a 19.

Dalším omezením je, že z predikce nemůžeme vědět, o jaké konkrétní značky a zboží bude zájem, protože se jedná o predikci souhrnných prodejů za kategorii. Jistě by šel vytvořit komplikovanější a přesnější model s použitím mnoha metod strojového učení, zohledňující mnohem více faktorů, ale to by bylo z hlediska přiměřené obtížnosti bakalářské práce příliš náročné. Proto byla zvolena cesta lineární regrese.

## 5 OPTIMALIZACE SKLADOVÝCH ZÁSOb

Tato kapitola se zabývá analýzou a optimalizací současného stavu zásob vybraného sortimentu. Kapitola vychází nejen z vytvořené prognózy, ale také z přijatých dat o prodeji. Již z předešlé kapitoly je známé, jak se budou prodeje vyvíjet do budoucna, a že pravděpodobně jedna kategorie produktů bude v prodejnosti stoupat, zatímco druhá stagnovat. Tudíž je možné doporučit, jak se na tuto změnu z pohledu firmy připravit. Dále bude vypočítán další ukazatel doby obratu zásob. Také se vypočítá, jak by společnost měla pracovat se zásobami, aby zůstala pokud možno co nejvíce efektivní. Dále proběhne segmentace dodavatelů podle metody ABC. Nakonec budou navržena vhodná opatření, jak by firma měla k těmto změnám na trhu přistoupit a jaké kroky podniknout.

### 5.1 Analýza současného stavu zásob sortimentu nářadí

Společnost měla k datu 1.5.2023 dohromady naskladněných 1 776 položek elektrického nářadí a 1 078 položek aku nářadí. Protože zveřejnění přesných nákupních cen zboží by poškodilo budoucí vyjednávací pozici s dodavateli, tyto ceny mi nebyly poskytnuty. Nákupní ceny se tedy spočítají pomocí již dříve spočítané průměrné marže. Bude se tedy předpokládat stejná 25 % marže na všechny produkty i přesto, že je to nepřesný předpoklad.

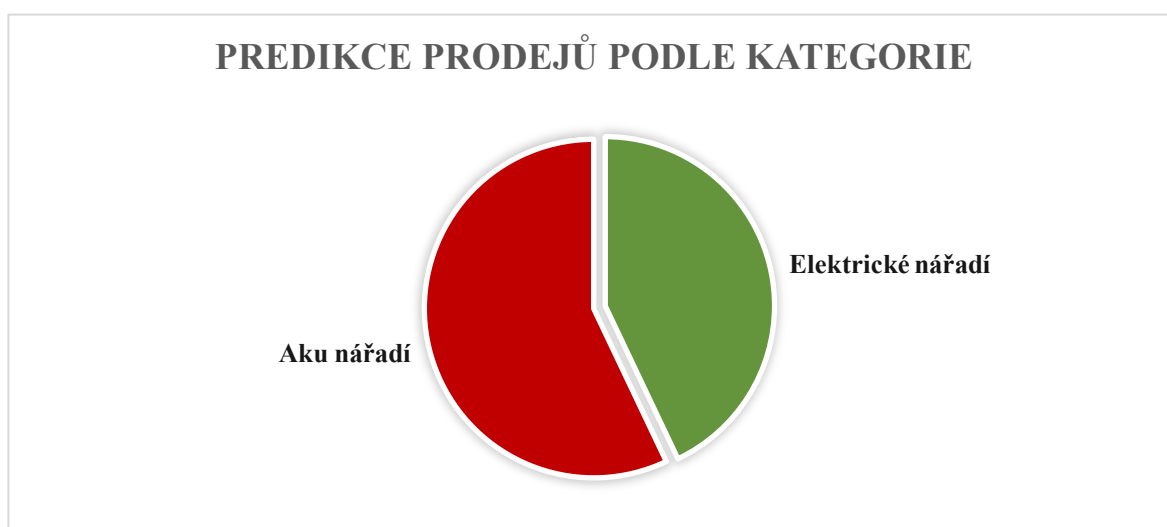
Průměrná prodejní cena položky naskladněného elektrického nářadí k datu 1.5.2023 je 18 260 Kč. Průměrná prodejní cena položky naskladněného aku nářadí k datu 1.5.2023 je 11 941 Kč. Poté nákupní cena po odečtení marže jednotky položky je  $18\,260 / 125 \times 100 = 14\,608$  Kč v případě elektrického nářadí a  $11\,941 / 125 \times 100 = 9\,553$  Kč v případě aku nářadí. Vynásobením počtem kusů na skladě se získá poměr peněžních prostředků vázaných na skladě v tomto sortimentu. To je zobrazeno na výsečovém grafu 3. Je zřetelně vidět, jaký nepoměr má společnost v alokaci kapitálu v zásobách. Hodnoty zde vypočítané budou také použity pro výpočet doby obratu zásob později.

Celkové zásoby sortimentu nářadí jsou:  $17\,418\,687$  Kč +  $7\,479\,186$  Kč =  $24\,897\,873$  Kč

Na grafu (obrázek 22) je možné vidět predikci poměru počtu prodejů výrobků podle kategorie v období měsíce dubna 2024. Srovnáním grafů na obrázcích 21 a 22 vidíme, že se poměry výrazně rozcházejí, v čemž může spočívat problém. Společnost má stále výrazně více peněžních prostředků vázaných v nářadí kabelovém, ale prodeje nářadí na baterii jsou již tento rok vyšší a budou se nadále zvyšovat.



Obrázek 21 Peněžní prostředky vázané ve zboží k 1.5.2023 (vlastní zpracování)



Obrázek 22 Predikce prodejů zboží podle kategorie v dubnu 2024 (vlastní zpracování)

## 5.2 Aplikace ABC analýzy

Cílem této ABC analýzy je diferencovat zásoby nářadí (Tabulka 2) do třech kategorií, aby bylo patrné, v jakých značkách naskladněného nářadí je vázáno nejvíce kapitálu. Toto rozdělení umožňuje efektivněji řídit zásoby. Klasifikačním kritériem je nákupní kumulativní cena. Kategorie jsou A, B, a C. Pro stanovení mezí byla data seřazena sestupně dle nákupní ceny celkem, která se skládá z počtu naskladněných kusů nářadí daného dodavatele vynásobených jejich cenou. Z ceny je výpočtem odečtena již dříve spočítaná marže 25 %, jelikož ceny v tabulce jsou prodejní, nikoliv nákupní. Každé nářadí je naceněno jinak, a proto je zde uveden pouze celkový součet. Kategorii A budou tvořit položky značek, která tvoří 60 % finančních prostředků. Do kategorie B budou zařazeny položky značky, které tvoří 20 % a zbylé položky budou patřit do kategorie C. Rozdělení zásob do jednotlivých kategorií podle zvoleného kritéria a stanovených mezí znázorňuje Tab. 3.

Tabulka 2 Podklady pro ABC analýzu (vlastní zpracování)

pořadí	Značky na skladu	Počet položek	Kumulativní cena [Kč]	Kumulativní cena bez marže [Kč]
1.	aku. Stanley	207	951 032	760 826
2.	aku. Rothenberger	77	2 797 847	2 238 278
3.	aku. Rems	11	271 168	216 934
4.	aku. Milwaukee	195	2 231 535	1 785 228
5.	aku. DeWALT	351	2 146 010	1 716 808
6.	aku. Bosch	237	951 391	761 113
7.	Stanley	245	667 870	534 296
8.	Rothenberger	55	1 803 110	1 442 488
9.	Rems	578	14 850 884	11 880 707
10.	DeWALT	403	2 167 598	1 734 078
11.	Bosch	495	2 283 896	1 827 117

Tabulka 3 Aplikace ABC analýzy (vlastní zpracování)

pořadí	Značky na skladu	Kumulativní cena bez marže [Kč]	Kumulativní cena bez marže [%]	Σ [%]	Kategorie
9.	Rems	11 880 707	47,72	47,7	<b>A</b>
2.	aku. Rothenberger	2 238 278	8,99	56,7	<b>A</b>
11.	Bosch	1 827 117	7,34	64,1	<b>B</b>
4.	aku. Milwaukee	1 785 228	7,17	71,2	<b>B</b>
10.	DeWALT	1 734 078	6,96	78,2	<b>B</b>
5.	aku. DeWALT	1 716 808	6,9	85,1	<b>B</b>
8.	Rothenberger	1 442 488	5,79	90,9	<b>C</b>
6.	aku. Bosch	761 113	3,06	93,9	<b>C</b>
1.	aku. Stanley	760 826	3,06	97	<b>C</b>
7.	Stanley	534 296	2,15	99,1	<b>C</b>
3.	aku. Rems	216 934	0,86	100	<b>C</b>
Σ		24 897 873	100		

Výpočet procentuálního počtu položek dle kategorie:

Položky A:  $(578 + 77) / 2\,854 \times 100 \doteq 23 \%$

Položky B:  $(495 + 195 + 403 + 351) / 2\,854 \times 100 \doteq 51 \%$

Položky C:  $(55 + 237 + 207 + 245 + 11) / 2\,854 \times 100 \doteq 26 \%$

### 5.2.1 Závěr z ABC analýzy

Z provedené ABC analýzy vyplývá, že ačkoliv položky kategorie A tvoří pouze 23 % z veškerých položek, vážou na sobě 60 % celkového finančního kapitálu, který je alokovan

v sortimentu náradí. Do kategorie A jsou zařazeny značky Rems, od kterého je odebíráno elektrické náradí a Rothenberger, který je dodavatelem aku náradí. Položky kategorie B tvoří 51 % z veškerých položek a váží na sebe 20 % kapitálu. Kategorie C představuje 26 % zbylých položek a váží 10 % kapitálu.

### 5.2.2 Rychlost obratu zásob

Z obdržených dat se nejdříve zjistí počet prodaných kusů za poslední měsíc. V tomto případě je to v dubnu 2023 75 kusů prodaného elektrického náradí a 78 aku náradí. Nyní je známo dostatek hodnot pro spočítání důležitého ukazatele v řízení zásob, doby obratu zásob. Doba obratu zásob se v tomto případě vypočítá: zásoby / tržby za tento měsíc.

Výpočet pro elektrické náradí:

$$\frac{17\,418\,687}{18\,260 \times 75} \doteq 12,7 \text{ měsíce}$$

Výpočet pro aku náradí:

$$\frac{7\,479\,186}{11\,941 \times 78} \doteq 8 \text{ měsíců}$$

Z výsledků vyplývá, že doba obratu pro kategorii elektrického náradí je horší, než pro kategorii aku náradí. Výsledek 12,7 měsíce není moc dobrý. Pro přesnější výpočet by bylo potřeba znát přesné marže těchto výrobků, protože z nich je spočítána hodnota zásob. Doba obratu zásob se také mění s tím, jak se postupem času vyvíjejí tržby.

### 5.2.3 Závěry z výpočtů

Na skladě se v dubnu 2023 nacházelo více položek kabelového náradí, než položek náradí na baterii. Průměrná nákupní cena kabelového náradí, které bylo skladem, byla také vyšší, než průměrná nákupní cena náradí na baterii. Stav zásob je výrazně převážen na stranu elektrického náradí, ačkoliv jsou objemy prodejů obou kategorií podobné a můžeme předpokládat, že vyšší obrat bude v budoucnu dělat především náradí na baterii. Výrazně nižší doba obratu zásob aku náradí znamená, že je pro společnost výhodnější. Společnost by měla hledat řešení, jak snížit dobu obratu zásob u elektrického náradí. Doba obratu zásoby 12,7 měsíce znamená velké množství držených zásob, což může vést k vysokým nákladům na skladování a riziku zastarávání zásob. Aku náradí s doby obratu zásoby 8 měsíců představuje mnohem lepší výsledek, který by se mohl ještě zlepšit, pokud bude prodejnost tohoto zboží narůstat.



### 5.3 Návrh doporučení

Posledním krokem je návrh opatření, které by společnosti pomohlo ušetřit náklady. Na základě provedeného zhodnocení současného stavu zásob bude v této závěrečné kapitole uvedeno, jak optimalizovat hladinu skladových zásob a tím ušetřit společnosti náklady spojené především s vázaností prostředků v zásobách. Jelikož se poptávka po zboží mění, je nejvyšší čas na to společnost připravit.

Prodej jak elektrického, tak aku náradí je pro společnost výhodný a v udržování tohoto sortimentu je vhodné doporučit pokračovat. Tímto společnost rozšiřuje svou komplexní nabídku zboží a zákazníci si mohou vše koupit na jednom místě. To vede k větší spokojenosti zákazníků. Celý sortiment náradí se prodává a poptávka po těchto produktech je konzistentní, nebo se spíše zvyšuje.

Z pohledu objemu prodeje by celkové finanční prostředky vynaložené na nákup nového náradí měla společnost spíše snižovat, nebo zachovat, pokud se jí daří a má volné finanční prostředky. V případě nedostatku finančních prostředků by stálo za zvážení, zda nejdříve nevyprodat již vlastněné zásoby náradí, například pomocí marketingové kampaně. Mělo by se také cílit na vyrovnaní stavu finančních prostředků vázaných v zásobách těchto dvou kategorií náradí, se kterými se v predikci pracovalo. Objem celkových prostředků vynaložených na nákup nového náradí napájeného baterií by měla společnost zvyšovat v poměru prostředků vynaložených na nákup náradí napájeného kabelem.

Z pohledu doby trvání obratu zásob by bylo vhodné doporučit snížení finančních prostředků vázaných v zásobách náradí. Obrátka zásob především u náradí kategorie elektrické by mohla být lepší. Doporučením je tedy zároveň snížit počet naskladněného zboží a tím snížit dobu obratu, nebo zvýšit prodeje pomocí marketingových kampaní a dalších speciálních nabídek. Společnost může dále snížit objem celkových prostředků uložených v zásobách tím, že tyto produkty omezí v dalším nakupování.

Celkově by se mělo cílit na snížení celkové úrovně zásob, pokud je to možné, aniž by se riskoval jejich nedostatek a projevilo by se to negativně na prodejích. K tomu by pomohla přesnější predikce poptávky zaměřená na jednotlivé značky, zkrácení dodacích lhůt od dodavatelů, zlepšení efektivity skladování, distribuce a pravidelné revize úrovně zásob.

Nakonec z provedené analýzy ABC vyplývá, že aby společnost ještě více zvýšila svou konkurenceschopnost, úsilí o zvýšení prodeje by mělo směřovat především na nejdůležitější kategorii „A“, tedy značky zboží na skladě od dodavatele Rothenberger a Rems.

## ZÁVĚR

Bakalářská práce se věnovala velkým objemům dat a aplikaci strojového učení pro predikci prodejů vybraného zboží společnosti Ptáček – velkoobchod, a.s. Práce byla rozdělena na teoretickou a praktickou část.

V teoretické části byla provedena rešerše teoretických poznatků. V této části se soustředila pozornost na objasnění a definování nejdůležitějších pojmů, týkajících se hlavního tématu práce. Jednalo se o big data, strojové učení a podnikové procesy. Bylo obtížné takto rozsáhlé téma vhodně pojmut a shrnout jen to nejdůležitější. Tato část práce byla již podrobněji shrnuta na jejím konci.

Praktická část začala uvedením společnosti a stručnou analýzou účetních výkazů. Pokračovala predikcí počtu kusů, které se prodají v následujícím roce. Bylo podrobně a přehledně vysvětleno, jak se k výsledkům predikce došlo a z jakých důvodů se vybrala použitá metoda. Z vypočtených budoucích hodnot prodejů, týkajících se skladových zásob sortimentu náradí, byla vytvořena analýza současných skladových zásob tohoto sortimentu. Výsledky predikce byly poté porovnány se současnou situací skladových zásob. Jsou předložena vhodná doporučení, jejichž dodržování by mohlo vést ke zlepšení situace ve společnosti. V neposlední řadě byly zhodnoceny zásoby podle ABC analýzy.

Byla vytvořena literární rešerše na témata související se zadáním práce. Aplikován algoritmus strojového učení na podnikový proces tak, aby pomohl firmě ušetřit náklady. Z uvedených důvodů lze tvrdit, že jak teoretická, tak praktická část byla úspěšně splněna a cíle stanovené na začátku práce byly dosaženy. Důležité v práci bylo nejen splnit stanovené cíle, ale také prezentovat principy této metodiky.

**SEZNAM POUŽITÉ LITERATURY**

BERKA, Petr, 2003. *Dobývání znalostí z databází*. Praha: Academia. ISBN 80-200-1062-9.

EMMETT, Stuart, 2008. *Řízení zásob: jak minimalizovat náklady a maximalizovat hodnotu*. Brno: Computer Press. Praxe manažera (Computer Press). ISBN 978-80-251-1828-3.

ERL, Thomas, Wajid KHATTAK a Paul BUHLER, 2016. *Big data fundamentals: concepts, drivers & techniques*. Vancouver, BC: Service Tech Press. Prentice Hall service technology series from Thomas Erl. ISBN 9780134291079.

HENDL, Jan, 2021. *Big data: věda o datech - základy a aplikace*. Praha: Grada Publishing. Průvodce (Grada). ISBN 9788027130313.

HOLUBOVÁ, Irena et al., 2015. *Big Data a NoSQL databáze*. Praha: Grada. Profesionál. ISBN 9788024754666.

JUROVÁ, Marie, 2016. *Výrobní a logistické procesy v podnikání*. Praha: Grada Publishing. Expert (Grada). ISBN 978-80-247-5717-9.

KALOUDA, František, 2015. *Finanční analýza a řízení podniku*. Plzeň: Vydavatelství a nakladatelství Aleš Čeněk. ISBN 978-80-7380-526-5.

MARR, Bernard, 2015. *Big data: using smart big data, analytics and metrics to make better decisions and improve performance*. Chichester: Wiley. ISBN 9781118965832

MARR, Bernard, 2016. *Big data in practice: how 45 successful companies used big data analytics to deliver extraordinary results*. Chichester: Wiley. ISBN 9781119231387

MAYERSCHÖNBERGER, Viktor a Kenneth CUKIER, 2014. *Big data: revoluce, která změní způsob, jak žijeme, pracujeme a myslíme*. Brno: Computer Press. ISBN 9788025141199

OTYEPKA, Michal, Pavel BANÁŠ a Eva OTYEPKOVÁ, 2013. *Základy zpracování dat*. Olomouc: Univerzita Palackého v Olomouci. ISBN 978-80-244-3636-4.

SVOZILOVÁ, Alena, 2011. *Zlepšování podnikových procesů*. Praha: Grada. Expert (Grada). ISBN 978-80-247-3938-0.

VÁCHAL, Jan a Marek VOCHOZKA, 2013. *Podnikové řízení*. Praha: Grada. Finanční řízení. ISBN isbn978-80-247-4642-5.

VYSKOČIL, Vlastimil K. a Ondřej ŠTRUP, 2003. *Podpůrné procesy a snižování režijních nákladů: (facility management)*. Praha: Professional Publishing. ISBN 80-86419-45-2.

**Zdroje dostupné online:**

*Co jsou big data?*, 2023. Oracle [online]. [cit. 2023-05-16]. Dostupné z: <https://www.oracle.com/cz/big-data/what-is-big-data/>

*Gartner: Gartner Glossary* [online], 2023. [cit. 2023-05-16]. Dostupné z: <https://www.gartner.com/en/information-technology/glossary>

Ptáček: velkoobchod [online], © 2023. [cit. 2023-05-15]. Dostupné z: <https://www.ptacek.cz/>

HRTÚSOVÁ Tereza a NOVÁK, Radek, 2018. *Big data v ČR: pojem vs. realita*. In: Česká spořitelna [online]. [cit. 2023-05-15]. Dostupné z: [https://www.csas.cz/content/dam/cz/csas/www\\_csas\\_cz/Dokumenty-korporat/Dokumenty/Analytici/Big\\_Data\\_v\\_CR.pdf](https://www.csas.cz/content/dam/cz/csas/www_csas_cz/Dokumenty-korporat/Dokumenty/Analytici/Big_Data_v_CR.pdf)

COX, Michael a ELLSWORTH, David, 1997. *Application-Controlled Demand Paging for Out-of-Core Visualization*. [online]. [cit. 2023-05-10]. Dostupné z: <https://www.nas.nasa.gov/assets/pdf/techreports/1997/nas-97-010.pdf>

ČESKO, 2019. *Zákon č. 110/2019 Sb. Zákon o zpracování osobních údajů*. In: Sběrka zákonů České republiky. Dostupné také z: <https://www.zakonyprolidi.cz/cs/2019-110>

ERIKSSON, Theresa, Alessandro BIGI a Michelle BONERA, 2020. *Think with me, or think for me? On the future role of artificial intelligence in marketing strategy formulation*. The TQM Journal. [online]. [cit. 2023-05-16]. Dostupné z doi: <https://doi.org/10.1108/TQM-12-2019-0303>

KOŘOUSKOVÁ, Barbora, 2023. *Internet věcí (IoT): definice, příklady využití, produkty*. Rascasone [online]. [cit. 2023-05-16]. Dostupné z: <https://www.rascasone.com/cs/blog/iot-internet-veci-definice-produkty-historie>

KUBINA, Milan, Michal VARMUS a Irena KUBINOVA, 2015. *Use of big data for competitive advantage of company*. 5. Dostupné z: doi:[https://doi.org/10.1016/S2212-5671\(15\)00955-7](https://doi.org/10.1016/S2212-5671(15)00955-7)

*Strojové učení u systému ERP*, 2022. GatemaIT [online]. [cit. 2023-05-16]. Dostupné z: <https://www.gatemait.cz/strojove-uceni-u-systemu-erp/>

*Velká data (big data): definice, výhody a výzvy (infografika)*, 2023. Zpravodajství Evropský parlament [online]. [cit. 2023-05-16]. Dostupné z: <https://www.europarl.europa.eu/news/cs/headlines/society/20210211STO97614/velka-data-big-data-definice-vyhody-a-vyzvy-infografika>

PIETRAŠ, Petr, 2020. *Čínská digitální diktatura je pro Západ obrovskou výzvou* [online]. [cit. 2023-05-18]. Dostupné z: [https://www.lidovky.cz/ceska-pozice/cinska-digitalni-diktatura-je-pro-zapad-obrovskou-vyzvou.A200105\\_012803\\_pozice-recenze\\_lube](https://www.lidovky.cz/ceska-pozice/cinska-digitalni-diktatura-je-pro-zapad-obrovskou-vyzvou.A200105_012803_pozice-recenze_lube)

ZUBAIR, Md., 2023. *Deep Understanding of Simple Linear Regression: Linear Regression from Scratch: Detailed Explanation* [online]. [cit. 2023-05-18]. Dostupné z: <https://towardsdatascience.com/deep-understanding-of-simple-linear-regression-3776afe34473>

**SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK**

BI	Business intelligence
ERP	Plánování podnikových zdrojů (z angl. Enterprise Resource Planning)
ICT	Informační a komunikační technologie (z angl. Information and Communication Technologies)
IČO	Identifikační číslo osoby
IoT	Internet věcí (z angl. Internet of Things)
SQL	Standardizovaný strukturovaný dotazovací jazyk (z angl. Structured Query Language)
VMI	Zásoby řízené dodavatelem (z angl. Vendor Managed Inventory)

Pozn.: V seznamu nejsou uvedeny symboly a zkratky všeobecně známé.

**SEZNAM OBRÁZKŮ**

Obrázek 1 Model analytik (vlastní zpracování podle Maoz 2013).....	18
Obrázek 2 Průběžné zlepšování procesu (vlastní zpracování podle Řepa, 2007).....	22
Obrázek 3 Logo společnosti Ptáček – velkoobchod, a.s. (Ptáček, 2023).....	31
Obrázek 4 Rozložení poboček v roce 2023 v České Republice (Ptáček, 2023).....	32
Obrázek 5 Graf souhrnných objemů prodeje nářadí podle kategorie (vlastní zpracování).	34
Obrázek 6 Obdržená data (vlastní zpracování).....	35
Obrázek 7 Import knihoven a dat (vlastní zpracování).....	37
Obrázek 8 Úprava dat (vlastní zpracování).....	37
Obrázek 9 Volba kategorie produktu (vlastní zpracování).....	38
Obrázek 10 Volba kategorie nářadí, převedení datumu do indexu (vlastní zpracování) ....	38
Obrázek 11 Rozdělení dat (vlastní zpracování).....	38
Obrázek 12 Volba modelu strojového učení (vlastní zpracování).....	39
Obrázek 13 Testování modelu a vyhodnocení (vlastní zpracování).....	39
Obrázek 14 Aplikace modelu a výsledek (vlastní zpracování).....	40
Obrázek 15 Tvorba Excelového souboru (vlastní zpracování).....	40
Obrázek 16 Výsledná tabulka v Excelu (vlastní zpracování).....	41
Obrázek 17 Tvorba grafu intervalu spolehlivosti (vlastní zpracování).....	41
Obrázek 18 Interval spolehlivosti pro nářadí elektrické aku. (vlastní zpracování) .....	42
Obrázek 19 Interval spolehlivosti pro elektrické nářadí (vlastní zpracování).....	42
Obrázek 20 Predikce prodeje na následujících 12 měsíců (vlastní zpracování).....	43
Obrázek 21 Peněžní prostředky vázané ve zboží k 1.5.2023 (vlastní zpracování).....	46
Obrázek 22 Predikce prodeje zboží podle kategorie v dubnu 2024 (vlastní zpracování) ...	46

**SEZNAM TABULEK**

Tabulka 1 Zaokrouhlené výsledky predikce (vlastní zpracování) .....	41
Tabulka 2 Podklady pro ABC analýzu (vlastní zpracování) .....	47
Tabulka 3 Aplikace ABC analýzy (vlastní zpracování) .....	47



## SEZNAM PŘÍLOH

Příloha P I: Použitý zdrojový kód

Soubor .zip obsahuje:

Export\_dat.csv

Použitý\_zdrojový\_kód.txt

Predikce\_prodejů.ipynb

## PŘÍLOHA P I: POUŽITÝ ZDROJOVÝ KÓD

```
# Import knihoven
import pandas as pd
import numpy as np
import openpyxl
import datetime
import statsmodels.api as sm
import scipy.stats as stats
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error, r2_score
# Import dat a pre-processing → rozdělení objemu prodejů do dvou kategorií po měsících
df = pd.read_excel('Export_dat.xlsx', sheet_name='Sheet1')
df = df.groupby('Vlastnost').sum()
# Zbavení se dat, která nebudou potřeba
df = df.drop(['ID', 'Zboží - popis', 'Jednotková cena', 'Zásoba - množství'], axis=1)
# Transponování tabulky a lepší formát času
df = df.transpose()
df.index = pd.to_datetime(df.index, format='%Y - %b')
# Takto nyní vypadá tabulka ve tvaru, kterém potřebujeme
df.head(10)
# Nejdříve si zvolíme pouze jednu kategorii produktů, například "elektrické aku." (nářadí na baterii)
elektrické_df = df.drop(['elektrické'], axis=1)
# Tabulka vypadá nyní takto
elektrické_df.head(10)
# Datum si převedeme do indexu, protože lineární regrese umí pracovat pouze s čísly a ne daty
X = np.array(range(len(elektrické_df)))[:, np.newaxis]
y = elektrické_df['elektrické aku.'].values
# Rozdělení dat do trénovací a testovací sady, zvolil jsem poměr 80% a 20%
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=41)
# Aplikace modelu strojového učení, lineární regrese
model = LinearRegression()
model.fit(X_train, y_train)
# Vytvoření predikce na testovací soubor
y_pred = model.predict(X_test)
# Vyhodnocení přesnosti modelu
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
print("Mean Squared Error: {:.2f}".format(mse))
print("R-squared: {:.2f}".format(r2))
# Budoucí předpověď
future_X = np.array(range(len(elektrické_df), len(elektrické_df)+12))[:, np.newaxis]
future_y = model.predict(future_X)
print("Future Predictions:", future_y)
# Vytvoření souboru Excel, kam se data uloží
workbook = openpyxl.Workbook()
# Výběr aktivního pracovního listu
worksheet = workbook.active
# Nastavení záhlaví sloupců
worksheet.cell(row=1, column=1, value="Produkt")
worksheet.cell(row=1, column=2, value="Měsíc")
worksheet.cell(row=1, column=3, value="Predikovaná hodnota")
# Cyklus pro vepsání předpokládaných hodnot pro každý měsíc
now = datetime.datetime.now()
for i in range(len(future_y)):
    next_month = now + datetime.timedelta(days=30*i)
    value = future_y[i]
    worksheet.cell(row=i+2, column=1, value='elektrické aku')
    worksheet.cell(row=i+2, column=2, value=next_month.strftime("%Y-%m"))
    worksheet.cell(row=i+2, column=3, value=value)
# Uložit .xlsx soubor do počítače
workbook.save("Předpověď_pro_elektrické_aku.xlsx")
model = LinearRegression()
model.fit(X_train, y_train)
```

```

y_train_pred = model.predict(X_train)
residuals = y_train - y_train_pred
residual_std = np.std(residuals)

future_y = np.array([79.08814726, 80.68759517, 82.28704308, 83.88649099, 85.48593889, 87.0853868,
                    88.68483471, 90.28428262, 91.88373052, 93.48317843, 95.08262634, 96.68207425])
future_X_range = range(len(y), len(y) + len(future_y))
confidence_interval = [residual_std * stats.t.ppf
                      ((1 + 0.95) / 2, len(y_train) - 1) * np.sqrt(i+1) for i in range(len(future_y))]

plt.figure(figsize=(10,6))
plt.plot(range(len(y)), y, 'o-', color='blue', label='Originální hodnoty')
plt.plot(future_X_range, future_y, 'o-', color='red', label='Predikované hodnoty')
plt.fill_between(future_X_range, future_y - confidence_interval, future_y + confidence_interval,
                 color='gray', alpha=0.5, label='95% interval spolehlivosti')

plt.grid(True)
plt.legend()
plt.title("95% interval spolehlivosti predikovaných hodnot pro elektrické aku.")
plt.xlabel("Měsíc")
plt.ylabel("Objem prodeju")
plt.show()

```