

EKONOMICKÁ UNIVERZITA V BRATISLAVE
FAKULTA HOSPODÁRSKEJ INFORMATIKY

Evidenčné číslo: 103006/B/2021/36122163605669380

AKTUÁLNE TRENDY V OBLASTI VEĽKÝCH DÁT
A V UMELEJ INTELIGENCII

Bakalárska práca

2021

Igor Šárkan

EKONOMICKÁ UNIVERZITA V BRATISLAVE
FAKULTA HOSPODÁRSKEJ INFORMATIKY

**AKTUÁLNE TRENDY V OBLASTI VEĽKÝCH DÁT
A V UMELEJ INTELIGENCII**

Bakalárska práca

Študijný program:	Manažérske rozhodovanie
Študijný odbor:	Ekonómia a manažment
Školiace pracovisko:	Katedra matematiky a aktuárstva
Vedúci záverečnej práce:	Mgr. Andrea Kaderová, PhD.

Bratislava 2021

Igor Šárkan



Ekonomická univerzita v Bratislave
Fakulta hospodárskej informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Igor Šárkan
Študijný program: manažérske rozhodovanie (Jednoodborové štúdium, bakalársky I. st., denná forma)
Študijný odbor: ekonómia a manažment
Typ záverečnej práce: Bakalárska záverečná práca
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický

Názov: Aktuálne trendy v oblasti veľkých dát a v umelej inteligencii

Anotácia: V záverečnej práci sa zameriame na definovanie pojmov veľké dáta a umelá inteligencia a budeme hľadať možnosti ako veľké dáta môžu pomôcť pri tvorbe umelej inteligencie. Budeme analyzovať konkrétne možnosti využitia umelej inteligencie v praxi s podrobnejším zameraním na jednu vybranú oblasť (napr. poisťovníctvo).

Big Data je koncept, ktorý vznikol na základe obrovského množstva neštruktúrovaných dát, ktoré denne vznikajú a je ich problém spracovať tradičnými metódami spracovania dát. Je to mega trend a zahŕňa v sebe veľmi veľa disciplín, databáz, cez strojové učenie, algoritmy, ekonómiu, sociálne vedy, a mnoho ďalších. Ale v podstate ide o nájdenie nejakých vzorov v dátach, ktoré sú platné, užitočné, neočakávané a dajú sa pochopiť.

Big Data v praxi majú primárny cieľ odhaliť niečo zaujímavé, čo sa dá použiť na zvýšenie zisku spoločnosti.

Umelá inteligencia je vlastne algoritmus, ktorý sa dokáže učiť z dát sám, bez toho, aby sa explicitne programoval.

Vedúci: Mgr. Andrea Kaderová, PhD.
Oponent: Mgr. Karina Chudá, PhD.
Katedra: KMA FHI - Katedra matematiky a aktuárstva FHI
Vedúci katedry: doc. Ing. Michal Páleš, PhD.

Dátum zadania: 16.03.2020

Dátum schválenia: 25.03.2020

doc. Ing. Michal Páleš, PhD.
vedúci katedry

Pod'akovanie

Touto cestou by som chcel vyjadriť moje poďakovanie vedúcej bakalárskej práce Mgr. Andrei Kaderovej, Phd. za odborné vedenie, pomoc, cenné pripomienky a rady pri vypracovávaní mojej bakalárskej práce.

Abstrakt

ŠÁRKAN, Igor: *Aktuálne trendy v oblasti veľkých dát a v umelej inteligencii*. – Ekonomická univerzita v Bratislave. Fakulta hospodárskej informatiky; Katedra matematiky a aktuárstva. – Vedúci záverečnej práce: Mgr. Andrea Kaderová, PhD. – Bratislava: FHI EU, 2021, 44 s.

Cieľom záverečnej práce „Aktuálne trendy v oblasti veľkých dát a v umelej inteligencii“ je priblížiť čitateľovi problematiku veľkých dát a umelej inteligencie, poskytnúť možné spôsoby riešenia problematiky a podrobnejšie sa zamerať na jednu z nich. Práca je rozdelená do štyroch kapitol. Prvá kapitola je venovaná definovaniu základných pojmov, rozdelení, princípu fungovania umelej inteligencie, historickému vývoju veľkých dát a umelej inteligencie a využitiu oboch oblastí v súčasnosti. V ďalšej časti sa charakterizujú hlavné ciele práce. Tretia kapitola popisuje metaheuristiky, ktoré predstavujú metodiku práce a metódy skúmania. Záverečná kapitola sa zaoberá rozšírením nadobudnutých poznatkov z predošlej kapitoly o podrobný popis metódy kolónie mravcov.

Kľúčové slová: veľké dáta, umelá inteligencia, metaheuristiky, kolónia mravcov

Abstract

The aim of the final thesis „Actual trends in the sphere of Big Data and in the artificial intelligence" has been to get closer to the reader the problematics of the Big Data and the artificial intelligence, to provide possible ways of the problematics solution and focus on one of them in more detail. The thesis has been divided into four chapters. The first part has been devoted to definition of the basic items, divisions, principle of working of the artificial intelligence, the historical development of the Big Data, the artificial intelligence and the using of both spheres nowadays. There are characterized the main aims of the thesis in the next part. The third part has described the metaheuristics, which represent methodics of the thesis and the methods of research. The final part has been dealt with the extension of received knowledge from the last part with a detailed description of the ant colony method.

Key words: Big Data, artificial intelligence, metaheuristics, ant colony

Obsah

Úvod	7
1 Súčasný stav riešenej problematiky doma a v zahraničí	8
1.1 Veľké dáta.....	8
1.1.1 Charakteristika veľkých dát.....	9
1.1.2 Štruktúrované a neštruktúrované dáta.....	11
1.1.3 Historický vývoj veľkých dát	12
1.1.4 Využitie veľkých dát v súčasnosti	13
1.2 Umelá inteligencia	15
1.2.1 ANI, AGI, ASI.....	16
1.2.2 Princípy fungovania umelej inteligencie	18
1.2.3 Historický vývoj umelej inteligencie.....	19
1.2.4 Využitie umelej inteligencie v súčasnosti.....	22
2 Cieľ práce	25
3 Metodika práce a metódy skúmania	26
3.1 Simulated Annealing	27
3.2 Tabu Search	28
3.3 Genetické algoritmy.....	29
3.4 Kolónia mravcov	29
4 Výsledky práce	31
4.1 Kolónia mravcov	31
4.1.1 Kolónia mravcov a problém obchodného cestujúceho	32
Záver	35
Zoznam použitej literatúry.....	36

Úvod

V tejto bakalárskej práci si popíšeme dve problematiky modernej doby, ktorými sú veľké dáta a umelá inteligencia. Dáta ako také tu existovali odjakživa, Ľudia sa naprieč storočiami pokúšali o analýzu dát a takisto o rôzne analytické techniky na podporu ich rozhodovania. Ako príklad môžeme uviesť starovekých Egypťanov, ktorí sa už 300 rokov pred našim letopočtom pokúsili o zachytenie všetkých existujúcich dát a taktiež Rímsku ríšu a ich štatistickú analýzu ozbrojených síl na určenie optimálnej distribúcie pre ich armádu. Od čias prvého počítača sa forma a spolu s ňou aj objem dát zmenil. Veľké dáta sú konceptom vzniknutým v dôsledku návalu obrovského množstva neštruktúrovaných dát, ktoré vďaka počítačom vznikajú. Tieto dáta je problematické spracovať tradičnými metódami, a tak sa hľadajú spôsoby, algoritmy a rôzne vzorce ako ich najefektívnejšie spracovať.

Umelá inteligencia je algoritmom schopným učiť sa z dát sám, bez toho, aby bol explicitne naprogramovaný. Počítače napriek tomu, že prekonávajú ľudí pri rozsiahlych výpočtových úlohách, majú užšie odborné znalosti a taktiež ich schopnosti v inej oblasti, v akej boli naprogramované, zaostávajú za ľudskou inteligenciou. Ľudská inteligencia sa prejavuje rôznymi spôsobmi, či už ide o logické a priestorové schopnosti alebo emočné rozpoznávanie. Človek musí využívať svoje kognitívne schopnosti, ktorými sú napríklad pracovná pamäť, neustála pozornosť, vytváranie rôznych kategórií a rozpoznávanie vzorov. Koncept umelej inteligencie sa snaží replikovať ľudskú inteligenciu a okrem replikácie ju v možnom prípade aj prekonať. Potenciál disponovať tak výkonnými strojmi môže byť lákavý, ale samotný koncept má množstvo neznámych dôsledkov.

Cieľom záverečnej práce je priblížiť problematiku veľkých dát a umelej inteligencie, uviesť základné metódy riešenia, ktorými sú metaheuristiky. V praktickej časti je potrebné jednu z metaheuristik podrobne popísať a uviesť princíp a spôsob riešenia problému.

1 Súčasný stav riešenej problematiky doma a v zahraničí

1.1 Veľké dáta

Všetko, čo robíme v digitálnom svete, zanecháva digitálne stopy, ktoré voláme dáta. Centrom tohto sveta je web, ktorý nás zaplavuje dátami každú sekundu. S viac ako jedným biliónom webových stránok, z ktorých je 50 miliárd indexovaných a je možné ich vyhľadávať prostredníctvom vyhľadávacích nástrojov, nám web ponúka bezkonkurenčné prepojenie umožňujúce komunikovať s kýmkoľvek v rámci jednej siete. Každá takáto interakcia generuje dáta, ktoré sú vedené a zaznamenávané na webe, čo v konečnom dôsledku pridáva na nejasnosti už dosť nejasnému konceptu. Následkom je preplnenie masívnymi dátami, ktoré je takmer nemožné spracovať a rozobrať na menšie časti a nájsť pre nich praktické využitie. Tieto obrovské a stále rastúce dáta prechádzajúce a ukladané na webe spolu s vyvíjajúcimi sa technológiami, ktoré sú určené aby si s nimi poradili sa spoločne označujú ako veľké dáta.

Veľké dáta sú jeden z posledných technologických trendov, ktoré výrazne ovplyvňujú spôsob, akým spoločnosti využívajú informácie. Sú rýchlo sa rozvíjajúcim konceptom v oblasti informačných technológií a správy údajov, ktorý predstavuje revolúciu v spôsobe riadenia spoločností. Tento trend im pomáha k zlepšeniu skúseností ich zákazníkov, vylepšujú ich produkty či služby, vytvára nevyužitý zdroj príjmu, pretvára obchodné modely a takisto dopomáha efektívne riadiť služby zdravotnej starostlivosti.

Aj keď je tento pojem tak populárny, presná a jednotná definícia neexistuje. Veľa IT expertov má problém sa zhodnúť na jednej a tej istej definícii, a tak sa môžeme stretnúť s viacerými pohľadmi na problematiku veľkých dát. Napriek tomu sa v definíciách objavuje aj veľa spoločného. Odborníci sa zhodujú v tom, že veľké dáta sú pojem, ktorý popisuje súbory dát tak obrovské, rýchle a komplexné, že je náročné ich spracovať pomocou tradičných metód. [1]

1.1.1 *Charakteristika veľkých dát*

Na začiatku roka 2001, Doug Laney, analytik spoločnosti Mega Group, publikoval výskumný príspevok pomenovaný 3D Data Management: Controlling Data Volume, Velocity, and Variety. Pojmy volume, velocity a variety predstavujúce 3V sa o desaťrošie neskôr stali všeobecne prijímanými definujúcimi rozmermi veľkých dát aj napriek tomu, že pojem veľké dáta sa v samotnom príspevku neobjavuje. Ide o prvú a najzákladnejšiu charakteristiku veľkých dát. [6]

Volume, predstavuje základ pyramídy veľkých dát. Objem popisuje veľkosť takých dátových súborov, ktoré sú veľké niekoľko stoviek terabajtov či petabajtov a je potrebné ich analyzovať a spracovať. Samotný objem dát vyžaduje odlišný prístup a taktiež odlišné technológie spracovania od tých tradičných. Inak povedané, dátové súbory sú príliš veľké na to, aby sme boli schopní ich spracovať bežnými počítačmi. Organizácie zozbierajú dáta z rôznych zdrojov, ktorými môžu byť rôzne obchodné transakcie, inteligentné zariadenia, priemyselné vybavenie, videá, sociálne médiá a mnoho ďalšieho. [7] [8]

Velocity predstavuje rýchlosť akou sa dáta generujú. Dáta generované vysokým tempom si taktiež vyžadujú odlišný prístup. Tie, ktoré sú určené pre analýzu a sú spracúvané klasickým spôsobom v dátových skladoch je nutné najprv transformovať a potom je možné ich pridať do analytických databáz. Okrem spravovania dát, firmy potrebujú aby bol prísun informácií rýchly, najlepšie čo najbližšie k reálnemu času. V tom prípade je potrebné, aby veľké dáta vstupovali do analytických procesov v surovom stave tak ako prichádzajú. Výkonný riaditeľ spoločnosti MetLife zdôrazňoval, že rýchlosť môže byť veľakrát podstatnejšia ako objem, keďže dokáže poskytnúť väčšiu konkurenčnú výhodu. [7] [9] [10]

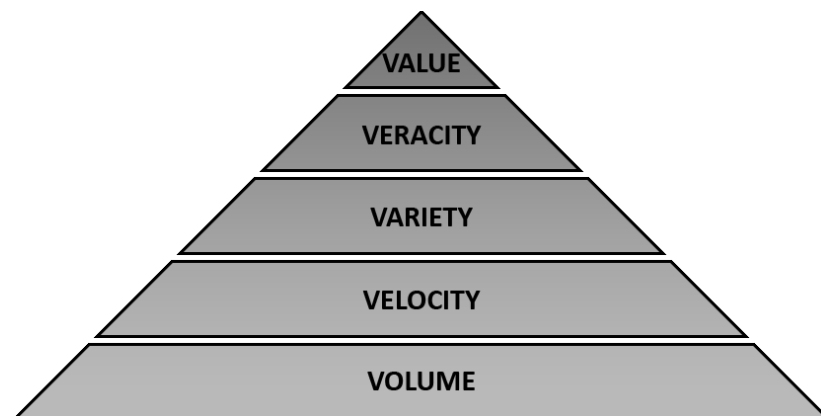
Variety znamená v preklade rozmanitosť a práve vďaka nej sú veľké dáta skutočne veľké. Dáta pochádzajú z rôznych zdrojov a všeobecne sa rozdeľujú na tri typy – štruktúrované, pološtruktúrované a neštruktúrované dáta. Štruktúrované dáta si môžeme predstaviť ako číselné údaje v tradičných databázach, tie neštruktúrované sú rôzne textové dokumenty, videá, audio súbory, finančné transakcie. Rozmanitosť typov dát veľakrát vyžaduje odlišné spracovanie a špeciálne algoritmy. Dôležitosť zdrojov informácií sa odlišuje v závislosti od povahy podnikania. [7] [8] [10]

S rozvojom technológie veľkých dát a takisto zložitou dát bola navrhnutá nová téza, podľa ktorej sa rozširuje samotná Laneyho charakteristika 3V o ďalšie pojmy. Prvým rozširujúcim pojmom je veracity, čo môžeme preložiť ako pravdivosť, presnosť. [11]

Veracity, odkazuje na kvalitu dát, ktoré sa analyzujú. Keďže dáta pochádzajú z obrovského množstva zdrojov, je ťažké ich prepojiť, spojiť, vyčistiť a transformovať medzi systémami. Dáta s vysokou pravdivosťou obsahujú veľa cenných záznamov, ktoré sú vhodné na analýzu a ktoré zmyselným spôsobom prispievajú k celkovým výsledkom. Naopak, dáta s príliš nízkou pravdivosťou obsahujú vysoké percento nezmyselných dát. Tie, ktoré neprinášajú žiadnu hodnotu sa označujú ako šum. [8]

Value, v preklade hodnota, uzatvárajúca skupinu 5V je vrcholom pyramídy veľkých dát. Hodnota poukazuje na schopnosť transformovať obrovské masy dát do podnikania. Po značnej investícii času a zdrojov, ak spoločnosť dokáže správne využívať veľké dáta, jej schopnosť spoznávať zákazníkov a monetizovať všetky tieto informácie je obrovská, keďže môžu ich zákazníkom ponúknuť v pravý čas to, čo chcú alebo potrebujú. Spoločnosť vďaka veľkým dátam dokáže lokalizovať zákazníkov, ktorí sú pre ňu najdôležitejší a vyhnúť sa zákazníkom, ktorí majú tendenciu rušiť objednávky, vyhľadávať zľavy alebo sa pokúšať o podvody. [10]

Obrázok 1: Pyramída 5V



Zdroj: Vlastné spracovanie

1.1.2 Štruktúrované a neštruktúrované dáta

Štruktúrované dáta označujú všetky dáta, ktoré sú obsiahnuté v relačných databázach a tabuľkách. Každá takáto tabuľka je svojou samostatnou entitou a je zvyčajne spojená aspoň s jednou ďalšou. Dobrá organizovanosť ich pomáha ľahko vyhľadať aj pomocou najjednoduchšieho vyhľadávacieho algoritmu. Takisto je ich možné jednoducho zadať, uložiť, analyzovať a dopytovať v relačných databázach. Ako príklad štruktúrovaných dát môžeme uviesť rôzne číselné, menové, abecedné dáta, ale taktiež aj mená, adresy a iné údaje. Štruktúrované dáta boli vítanou alternatívou k tradičným papierovým systémom správy dát, ktoré sú veľmi neštruktúrované a ich spravovanie je skrz ich objem pomalé a veľmi náročné. Problémom ale naďalej zostávala obmedzená úložná kapacita a tak štruktúrované dáta bolo potrebné rozšíriť o papierové alebo mikrofilmové úložiská. Spravovanie štruktúrovaných dát si vyžaduje využitie programovacieho jazyka, ktorý bol pôvodne vyvinutý spoločnosťou IBM v 70. rokoch pod názvom Structured Query Language, známejší pod skratkou SQL.

Neštruktúrované dáta označujú množiny dát, ktoré sú veľké, nesúvisiace, chaotické a neusporiadané do špecifických polí, čo znamená, že sa nedajú ľahko zobrazit'. Z tohto dôvodu je ťažké ich interpretovať tradičnými databázami a dátovými modelmi. Podľa niektorých odhadov neštruktúrované dáta tvoria viac než 80% všetkých dát v organizáciách a rastú desaťkrát až päťdesiatkrát rýchlejšie ako štruktúrované dáta. Na vyhľadanie neštruktúrovaných dát je potrebné použiť algoritmus sémantického vyhľadávania. Príkladom sú metadáta, fotografie a iné grafické obrázky, webové stránky, PDF súbory, dokumenty, videá, príspevky na sociálnych sieťach a podobne.

Pološtruktúrované dáta sú mixom štruktúrovaných a neštruktúrovaných. Dokonalým príkladom sú e-maily, keďže samotný text v e-maili reprezentuje neštruktúrované dáta, zatiaľ čo tie štruktúrované sú meno príjemcu, čas odoslania alebo e-mailová adresa odosielateľa. [1] [2]

1.1.3 *Historický vývoj veľkých dát*

V štyridsiatych rokoch minulého storočia sme mohli byť svedkami prvých pokusov o kvantifikovanie rýchlosti rastu objemu dát. Tento náhly rast môžeme poznať aj pod pojmom informačná explózia. Jeden z prvých autorov, ktorý sa k téme vyjadril, bol univerzitný knihovník Fremont Rider. Ten vo svojom diele *The Scholar and the Future of the Research Library* z roku 1944 odhadoval, že veľkosť amerických univerzitných knižníc sa bude zdvojnásobovať každých 16 rokov.

V roku 1967, B.A.Marron a P.A.D. de Maine napísali pre magazín *Communications of the ACM* článok s názvom *Automatic data compression*. V tomto článku autori písali o tom, že informačná explózia, ktorá bola zaznamenaná v posledných rokoch, vyžaduje, aby boli požiadavky na ukladanie všetkých informácií minimálne. Článok takisto popisoval, ako je možné znížiť pomalé externé požiadavky na uchovávanie dát a tiež zvýšenie rýchlosti prenosu informácií cez počítač.

O 2 roky neskôr, Ministerstvo pôšt a telekomunikácií v Japonsku prišlo s myšlienkou sčítania toku informácií. Toto sčítanie sa nakoniec uskutočnilo v roku 1975 a jeho hlavnou myšlienkou bolo sledovať objem informácií, ktoré kolovali v Japonsku. Jednotkou merania bol počet slov, ktorá bola jednotná pre všetky médiá. Sčítaním sa zistilo, že ponuka informácií rastie oveľa rýchlejšie ako ich spotreba. Spoločnosť sa posúvala do doby, v ktorej sa kládol dôraz na podrobnejšie informácie vyhovujúce individuálnym potrebám. O pár rokov neskôr, Maďarský ústredný štatistický úrad začal výskumný projekt, ktorý bol zameraný na informačné odvetvia krajiny, vrátane merania objemu informácií.

Pojem veľké dáta, ako ho poznáme dnes, sa začal používať v deväťdesiatych rokoch. Digitálne ukladanie dát sa stáva nákladovo efektívnejším ako ukladanie v papierovej forme. Dvojica Michael Cox a David Ellsworth publikovali v roku 1997 článok, v ktorom sa písalo o tom, že súbory dát sú dosť veľké a zaťažujú kapacity hlavnej pamäte, miestneho disku ale aj vzdialeného disku, čo taktiež môžeme nazvať problémom veľkých dát. Ide o prvý článok, ktorý sa nachádza v digitálnej knižnici Asociácie pre výpočtovú techniku, v ktorom sa vyskytuje výraz veľké dáta. Napriek tomu sa všetky zásluhy pripisujú Johnovi R. Masheyemu, bývalému hlavnému vedeckému pracovníkovi spoločnosti Silicon Graphics. Práve on spopularizoval tento pojem prejavom o problematike veľkých dát na stretnutí USENIX.

V roku 2000 vznikla prvá štúdia svojho druhu zameraná na kvantifikovanie celkového množstva nových a originálnych informácií uložených za rok 1999 v štyroch formách: papier, film, optické disky (CD, DVD) a magnetické úložiská. Štúdia zisťuje, že celý svet vyprodukoval približne 1,5 exabajtu (1 exabajt sa rovná 10^{18} bajtom dát) jedinečných informácií, čo predstavuje 250 megabajtov na každého jedného človeka na Zemi. Podobná štúdia, ktorá bola vykonaná v roku 2003 tými istými výskumníkmi hovorila o tom, že za rok 2002 sa po celom svete vyprodukovalo 5 exabajtov nových informácií, z čoho bolo 92% uložených na magnetických médiách, hlavne na pevných diskoch. V marci roku 2007 skupina vedcov publikuje ďalšiu štúdiu, ktorá vďaka údajom za predošlý rok odhadovala, že medzi rokmi 2006 a 2010 sa informácie v digitálnom vesmíre zväčšia 6-krát.

Do roku 2008 sa pojem veľké dáta používal veľmi málo, len malou skupinou ľudí v akademickej alebo priemyselnej oblasti. Vyhľadávače postupne menia spôsob prístupu k informáciám a takisto iné formy computingu veľkých dát pomáhajú transformovať aktivity rôznych spoločností, vedeckých výskumníkov, lekárov, ale aj obranné a spravodajské operácie národa. Computing veľkých dát je pravdepodobne najväčšou inováciou vo výpočtovej technike za prvé desaťročie 21. storočia. Od tejto doby sa pojem veľké dáta stal populárnym, začal sa bežne používať v obchodných kruhoch a médiách s pravidelnými komentármi v populárnych a vedeckých časopisoch, ale takisto aj v novinách. [6]

1.1.4 Využitie veľkých dát v súčasnosti

Nie je prehnané tvrdiť, že veľké dáta sú v súčasnosti prítomné na každom kroku. Bez prítomnosti veľkých dát v normálnej spoločnosti je stále ťažšie a ťažšie fungovať, aj keď mnohé zo zmien sú také drobné, že si ich len sotva všimneme. Väčšina organizácií si je vedomá, že ak zhromažďujú všetky dáta, ktoré plynú do spoločnosti, tak je ich možné využiť na analýzu a získať tak významné informácie, ktoré môžu výrazne dopomôcť k zvýšeniu efektívnosti.

Prvou oblasťou, v ktorej sa každý deň stretávame s veľkými dátami sú mobilné mapy a GPS. Pomocou satelitnej technológie GPS sme schopní zadať naše počiatočné a koncové body odkiaľkoľvek a získať tak cestovné pokyny v reálnom čase vo väčšine priemyselných miest na svete. So zavedením väčšieho počtu mobilných aplikácií a väčšieho počtu

inteligentných telefónov, ako aj s liberalizáciou telekomunikačných sietí, je potrebné obrovské množstvo veľkých dát na zabezpečenie presných smerov v reálnom čase. Môžeme si všimnúť na našich inteligentných zariadeniach, že časom sú poskytované pokyny službou Google Maps vždy lepšie a podrobnejšie, s možnosťou trasy pre cyklistov alebo verejnú dopravu. Tie rozvinutejšie služby a aplikácie môžu dokonca poskytovať trasy, ktoré budú najrýchlejšie z bodu A do bodu B, či poskytovať informácie o udalostiach na trase.

Od rastu veľkých dát v maloobchode sa celé odvetvie mení. Maloobchodníci používajú veľké dáta od momentu, keď ľudia začnú hľadať produkty prostredníctvom cielených reklám až po samotné doručenie ich zásielok. Veľké dáta sú prítomné pri online nakupovaní v podobe zaznamenávania našich úmyslov pomocou webových prehliadačov, takisto aj pomocou GPS, ktoré poskytujú efektívnejšie cielenie reklám. Vďaka týmto záznamom sa vytvára online profil kupujúceho, ktorý ale poskytuje maloobchodníkovi len zlomok informácii, ktorý mu umožňuje optimalizovať ponuku. Podstatne užitočnejší profil nákupu môže maloobchodník vytvárať pomocou sledovania kliknutí na webstránke predajcu. Spolu s demografickými a lokalizačnými informáciami umožňuje stabilizovať veľké množstvo dát od nakupujúcich prostredníctvom komplexného spracovania udalostí, čo umožňuje dynamickú segmentáciu a úspech maloobchodného predajcu v online priestore. Veľké dáta ovplyvňujú maloobchod aj v reálnom priestore, keďže dávajú možnosť rýchlo a presne prepraviť tovar v reálnom čase.

Veľké dáta sa neustále používajú v súvislosti s inteligentnými mestami na plánovanie mestských centier. Umožňujú plánovačom mesta rozvinúť nové chápanie fungovania miest a taktiež v oveľa kratších časových harmonogramoch, ako to bolo niekedy k dispozícii. Plánovačom miest otázka samotného plánovania zaberá niekoľko minút, hodín či prinajhoršom niekoľko dní na rozdiel od rokov až desaťročí. Ako príklad využitia veľkých dát v mestskom plánovaní môžeme uviesť ovplyvnenie funkčnosti verejnej dopravy dátami. Dopravné prostriedky verejnej dopravy dokážu sledovať tok cestujúcich a zaznamenávať dáta pomocou lístkových systémov v reálnom čase. Dopravcovia môžu tieto informácie využiť na určenie času špičky cestujúcich. Takisto vďaka dátam dokážu ohlásiť meškanie či nehody na trase v reálnom čase, ktoré dopomôžu cestujúcim upraviť svoje cestovné plány včas.

S využitím veľkých dát sa takisto stretáme pri otázke spotreby energie. Veľké dáta umožňujú inteligentným meračom, ktoré zhromažďujú dáta zo senzorov v celom mestskom

priestore, samoregulovať spotrebu energie pre čo najvyššiu efektivitu využitia energie. Dokážu určiť, kde sú prílivy a odtoky energie v určitom okamihu najvyššie a následne ju redistribuovať do najviac potrebných miest v celej sieti. Automaticky sa prispôbia, aby zabezpečili efektívnu distribúciu energie v sieti.

V dnešnej dobe je čoraz viac populárnejšia nositeľná elektronika, ktorá taktiež poskytuje veľké dáta. Nosenie takéhoto zariadenia, ktoré monitoruje aktivitu a spánok, ľudia využívajú aby udržali svoje zdravie a kondíciu. Poskytuje používateľom možnosť sledovať, či udržiavajú zdravé návyky, zapisovať si poznámky a mnoho ďalších funkcií, ktoré menia ich životný štýl. Všetky tieto funkcie vytvárajú dáta, ktoré poskytujú cennú spätnú väzbu pre ďalšie vylepšovanie. [13]

1.2 Umelá inteligencia

Ľudská inteligencia pokrýva široké spektrum spôsobov, akými sa prejavuje, či sa jedná o logické, priestorové schopnosti alebo emočné rozpoznávanie. Aj keď počítače prekonávajú ľudí pri rozsiahlych výpočtových úlohách, ich odborné znalosti sú užšie a schopnosti strojov v iných oblastiach zaostávajú za ľudskou inteligenciou. [3]

Tisíce rokov sa ako ľudstvo snažíme pochopiť, ako dokážeme myslieť, vnímať, chápať, predvídať a hlavne manipulovať svet. Oblasť umelej inteligencie, ako jedna z najnovších odborov v oblasti vedy a techniky, sa pokúša nielen porozumieť, ale aj budovať inteligentné entity. Práca v tejto oblasti začala krátko po druhej svetovej vojne a samotný názov vznikol v roku 1956. Zahŕňa širokú škálu rôznych oblastí a podoblastí, od tých všeobecných, ktorými môžu byť učenie alebo vnímanie, až po konkrétne, napríklad hranie šachu, dokazovanie matematických viet, ale aj riadenie auta na preplnenej ulici a diagnostika chorôb. [4]

Umelá inteligencia je podstatná pre každú intelektuálnu úlohu, čím dokazuje, že je skutočne univerzálna. Umožňuje strojom učiť sa z nadobudnutých skúseností, prispôbovať sa novým vstupom a vykonávať úlohy podobné tým ľudským. Vo veľkej miere sa spolieha na deep learning a spracovanie prirodzeného jazyka. Pomocou týchto a mnohých ďalších technológií môžu byť počítače trénované na vykonávanie konkrétnych

úloh spracovaním veľkého množstva dát a rozpoznávaním vzorcov, bez toho aby k tomu museli byť špeciálne naprogramované. [14]

1.2.1 ANI, AGI, ASI

Technológie umelej inteligencie sú kategorizované podľa ich schopnosti napodobňovať ľudské vlastnosti, technológie, ktoré na to využívajú, ich aplikácií v skutočnom živote a podľa teórie mysle. Ak použijeme tieto charakteristiky na porovnanie, všetky systémy umelej inteligencie, spadajú do jedného z troch typov umelej inteligencie, ktorými sú úzka umelá inteligencia, ktorá má úzky rozsah schopností, všeobecná umelá inteligencia, ktorá je na rovnakej úrovni ako ľudské schopnosti a umelá superinteligencia, ktorá je schopnejšia ako človek.

Úzka umelá inteligencia alebo inak aj slabá umelá inteligencia, ktorá sa označuje ako ANI z jej anglického názvu artificial narrow intelligence, je jediný typ umelej inteligencie, ktorý sme do dnešného dňa úspešne zrealizovali. Je zameraná na konkrétny cieľ, a tým pádom je navrhnutá na vykonávanie len samostatných úloh, ktorými môžu byť rozpoznávanie tváre, rozpoznávanie reči, hlasoví asistenti, riadenie vozidla alebo prehľadávanie internetu. Nenapodobňuje ani nereplikuje ľudskú inteligenciu, iba simuluje ľudské správanie na základe úzkeho rozsahu parametrov a súvislostí. Pri plnení konkrétnej úlohy, na ktorú je naprogramovaná, je vysoko inteligentná a účinná. Napriek tomu fungujú pod úzkou skupinou obmedzení, vďaka čomu sa tento typ umelej inteligencie označuje ako slabá. Môže byť reaktívna alebo môže mať obmedzenú pamäť. Reaktívna umelá inteligencia je neuveriteľne základná, čo znamená, že nemá schopnosť pamäte alebo ukladania dát a napodobňuje schopnosť ľudskej mysle reagovať na rôzne druhy stimulov bez predchádzajúcich skúseností. Umelá inteligencia s obmedzenou pamäťou je pokročilejšia a je vybavená schopnosťami ukladania dát a učenia sa, ktoré umožňujú strojom využívať historické údaje, na základe ktorých udávajú rozhodnutia. Väčšina umelej inteligencie je práve umelou inteligenciou s obmedzenou pamäťou, kde stroje využívajú veľké objemy dát na deep learning, ktorý umožňuje personalizovanú skúsenosť z umelej inteligencie, ako napríklad virtuálnych asistentov alebo vyhľadávače, ktoré ukladajú údaje o užívateľovi a prispôbujú jeho budúce skúsenosti. Systémy umelej inteligencie sa dnes používajú

napríklad v medicíne na diagnostiku rakoviny a iných chorôb s extrémnou presnosťou prostredníctvom replikácie ľudského uvažovania a poznania.

Všeobecná umelá inteligencia, tiež známa pod pojmmami silná alebo hlboká umelá inteligencia, z anglického názvu artificial general intelligence označovaná ako AGI, je koncept stroja so všeobecnou inteligenciou, ktorý napodobňuje ľudskú inteligenciu a správanie, so schopnosťou učiť sa a využívať svoju inteligenciu na vyriešenie akéhokoľvek problému. Tento typ umelej inteligencie môže myslieť, rozumieť a konať spôsobom, ktorý je v každej situácii na nerozoznanie od človeka. Vedci a výskumníci v oblasti umelej inteligencie zatiaľ nedosiahli túto úroveň. Aby to bolo vôbec možné, museli by nájsť spôsob, ktorým by dosiahli, že stroje by si svoju činnosť plne uvedomovali, k čomu by sa dalo dopracovať naprogramovaním celej sady kognitívnych schopností. Stroje by museli posunúť učenie zo skúseností na vyššiu úroveň, nielen zvýšiť efektivitu pri jednotlivých úlohách, a takisto aj získať schopnosť aplikovať vedomosti z nadobudnutých skúseností na širšiu škálu rôznych problémov. Silná umelá inteligencia využíva teóriu mysle umelej inteligencie, ktorá sa vzťahuje na schopnosť rozlišovať potreby, emócie, názory a myšlienkové procesy iných inteligentných subjektov. S pokrokom v technológii rozpoznávania obrazu a tváre je pravdepodobné, že sa schopnosti strojov učiť sa a vidieť zlepšia.

Umelá superinteligencia, označovaná ako ASI z anglického artificial super intelligence, je hypotetická umelá inteligencia, v ktorej stroje nielen napodobňujú a rozumejú ľudskej inteligencii, ale takisto si uvedomujú samých seba a prekonávajú kapacitu ľudskej inteligencie a schopností. Koncept umelej superinteligencie vidí umelú inteligenciu vyvíjať sa tak, že bude podobná ľudským emóciám a skúsenostiam, bude im rozumieť a takisto bude vyvolávať emócie, potreby, presvedčenia a túžby. Okrem replikovania všestrannej ľudskej inteligencie by umelá superinteligencia mala byť teoreticky lepšou vo všetkom čo ľudia vykonávajú, ako napríklad matematika, prírodoveda, medicína, ale aj umenie, koníčky, šport a vzťahy, takisto by mala mať väčšiu pamäť a rýchlejšiu schopnosť spracovávať a analyzovať dáta či rôzne podnety. [15]

1.2.2 Princípy fungovania umelej inteligencie

Budovanie systému umelej inteligencie je starostlivý proces reverzného inžinierstva ľudských vlastností a schopností v stroji a jeho výpočtová zdatnosť prekonáva to, čoho sú ľudia schopní. Aby sme porozumeli, akým spôsobom funguje umelá inteligencia, je potrebné sa zamerať na rôzne podoblasti umelej inteligencie a vďaka nim pochopiť, ako je možné tieto oblasti uplatniť v rôznych priemyselných odvetviach.

Machine learning, v preklade strojové učenie, je prvou podoblasťou umelej inteligencie. Strojové učenie poskytuje počítačovým systémom schopnosť automaticky sa učiť a zdokonaľovať z predošlých skúseností bez toho, aby na to boli vyslovene naprogramované. Dokáže identifikovať vzorce, analyzovať dáta z predchádzajúcich období a tak dospieť k možnému záveru bez toho, aby musel zahrnúť ľudskú skúsenosť. Takáto automatizácia šetrí firmám obrovské množstvo času a pomáha pri lepšom rozhodovaní.

Ďalšou oblasťou je deep learning, hlboké učenie, ktoré je technikou strojového učenia. Učí stroj spracovávať viaceré vstupy pomocou spolupracujúcich vrstiev s cieľom ich zaradiť, odvodiť a predvídať jeden konkrétny výsledok, na čo využíva umelé neurónové siete, ktoré napodobňujú biologické neurónové siete v ľudskom mozgu.

Neurónové siete fungujú na podobných princípoch ako ľudské nervové bunky. Ide o sériu algoritmov zachytávajúce vzťah medzi rozličnými základnými premennými a spracovávajú dáta tak, ako to robí ľudský mozog. Tieto siete sa učia spracovaním tréningových príkladov, ktoré sú vo forme obrovských dátových súborov. Dáta analyzuje mnohokrát, aby našiel asociácie a dal zmysel predtým nedefinovaným dátam.

Spracovanie prirodzeného jazyka, je vedná disciplína o čítaní, porozumení a interpretácii jazyka strojom. Akonáhle stroj pochopí, čo má užívateľ v úmysle povedať, odpovie zodpovedným spôsobom. Konečným cieľom spracovania prirodzeného jazyka je umožniť plynulú interakciu so strojmi, a to výučbou systémov na porozumenie ľudskej reči v kontexte a vytváraním logických odpovedí.

Algoritmy počítačového videnia sa snažia porozumieť obrazu takým spôsobom, že ho najprv rozoberú na viaceré časti a následne tieto časti študujú. Toto pomáha stroju zatriediť a učiť sa zo sady obrázkov aby mohol následne urobiť lepšie výstupné rozhodnutie na základe predchádzajúcich pozorovaní.

Kognitívne výpočty sa ich algoritmami snažia napodobniť ľudský mozog analýzou textu, reči, obrázkov a objektov spôsobom, ktorý robí a snaží sa poskytnúť požadovaný výstup a taktiež zlepšovať interakciu medzi ľuďmi a strojmi. Snažia sa pretvoriť proces myslenia človeka do počítačového modelu pochopením ľudského jazyka a významu obrázkov. Spoločne sa spolu s umelou inteligenciou snažia o vybavenie strojov chovaním podobným človeku a schopnosťami spracovania informácií. [16] [17]

1.2.3 Historický vývoj umelej inteligencie

Od konca staroveku až do konca 16. storočia, rôzni matematici, teológovia, filozofi, profesori a autori uvažovali o mechanických technikách, počítačích prístrojoch a o numerických systémoch, ktoré nakoniec viedli ku koncepcii mechanizovaného ľudského myslenia u neľudských bytostí. [19]

V roku 1763 Thomas Bayes vyvinul štruktúru pre úvahu o pravdepodobnosti udalostí. Hovorilo sa, že Bayesovský logický úsudok sa stane vedúcim prístupom v strojovom učení. O takmer sto rokov neskôr prišiel George Boole s tvrdením, že logické uvažovanie je možné vykonávať systematicky rovnakým spôsobom ako pri riešení sústavy rovníc. Nikola Tesla predviedol na výstave elektriny, ktorá sa konala v roku 1898 v záhrade Madison Square Garden, prvé rádiové riadené plavidlo na svete. V roku 1914 predstavil španielsky stavebný inžinier prvý šachový stroj, ktorý bol schopný koncovky pomocou kráľa a veže proti kráľovi bez ľudského zásahu.

Spoločnosť Houdina Radio Control v roku 1925 uviedla na trh rádiové riadené auto bez vodiča, ktoré jazdilo v uliciach New Yorku. O pár rokov neskôr, Makoto Nishimura navrhol robota pomenovaného Gakutensoku, čo je v japončine pojem pre „učenie sa zo zákonov prírody“. Tento robot mohol meniť jeho výraz tváre a pohybovať hlavou a rukami pomocou mechanizmu tlaku vzduchu.

Warren S. McCulloch a Walter Pitts publikovali v roku 1943 článok s názvom A Logical Calculus of the Ideas Immanent in Nervous Activity pre spravodajstvo Bulletin of Mathematical Biophysics. V tomto vplyvnom príspevku diskutovali o sieťach idealizovaných a zjednodušených umelých „neurónov“ a taktiež o tom, ako môžu vykonávať jednoduché logické funkcie. Príspevok sa stal inšpiráciou pre počítačové

„neurónové“ siete a ich populárny popis ako napodobňovanie mozgu. Neskôr sa stal inšpiráciou aj pre deep learning. Na začiatku päťdesiatych rokov sa stal Claude Shannon autorom prvého zverejneného článku o vývoji počítačového programu na hranie šachu. V tom istom roku publikoval Alan Turing knihu s názvom Computing Machinery and Intelligence, v ktorej navrhol takzvanú napodobňovaciu hru, ktorá sa neskôr stala známou ako „Turingov test“.

Marvin Minsky a Dean Edmonds zostrojili v roku 1951 SNARC, celým názvom Stochastic Neural Analog Reinforcement Calculator, prvú umelú neurónovú sieť, ktorá bola nasimulovaná pomocou 3000 vákuových trubíc. V roku 1955, John McCarthy sa vtedy ešte ako mladý odborný asistent matematiky rozhodol zorganizovať skupinu na objasnenie a rozvoj myšlienok o mysliacich strojoch. Nová oblasť dostala pomenovanie „umelá inteligencia“, pre ktorú sa stal všeobecne považovaným dátumom vzniku pracovný seminár konajúci sa o rok neskôr.

Koncom päťdesiatych rokov 20. storočia bolo predstavených viacero inovácií v oblasti umelej inteligencie. V roku 1957 Frank Rosenblatt vyvinul Perceptron, ranú umelú neurónovú sieť, ktorá umožňovala rozpoznávať vzory na základe dvojvrstvovej počítačovej vzdelávacej siete. Túto sieť nazval denník New York Times zárodkom elektronického počítača, od ktorého sa očakávalo, že bude schopný chodiť, rozprávať, vidieť, písať, reprodukovat' sa a byť si vedomý svojej inteligencie. Nasledovný rok bol Johnom McCarthym vyvinutý programovací jazyk Lisp, ktorý sa stal najpopulárnejším programovacím jazykom používaným pri výskume umelej inteligencie. Taktiež sa prvýkrát použil pojem machine learning hovoriaci o umelej inteligencii, ktorá sa dokáže sama učiť z dát a potom tieto nadobudnuté vedomosti aplikovať bez ľudského zásahu.

Začiatkom šesťdesiatych rokov bol zostrojený prvý priemyselný robot s menom Unimate, ktorý začal pracovať na montážnej linke v závode General Motors v New Jersey. V roku 1966 bol predstavený robot Shakey, prvý univerzálny mobilný robot, ktorý dokáže rozmýšľať o svojich činoch. Pre časopis Life sa v článku vyjadril Marvin Minsky, v ktorom s istotou vyslovil názor, že v nasledujúcich troch až ôsmich rokoch by všeobecná inteligencia strojov mohla dosiahnuť úroveň priemerného človeka. Koncom desaťročia bol vyvinutý aj počítačový program pre pochopenie prirodzeného jazyka.

Rok 1970 predstavoval zostrojenie prvého robota s názvom WABOT-1, ktorý mal ľudské vlastnosti. Robot bol vyrobený v Japonsku a pozostával zo systému na kontrolu

končatín, systému videnia a systému konverzácie. Neskôr bol na Stanfordskej univerzite vyvinutý systém MYCIN, prvotný expertný systém na identifikáciu baktérií spôsobujúcich ťažké infekcie a na odporúčanie antibiotík. V roku 1979 bol uvedený jeden z prvých príkladov autonómneho vozidla, Stanfordský vozík, ktorý úspešne prešiel miestnosťou naplnenou stoličkami bez ľudského zásahu.

Prvý automobil bez vodiča bol zostrojený v roku 1986 pod vedením Ernsta Dickmansa na univerzite Bundeswehr v Mníchove. Šlo o dodávku značky Mercedes-Benz vybavenú kamerami a senzormi, ktorá dokázala jazdiť po prázdnych uliciach rýchlosťou až 55 míľ za hodinu. Publikácia Judea Pearla z roku 1988, ktorá nesie názov *Probabilistic Reasoning in Intelligent Systems*, vytvorila reprezentatívny a výpočtový základ pre spracovanie informácií za neistoty. Autor sa v diele zaslúžil o vynález Bayesovských sietí, matematický formalizmus pre definovanie zložitých pravdepodobnostných modelov, ale aj hlavné algoritmy použité na odvodenie týchto modelov. Táto práca bola nielen revolučnou v oblasti umelej inteligencie, ale taktiež sa stala dôležitým nástrojom pre mnoho ďalších odvetví strojárstva a prírodných vied.

V publikácii *The Coming Technological Singularity* z roku 1993 od Vernora Vinga sa predpovedalo, že v najbližších tridsiatich rokoch nadobudneme technologické prostriedky na vytvorenie nadľudskej inteligencie, čím sa krátko na to skončí ľudská éra. Koniec tisícročia rovnako priniesol niekoľko inovácií a vynálezov. Sepp Hochreiter a Jürgen Schmidhuber navrhli sieť Long Short-Term Memory, označujúcu sa skratkou LSTM, čo je typom rekurentnej neurónovej siete, ktorá sa dnes používa na rozpoznávanie rukopisu a rozpoznávanie reči.

V Mohavskej púšti ležiacej v Kalifornii sa v roku 2004 konala prvá súťaž pre autonómne vozidlá s názvom DARPA Grand Challenge. Trať dlhú 150 míľ nedokončilo žiadne zo súťažiacich autonómnych vozidiel. Neskôr sa zaviedol pojem „machine reading“, ktorý môžeme definovať ako nekontrolované autonómne chápanie textu. Kniha Geoffreyho Hintona s názvom *Learning Multiple Layers of Representation* sumarizovala nápady podnecujúce k viacvrstvovým neurónovým sieťam, ktoré obsahujú spojenie zhora nadol a ktoré sú trénované, aby boli schopné generovať senzorické údaje, radšej ako ich klasifikovať. V roku 2009 začala spoločnosť Google tajne vyvíjať auto bez vodiča a o päť rokov nato prešlo ako prvé skúšku samoriadenia v americkej Nevade. Od roku 2010 sa každoročne koná súťaž v rozpoznávaní objektov umelou inteligenciou, ktorá nesie názov

ImageNet Large Scale Visual Recognition Challenge, skrátene ILSVRC. Výborný výsledok dosiahla konvolučná neurónová sieť, ktorá vyhrala súťaž v rozpoznávaní nemeckých dopravných značiek s presnosťou 99,46%, pričom porazila ľudí a ich výsledok 99,22%. Výskumníci z IDSIA vo Švajčiarsku uviedli len 0,27 percentnú chybovosť pri rozpoznávaní rukopisu pomocou konvolučných neurónových sietí, čo predstavovalo výrazné zlepšenie oproti predchádzajúcim rokom. Taktiež konvolučná neurónová sieť navrhnutá výskumníkmi Torontskej univerzity dosiahla chybovosť iba 16% v súťaži ILSVCR, čo bolo lepším výsledkom v porovnaní s predošlým rokom. [18]

1.2.4 Využitie umelej inteligencie v súčasnosti

Podľa písomnej správy Texaského dopravného inštitútu, v roku 2014 sa v USA časy dochádzania jedného dochádzajúceho predĺžili o 42 hodín dopravného zdržania v dopravnej špičke. Pomocou anonymizovaných údajov o polohe zo smartfónov môžu Google Mapy analyzovať rýchlosť pohybu dopravy v danom okamihu. Vďaka nadobudnutiu dopravnej aplikácie Waze, ktorá funguje na základe získavania hlásení rôznych udalostí na ceste od jej používateľov, môže aplikácia Google Mapy ľahšie začleniť tieto dopravné udalosti. Prístup k obrovskému množstvu dát prenášajúceho sa do jej vlastných algoritmov znamená, že Google Mapy môžu skrátiť čas dochádzania pomocou navrhnutia najrýchlejšej trasy. Takisto spoločnosť Uber využíva machine learning pre výpočet času príjazdu, výpočet optimálnych miest vyzdvihnutia a taktiež na detekciu podvodov. Vďaka tomu môžu vopred vypočítať cenu vašej jazdy a minimalizovať čas čakania. Umelá inteligencia využívaná v autopilotoch v komerčných leteckých spoločnostiach predstavuje prekvapivo skoré použitie tejto technológie, ktoré sa datuje späť už v roku 1914. V článku pre New York Times sa uvádza, že priemerný let lietadla Boeing zahŕňa iba sedem minút letu riadeného človekom, ktoré sú zvyčajne vyhradené iba na vzlet a pristátie.

E-mailová schránka vyzerá na prvý pohľad ako nepravdepodobné miesto pre umelú inteligenciu, ale práve umelá inteligencia je technológiou, ktorá vo veľkej miere poháňa jednu z najdôležitejších funkcií e-mailovej schránky, spamový filter. Filtre založené na jednoduchých pravidlách nie sú proti spamu účinné, pretože spameri dokážu rýchlou aktualizáciou správ tieto filtre obísť. Namiesto toho sa používajú spamové filtre, ktoré sa neustále učia z rôznych signálov, ako sú napríklad odosielateľ správy, odkiaľ správa

pochádza alebo konkrétne slová v správe. Je potrebné, aby filtre personalizovali výsledky, pretože pre každého používateľa spam predstavuje niečo iné. Služba Gmail vďaka použitiu algoritmov strojového učenia úspešne filtruje 99,9% spamu.

S umelou inteligenciou sa môžeme stretnúť aj v oblasti bankovníctva a osobných financií. Väčšina veľkých bánk ponúka možnosť vkladať šeky priamo prostredníctvom aplikácií pre smartfóny, čo odstraňuje potrebu doručovať šeky fyzicky do banky. Veľká časť týchto bánk sa spolieha na technológiu vyvinutú spoločnosťou Mitek, ktorá využíva umelú inteligenciu a machine learning na dešifrovanie a prevod rukopisu na šekoch na text pomocou optického rozoznávania znakov. Vo veľkom množstve prípadov je denný objem transakcií príliš vysoký na to, aby bolo možné skontrolovať každú transakciu manuálne. Namiesto toho sa umelá inteligencia používa na vytváranie systémov, ktoré sa učia, aké typy transakcií sú podvodné. Popredná spoločnosť v oblasti analytiky FICO, ktorá vytvára známe úverové ratingy používané na stanovenie úverovej schopnosti, používa neurónové siete na predpovedanie podvodných transakcií. Faktory, ktoré môžu konečné výstupy tejto neurónovej siete ovplyvniť, zahŕňajú frekvenciu transakcií, veľkosť transakcie a druh zapojeného maloobchodníka. Pri žiadosti o pôžičku alebo kreditnú kartu musí finančná inštitúcia zistiť, či žiadateľovi vyhoví, a ak áno, aké konkrétne podmienky mu ponúknuť. Spoločnosť FICO používa machine learning pri tvorbe skóre, ktoré potom väčšina bánk využíva na prijímanie úverových rozhodnutí a taktiež pri určovaní konkrétneho hodnotenia rizika pre jednotlivých klientov.

Facebook využíva umelú inteligenciu na rozpoznávanie tváří na fotografiách. Akonáhle používateľ nahrá fotografiu, služba automaticky zvýrazní tváre a navrhne priateľov na označenie. Na túto funkciu využíva umelé neurónové siete – algoritmy machine learningu, ktoré napodobňujú štruktúru ľudského mozgu poháňajúce softvér na rozpoznávanie tváre. Táto sociálna sieť používa umelú inteligenciu taktiež na prispôsobenie noviniek, príspevkov a reklám, ktoré sa používateľovi zobrazia a môžu ho zaujímať. Od roku 2016 využíva DeepText, nástroj na porozumenie textu, o ktorom spoločnosť tvrdí, že dokáže s takmer ľudskou presnosťou porozumieť textovému obsahu niekoľkých tisíc príspevkov za sekundu vo viac ako 20 jazykoch. Tento nástroj sa využíva na zisťovanie zámerov a automatizáciu odstraňovania spamu v aplikácii Messenger.

Umelá inteligencia pri online nakupovaní dokáže v momente ukázať zoznam najrelevantnejších produktov, ktoré súvisia so zadanými parametrami zákazníka. Algoritmy

sa automaticky učia kombinovať viaceré znaky dôležitosti. Štruktúrované dáta poskytujú mnoho dôležitých funkcií, ktoré pomáhajú učiť sa z minulých vzorcov vyhľadávania. Zákazníci majú možnosť vidieť odporúčané produkty, ktoré si predošlí zákazníci zobrazili alebo kúpili spolu s ich vyhľadávaným produktom a to podľa štúdií môže spoločnosti priniesť až 30 percentný nárast tržieb.

Dnes už štandardnou funkciou smartfónov je prevádzanie hlasových povelov na text a to jednoduchým stlačením tlačidla alebo vyslovením konkrétnej frázy. Tento automatizovaný prepis v minulosti presahoval možnosti aj tých najpokročilejších počítačov. V dnešnej dobe, keď je táto technológia dostatočne presná a spoľahlivá, stala sa ovládacím rozhraním pre novú generáciu inteligentných osobných asistentov. Prvými iteráciami boli jednoduchší telefónni asistenti ako Siri a Google Now, ktorí boli schopní vyhľadávať na internete, nastavovať pripomienky a byť prepojený s kalendárom. Spoločnosť Amazon uviedla na trh Alexu, svojho osobného hlasového asistenta využívajúceho umelú inteligenciu, ktorý prijíma hlasové príkazy na vytváranie zoznamov úloh, objednávanie produktov online, nastavovanie pripomienok a takisto hľadanie odpovedí na rôzne otázky online. Spoločnosť Microsoft vyvinula svojho vlastného asistenta pomenovaného Cortana, ktorý je predinštalovaný v počítačoch so systémom Windows a smartfónoch spoločnosti Microsoft. [20]

2 Cieľ práce

Cieľom tejto bakalárskej práce je priblížiť čitateľovi problematiku veľkých dát a umelej inteligencie, poskytnúť možné spôsoby riešenia problematiky a následne sa podrobnejšie zamerať na jednu z nich.

V prvej kapitole sme si bližšie definovali základné pojmy veľkých dát a umelej inteligencie a taktiež ich základné rozdelenia. Ďalej sme sa sústredili na popis princípu fungovania umelej inteligencie. Pre správne pochopenie bolo potrebné priblížiť historický vývoj oboch problematík. Časť kapitoly sme venovali aj využitiu veľkých dát a umelej inteligencie v súčasnosti.

V ďalšej kapitole je potrebné sa oboznámiť so základnými metódami, ktorými sú metaheuristiky. Metaheuristiky tvoria pomyselnú hranicu medzi umelou inteligenciou a operačným výskumom. Vymedzíme si základné informácie a ich delenie. Objasníme si jednotlivé metaheuristiky, ku ktorým si uvedieme ich základné princípy.

Nadobudnuté poznatky rozšírime v poslednej kapitole, kde sa zameriame na jednu konkrétnu metaheuristicu, ku ktorej si podrobne popíšeme jej princíp, jej všeobecný postup riešenia a takisto postup zameraný na jednu vybranú oblasť, ktorou je operačný výskum, konkrétne problém obchodného cestujúceho.

3 Metodika práce a metódy skúmania

Heuristiky môžeme definovať ako kritéria, metódy ale aj ako princípy pre rôzne rozhodnutia vyzerajúce ako najefektívnejšie pre dosiahnutie cieľa. Sú špecificky navrhnuté pre konkrétny problém a sú založené na odhade. Vďaka pokroku výpočtového výkonu počítačov sa vo veľkom začali používať metaheuristické metódy. Metaheuristiky sú všeobecné postupy aplikované na daný problém inšpirované prírodnými javmi a procesmi, ktoré sú založené na stochastickom prehľadávaní. Heuristiky s podstatou založenou na podobnosti s prírodnými procesmi tvoria pomyselnú hranicu medzi umelou inteligenciou a operačným výskumom. Ich hlavnou výhodou oproti klasickým heuristikám je schopnosť opustiť nájdený lokálny extrém účelovej funkcie a pomocou následnej postupnosti iteračných krokov sa dostať k riešeniu s lepšou hodnotou účelovej funkcie. Napriek tomu, rovnako ako heuristiky, ani metaheuristiky nám neposkytujú istotu, že nájdeme optimálne riešenie.

Rozlišujeme 2 typy metaheuristík, a to metaheuristiky využívajúce prehľadávanie založené na jednom riešení a metaheuristiky využívajúce prehľadávanie založené na populácii riešení. Toto rozdelenie spočíva v miere využitia princípu intenzifikácie, resp. diverzifikácie. Princíp intenzifikácie hovorí o zacielení sa na jedno dobré riešenie, ktoré sa miernymi úpravami vylepší. Svoju pozornosť sústreďujú na sľubné oblasti, ktoré sa určujú na základe už získaných dobrých riešení. Princíp diverzifikácie hovorí o preskúmaní čo najväčšieho množstva rôznych riešení v zatiaľ nepreskúmaných oblastiach.

Metaheuristiky využívajúce prehľadávanie založené na jednom riešení sú:

- Hill Climbing
- Simulated Annealing
- Tabu Search
- Iteračné lokálne prehľadávanie

Metaheuristiky využívajúce prehľadávanie založené na populácii riešení sú:

- Evolučné algoritmy
- Inteligencia roja

V nasledujúcej časti si podrobnejšie uvedieme tie najznámejšie z nich. [21]

3.1 Simulated Annealing

Simulated Annealing, v preklade simulované chladenie (žíhanie), je metaheuristika patriaca medzi metaheuristiky s prehľadávaním založeným na jednom riešení. Táto metóda bola inšpirovaná Hill Climbing algoritmom, ktorý prehľadáva množinu všetkých riešení, a fyzikálnymi dejmi prebiehajúcimi pri žíhaní telesa, ktoré slúžili k odstráneniu vnútorných defektov. Zahriatím telesa na vysokú teplotu a postupným ochladzovaním sa rovnovážne polohy všetkých atómov stabilizujú, a pri konečnej teplote nižšej ako počiatočnej teleso neobsahuje žiadne vnútorné defekty. Na začiatku osemdesiatych rokov prišli výskumníci na to, že by tieto deje bolo možné využiť na hľadanie globálneho minima.

Keďže táto metóda bola inšpirovaná Hill Climbing algoritmom, je potrebné ho v krátkosti priblížiť. Tento algoritmus začína náhodným vygenerovaním počiatočného riešenia, za čím nasleduje postupné generovanie konečného počtu riešení, ktoré ležia v určitom okolí počiatočného riešenia. Z nadobudnutých riešení sa potom vyberie najlepšie riešenie z hľadiska účelovej funkcie. Nevýhodou algoritmu je akceptovanie rovnako dobrého alebo lepšieho riešenia, ako bolo počiatočné riešenie, čo môže mať za následok, že sa dostaneme k nevýraznému lokálnemu minimu a optimálne riešenie nikdy nedosiahneme. Opakovaným spúšťaním algoritmu je možné tento nedostatok odstrániť, čo znamená, že stochastickosť tejto metódy spočíva len v hľadaní počiatočného riešenia.

Metóda simulovaného chladenia rieši spomínané uviaznutie v lokálnom minime tým, že s určitou pravdepodobnosťou prijíma aj riešenia, ktoré sú horšie ako tie počiatočné. Východiskom metódy je určité počiatočné riešenie, ktoré postupne obohacujeme o ďalšie riešenia hľadané v iteráciách a ktoré porovnávame s aktuálne najlepším riešením. Obsahom porovnania je tzv. stochastický operátor, ktorý premieňa aktuálne riešenie na nové. O prijatí nového riešenia rozhoduje Metropolisove kritérium, ktoré určuje pravdepodobnosť nahradenia riešenia. Hodnota pravdepodobnosti prijatia nového riešenia je vo výraznej miere ovplyvnená teplotou v danej iterácii. Veľké hodnoty teploty značia pravdepodobnosť blízku jednej, čo znamená prijatie takmer všetkých horších riešení. Naopak, teplota blízka nule hovorí o výnimočne prijatých horších riešeniach. Vo väčšine prípadov je teplota stupňovito klesajúcou funkciou, kde sa počiatočná hodnota po každej iterácii vynásobí hodnotou α , ktorá nadobúda hodnoty z intervalu $(0 ; 1)$. To znamená, že časom pravdepodobnosť prijatia horšieho riešenia klesá. V praxi α najčastejšie nadobúda hodnoty 0,8 až 0,99. Nevýhodou

tejto metódy je množstvo nastaviteľných parametrov - α , teplota, maximálny počet iterácií. [21]

3.2 Tabu Search

Tabu Search, alebo aj zakázané prehľadávanie, je jednou z najviac rozšírených metód založených na lokálnom prehľadávaní. Rovnako ako metóda simulovaného chladenia, aj metóda zakázaného prehľadávania vychádza z metódy Hill Climbing. Je jeho vylepšenou verziou, čo spočíva v zavedení mechanizmu krátkodobej pamäte.

Krátkodobá pamäť je jeho charakteristickou črtou a zabraňuje zacykleniu iterovania, inými slovami, zabraňuje návratu k riešeniu, ktoré už bolo predtým preskúmané. Počiatočné riešenie sa postupne vylepšuje lokálnymi úpravami. Aplikáciou úprav alebo posunmi na aktuálne riešenie získavame množinu susedov predstavujúcu množinu riešení. Riešenia, ktoré obsahujú niektorú zo zapamätaných vlastností, nie sú prípustné pre určitý počet iterácií. Inými slovami, algoritmus pracuje s tzv. tabu listom, do ktorého sú vo väčšine prípadov ukladané transformácie inverzné k zakázaným činnostiam. Napríklad, ak riešime problém okružnej cesty, v prípade, že bol zákazník presunutý z cesty A na cestu B, návrat na cestu A je na určitý počet iterácií zakázaný. Ak nastane situácia, že tabu list dosiahol svoju maximálnu kapacitu, najstaršia transformácia sa z neho automaticky vylúči, teda môžeme povedať, že sa tabu list cyklicky obnovuje. Celý výpočet môžeme zastaviť nastavením časového obmedzenia alebo počtu transformácií bez zlepšenia aktuálneho riešenia.

Ďalšími vlastnosťami obsiahnutými v algoritme sú rôznotvárnosť a zosilnenie. Zosilnenie je výraznejším prehľadávaním priestoru okolo najlepších poznaných riešení. Rôznotvárnosť má za úlohu zaistiť, aby proces prehľadávania nezanechal príliš veľké nepreskúmané podmnožiny blízkych riešení, čo je možné docieľiť zapamätávaním postupnosti posledných riešení. [21]

3.3 Genetické algoritmy

Určitú postupnosť krokov, ktorá vedie k nájdeniu optimálneho riešenia a využíva pri tom počítačovú simuláciu darwinovskej evolúcie, nazývame genetický algoritmus. Tento typ algoritmov, simulujúci proces neustáleho vývoja prírody založeného na výbere, krížení a mutáciách v priebehu jednotlivých generácií, patrí do kategórie evolučných algoritmov. Evolučné algoritmy sú metaheuristikami založenými na populácii, preto sa nazývajú aj populačnými metaheuristikami.

Základná idea tohto prístupu spočíva v tom, že jedinec, ktorý má lepšie schopnosti ako obyčajná populácia, má vyššiu šancu aby prežil a stal sa rodičom. Potomkovia týchto jedincov majú ešte lepšie schopnosti a vyššiu šancu prežiť. Chromozóm, pojem prevzatý z genetickej terminológie, predstavuje každé jedno riešenie zakódované do genetického reťazca. Zakódovanie riešenia je potrebné, aby bolo možné genetický algoritmus použiť. Samotné položky vyskytujúce sa v genetickom reťazci sa nazývajú gény a vo väčšine prípadov nadobúdajú hodnoty 0 a 1. Počiatočná populácia vznikne náhodným vygenerovaním množiny chromozómov, ktoré sa každou generáciou zlepšujú do momentu, kým nie je splnená určitá podmienka zastavenia. [21]

3.4 Kolónia mravcov

Metóda kolónie mravcov patriaca medzi metaheuristiky inteligencie roja má taktiež svoj základ v biológii. Ako z názvu vyplýva, táto metóda je založená na odpozorovanom správaní mravcov v kolónii, ktoré hľadajú najkratšiu cestu z mraveniska k potrave a späť. Jednotlivé mravce si na začiatku hľadajú cestu náhodne, no následne, po tom ako potravu objavia, zanechávajú za sebou pachovú stopu, vďaka čomu sú ďalšie mravce tiež schopné nájsť potravu. Pachové stopy, konkrétnejšie čiastočky feromónov, svojou kumuláciou množstva určujú hodnotu výhodnosti cesty. Viac pachovej stopy mravec zanechá, ak je cesta kratšia, teda na jej prejdanie je potreba menej času. Čím väčšie množstvo mravcov prejde po danej ceste, tým má vyššiu feromónovú hodnotu – je atraktívnejšia. Časom ale tento feromón vypchá, čo znamená, že príslušné menej navštevované cesty sa stávajú menej atraktívnymi, v dôsledku čoho je možné preskúmať aj iné alternatívne cesty. V prípade, že by feromón

nevypíchal, pohyb mravcov by bol uskutočňovaný len po jednej ceste, ktorá by nebola optimálna.

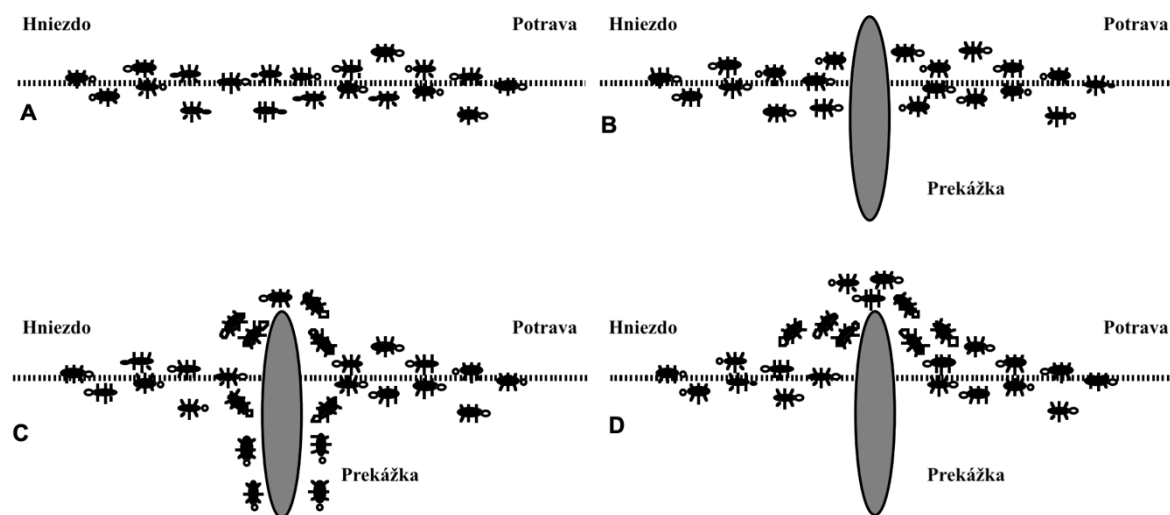
Cesta je iteratívne vytváraná algoritmom pre každého mravca. Mravce uskutočňujú svoj pohyb po vrcholoch a všetky navštívené vrcholy je potrebné si zapamätať. Prechádzaním vrcholov disponujú informáciami o jednotlivých vzdialenostiach medzi susednými vrcholmi a o množstve feromónu nachádzajúceho sa na hrane medzi určitými vrcholmi. Na základe týchto informácií si mravec zvolí hranu z vychádzajúceho vrcholu do susedného nenavštíveného vrcholu na základe istej pravdepodobnosti, ktorá je vypočítaná pomocou množstva feromónu na príslušnej hrane, atraktívnosti a vplyvu tejto atraktívnosti a množstva feromónu na výber hrany. Algoritmus pokračuje v iteráciách, kým nie je splnená ukončovacia podmienka. [21]

4 Výsledky práce

4.1 Kolónia mravcov

Skutočné mravce sú schopné nájsť najkratšiu cestu k zdroju jedla a späť bez akýchkoľvek vizuálnych stôp. Pri výskyte novej prekážky, vďaka ktorej nebude stará cesta realizovateľná, majú schopnosť prispôbiť sa zmenám v prostredí a tak nájsť novú najkratšiu cestu.

Obrázok 2: Hľadanie najkratšej cesty kolóniou mravcov



Zdroj: <https://people.idsia.ch/~luca/acs-bio97.pdf>

Na obrázku 2, ktorý znázorňuje hľadanie najkratšej cesty kolóniou mravcov, môžeme v kroku A vidieť mravce pohybujúce sa po priamke spájajúcej ich hniezdo a zdroj potravy. Ako bolo spomenuté v predchádzajúcej kapitole, primárnym prostriedkom mravcov na tvorbu a udržiavanie najkratšej cesty je feromónová stopa. Pri chôdzi mravcov je ukladané určité množstvo feromónu a každý mravec si s určitou pravdepodobnosťou vyberie cestu bohatú na túto látku.

Týmto základným správaním skutočných mravcov je možné vysvetliť ich schopnosť opätovne nájsť najkratšiu cestu aj v prípade, že sa náhle vyskytne neočakávaná prekážka. V kroku B je znázornený práve tento výskyt prekážky. Mravce nachádzajúce sa priamo pred prekážkou nemôžu pokračovať v sledovaní stopy a tak si musia vybrať medzi cestou doľava

a doprava, pričom predpokladáme, že sa polovica rozhodne pre cestu doprava a druhá polovica pre cestu doľava.

Podobná situácia nastáva na druhej strane prekážky, čo vyobrazuje krok C. Mravce, ktoré si zvolili kratšiu cestu okolo, obnovujú prerušenú feromónovú stopu rýchlejšie ako mravce, ktoré si zvolili dlhšiu cestu. Kratšia cesta prijme väčšie množstvo feromónu za istú časovú jednotku, čo má za následok väčší počet mravcov, ktoré si túto cestu zvolia.

Vďaka takejto pozitívnej reakcii, ktorú môžeme nazvať aj autokatalytický proces, si v konečnom dôsledku každý mravec vyberie kratšiu cestu, čo je zobrazené v obrázku 2 ako krok D. Nájdenie najkratšej cesty okolo prekážky je nevyhnutnou vlastnosťou interakcie medzi tvarom prekážky a správaním mravcov. Aj keď majú všetky mravce približne rovnakú rýchlosť a taktiež ukladajú feromónovú stopu rovnakou rýchlosťou, obrysovanie prekážok na dlhšej strane trvá dlhšie ako na kratšej strane, čo spôsobí rýchlejšie nahromadenie feromónovej stopy. Mravce uprednostňujú vyššiu úroveň feromónovej stopy, vďaka čomu je hromadenie na kratšej strane rýchlejšie. [22]

4.1.1 Kolónia mravcov a problém obchodného cestujúceho

V našej práci popíšeme podrobnejšie metódu kolónie mravcov ako jednu z možností riešenia problému obchodného cestujúceho. Problém obchodného cestujúceho, alebo aj hľadanie najkratšej okružnej cesty, je jednou z najznámejších optimalizačných úloh. Jeho podstata spočíva vo výbere určitého miesta, návštevy každého zaumieneného miesta práve raz a nakoniec vo vrátení sa do počiatočného miesta. Takúto cestu sa snažíme absolvovať tak, aby vzdialenosť, ktorú precestujeme, bola najmenšia. Každá cesta medzi dvomi navzájom prepojenými miestami musí byť ohodnotená číslom, ktoré predstavuje napríklad vzdialenosť, čas, cenu a podobne. S problémom obchodného cestujúceho sa môžeme stretnúť v množstve praktických aplikácií, napríklad pri plánovaní dopravy, výroby ale aj riadení dodávateľského reťazca.

Pri riešení tohto problému pomocou metódy kolónie mravcov je obchodným cestujúcim umelý mravec. Ten sa pohybuje z miesta na miesto na grafe TSP a to tak, že si vyberie miesto, ktoré prejde pomocou pravdepodobnostnej funkcie stopy nahromadenej na hranách a heuristickej hodnoty zvolenej ako funkcia dĺžky hrán. Umelé mravce

uprednostňujú mestá, ktoré sú spojené hranami s množstvom feromónových stôp a ktoré sú blízko. Na začiatku sa do náhodne vybraných miest umiestni m umelých mravcov, ktoré sa každým krokom presunú do nového miesta a upraví feromónovú stopu na využitých hranách – to nazývame aktualizáciou lokálnej stopy. Po absolvovaní cesty všetkými mravcami, mravec, ktorý prešiel najkratšiu cestu, upraví hrany patriace k jeho ceste – to nazývame aktualizáciou globálnej stopy.

Tri myšlienky, ktoré boli prenesené z prirodzeného chovania mravcov do kolónie umelých mravcov, sú preferencia ciest s vysokou hladinou feromónov, vyššia rýchlosť rastu množstva feromónu na kratších cestách a komunikácia medzi mravcami sprostredkovaná stopami. Kolónia umelých mravcov má taktiež niekoľko schopností, ktoré nepochádzajú z prirodzeného chovania mravcov, ale sú vhodné pre problém obchodného cestujúceho. Medzi ne patrí schopnosť určiť vzdialenosť medzi miestami a vybavenie funkčnou pamäťou využívanou na zapamätanie si navštívených miest. Pamäť je vyprázdnená na začiatku každej novej cesty a je aktualizovaná pridaním novo navštíveného mesta po každom kroku.

Všetky uvedené princípy je možné previesť do výpočtového systému vhodného na riešenie problému obchodného cestujúceho viacerými spôsobmi. V systéme kolónií mravcov si umelý mravec k v meste r vyberie mestá s na to, aby sa presunul medzi tie, ktoré nepatria do jeho pracovnej pamäte M_k a to použitím nasledujúceho pravdepodobnostného vzorca:

$$s = \begin{cases} \arg \max_{u \notin M_k} \{[\tau(r, u)] * [\eta(r, u)]^\beta\} & \text{ak } q \leq q_0 \\ S & \text{inak} \end{cases} \quad (1.1)$$

kde $\tau(r, u)$ je množstvo feromónovej stopy na hrane (r, u) , $\eta(r, u)$ je heuristická funkcia zvolená ako inverzná vzdialenosť medzi mestami r a u , β je parameter, ktorý váži relatívny význam feromónovej stopy a blízkosti, q je náhodne zvolená hodnota s jednotnou pravdepodobnosťou $[0,1]$, q_0 je parameter s hodnotami z intervalu $\langle 0;1 \rangle$, S je náhodná premenná vybraná podľa nasledujúceho rozdelenia pravdepodobnosti, ktorá uprednostňuje kratšie hrany s vyššou úrovňou feromónovej stopy:

$$p_k(r, s) = \begin{cases} \frac{[\tau(r, u)] * [\eta(r, u)]^\beta}{\sum_{u \notin M_k} [\tau(r, u)] * [\eta(r, u)]^\beta} & \text{ak } s \notin M_k \\ 0 & \text{inak} \end{cases} \quad (1.2)$$

kde $p_k(r, s)$ je pravdepodobnosť, s akou sa mravec k rozhodne presunúť z mesta r do mesta s .

Feromónová stopa je menená lokálne aj globálne. Globálna aktualizácia slúži na odmenu hran, ktoré patria ku kratším cestám. Akonáhle všetky umelé mravce absolvujú cestu, najlepší mravec uloží feromón na ním navštívené hrany, pričom ostatné hrany zostávajú nezmenené. Množstvo feromónu $\Delta\varphi(r, s)$ uloženého na každej hrane (r, s) navštívenej najlepším mravcom, je nepriamo úmerné dĺžke cesty, čo znamená, že čím kratšia je cesta, tým je väčšie množstvo feromónu uloženého na hranách. Spôsob, akým je feromón ukladaný, je určený na emuláciu vlastnosti diferenciálnej akumulácie feromónových stôp, ktorá v prípade skutočných mravcov bola spôsobená súhrou medzi dĺžkou cesty a kontinuitou času. Vzorec pre globálnu aktualizáciu stôp má tvar

$$\varphi(r, s) \leftarrow (1 - \alpha) * \varphi(r, s) + \alpha * \Delta\varphi(r, s) \quad (1.3)$$

kde $\Delta\varphi(r, s) = (\textit{najkratšia cesta})^{-1}$.

Cieľom lokálnej aktualizácie je zabrániť situácii, v ktorej by si všetky mravce vybrali veľmi silnú hranu. Zakaždým, ako si mravec vyberie hranu, jeho množstvo feromónu sa zmení pomocou vzorca na aktualizáciu lokálnej stopy, ktorý je v tvare

$$\tau(r, s) \leftarrow (1 - \alpha) * \tau(r, s) + \alpha * \tau_0 \quad (1.4)$$

kde τ_0 je parameter. Aktualizácia lokálnej stopy je taktiež inšpirovaná vyparovaním stopy skutočných mravcov.

Kolóniu mravcov je možné interpretovať ako posilňovací systém učenia, v ktorom posilnenia upravujú silu (feromónovú stopu) prepojení medzi mestami. Vzorce 1.1 a 1.2 určujú, že mravec môže s pravdepodobnosťou q_0 využiť skúsenosti nahromadené kolóniou mravcov vo forme feromónovej stopy, ktorá má tendenciu rásť na krátkych hranách patriacich krátkym cestám, vďaka čomu sú žiadanejšie. Taktiež môže s pravdepodobnosťou $1-q_0$ uplatniť zaujatý prieskum voči krátkym a vysokým hranám nových chodníkov a to náhodným výberom miesta, do ktorého sa chce presunúť. [22]

Záver

V tejto bakalárskej práci sme využili odbornú zahraničnú a domácu literatúru, aby sme čitateľom priblížili problematiku veľkých dát a umelej inteligencie. V prvej kapitole sme obsiahli ich definície a základné rozdelenia. Taktiež bolo potrebné uviesť popis princípu umelej inteligencie. Aby sme problematiku správne pochopili, bolo potrebné sa pozrieť na obe problematiky z historického hľadiska a na ich vývoj v čase. Na záver sme uviedli niekoľko príkladov využitia oboch oblastí v súčasnosti.

V kapitole popisujúcej metodiku práce a metódy skúmania sme sa oboznámili so základnými metaheuristikami. Metaheuristiky tvoria pomyselnú hranicu medzi umelou inteligenciou a operačným výskumom. Uviedli sme si ich základné informácie a ich základné delenie na metaheuristiky využívajúce prehľadávanie založené na jednom riešení a metaheuristiky využívajúce prehľadávanie založené na populácii riešení. Z oboch týchto kategórií sme vybrali dve metaheuristiky, ktoré boli bližšie popísané. Pre jednotlivé metaheuristiky sme popísali základné princípy a ich vlastnosti.

Poznatky týkajúce sa metódy kolónie mravcov nadobudnuté v predchádzajúcej kapitole sme rozšírili v poslednej kapitole. Hlbšie sme si uviedli a priblížili jej podstatu a princíp fungovania. Taktiež sme popísali, ako sa správanie skutočných mravcov prenáša do oblasti umelej inteligencie. Metóda kolónie mravcov môže byť jednou z metód na riešenie problému obchodného cestujúceho.

Záverečná práca neobsahuje praktickú aplikáciu, ktorú nebolo možné uskutočniť z dôvodu rozsahu a náročnosti problematiky. Praktická aplikácia tak môže byť námetom na diplomovú prácu.

Zoznam použitej literatúry

- [1] REYNOLDS, Vince. 2016. *Big Data for Beginners : Understanding SMART Big Data, Data Mining & Data Analytics For Improved Business Performance, Life Decisions & More!*. Scotts Valley : CreateSpace, 2016. 150 s. ISBN 9781530412044
- [2] SIMON, Phil. 2013. *Too Big to Ignore : The Business Case for Big Data*. New Jersey : Wiley, 2013. 231 s. ISBN 9781118642108
- [3] YAO, Marina. - JIA, Marlene. - ZHOU, Adelyn. 2018. *Applied Artificial Intelligence : A Handbook For Business Leaders*. Middletown : TOPBOTS, 2018. 181 s. ISBN 9780998289052
- [4] RUSSELL, Stuart Jonathan. – NORVIG, Peter. 2010. *Artificial Intelligence: A Modern Approach*. Harlow : Pearson Education Limited. 2010. 1132 s. ISBN 9781292153964
- [5] PEKÁR, Juraj. – BREZINA, Ivan. 2019. *Úvod do operačného výskumu II.* Bratislava : Letra Edu. 2019. 134 s. ISBN 9788089962297
- [6] FORBES : *A Very Short History Of Big Data*. [online]. [cit. 2021-25-2] Dostupné na: <<https://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/?sh=14a030f265a1>>
- [7] BIG DATA FRAMEWORK : *The Four V's of Big Data*. [online]. [cit. 2021-2-3] Dostupné na: <<https://www.bigdataframework.org/four-vs-of-big-data/>>
- [8] SAS : *Big Data:What it is and why it matters*. [online]. [cit. 2021-2-3] Dostupné na: <https://www.sas.com/en_us/insights/big-data/what-is-big-data.html>
- [9] DENNIK N : *Čo je to big data?*. [online]. [cit. 2021-6-3] Dostupné na: <<https://dennikn.sk/blog/278593/co-je-to-big-data/>>
- [10] BBVA : *The five V's of big data*. [online]. [cit. 2021-2-3] Dostupné na: <<https://www.bbva.com/en/five-vs-big-data/>>
- [11] PROGRAMMERSOUGHT : *The question is, characteristics of big data to see how many V?*. [online]. [cit. 2021-8-3] Dostupné na: <<https://www.programmersought.com/article/80633123404/>>
- [12] BIG DATA FRAMEWORK : *A short history of Big Data*. [online]. [cit. 2021-16-3] Dostupné na: <<https://www.bigdataframework.org/short-history-of-big-data/>>

- [13] BISMART : *Big Data is Everywhere:5 Ways It's Used in Your Everyday Life*. [online]. [cit. 2021-24-3] Dostupné na: <<https://blog.bismart.com/en/big-data-is-everywhere>>
- [14] SAS : *Artificial Intelligence:What it is and why it matters*. [online]. [cit. 2021-29-3] Dostupné na: <https://www.sas.com/en_us/insights/analytics/what-is-artificial-intelligence.html>
- [15] CODEBOTS : *What are the 3 types of AI? A guide to narrow, general, and super artificial intelligence*. [online]. [cit. 2021-5-4] Dostupné na: <<https://codebots.com/artificial-intelligence/the-3-types-of-ai-is-the-third-even-possible>>
- [16] MYGREATLEARNING : *What is Artificial Intelligence? How does AI work, Types and Future of it?*. [online]. [cit. 2021-11-4] Dostupné na: <<https://www.mygreatlearning.com/blog/what-is-artificial-intelligence/>>
- [17] INNOPLEXUS : *How does Artificial Intelligence work?*. [online]. [cit. 2021-13-4] Dostupné na: <<https://www.innoplexus.com/blog/how-artificial-intelligence-works/>>
- [18] FORBES : *A Very Short History of Artificial Intelligence (AI)*. [online]. [cit. 2021-16-4] Dostupné na: <<https://www.forbes.com/sites/gilpress/2016/12/30/a-very-short-history-of-artificial-intelligence-ai/?sh=2b3ab9bd6fba>>
- [19] G2 : *A Complete History of Artificial Intelligence*. [online]. [cit. 2021-15-4] Dostupné na: <<https://learn.g2.com/history-of-artificial-intelligence>>
- [20] EMERJ : *Everyday Examples of Artificial Intelligence and Machine Learning*. [online]. [cit. 2021-18-4] Dostupné na: <<https://emerj.com/ai-sector-overviews/everyday-examples-of-ai/>>
- [21] MUNI : *Metaheuristické metódy na riešenie vybraných dopravných problémov*. [online]. [cit. 2021-14-6] Dostupné na: <https://is.muni.cz/el/econ/podzim2016/MPM_OMVE/um/35333334/Vojtekova-Blazekova.pdf>
- [22] IDSIA : *Ant colonies for the traveling salesman problem*. [online]. [cit. 2021-11-7] Dostupné na: <<https://people.idsia.ch/~luca/acs-bio97.pdf>>