

UNIVERZITET U SARAJEVU

EKONOMSKI FAKULTET

ZAVRŠNI RAD

**RUDARENJE I ANALIZA VELIKIH PODATAKA U CILJU
POBOLJŠANJA PERFORMANSI U POSLOVANJU**

Sarajevo, novembar 2023.godine

MELISA JUSUFI

U skladu sa članom 54. Pravila studiranja za I, II ciklus studija, integrisani, stručni i specijalistički studij na Univerzitetu u Sarajevu, daje se

IZJAVA O AUTENTIČNOSTI RADA

Ja, **Melisa Jusufi**, studentica drugog (II) ciklusa studija na Odsjeku **Menadžment**, Smjer **Menadžment i informacione tehnologije (MIT)**, pod naslovom:

Rudarenje i analiza velikih podataka u cilju poboljšanja performansi u poslovanju

izjavljujem da sam završni rad izradila samostalno i da se zasniva na rezultatima mog vlastitog istraživanja. Svjesna sam činjenice da svaki oblik plagijarizma podliježe sankcijama u skladu sa relevantnim pravilima Univerziteta u Sarajevu i Ekonomskog fakulteta.

Ovom izjavom potvrđujem i da sam za potrebe arhiviranja predao/la elektronsku verziju rada koja je istovjetna štampanoj verziji završnog rada.

Dozvoljavam objavu ličnih podataka vezanih za završetak studija (ime, prezime, datum i mjesto rođenja, datum odbrane rada, naslov rada) na web stranici i u publikacijama Univerziteta u Sarajevu i Ekonomskog fakulteta.

U skladu sa članom 34. 45. i 46. Zakona o autorskom i srodnim pravima (Službeni glasnik BiH, 63/10) dozvoljavam da gore navedeni završni rad bude trajno pohranjen u Institucionalnom repozitoriju Univerziteta u Sarajevu i Ekonomskog fakulteta i da javno bude dostupan svima.

Sarajevo, 10.11.2023.g.

Studentica

Potpis studentice

SADRŽAJ

POPIS TABELA.....	III
POPIS ILUSTRACIJA	III
SAŽETAK	V
ABSTRACT	VI
KLJUČNE RIJEČI.....	VII
1. UVOD	1
1.1. Predmet i problem istraživanja	1
1.2. Ciljevi istraživanja.....	7
1.3. Istraživačka pitanja	7
1.4. Metodologija istraživanja.....	7
1.5. Struktura rada	7
2. PREGLED LITERATURE.....	9
2.1. Rudarenje podataka	9
2.1.1. Definicija i uloga.....	9
2.1.2. Metode i okviri rudarenja	18
2.1.3. Analiza podataka na osnovu koncepta rudarenja	23
2.2. Važnost podataka u poslovanju	26
2.2.1. Prednosti i nedostaci tehnika rudarenja poslovnih podataka.....	26
2.2.2. Segmentacija dobivenih podataka.....	27
2.2.3. Korisnost analize podataka za poslovanje kompanije.....	30
3. METODOLOGIJA ISTRAŽIVANJA	34
3.1. Identifikacija i selekcija izvorne literature	34
3.2. Bibliometrijska analiza podataka	38
3.3. Interpretacija rezultata i diskusija.....	45
3.3.1. Najznačajniji autori	45
3.3.2. Najcitiraniji radovi	47
3.3.3. Ključni pojmovi istraživanja.....	48
3.3.4. Geografska područja istraživačkih radova	50
3.3.5. Diskusija rezultata.....	51
3.4. Ograničenja i preporuke	52
3.4.1. Ograničenja pregleda literature.....	52
3.4.2. Prijedlozi za buduća istraživanja.....	56
4. ZAKLJUČAK	58

REFERENCE.....	59
-----------------------	-----------

POPIS TABELA

Tabela 1 Razlike rudarenja i analize podataka	25
Tabela 2 Najznačajniji autori odabrane literature	45

POPIS ILUSTRACIJA

Ilustracija 1 Karakteristike velikih podataka.....	11
Ilustracija 2 Data mining u pronalasku informacija.....	13
Ilustracija 3 Tehnologije Data mining-a.....	15
Ilustracija 4 Algoritmi mašinskog učenja.....	17
Ilustracija 5 Prikaz metode klasifikacije	19
Ilustracija 6 Prikaz metode klastera	20
Ilustracija 7 Metode pravila pridruživanja	21
Ilustracija 8 Data mining pri korištenju metode neuronskih mreža	22
Ilustracija 9 Stablo odlučivanja	23
Ilustracija 10 Rudarenje podataka i upravljanje životnim vijekom kupca	29
Ilustracija 11 Modeli klasifikacije	30
Ilustracija 12 Pretraga i selekcija literature	36
Ilustracija 13 Vrste analize literature	39
Ilustracija 14 Kocitirani radovi.....	41
Ilustracija 15 Dijagram toka istraživanja	43
Ilustracija 16 Definisane 5 klastera literature.....	44
Ilustracija 17 Najznačajniji autori odabrane literature.....	46
Ilustracija 18 Najcitiraniji radovi odabrane literature	47
Ilustracija 19 Najcitiraniji radovi odabrane literature	48
Ilustracija 20 Ključni pojmovi istraživanja odabrane literature	49
Ilustracija 21 Ključni pojmovi istraživanja - vremenski prikaz	50
Ilustracija 22 Pregled po državama odabrane literature.....	51

SAŽETAK

Budući da je razvoj tehnologija imao direktan utjecaj na način poslovanja kompanija i bitnost informacija na kojima se bazira, tema istraživačkog rada „Rudarenje i analiza velikih podataka u cilju poboljšanja performansi u poslovanju“ jeste zasigurno bitan aspekt na današnjem tržištu. S obzirom na činjenicu da informacione tehnologije u poslovanju imaju rastući trend bitnosti, također je i širok spektar njihove primjenjivosti, ali ista nije dovoljno istražena. Odnosno, bitnost informacija u poslovanju kao krucijalni element svih kompanija se tako reći podrazumijeva, zbog čega se ujedno i olakšo shvata.

U današnjici, podaci i informacije imaju signifikantan utjecaj u svim sferama poslovanja. To direktno utječe na sve povezane procese koji nastaju manipulacijom istih, pri čemu i oni dobijaju na značaju. Rudarenje podataka, kao proces koji podrazumijeva otkrivanje znanja na osnovu analize velike količine podataka iz različitih izvora, pretvarajući ih u korisne informacije je zasigurno temelj poslovanja svih kompanija.

Shodno tome, izdvajanje informacija i kreiranih znanja u različite baze i spremišta predstavlja suštinu napretka statističkih obrada, mašinskog učenja i vještačke inteligencije, gdje se analizama i različitim procesima prepoznaju uzorci koji omogućavaju različite vrste predikcija i poslovnih procesa koji olakšavaju poslovanje.

Kako različita polja poslovanja primjenjuju rudarenje podataka, postoje mnoga istraživanja o važnosti samog rudarenja podacima, ali i samih podataka, te njihove uloge u kompanijama. Iako globalno postoje dosta slični stavovi o informacijama, ipak je važno primjeniti sveobuhvatnu analizu, te uporediti zaključke koji su ponuđeni.

Daljne istraživanje u suštini predstavlja sistemsku analizu literature, odnosno prikupljanje istraživačkih radova čija se tema odnosi na rudarenje podataka i utjecaj na poslovanje, pri čemu su detaljno obrađeni radovi sa ciljem pronalaska cjelokupnog zaključka o datoj temi.

Prvenstveno, rad se fokusira na objašnjenju same teme, pri čemu se pobliže objašnjavaju pojmovi i procesi koji su vezani za istu, a zatim je izvršeno bibliografsko istraživanje. Istraživački radovi su generisani kroz bazu Web of Science na osnovu koje je sprovedena analiza.

Sama sistemaska analiza literature je koncept primjenjen u radu koji se zasniva na stavovima istraživača prema potrebama rudarenja podataka i primjeni u poslovanju kompanija, pri čemu se identifikuje, selektira i kritički evaluiraju istraživački radovi iz Web of Science baze. Birani radovi su zasnovani na kvalitativnom ili kvantitativnom izvornom istraživanju u posljednjih 20 godina, izdati na engleskom jeziku, bez dodatnih kriterija osim da je osnovna tematika svih radova i fokus istraživanja rudarenje podataka i primjena u poslovanju.

Podaci iz radova su detaljno obrađeni i izdvojeni, tražeći zajedničke komponente kako bi se izveo adekvatan zaključak o istim. Ključni pojmovi korišteni za pretragu radova su u ovom radu navedeni u nastavku. Budući da pregledi literature generalno imaju određene nedostatke, tako su isti uočeni i diskutovani, pri čemu su razmotrili način za buduća unaprijeđenja.

Generalno, važnost obrade sistemske literature ogleda se u tome da mnoga istraživanja su bazirana na izdvojenim stavovima ili ograničenim pristupima istraživača, zbog čega je bitno uporediti ista, kako bi se stvorila utemeljenija slika. Shodno tome, u radu se diskutuju različiti vidovi poboljšanja poslovnih procesa, na osnovu rudarenja podatka, pri čemu se obrađuje samo evoluiranje *data mining-a* kroz vrijeme, kao i u budućnosti.

ABSTRACT

Since the development of innovation had a direct effect on the company's way of doing business and the significance of the data on which it is based, the subject of the term paper "Mining and investigation of enormous information in arrange to make strides execution in business" is certainly an imperative perspective in today's showcase. Considering the reality that data innovations in business have a developing slant of significance, there's too a wide range of their relevance, but it has not been adequately investigated. That's, the significance of data in business as a vital component of all companies is hence caught on, which is why it is additionally effectively caught on.

Nowadays, information and data have a noteworthy effect in all circles of business. This straightforwardly influences all related forms that are made by controlling them, and they moreover pick up significance. Information mining, as a preparation that suggests the disclosure of information based on the investigation of huge sums of information from different sources, turning them into useful information, is certainly the premise of the business of all companies.

Thus, extricating data and made information totally different databases and stores speaks to the quintessence of the advance of factual preparing, machine learning and fake insights, where designs are recognized through examination and distinctive forms that empower diverse sorts of expectations and business forms that encourage business.

As diverse areas of commerce apply information mining, there are numerous thoughts about the significance of information mining itself, but also of information itself, and its part in companies. In spite of the fact that all inclusive there are numerous comparative demeanors around data, it is still important to apply a comprehensive investigation and compare the conclusions that have been advertised.

Encourage inquire about basically speaks to an orderly investigation of the writing, that's , the collection of inquire about papers whose theme is related to information mining and the affect on commerce, where the papers are handled in detail with the point of finding an by and large conclusion on the given topic. Fundamentally, the paper centers on the clarification of the same point, where the concepts and forms related to the same are clarified in more detail, and after that a bibliographic research was carried out. The research papers were created through the Internet of Science database, on the premise of which the investigation was carried out.

The efficient investigation of writing itself may be a concept connected within the work based on the researcher's states of mind towards wants of information mining and application within the company's commerce, whereby inquire about works from the *Web of Science* database are distinguished, chosen and basically assessed.

Chosen papers are based on subjective or quantitative unique inquire about within the final 20 a long time, distributed in English, without extra criteria but that the essential point of all papers and the center of inquire about is information mining and application in commerce.

The information from the papers were prepared and isolated in detail, searching for common components in order to draw a satisfactory conclusion almost them. The key terms used to look for papers in this paper are recorded below. Since writing surveys for the most part have certain inadequacies, these are famous and examined, with ways for future advancement talked about.

In general, the importance of handling systemic writing is reflected within the reality that numerous considerations are based on disconnected views or constrained approaches of analysts, which is why it is critical to compare them, in order to form a more grounded picture. In like manner, the paper talks about various types of commerce preparing change, based on information mining, where as it were the evolution of information mining over time, as well as within the future, is tended to.

KLJUČNE RIJEČI

Rudarenje podataka, analiza podataka, veliki podaci, predikcija, poslovna vrijednost, poslovni razvoj, poslovna inteligencija, mašinsko učenje, poslovne performanse, uvid u poslovanje.

Data mining, data analysis, big data, prediction, business value, business development, business intelligence, machine learning, business performance, business insight.

1. UVOD

1.1. Predmet i problem istraživanja

Značaj velikih količina podataka iz dana u dan raste. Kao što autori Hao *et al.* analiziraju, vizuelno rudarenje podataka predstavlja značajan alat baziran na algoritamskim tehnikama. Navode da je vizualizacija podataka nastalih rudarenjem sve popularnija. Međutim, postoji nekoliko izazova koji osporavaju korisnost vizualizacije, pri čemu se ističe veliki obim podataka i navigacija kroz iste i visoka dimenzionalnost podataka. Cilj autora jeste da otkriju efektivnost infrastrukture, te na koji način on minimizira navedene izazove. Nakon detaljne analize autora, zaključak je da uz adekvatnu upotrebu softvera ili infrastruktura, moguće je lakše kretanje kroz različite sinkronizirane poglede bez preplavljenosti velikim brojem nepovezanih prezentacija (Hao *et al.*, 2000).

Današnjica poslovnog svijeta prema autorima Khder *et al.* ističu da bitni i ključni pokazatelji za uspjeh kompanija jeste da se u pravo vrijeme izvrši prikupljanje i analiziranje informacija i podataka. Poslovna preduzeća trebaju adekvatno/ciljano razumjeti kupce, samim tim pratiti i konkurenciju i njegovo okruženje, te donositi detaljne odluke, kako bi uspjeli u istom. Također, smatraju da je od velikog značaja za uspjeh na tržištu implementacija tehnologija kao što su rudarenje podataka i poslovna inteligencija.

Odnos poslovnih subjekata i kupaca u današnjici se mnogo razlikuje u odnosu na prošlost, te samim tim se dovodi do toga da preduzeća trebaju stvoriti nove i bolje strategije i novim promjenama se prilagoditi kupcima kako bi došlo do boljeg profita.

Kako bi se u poslovnom svijetu ispunile potrebe i želje kupaca nužna je primjena tehnikke rudarenje podataka, skladištenje podataka i upravljanje odnosima sa kupcima u kojima nova tržišta diktiraju. Nakon istraživanja analize, kao zaključak autori smatraju da uz rudarenje podataka se može poboljšati i poslovna inteligencija i profit poslovanja, te da se nakon donošenja adekvatnih odluka može smanjiti rizik (Khder *et al.*, 2021).

Metodologije, procesi i tehnologije koje pretvaraju sirove podatke u informacije koje mogu pomoći u izbjegavanju prijetnji se uključuju u donošenje pravilnih odluka, te samim time dolazi do pronalaska novih poslovnih ideja i prilika, koji kao rezultat prikazuje poboljšani proces donošenja odluka tvrdi autor Llave M., 2018. U nastavku spominje da se može dovesti i do donošenja nepotpunih i nepravilnih odluka koje neće bit optimalne ako se budu koristili prikupljeni podaci koji su samo nestrukturirani (Llave, 2018).

Sami podaci mogu biti strukturirani i nestrukturirani. Jedna od metoda analize u kojima se mogu ovi podaci analizirati je data mining tj. rudarenje podataka, te smatraju da ova metoda se ne odnosi samo na analizu izvještaja nego pruža mogućnost analize različitih vrsta sadržaja, kao što su videozapisi, slike, e-pošta i puno toga, te autori smatraju da ova analiza ima i mogućnost da predvidi ishode u kojima se pronalaze uzorci i korelacija u bazi podataka. Pored toga, rudarenje podataka je zahtjevnija metoda od poslovne inteligencije, jer podaci kada se prikupljaju su većinom nestrukturirani i mnogo ih je teže analizirati. Autori ističu da će rudarenje podataka pomoći u analizi sirovih podataka i pružiti poslovnom inteligencijom informacije za donošenje neophodnih zaključaka (Shmueli *et al.*, 2016).

Upravljanjem znanjem koja se dovodi u svrhe poslovne inteligencije omogućava vezu rudarenja podataka, te prikazuje identičan proces otkrića podataka koji su nestrukturirani (Grbavac, 2018). Autori smatraju da je sam koncept rudarenja podataka jako važan i bitan alat koji omogućava kratko i jasno rješenje. Iako se smatra kao potencijalni alat, u kontekstu poslovne inteligencije koristi rudarenja podataka nisu dovoljno istražene. Samim tim, oni vjeruju da je efektivno rudarenje podataka od bitnijih aspekata za samu poslovnu inteligenciju jer predstavlja fazu razmjene znanja i bolja planiranja koja se dovode do povezanosti u kompanijama (H. Wang i Wang, 2008).

Preduzećima se stvara mogućnost da prikupljaju i čuvaju ogromne količine podataka, te kao razlog navode sam napredak informacijskih tehnologija i rast interneta u današnjici. Samim tim, autori Seng i Chen, 2010 ističu da su ljudi počeli shvatati da podaci nisu isto što i informacija, te da se podaci trebaju detaljnije izdvajati kao i dalje analizirati. U ovom radu autori objašnjavaju da su stručnjaci naučeni da tumače podatke i da sprovode analizu, ali dolazi istovremeno i do povećanja količine, vrste podataka i analitičke dimenzije. Također, u današnjem vremenu informacijska tehnologija je postala još širi pojam od same pohrane, prijenosa i obrade podataka. Podaci se trebaju pretvoriti u informacije i znanja kako bi se podržalo donošenje odluka (Seng i Chen, 2010).

Količina podataka idalje raste rapidno, iako su skladišta podataka već ogromna. Jedan od glavnih izazova jeste kako pretvoriti bazu podataka u konkurentnu inteligenciju tako što će se naizgled beskorisni podaci pretvoriti u korisne informacije. Za autore Han i Kamber ključno je na koji način će se riješiti ovaj izazov jer preduzeća sve više ovise o efektivnoj analizi informacija kako bi ostala konkurentna. Pomoću otkrivanja znanja u bazama podataka, zanimljiva znanja i pravilnosti informacije mogu se izvući iz relevantnih skupova podataka u bazama podataka i istražiti iz različitih uglova (Seng i Chen, 2010).

Iz perspektive skladišta podataka, rudarenje podataka se može smatrati naprednom fazom on-line analitičke obrade (OLAP). Svakako, rudarenje ide daleko iznad ograničenog opsega pripreme informacija i skladištenja informacija, konsolidacijom naprednih strategija, te razumijevanja informacija (Han *et al.*, 2012).

U današnje vrijeme, savremeni obrasci, kao što su dostupnost ogromnih količina informacija, zahtjev za brzim izvođenjem analitike i uređenje ekspozitorija dolazi do klijenata na razuman način, detaljno su razradili autori Apte *et al.*, 2002. širenju primjene procedura otkrivanja informacija u bazama podataka (KDD). Dopunjavanje ili zamjena postojećih analitičkih tehnika ljudskih stručnjaka se temelji na tehnikama skalabilnosti, pouzdanosti i strukturama koje su automatizirane. Štaviše, tradicionalni pristup ispitivanju informacija za povratni izbor kombinuje informacije o prostoru sa činjeničnim modeliranjem kako bi se stvorio aranžman za određena pitanja (Apte *et al.*, 2002).

Za analizu ogromnih skupova informacija, rudarenje informacija koristi i tradicionalne mjerljive aparate i vještačke proračune uvida, koji su se pokazali izuzetno imperativnim u brojnim krugovima, računajući samo poslovanje.

Autor Olsona L.D., 2006 ističe u svom radu da primjene rudarenje podataka koje su relevantne za uslužnu industriju prikazane osnovne poslovne funkcije i metode rudarenje podataka. Navodi se da je rudarenje podataka također u današnjici postao bitan alat i u medicini, inženjerstvu i znanosti, jer svim ovim oblastima je potreban za obrađivanje masovnih podataka koje svakako generira savremeno društvo.

U poslovanju se *data mining* koristi za otkrivanje prevara, analizu kredita i segmentaciju kupaca, što utječe na uslužnu industriju.

Rudarenje podataka daje kapacitet da se brzo prikupi neiskorišćeno razumijevanje kroz ispitivanje ekspanzivnih suma informacija (Olson, 2007).

Poštovanje mjerljivog izvršenja i širenje razvoja oglašavanja su zone u kojima se redovno nalazi napredak u poslovanju. Rudarenje podataka, kao što navode Gazzawe i Alturki, 2022 je proces kojim se sirovi podaci pretvaraju u značajne informacije koje firme koriste za donošenje dugoročnih odluka i postizanje konkurentnosti na tržištu. Primjena rudarenja podataka uključuje prognoziranje, grupiranje, pronalaženje sekvenci, rizik i vjerojatnost.

To omogućava kompanijama da efikasno razumiju potrebe klijenata i donose odluke koje utječu na vrijednosti grupe. Jedna od utjecajnijih perspektiva koja pruža mogućnost preduzećima da svoj izlaz na tržištu efikasno povećaju jeste rudarenje podataka. To znanje pruža preduzećima veliku kontrolu, kako bi mogli da upravljaju sa tim da njihovi klijenti mogu doživljavati njihove poslovne funkcije, što dovodi do cjelokupnog preuzimanja svih kontrola koje utiču na same klijente.

Postoji i ogromna razlika koja omogućava svakoj kompaniji koja primjenjuje ovu perspektivu da se potpuno prilagodi različitim šemama. Stoga, kada se tržište promijeni da bi odražavalo druge perspektive koje utiču na društvo, postaje lahko ostati potpuno konkurentan bez suočavanja s izazovima propadanja ili pada u poslovnim aktivnostima (Gazzawe i Alturki, 2022).

Budući da je upotreba rudarenja podataka u poslovnim procesima postala češća, nije dosegla razinu koja bi proizvela potencijalne koristi. Također, može se reći da se rudarenje podataka većinom koristi za podršku donošenju odluka na taktičkoj razini u procesima odlučivanja. Autori Chung i Gray, tvrde da je provedeno mnogo istraživanja u području stvaranja modela rudarenja podataka, ali da nije posvećena dovoljna pažnja upotrebi tih modela u poslovnim procesima (Chung i Gray, 1999).

Kohavi i Provost tvrde da je važno omogućiti upotrebu rudarenja podataka u poslovnim procesima putem automatiziranih rješenja (Kohavi i Provost, 2001). Samim tim, autori raspravljaju o važnosti jednostavnosti integracije rudarenja podataka u poslovne procese. U svom radu autori smatraju integraciju rudarenja podataka u poslovne procese posljedicom potrebe za uključivanjem pozadinskog znanja u poslovne procese (Pivk *et al.*, 2013).

Informacije koje se prikupljaju u svakodnevnim aktivnostima je poslovna inteligencija, te preduzeća posjeduju potencijal korištenja istih (Vidal-García *et al.*, 2019). U optimizaciji organizacijske učinkovitosti u pronalasku novih prilika i izazova poslovna inteligencija ima važnu ulogu.

Ta uloga se odnosi i na otkrivanje mogućih rizika i novih uvida u industriji, te jačanjem okvira za bolje donošenje odluka, navode autori (Zhang *et al.*, 2022).

Poslovna inteligencija koristi aktuelne i historijske podatke, ali ne i poslovna analitika. Ona koristi prethodne podatke za uvide i provođenje svih aktivnosti kako bi se ostvarili zahtjevi kupaca i povećala produktivnost. BI sistemi se trenutno oslanjaju na centralizirane i interne poslovne podatke.

Velika količina korisnog znanja u neorganiziranom i postojećem informacijskom okruženju mogla bi pridonijeti nepotpunom pregledu činjenica i djelomičnom procesu donošenja odluka, navode autori (Gao *et al.*, 2022).

Ukoliko podaci nisu analizirani i obrađeni, smatraju se beskorisnima. Potrebni su softveri poput Hadoop-a koji će se brinuti o njima. Programeri i inženjeri trebaju više mogućnosti zapošljavanja za razvoj tih rješenja. Računarna i internetska tehnologija stvorile su nove potrebe i izazove za BI kako bi se prikupili bitni podaci iz različitih izvora. Ti podaci mogu biti organizirani i neorganizirani, složeni i ključni (Niu *et al.*, 2021).

Za autore Zong *et al.*, 2021 ključno je da veliki podaci sada omogućuju BI-u da pruži uvide koji bi kompanijama omogućili bolje razumijevanje svojih klijenata, poboljšanje marketinške tehnologije, omogućavanje personalizacije i prepoznavanje problema i prilika u stvarnom vremenu (Zong *et al.*, 2021).

Veliki podaci su u posljednje vrijeme dobili mnogo zanimanja zbog njihove sposobnosti stvaranja tržišne vrijednosti (Niu *et al.*, 2021). Podaci iz društvenih mreža o analizi velikih podataka za poslovnu inteligenciju su jedinstveni.

Kao što su autori Shamsi i Khojaye, 2018 istakli da analiza podataka omogućava menadžerima donošenje odluka utemeljenih na statističkim činjenicama (Shamsi i Khojaye, 2018).

Izazov pravovremene analize velikih podataka čini izvršavanje poslovnih procesa velikih i kompleksnih lanaca opskrbe koja rezultira generisanjem velikih nestrukturiranih podataka o događajima. Također, autori navode da arhitektura za integraciju koja se odnosi na analizu podataka u upravljanje poslovnim performansama direktno pomaže korisnicima da poboljšaju i analiziraju iste. Preduzeća koriste i rudarenje podataka i poslovnu inteligenciju kako bi mogli primjenjivati različite i statističke i umjetne inteligencije tehnike za analizu velikih podataka koje su povezane s procesima. Tri vrste analize poslovnih procesa (BPA) čine validacija, verifikacija i performanse, te sve zahtijevaju prikupljanje i pohranjivanje velikih količina podataka o procesima i događajima (Vera-Baquero *et al.*, 2013).

U svom istraživanju Ardito i Scuotto, analiziraju da nove tehnologije poput Web 2.0 i Web 3.0, Interneta stvari (eng. *Internet of things*) i Industrije 4.0 omogućile su ogromnu količinu podataka o životnom ciklusu proizvoda, logistici i potrebama tržišta. To je dovelo do tzv. revolucije *Big Data*, gdje su podaci karakterizirani velikim volumenom, raznolikošću i brzinom. Iako postoje tehnike i algoritmi za analizu *Big Data*, istraživanja poslovanja i upravljanja tek nedavno su počela proučavati prednosti analitike *Big Data* (Ardito *et al.*, 2019).

Velika analitika podataka postala je sve važniji element za kompanije širom razvijenih ekonomija. U ovom istraživanju autori Ren *et al.*, ispituju dinamiku kvaliteta u okruženju velikih podataka koja su povezana s poboljšanjem poslovne vrijednosti i performansi kompanije (FPER). Istraživanje identificira da je kvaliteta sistema i kvaliteta informacija veoma bitna i ključna za poboljšanje i poslovne vrijednosti i performansi preduzeća u velikom krugu podataka.

Kao kvalitet sistema autori smatraju da se to odnosi na pouzdanosti, dostupnost podataka, prilagodljivost i slično, a kao kvalitet informacija sadrži kompletnost, tačnost i sl.

Također, samo istraživanje predstavlja prijedlog da su kvaliteta i performanse preduzeća posredovani vrijednošću ogromnih podataka. Polazeći od teorije zasnovane na resursima i literature o uspjehu informacionih sistema, ovaj rad proširuje znanje u ovoj oblasti povezujući kvalitet sistema, kvalitet informacija, poslovnu vrijednost i performansi kompanije (Ji-fan Ren *et al.*, 2017).

U posljednjim godinama svjedočimo trendu revolucije vođene podacima u poslovanju. Pojam velikih podataka (*Big Data*), privukao je široku pažnju i donio, ne samo tehnološke, već i menadžerske probleme (Huang *et al.*, 2020). Veliki podaci se smatraju izvorom konkurentske prednosti i neophodnom stavkom za svakodnevne poslovne aktivnosti.

Na primjer, autori Manyika *et al.*, 2011 tvrde da su veliki podaci "sljedeća granica za inovacije, konkurenciju i produktivnost" (McKinsey Global Institute, 2011).

Autori Kiron *et al.*, 2013 u svom radu spominju da operativna efikasnost i učinkovitost poslovanja može dovesti do bolja i efikasna donošenja odluka koje se samim tim temelje na podacima (Bean i Kiron, 2013). Duži niz kao glavni fokus se predstavlja područje informacijskog sistema za upravljanje, kao i vrijednost poslovanja informacijskih tehnologija (Sabherwal i Jeyaraj, 2015).

Također, jedan od primjera jeste da autori Hitt i Brynjolfsson, 1996 istražuju vrijednost IT-a iz perspektive produktivnosti i profitabilnosti poslovanja (Brynjolfsson i Hitt, 1998).

Istraživači tokom razvoja velikih količina podataka u posljednjih nekoliko godina razmatraju metode koje su prihvatljive i koje pružaju mogućnost uključivanja novih naprednijih tehnologija u konkurentskim strategijama. Big data predstavlja pomoć kompanijama u odluci donošenja boljih i ispravnijih odluka za bolje poslovanje u budućnosti, te samim tim privlače veliku pažnju jer kompanije se bave ulaganjem u analitiku ogromnih podataka radi boljih performansi i konkurentske prednosti na dnevnoj bazi sve više rastu.

Također, primjećuje se da se daje malo pažnje istraživanju za buduća poboljšanja koristeći i usvajajući novih tehnologija.

Veliki podaci kao visoki obim, brzina i visoka raznovrsnost sirovih informacija, zahtijevaju ekonomičnu i inovativnu tehniku analize informacija za prikupljanje podataka za donošenje odluka.

Kada je riječ o analizi velikih podataka koji samim tim pojavljuju se podaci koji nisu obrađeni za dalju upotrebu i iz kojih nije iskorištena skrivena informacija. U današnjici dominantnijom metodom se smatra analitika velikih podataka koja je superiornija u prikupljanju masovnih količina sirovih informacija i primjene najboljih analitičkih praksi za mjerenje istih. Postao je alat kojim kompanije prikupljaju različite podatke i koriste automatsku analitiku podataka kako bi informirale odgovarajuće odluke koje su se prije oslanjale na prosudbu i percepciju donositelja odluka.

Dakle, analitika velikih podataka okreće se oko tri ključne značajke: samih podataka, analitike primijenjene na podatke i prezentacije rezultata na način koji omogućava stvaranje poslovne vrijednosti za kompanije i njihove kupce (Maroufkhani *et al.*, 2019).

Kako bi se dovelo do učestale analize velikih podataka i poboljšanja svojih proizvoda i usluga, te i podržale donošenje odluka, sve više preduzeća sa razvojem digitalizacije razmatra tu mogućnost korištenja velikih podataka. Preduzeća trebaju samim tim, da iskoriste potencijal koji posjeduju veliki podaci i poslovne analitike kako bi stekle bolju konkurentsku prednost na tržištu.

Postoje istraživanja kako analizu velikih podataka koristiti i iskoristiti na najbolji način navode autori Rialti *et al.*, 2019, jer smatraju da je analiza velikih podataka i dalje brzo razvijajuća tehnološka i poslovna praksa (Rialti *et al.*, 2019).

U svom radu autori također ističu da ranija istraživanja, iako su pokazala dosta prednosti o usvajanju analize velikih podataka u različitim aspektima, nedostaje dokaza o tome kako i na koji način primjeniti ovu soluciju kako bi se uspjela stvoriti konkurentska prednost. U području poslovanja i upravljanja postoje sistematska istraživanja usmjerena na analizu velikih podataka, autori Rialti *et al.*, navode da je "minimalna pažnja posvećena sistematiziranju literature o velikim podacima i dinamičkim sposobnostima".

Godine 2016. autori Suoniemi *et al.*, ističu da se 90% svjetskih podataka kreiralo samo u prethodne dvije godine, brzinom od između 2 do 3 kvintiliona bajta podataka na dan. Ta količina podataka enormno raste i samim tim se i očekuje da će do 2020. godine iznositi više od 16 zetabajta (16 triliona GB) korisnih i važnih podataka koje su preduzećima potrebne.

Danas u novi trend spadaju mobilni uređaji, internet stvari i društvene mreže, te u ovom istraživanju od strane autora Suoniemi *et al.*, veliki podaci se konkretno odnose ne tehnike i tehnologije, na sisteme, metodologije i primjene povezane sa analizom i prikupljanjem, te skladištenjem ogromnih količina i strukturiranih i nestrukturiranih podataka u podršci odlučivanja u poslovanju preduzeća. Svakako da je tržište velikih podataka značajno poraslo u zadnjih par godina (Suoniemi *et al.*, 2020).

Interes za velikim podacima i poslovnom analitikom je značajno rastao posljednjih deset godina, čak od prvih ključnih članaka koji su raspravljali o potencijalu velikih podataka da revolucioniziraju način na koji radimo, živimo i obavljamo poslovanje (Mikalef *et al.*, 2020).

Od tada, istraživači informacijskih sistema su pokušavali istražiti fenomen velikih podataka i poslovne analitike te razumjeti kako organizacije mogu stvarati i iskoristiti vrijednost iz svojih resursa podataka.

1.2. Ciljevi istraživanja

Osnovni ciljevi istraživanja obuhvataju:

1. Približiti osnovni koncept rudarenja i analize podataka u kompanijama
2. Istražiti i pojasniti značaj rudarenja podataka u poslovanju kao i napretku poslovnog performansa kompanije
3. Identifikovati prednosti i nedostatke koje rudarenje podataka nosi
4. Pojasniti modele koji se koriste za rudarenje podataka u poslovanju
5. Analizirati koncept „velikih podataka“ i njihove važnosti u kompaniji

1.3. Istraživačka pitanja

Sa navedenim predmetom istraživanja, navode se istraživačka pitanja:

1. Kako se rudarenje podataka primjenjuje u poslovnom okruženju i koje su njegove prednosti i nedostaci?
2. Da li i na koji način se rudarenje podataka može primijeniti za poboljšanje poslovnih performansi i povećanje efikasnosti kompanije?
3. Kako se korištenjem tehnika rudarenja podataka i analize mogu postići bolji rezultati i poboljšati validnost postojećih poslovnih procesa?

1.4. Metodologija istraživanja

Za izradu ovog rada, u cilju analize koncepta rudarenja podataka u poslovanju, koristit će se sistemski pregled literature kao osnovni metodološki pristup. Isti predstavlja metodološki princip identifikovanja, evaluiranja i interpretiranja dostupnih empirijskih istraživanja za datu istraživačku temu (Ghapanchi i Aurum, 2011). Osim toga, predstavlja način na koji se vrši sinteza dokaza postojećih istraživanja kojih karakteriše zajednička osobina, tj. tema, a za koju se daje kritički osvrt i analiza. Dodatno, sistemski pregled literature pridodaje veću naučnu vrijednost, zbog čega će se i koristiti.

Shodno tome, sistemski pregled literature će se koristiti kako bi se utvrdile zajedničke tačke istraživanja na temu rudarenja podataka u poslovanju, uz deskriptivnu analizu koja podrazumijeva analizu citiranosti određenih radova ili časopisa u kojim su objavljeni. Također, podrazumijeva se kociatna analiza i grafički prikazi iste (Rizvić, 2016).

Na osnovu navedene metodologije, istraživanje će približiti značaj i sličnosti koje se dovode u vezu sa temom, pri čemu će se izvršiti teoretski osvrt oblasti.

1.5. Struktura rada

Cilj koji identifikuje, vrši odabir i kritički procjenjuje istraživanja kako be se samim tim pružili odgovori na jasna postavljena istraživačka pitanja jeste sistemski pregled literature koja je poznata kao i sistemski pregled literature (SLR - Systematic Literature Reviews - Literature Review - Library Guides at Charles Sturt University, n.d.).

Sistemska revizija treba slijediti jasno definirani plan u kojem su kriteriji jasno navedeni prije provođenja revizije. To je sveobuhvatna i transparentna pretraga koja se provodi na više baza podataka koji čine literaturu, te može biti ponovljena i reprodukovana od strane drugih istraživača. Uključuje planiranje dobro promišljene strategije pretraživanja koja ima specifičan fokus ili odgovara definisanom pitanju. Revizija identifikuje vrstu informacija koje se traže, kritizira i izvještava unutar poznatih vremenskih okvira. Pretraživački pojmovi, strategije pretraživanja (uključujući nazive baza podataka, platforme, datume pretraživanja) i ograničenja sve trebaju biti uključeni u reviziju.

Pittway identifikuje sedam ključnih principa koji se primjenjuju u sistemskim literaturnim revizijama:

1. Transparentnost
2. Jasnoća
3. Integracija
4. Fokus
5. Jednakost
6. Dostupnost
7. Pokrivenost

Autor Smith, 2018 u svom radu ističe da izvršavanje pregleda relevantne literature posjeduje ključnu ulogu u svim disciplinama, te ovaj proces pruža pomoć u detaljnom mapiranju i samoj procjeni znanja koja su već postojeća o određenoj temi, identifikaciji nedostataka i praznine u istraživanju koja samim tim, doprinosi daljem razvoju baze znanja (Smith, 2018).

Također, analiza literature dalje pruža mogućnost pregleda relevantnih istraživanja, te nakon dobivenih rezultata i zaključaka može graditi temelj za buduća istraživanja i donošenja informiranih odluka (Tranfield *et al.*, 2003).

Sistematska analiza literature (SLR) se razlikuje od tradicionalnih narativnih pregleda usvajanjem znanstvenih i transparentnih pristupa (Mengist *et al.*, 2020). Sam proces pozitivno utječe na prikupljanje odgovarajućih dokumenata, kako bi se moglo odgovoriti na zadato istraživačko pitanje.

Sa ciljem smanjenja pristranosti prilikom različitih procesa koji podrazumijevaju pretragu, identifikaciju, procjenu, sintezu, analizu i sažetak studija, sistemska analiza literature zasnovana je na jasnim i sistematičnim postupcima. Studija pruža pouzdane rezultate samo ako se postupak pravilno izvede sa minimalnim greškama, a samim tim i pouzdan zaključak na osnovu kojeg je moguće donijeti odluke za proporcionalne postupke znanstvenih praktičara (Lau *et al.*, 1992).

Dobro izveden postupak za SLR proces je ključan i osigurava da se rad pažljivo planira prije početka stvarnog pregleda, ali pored toga potrebno je svakako primijeniti i osnovne statističke tehnike koji se nakon primijene dolazi do dobivenih rezultata nastalih na kombinaciji podataka iz drugačijih povezanih studija slučaja, što proizvodi nastanak meta-analize. Primjena ove metode može doprinijeti generisanju preciznijih procjena vezanih za istraživačku temu.

Glavne karakteristike SLR-a i pripadajućeg postupka, meta-analize, su:

- jasno postavljeno istraživačko pitanje na koje će studija odgovoriti,
- jasno definirani ciljevi s eksplicitnom i ponovljivom metodom,
- niz pretraživanja koji uključuje sve relevantne studije koje zadovoljavaju kriterije za uključivanje,
- procjena kvalitete/valjanosti odabranih studija (npr. procjena rizika od pristranosti i povjerenje u kumulativne procjene),
- sistematično prezentiranje i sinteza izvučenih podataka iz odabranih studija i
- dostupnost rezultata studije u svrhu znanstvenog istraživanja i donošenja odluka (Shamseer *et al.*, 2015).

2. PREGLED LITERATURE

2.1. Rudarenje podataka

2.1.1. Definicija i uloga

Ključna sintagma koja je nastala razvojem tehnologija u posljednjem stoljeću predstavlja *razvoj podataka*. Podaci predstavljaju normativni mehanizam iz kojeg nastaju različiti poslovni procesi.

Sam tehnološki napredak jeste katalizator napretka velikih količina podataka nastalih iz različitih izvora, a koji su danas globalno zastupljeni, te lahko dostupni, zbog čega dolazi do ubrzanog rasta podataka, što dovodi do koncepta „velikih podataka“ (eng. *Big data*). Isti predstavlja kompleksni i sveobuhvatni skup podataka prikupljenih u realnom vremenu, a procesiranih u velikim količinama (Ishwarappa i Anuradha, 2015). Potrebe za velikim podacima prvenstveno dolaze razvojem poslovanja kompanija na tržištu, a i samog tržišta.

Veliki podaci se definišu na osnovu sljedećih karakteristika:

Volumen

Mnoge kompanije već posjeduju velike količine arhiviranih podataka u obliku logova. Shodno tome, suočavaju se s novim izazovom jer nemaju dovoljno kapaciteta ili resursa da se ti podaci efikasno obrade. Posebna atrakcija analitike velikih podataka se nalazi u prednosti koja se dobiva sposobnošću, kako bi se mogla izvršiti detaljna obrada masovnih količina informacija.

Brzina

Brzina se odnosi na povećanje brzine u kome se podaci kreiraju i stvaraju, samim tim na svu veću brzinu podaci se mogu obraditi i pohraniti, te i analizirati putem baze podataka koji su relacioni podaci. Kao primjeri u današnjici se navode, da se na platformi YouTube učita stotine sati video sadržaja, a preko Gmail-a se šalje preko 500 miliona emailova.

Varijabilnost

Varijabilnost podataka podrazumijeva raznovrsne oblike i forme podataka koje zauzimaju veliki podaci. Cijeli koncept velikih podataka bazira se na podacima koji mogu biti strukturirani odnosno podaci koji su organizovani, te imaju određenu formu poput podataka u tabelama ili bazama i nestrukturirani koji nemaju navedeno, odnosno prikazani su u obliku teksta, audio ili video zapisa.

Dakle, varijabilnost podrazumijeva da veliki podaci imaju različite vrste, oblike, formate i strukture. Kako bi se dobile važne i skladne informacije iz takvih podataka, veoma je zahtjevno sprovesti detaljnu obradu i analizu.

Vjerodostojnost

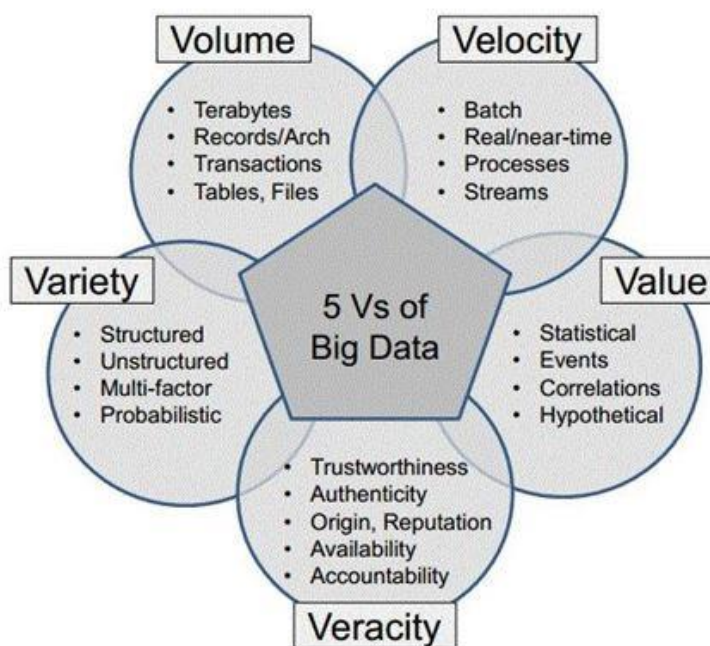
Pouzdanost i ispravnost velikih podataka je vrijednost, te budući da se susreće sa ogromnim i intenzitetom i brzinom, također i sa raznim podacima nije moguće da u datom momentu mogu postići potpunu tačnost podataka. To je razlog zbog čega se pojavljuje ta pojava tzv. „prljavih“ ili nepotpunih podataka.

Dodatno, ističe se da je u okviru razumijevanja za realizaciju polaznog i preciznog ispitivanja podataka, treba se kao prvo i osnovno osigurati izvor informacija koji uključuje upotrijebljenu pažnju, kako bi se cilj realizovao sa onoliko tačnosti i nepokretnosti koliko je to moguće u donošenju odluka u poslovanju.

Vrijednost

Pojam vrijednosti podrazumijeva onu vrijednost podataka koja je nastala iz analize i interpretacije samih podataka, a zauzima različite oblike poput otkrivanja novih informacija, procesa donošenja odluka, optimizacije poslovnih procesa, unaprijeđenja procesa i poslovanja ili kreiranja konkurentske prednosti. Ukoliko se podaci ne mogu pretvoriti u vrijednost, oni se smatraju beskorisnim, jer vrijednost velikih podataka jeste enormna.

Ilustracija 1 Karakteristike velikih podataka



Izvor: [The 5V of Big Data Characteristics | Download Scientific Diagram \(researchgate.net\)](#)

Značaj podataka je enorman. Sva poslovanja kompanija zavise od podataka kojima raspolažu. Kako bi se postiglo kvalitetno i jednostavnije donošenje odluka u kompaniji, podaci moraju biti pravovremeni i tačni jer su u suštini temelj za poslovanje, a podrazumijevaju i oblik znanja. Podaci mogu biti interni i eksterni.

Ukoliko su interni, oni se vežu direktno za kompaniju, te ne mogu biti primjenjeni van iste, dok eksterni podaci su oni koji menadžmentu daju informaciju realnoj slici tržišta i egzogenih faktora koji bi potencijalno utjecali na poslovanje.

Podatke je moguće koristiti za rješavanje nastalih poslovnih problema. Na osnovu podataka, moguć je uvid u problem sa ciljem razumijevanja, te određivanja potencijalnih rješenja. Dodatno, moguće je sprovesti implementaciju novih rješenja i napretka poslovnih procesa, kako bi se postigla efikasnost i kvalitetnije poslovanje. Zasiurno analizom podataka, menadžeri imaju bolji pogled koji ispisuje nedostatke, prepreke ili manjkavosti sistema.

Razumijevanje poslovanja u kompaniji se također temelji na adekvatnim podacima koji se interpretiraju menadžerima. Shodno tome, menadžeri su u prilici da bolje razumiju samu kompaniju, ali i okruženje u kom se nalazi (Saha, 2020).

Shodno tome nastao je koncept rudarenja podataka (eng. *Data mining*).

Razne tehnološke inovacije su u današnjici osnovni alat za ekstrakciju i sprovedbu potpune analize velikih podataka koje nose osnovne informacije za korisnika. Pomoću njih, mogu se otkriti bitne veze između različitih varijabli. Kao što je navedeno, rudarenje podataka sadrži metodologiju kako bi se došlo do navedenih procesa. Autori tvrde da se rudarenje podataka može koristiti i u drugim oblastima kao što su bankarstvo, industrija, medicina, marketing, prodaja i sl. (Akćetin *et al.*, 2015).

Data mining, poznat kao otkrivanje znanja u bazama podataka (eng. *Knowledge discovery in databases - KDD*), predstavlja proces pronalaženja uzoraka u velikim skupovima podataka koji su korisni za kompanije. Sam proces čini otkrivanje znanja iz dostupnih podataka koji dovode do njihove obrade, pretvarajući ih u strukturane, čitljive podatke. Navedeni uzorci omogućavaju kompanijama da donose kvalitetne poslovne odluke koje podstiču rast i razvoj (Srića, 2018).

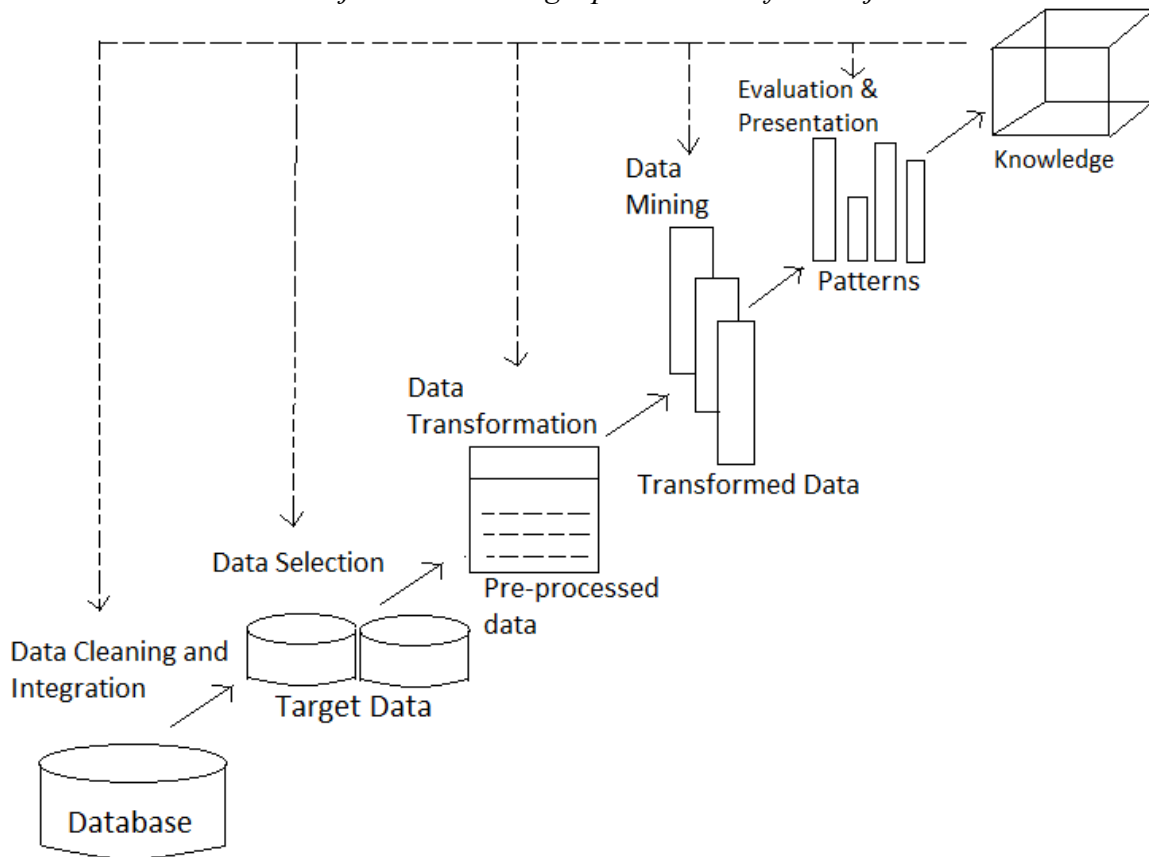
Uz tehnološke napretke, koji su bili vrijedni pažnje u posljednjoj deceniji, informacije koje su odložene i dostupne u elektronskim oblicima brzo se razvijaju. Sve kompanije na oglasu imaju zajedničku osobinu, koji podrazumijeva prikupljanje različitih vrsta i oblika podataka približno dostupnih poslovnih formi, pri čemu se kontrolišu monetarne informacije, podaci oko oglasa, klijenti i slično. Takve informacije su zasigurno jedni od najbitnijih resursa kompanije.

Kako bi se takve dostupne informacije analizirale iz karakterističnih izvora, brojne kompanije kontinuirano pokušavaju pronaći različite zadovoljavajuće strategije za primjenu istraživanja, kao i inovativne instrumente (Fayyad *et al.*, 1996).

Smatra se da je definicija diskutabilna zbog činjenice da uključuje širok spektar strategija i kalkulacija iz različitih oblasti kao što su baze podataka, mašinsko učenje i uvidi, ali naknadno rudarenje informacija aludira na jedan od koraka, ali također i "ključ" pronalaženja informacija.

Kako rudarenje informacija raste, tako se razvijaju i drugi koraci kao što su čišćenje informacija i vizualizacija, što su koraci koje bi trebalo izvršiti kako bi se dobili imperativni podaci iz ogromnih količina informacija koje su izuzetno kritične za buduće poslovanje.

Ilustracija 2 Data mining u pronalasku informacija



Izvor: Knowledge and Discovery Process Steps [5]. / Download Scientific Diagram (researchgate.net)

U svijetu u kojem se velika količina podataka pouzdano stvara i širi iz dana u dan, gdje se kompanije suočavaju s problematičnim izazovom, budući da se njihovo poslovanje temelji na gledanju trenutno dostupnih podataka i analiziranju prikupljenih informacija kako bi došli do cilja. Na dnevnoj bazi se bilježe terabajti ili petabajti podataka koji se generišu iz različitih sfera poslovanja, društva, nauke, inženjerstva, medicine i drugih, a ti podaci se skladište putem *www*-a (eng. *World Wide Web*).

U ovakvim okolnostima, kada svijetom kruže ogromne informacije, izuzetno je problematično izvući vitalne i relevantne podatke iz ogromnih količina informacija, kompanije prepoznaju značaj rudarenja informacija za napredak trgovine, jer omogućava preciznije istraživanje klijenata, pravljenje profila i prikupljanje kupaca.

Shodno tome, proces koji se sastoji od sortiranja, organizovanja i/ili grupisanja podataka poznat kao rudarenje podataka, ima ključnu važnost za ovakve kompanije.

Postoje mnoge kompanije koje podatke danas tretiraju kao sinonim za otkrivanje znanja iz podataka ili KDD, dok drugi smatraju da je Data mining bitan korak u procesu otkrivanja znanja (Han, Kamber, Pei, *et al.*, 2012).

Proces otkrivanja znanja se sastoji od:

1. Čišćenje podataka
2. Integracija podataka
3. Izbor podataka
4. Transformacija podataka
5. Data Mining
6. Procjena obrasca
7. Przentacija znanja

U prvom koraku unutar ovog procesa vrši se čišćenje podataka što predstavlja uklanjanje smetnji i neskladnih podataka, te samim tim zahtjeva određene radnje. Kao što autor Stephane Tuffery ističe, potrebno je prvo definisati problem koji će se riješiti nakon završetka procesa, a zatim započeti analizu gdje je Data mining već prethodno uspješno korištena. Drugi korak, koji je vremenski najduži i gdje se mogu kombinirati više izvora podataka, predstavlja ključnu fazu za konačni izlaz iz procesa rudarenja podataka. Shodno tome u trećem koraku, neophodno je adekvatno procijeniti podatke iz baze kako bi se dobio što precizniji model. Transformacija podataka je četvrti korak u ovom procesu, te podaci se transformiraju i objedinjuju u obrasce koji su prikladni za rudarenje. Peta faza je Data mining i u ovoj fazi se primjenjuju inteligentne metode za izdvajanje obrazca iz velike količine podataka, te shodno tome različite tehnike, statističke metode ili rudarenje, koriste se za otkrivanje skrivenih obrazca, povezanosti ili trendova u podacima.

U narednom šestom koraku je procjena obrazaca kako bi se identificirali zanimljivi i relevantni obrasci na temelju mjera zanimljivosti.

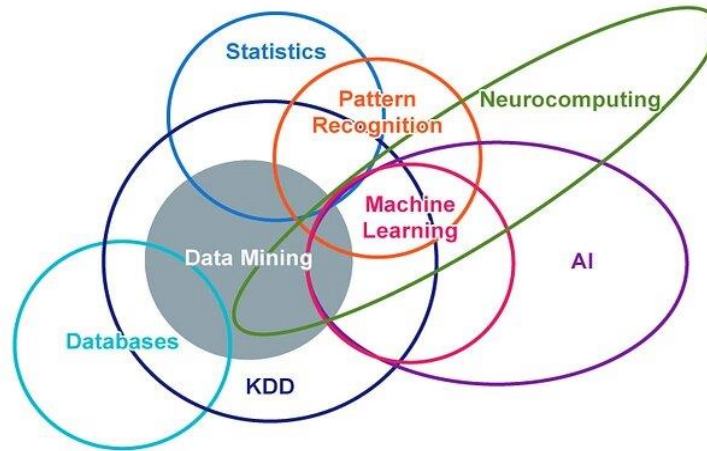
Zadnji korak jeste prezentacija znanja, te odnosi se na to da se predstave dobivena znanja korisnicima.

U ovoj sedmoj fazi se koriste ne samo predstavljanje znanja nego i tehnike vizualizacije kako bi se prikazali rezultati Data mininga na razumljiv i interpretirajući način, te se to odnosi na prikazivanje grafičkih prikaza, tablica, izvještaja i sl. Ove faze su samo neke od koraka koji omogućavaju lakši proces korištenja Data Mininga (Tuffery, 2008).

Tehnologije unutar Data mininga

Data Mining je domena koja se temelji na aplikacijama i obuhvata mnoge tehnike iz različitih područja, a to su: statistika, mašinsko učenje, prepoznavanje uzorka, sistem baza podataka, pronalaženje informacije, vizualizacija, umjetna inteligencija i mnoge druge aplikacijske domene. U nastavku bit će navedeni primjeri nekoliko disciplina koje imaju snažan uticaj na razvoj metoda Data mininga (Cypher i Halbert, 1993).

Ilustracija 3 Tehnologije Data mining-a



Izvor: [Illustration for nlt blog Automated Machine Learning as it is 3-1.jpg](#)
(767×407) (newline.tech)

Statistika

Statistika, kao jedna od najkorištenijih nauka ima za zadatak da prikupi, analizira, tumači i prezentuje velike skupove podataka, predstavljajući skup u kom su dostupne matematičke funkcije kojima se opisuje ponašanje objekata u ciljnoj klasi koristeći varijable i raspodjele vjerojatnosti (Hill, 2012).

Statistički modeli se koriste za modeliranje podataka i klasifikaciju. Kao primjer u oblastima Data mininga, statistički modeli mogu se konstruisati za ciljne klase u zadacima kao što su karakterizacija i klasifikacija podataka. Metode rudarenja podataka podrazumijevaju složenije procese od deskriptivnih statističkih metoda, pri čemu se ovakve metode naslanjaju na alate vještačke inteligencije poput neuronskih mreža, teorije informacija, teorije mašinskog učenja, inferencijalne statistike i klasične analize podataka (Wilson, 1999).

Pri tome se podrazumijeva i faktorska analiza, klasterisanje, diskriminantna analiza i druge tehnike. Važno je pažljivo odrediti algoritme i prilagoditi ih kompaniji zajedno sa prilagodbom metoda na velikim skupovima podataka sa ciljem pozitivnog utjecaja na poslovanje kompanija u svim segmentima (Han, Kamber, Pei, *et al.*, 2012).

Istraživanje podataka se sve više implementira u svijetu poslovanja iz raznih uzroka. Razlozi koji su doprinjeli samom razvoju su: pritisak konkurencije i rastuća očekivanja potrošača, te regulatorni zahtjevi u određenim industrijama i mnogim drugim. Kao rezultat toga, ostvarena su tri (više) ključna tehnička napretka, koja su samim tim doprinjela napredovanju istraživanja podataka.

Prvo, moderna računarska oprema i metode pružaju velike kapacitete skladištenja i proračuna, te shodno tome skladišta podataka mogu sadržati desetine terabajta. Drugi napredak egzistira u sve većoj dostupnosti različitih paketa statističkih algoritama i algoritama za istraživanje podataka.

Ovi algoritmi se mogu automatski povezati, pružajući jednostavan pristup, visokokvalitetne rezultate i interaktivne mogućnosti koje su ranije bile nezamislive. Treći napredak se odnosi na promjene u procesu donošenja odluka, te istraživanje podataka se koristi u proizvodnim procesima gdje se periodično dostavljaju krajnjim korisnicima uz automatsko pokretanje događaja.

Smatraju da sa ovim napretcima su značajno unaprijedili područje istraživanja podataka, pružajući veće kapacitete, bolje alate i šire mogućnosti primjene. Uz tri prethodno spomenuta napretka, postoji i četvrti koji se odnosi na sposobnost obrade različitih vrsta podataka, uključujući nepotpune podatke, to znači oni kojima nedostaju vrijednosti. Peti napredak koji je odigrao imperativnu ulogu u napretku informacionog rudarenja je osnivanje ogromnih baza podataka za ispunjavanje zahtjeva administracije kompanije u poslovnim situacijama (Tuffery, 2008).

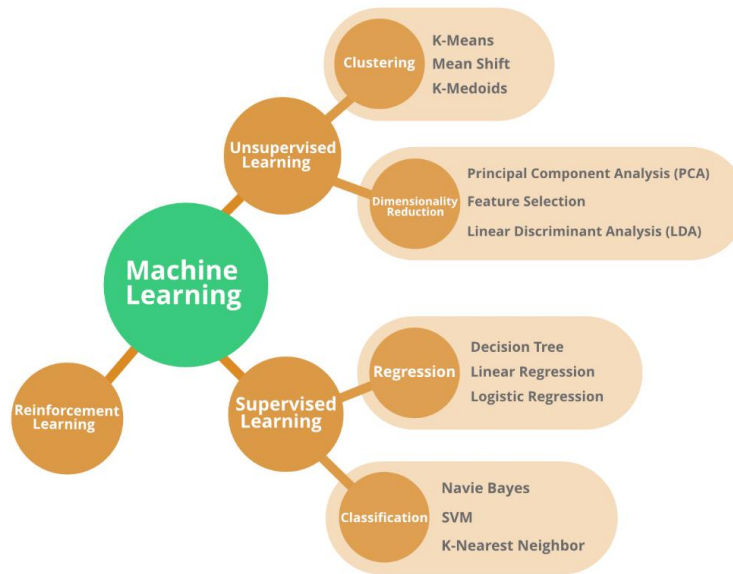
Inteligentni sistemi neprestano rastu, pa se danas očekuje da i najjednostavniji sistemi posjeduju sposobnosti koje su nekad bile smatrane inteligentnim.

Dinamična inteligencija se postiže kroz proces učenja iz ogromnih dostupnih informacija i interakcije sa okolinom. Način na koji funkcionišu jeste da se koriste i raspoloživim informacijama, te kao i baza promjenjivih činjenica koja će im pomoći da djeluju u složenim situacijama.

Također, mašinsko učenje predstavlja se kao podvrsta vještačke inteligencije koja stvara mogućnost računarima da programiraju bez prethodnog predznanja. Kao osnova mašinskog učenja podrazumijeva izučavanje i shvaćanje algoritama koji se prilagođavaju, te i učenja novim podacima tokom samog treninga.

Mašinsko učenje i rudarenje podataka imaju i zajedničku tačku, a to je činjenica da u oba područja istražuju obrasce u podacima. Shodno tome, mašinsko učenje djeluje inteligentno u pronalaženju obrasca prilagođavajući se novim skupovima podataka, a Data mining se bavi izdvajanjem podataka koji su izloženi interpretaciji od strane ljudi. Mašinsko učenje je disciplina koja brzo napreduje (Han, Kamber, Pei, *et al.*, 2012).

Ilustracija 4 Algoritmi mašinskog učenja



Izvor: <https://www.geeksforgeeks.org/top-10-algorithms-every-machine-learning-engineer-should-know/>

Učenje pod nadzorom proizilazi iz označenih primjera u sklopu podataka za obuku. Učenje bez nadzora se odnosi na klasteriranje, te klasteriranje se koristi za otkrivanje različitih klasa unutar podataka. Polunadgledno učenje obuhvata različite tehnike mašinskog učenja koje koriste kombinaciju označenih i neoznačenih primjera prilikom treniranja modela. Označeni primjeri se koriste za učenje klasifikacijskog modela, dok neoznačeni primjeri pomažu u razgraničavanju granica među klasa.

Sistem baza podataka

Posebna pažnja je posvećena istraživanju u oblastima sistema baza podataka koje su na razvoju, održivosti te korištenju istih za kompanije i korisnike. Uspostavljeni su posebni principi na ovom području koji se ne odnose samo na modeliranje podataka, nego i na samo skladištenje podataka i metodama pristupa. Baze podataka su poznate po svojoj visokoj skalabilnosti za obradu velikih, relativno strukturiranih skupova podataka. Data mining zahtjeva obradu velikih skupova podataka i brz prijenos podataka u stvarnom vremenu.

Kako bi postigli visoku učinkovitost i iskoristili skalabilnost na velikim skupovima podataka, Data mining koristi skalabilne tehnologije baza podataka. Shodno tome, ove tehnologije omogućuju proširenje sposobnosti postojećih baza podataka kako bi zadovoljile zahtjeve naprednih korisnika za analizom podataka.

Razvoji u području baza podataka integriraju skladišta podatka i alate za Data mining kako bi omogućili sistemsku analizu podataka.

Ovaj model ne samo da olakšava OLAP (eng. *Online Analytical Processing*) u višedimenzionalnim bazama podataka, već također potiče višedimenzionalno rudarenje podataka (K. Ojo i B. Adeyemo, 2017).

2.1.2. Metode i okviri rudarenja

Korištenje Data mininga rezultira dobivanjem korisnih informacija u cilju da kompanije donose preciznije i tačnije odluke u svom poslovanju. Kako bi se stvarali uspješni rezultati u analizi može postati značajno za preduzeća i poziciju na tržištu, identifikaciju kupaca i shvaćanje njihovih konkretnih navika, te kupovne moći i dr. Rudarenje podataka također ima korijene i u drugim sferama kao što su matematika, vještačka inteligencija, veliki podaci i statistika. Napredak tehnologije i informatizacije omogućio je pristup podacima koji se nalaze u različitim bazama, te mogu se analizirati u različitim svrhama (Garača i Jadrić, 2011).

Metode za pronalaženje informacija iz baze podataka podrazumijevaju:

1. Klasifikacija,
2. Klaster,
3. Regresija,
4. Pravila pridruživanja,
5. Neuronske mreže,
6. Stablo odlučivanja.

Klasifikacija

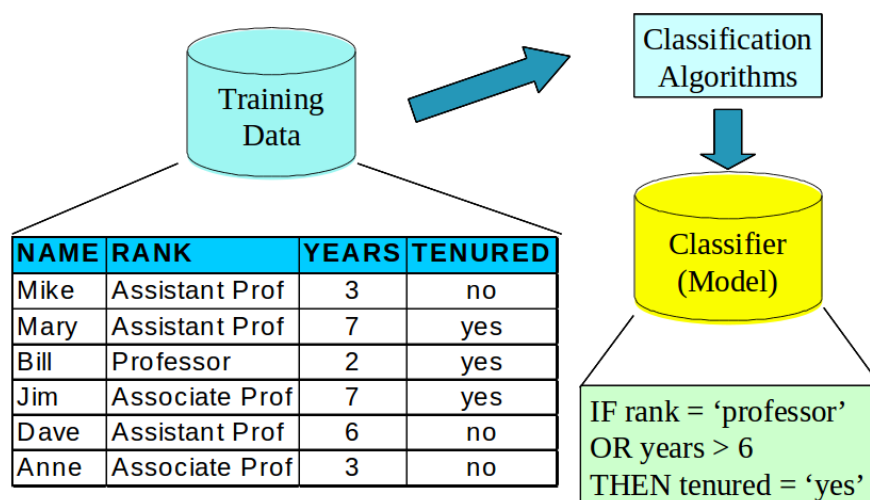
Klasifikacija je tehnika analize podataka koja se oslanja na prethodno klasificirane primjere kako bi razvila model sposoban za klasifikaciju cjelokupne populacije zapisa. Shodno tome, korisna je u detekciji prevara i procjeni kreditnog rizika. Stablo odlučivanja i neuronske mreže su također metode koje su u ovom pristupu kombinirajući različite tehnike korišćene. Proces klasifikacije podataka uključuje fazu učenja i fazu klasifikacije. U fazi učenja, podaci se analiziraju pomoću algoritama za klasifikaciju, dok se podaci za testiranje koriste za procjenu tačnosti i klasifikacijskih prava (Ramageri, 2014).

Postoje različite vrste klasifikacionih modela koje se koriste u analizi podataka a to su:

1. Klasifikacija po indukciji stabla odlučivanja – kako bi se klasificirali podaci na osnovu pitanja i uvijeta ova metoda koristi stablo odlučivanja.
2. Bayesova klasifikacija – ovom tehnikom se podrazumijeva direktno oslanjanje na Bayesovu statistiku kako bi se mogla izvesti klasifikacija na osnovu prethodnih informacija koje su dostupne i na osnovu vjerovatnoće.
3. Neuronske mreže – one su inspirisane strukturom ljudskog mozga i koriste se za klasifikaciju kroz mrežu međusobno povezanih neurona.
4. Podržavaju vektorske mašine (eng. SVM) – koristi se matematički pristup kako bi pronašao šta najbolje razdvaja različite klase podataka.
5. Klasifikacija zasnovana na udruženjima – metoda prikazuje da se potvrđuju obrasci ili pravila u podacima koji imaju tu mogućnost klasificiranja novih podataka.

Pored toga, svaka metoda posjeduje i svoje karakteristike i prednosti, te će biti iskorištene u zavisnosti od zahtjeva i same prirode podataka koje se trebaju klasificirati u budućnosti.

Ilustracija 5 Prikaz metode klasifikacije

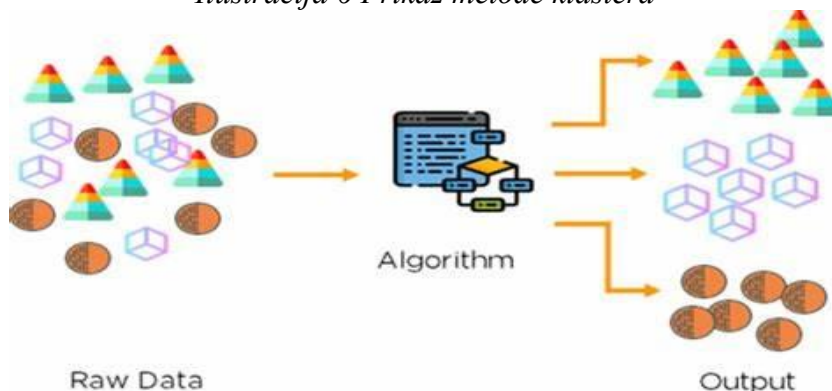


Izvor: <http://jblomo.github.io/datamining290/slides/2013-02-22-Decision-Trees.html>

Klaster

Klaster je tehnika identifikacije sličnih grupa objekata. Koristeći klasterizaciju, mogu se identifikovati i gusta i rijetko naseljena područja u prostoru objekta, otkriti ukupni obrazac distribucije i korelaciju između atributa podataka. Danas, klasterizacija je postala popularna metoda zbog široke dostupnosti i jednostavne primjene, te može također biti efikasno sredstvo za razlikovanje grupa ili klasa objekata. Koristi se i kao pristup za odabir i klasifikaciju podskupova atributa kako bi se primjetile sličnosti i razlike unutar podataka. Na primjer, klasterizacija se može iskoristiti, tj. da kompanije formiraju grupe kupaca na osnovu obrazca kupovine, samim tim omogućava se segmentiranje kupaca (Ramageri, 2014).

Ilustracija 6 Prikaz metode klastera



Izvor: <https://medium.com/analytics-vidhya/overview-of-collaborative-filtering-algorithms-9a76d2eb861b>

Postoje različite metode klasteriranja koje se koriste za grupisanje objekata, a to su:

1. Metode K-sredina
2. Hijerarhijsko aglomerativne (podjele) metode
3. Metode Two-step cluster analiza
4. Metode zasnovane na mreži
5. Metode zasnovane na gustini.

Regresija

Regresijska analiza se primjenjuje između dvije varijable, a to se odnosi između nezavisnih varijabli i zavisnih varijabli kako bi modelirali vezu koja se može koristiti za predikciju, te samim tim su neovisne varijable poznate i bitnije, a zavisne varijable se moraju predvidjeti. Samim tim, u današnjem stvarnom svijetu postoje mnogi problemi gdje nije u pitanju samo predviđanje. Regresija se odlikuje ovisnošću između varijabli i prihvata uticaj jedne varijable na reakciju druge varijable. U takvim slučajevima su potrebne napredne procedure, kao što su stabilno donošenje odluka ili neuronski sistemi za predviđanje budućih vrijednosti.

Metode regresije koje se primjenjuju u analizi su:

1. Linearna regresija – podrazumijeva pristup modeliranju linearnih odnosa između više varijabli sa mogućnošću predviđanja njihovih vrijednosti na osnovu kombinacije nezavisnih varijabli, zaviseći od nepoznatih aproksimacija parametara iz podataka.
2. Multivarijantna linearna regresija – podrazumijeva pristup gdje varijable zavise od nekoliko slobodnih faktora.
3. Nelinearna regresija – ova regresija se vrši kada postoji odnos koji je nelinearan između nezavisnih i zavisnih varijabli, te pri tome ova metoda se može koristiti za različite funkcije i zakrivljenosti u podacima kako bi se modelirali.
4. Multivarijantna nelinearna regresija – predstavlja model koji kombinira nelinearnu regresiju sa više nezavisnih varijabli.

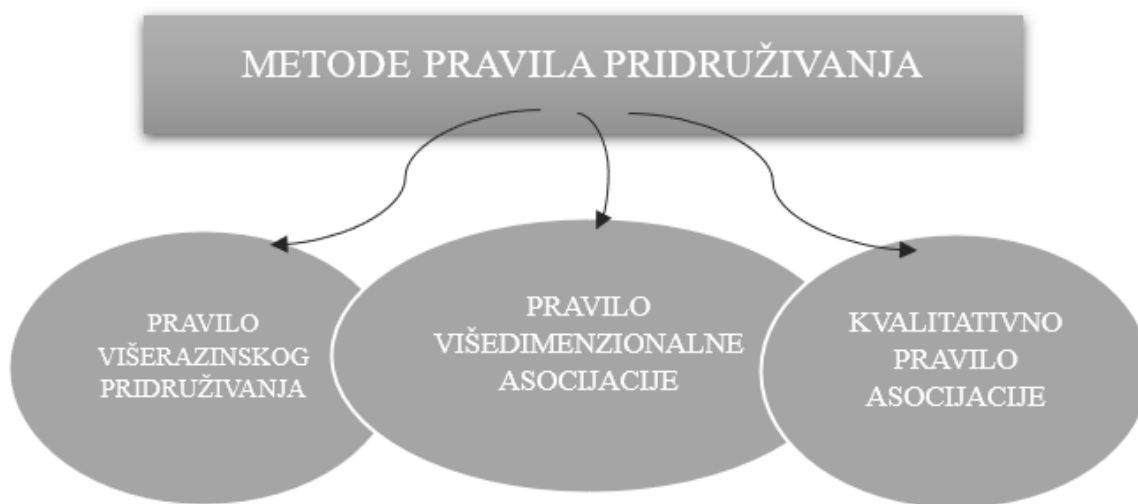
Svaka od ovih metoda regresije ima svoje karakteristike i primjenu ovisno o prirodi podataka i ciljevima analize (Rodríguez del Águila i Benítez-Parejo, 2011).

Pravila pridruživanja

Tehnike u kojima se otkrivaju sakriveni obrasci u podacima, te sa svrhom identifikacije veza među varijablama zove se pravila pridruženja. Shodno tome, ova analiza omogućava otkriće događaja i frekventne kombinacije različitih varijabli, te također mogućnost ovog pristupa tj. pravila pridruživanja prepoznaje uzorke i povezuje minimalno dvije varijable u podacima. Dakle, pravila pridruživanja imaju i praktičnu primjenu u poslovnim odlukama, a to podrazumijeva oblikovanje kataloga, unakrsni katalog i sl.

Algoritmi pravila pridružavanja moraju generirati pravila s određenom pouzdanošću, ali shodno tome broj mogućih pravila pridruživanja za određeni skup podataka je vrlo velik, te zbog toga samo mali dio pravila ima značajnu vrijednost (Ramageri, 2014).

Ilustracija 7 Metode pravila pridruživanja



Izvor: Autor

Neuronske mreže

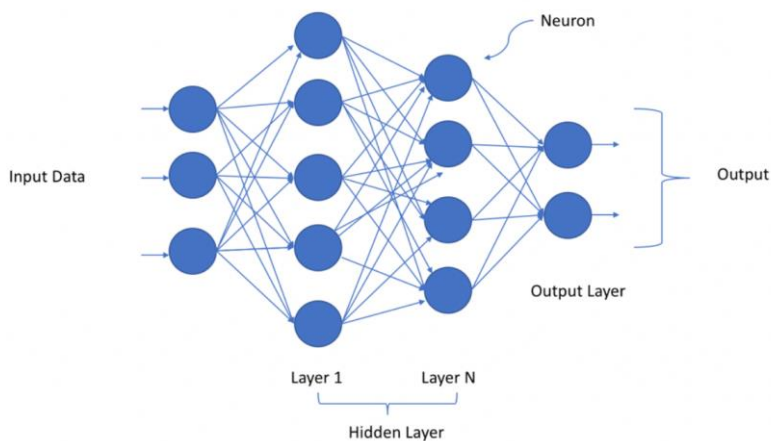
Neuronska mreža se sastoji od povezanih jedinica za ulaz i izlaz, pri čemu svaka veza ima težinu. Tokom trajanja faze učenja, mreža prilagođava težinu kako bi mogla precizno predviđati klasifikaciju ulaznih podataka.

Također, izuzetno su sposobne za izvlačenje značenja iz kompleksnih ili nepreciznih podataka koje se mogu koristiti za otkrivanje obrazca i trendova koji su preazhtjevni za ljude ili druge računarske tehnike.

Neuronske mreže identificiraju obrasce ili trendove u podacima, te su samim tim savršene za predviđanje potreba oblasti u kojima se primjenjuje i veoma su uspješna područja, a neke od tih oblasti su: obrazovanje, elektronika, poslovanje, zdravstvo i sl. Tolerisanje grešaka i jeste jedna od sposobnosti neuronskih mreža, jer ona se sastoji od mnogo elemenata koje su potrebne za obradu, te samim tim mogu u nastavku funkcionisati iako se dokazuje da je dio mreže oštećen.

Shodno tome, iako se prezentira nepotpun skup ulaznih podataka, mreža će i dalje biti sposobna da izvrši izlaznu aktivnost.

Ilustracija 8 Data mining pri korištenju metode neuronskih mreža



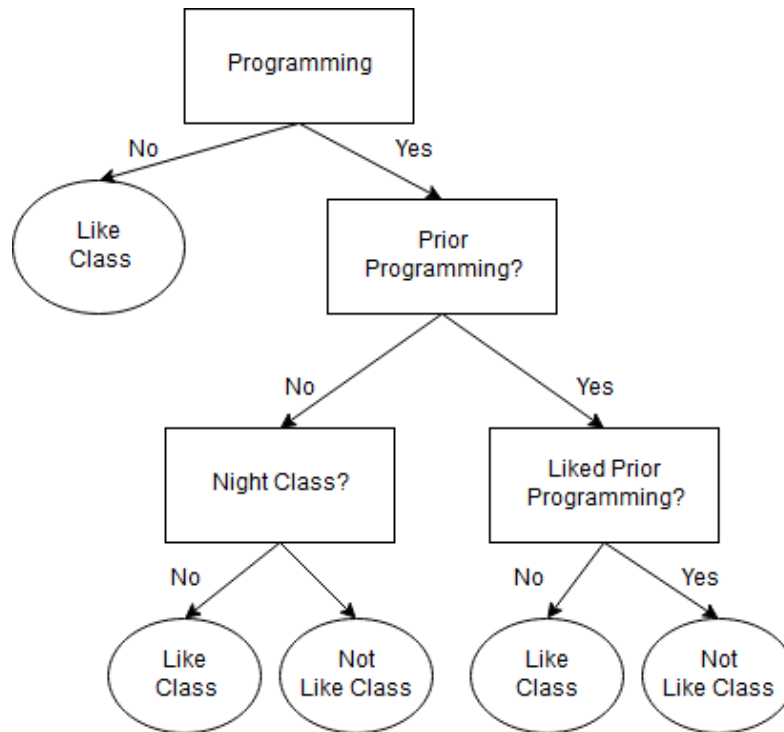
Izvor: <https://towardsdatascience.com/a-laymans-guide-to-deep-neural-networks-ddcea24847fb>

Stablo odlučivanja

Stablo odlučivanja se koristi za organizacije ili predviđanje podataka, te samim tim proces izgradnje stabla odlučivanja započinje pitanjem koji ima najmanje dva moguća odgovora. Svaki odgovor vodi ka novom pitanju koji se koristi za karakterizaciju ili razlikovanje podataka koji se mogu klasificirati.

Metodom stabla odlučivanja se dobivaju grafikoni koji prikazuju utjecaj ulaznih varijabli na izlazne varijable, koje su prikazane u obliku klasa. Svaki čvor u stablu predstavlja jednu ulaznu varijablu, a na granama čvora označavaju se novi „čvorovi“ za svaku moguću vrijednost te ulazne varijable. Stablo se gradi „učenjem“ na temelju dostupnih podataka, pri čemu se originalni skup podataka razgrana na manje skupove na temelju testiranja vrijednosti varijabli.

Ilustracija 9 Stablo odlučivanja



Izvor: <https://www.captechu.edu/blog/machine-learning-decision-trees-revisited>

2.1.3. Analiza podataka na osnovu koncepta rudarenja

U današnjem digitalnom dobu, količina dostupnih podataka konstantno raste, samim tim se otvaraju nove mogućnosti za dublje razumijevanje i donošenje odluka. Rizik koji nastaje među pojavom konstantnog rasta informacija jeste njihova netačnost. Mnogi podaci koji su lahko dostupni i imaju konstantan tok na tržištu su nevjerodostojni, a mogu narušiti poslovanje kompanije ili njenu reputaciju, budući da daju pogrešnu sliku o nekoj tematici zbog koje donesene odluke utemeljene na lažnim informacijama mogu odvesti poslovanje u pogrešan smjer.

Brojne kompanije prepoznaju poštovanje i važnost značajnih podataka, slažući se sa otvaranjima koja se pojavljuju, kao što je kreiranje različitih aranžmana za pitanja ili napredovanje oblika koji utiču na kompletno poslovanje.

Prilikom prikupljanja informacija, kompanije moraju jamčiti legitimitet koji podrazumijeva oštromnost i kvalitet informacija što je u osnovi ključna dionica istrage, koja sugerira preciznost, 'čistoću' i potpunost informacija na način da se otklone sve greške koje će biti pronađene, uzimajući u obzir potencijalni problem koji će se dogoditi ako informacija bude oštećena.

Dakle, može se zaključiti da informaciono rudarenje podataka sugerise način pronalaženja kvalitetnih podataka i informacija među ogromnim informacijama, pri čemu je pronalaženje strukture, asocijacija i obrazaca među informacijama cilj u organizovanju podsticanja trgovinskih formi zasnovanih na odlučivanju kompanija, koji moraju biti tačni i relevantni.

Sa razvojem i brzim razvojem zbira dostupnih književnih informacija u naprednom rasporedu, javljaju se nova otkrića istraživanja zamislivih ishoda kompjuterizovanim kanalima, budući da je istovremeno rukovanje sadržajem nesumnjivo jedan od bitnih oblika istraživanja. U slučaju da kompanija inkorporira složene upite za svoju trgovinu, njihovo istraživanje će također biti složeno, najvećim dijelom zbog utjecaja na napredovanje inovacija i primjene umjetnih uvida, što uključuje različite pripreme, mašinsko učenje ili priznavanje obrazaca u podacima.

Prepoznatljivi modeli trgovine, oblici i strukture koji se pojavljuju u okviru poslovanja kompanije, imaju sve veću potrebu za primjenom metoda rukovanja sadržajem kako bi organizirali zloupotrebu podataka i dobili informacije kako bi preživjeli unutar izložbenog prostora ili se suočili s konkurencijom na sljedećem nivou (Hassani *et al.*, 2020).

Takvo istraživanje angažuje pojedince da traže gotovo značajne informacije iz karakterističnih izvora koji uključuju različite vrste podataka – organizirane i nestrukturirane podatke. Uprkos činjenici da se konvencionalno ispitivanje podataka u osnovi fokusira na organizovane podatke koji se nalaze u bazama podataka u okviru traženih tabela i brojeva, dodatno je zamislivo analizirati štampane podatke u okviru zapisa, e-mailova ili drugih izvora nestrukturiranih podataka.

Kasnije, dugo vremena, došlo je i do povećanja količine dostupnih kompjuteriziranih književnih informacija, dajući neiskorištena iskustva koja otvaraju otvore za raspitivanje kroz neiskorištene kanale. U ovom brzom evoluirajućem području tehnika analize velikih podataka, obrada tekstova je dobila veliku pažnju u širokom spektru primjena. U svim oblastima postignut je napredak ka istraživanju poduhvata i složenijih istražnih pitanja koja zahtijevaju jednostavniji oporavak informacija. Zbog sve većeg uticaja lažnih uvida i njihove primjene na kompjuterizovanim fazama, ključna primjena je paralelno rukovanje, duboko učenje i priznanje dizajna na štampanim podacima. Svi oblici trgovačkih modela, raspitivanje o reklamama, izlaganje planova, političke kampanje ili vitalno donošenje odluka suočavaju se sa sve izraženijim potrebama za strategijama rukovanja riječima, tako da se kompanije mogu cjenkati s konkurencijom u okviru oglašavanja (Hassani *et al.*, 2020).

Velike količine tekstualnih podataka mogu se prikupiti kao dio istraživanja, kao što su naučna literatura, transkripti u marketingu, ekonomskim sektorima i mnogim drugim sektorima. Shodno tome, online izvori poput e-pošte, web stranica, blogova, objava na društvenim mrežama pružaju bogat izvor tekstualnih podataka za istraživanje.

Masovne količine podataka se također prikupljaju u polustrukturiranom obliku, poput dnevničkih datoteka koje sadrže bitne informacije. Samim tim, analiza obrade tekstova korisna je za nestrukturirane i polustrukturirane tekstualne podatke.

Data mining i text mining se razlikuju po vrsti podataka sa kojima upravljaju. Data mining se bavi strukturiranim podacima koji dolaze iz sistema kao što su baze podataka, proračunarske tablice, ERP, CRM i računovodstvene aplikacije, a text mining se bavi nestrukturiranim podacima koje se nalaze u dokumentima, e-pošti, društvenim medijima i webu (Bach *et al.*, 2019). Jedna od razlika jeste u tome što se u text miningu obrazci izvlače iz prirodnog jezičkog teksta umjesto iz strukturiranih baza podataka. Sve pisane ili izgovorene informacije mogu se predstaviti u tekstualnom obliku, shodno tome data mining zahtjeva sve vrste alata za text mining, kada je riječ o tumačenju i analizi rečenica, riječi, fraza, govora, tvrdnji, reklama i izjava (Hearst, 2012). Rudarenje podataka i analiza podataka su ključni koraci za kompanije u svakom projektu zasnovanom na odlukama temeljenim na podacima. Danas, analiza podataka i razvoj strategija igraju važnu ulogu u prikupljanju važnih informacija iz dostupnih skupova podataka (*Difference Between Data Mining and Data Analysis - Javatpoint*, n.d.).

Tabela 1 Razlike rudarenja i analize podataka

Rudarenje podataka	Analiza podataka
Rudarenje podataka je proces koji omogućava izdvajanje korisnih i bitnih informacija, obrazaca i trendova iz masovnih podataka.	Analiza podataka predstavlja metodu koja se koristi ne samo za istraživanje, nego i analizu i prikazivanje podataka zbog pronalaženja korisnih informacija koje su potrebne za donošenje odluka .
Rezultat rudarenja podataka daje uzorak podataka.	Rezultat analize podataka je potvrđena hipoteza ili uvidi zasnovani na podacima.
Uključuje bazu podataka, mašinsko učenje i statistiku.	Zahtjeva stručnost ne samo u računarstvu, nego i u matematici, statistici i umjetnoj inteligenciji.
Drugi naziv je KDD.	Različite vrste – analiza teksta, prediktivna analiza, rudarenje podataka, itd.
Odgovorno je za izvlačenje korisnih uzoraka i trendova u podacima.	Odgovorno je za razvoj modela, testiranje i predlaganje hipoteza koristeći analitičke metode.
Kao primjer rudarenja podataka navodi se sektor elektronske trgovine u kompanijama, gdje web stranice istih prikazuju opcije onih koji su kupili i pregledali određeni proizvod.	Kao primjer analize podataka navodi se popis stanovništva.

Izvor: Difference Between Data Mining and Data Analysis - Javatpoint

2.2. Važnost podataka u poslovanju

2.2.1. Prednosti i nedostaci tehnika rudarenja poslovnih podataka

Integracija informacija predstavlja način kombinovanja informacija iz datih izvora ili izvora informacija u jedinstvenu, kreirajući cjelokupnu sliku.

Za sređivanje informacija i izvođenje ovih strategija integracije informacija koriste se različiti neposlušni koji vrše prilagođavanje, kombinovanje i slično. Na osnovu toga, pravovremene i sadržajne informacije su dostupne kompaniji sa svih gledišta trgovine i administracije, kako bi ostala konkurentna i relevantna. Tri procesa manipulisanja podacima koja obuhvataju integraciju, otkrivanja znanja na osnovu rudarenja i formiranje podrške u odlučivanju se može grupisati u tri različite skupine, a to su:

- Dizajn i razvoj informacionih sistema
- Razvoj tehnologija integracije podataka
- Funkcije podrške odlučivanju (Kauffmann *et al.*, 2019).

Metode integracije podataka obično se definišu kao trojstvo GSM odnosno globalna šema (G), heterogeni izvori šema (S) i mapiranje među izvorima i kverijima (M).

Implementacija informacijskog sistema može biti skupa i vremenski zahtjevna, te samim tim zahtjeva redovno održavanje i nadogradnju.

Integracija podataka također podrazumijeva kombinaciju podataka iz različitih izvora sa ciljem kreiranja šire slike poslovanja, koristeći korporativnu inteligenciju i analitike podataka na osnovu kojih se vrši simulacija isporuka informacija, a dovodeći ih u centar za sve poslove izvještavanja, predikcije ili drugih procesa unutar organizacije.

Osim boljeg uvida u poslovanje, kompanijama su raspoloživi i ključni pokazatelji uspješnosti.

Kompanije moraju okarakterizirati solidnu tehniku integracije informacija, kako bi shvatile specijalizirane izazove koji se u osnovi pojavljuju na osnovu karakterističnih informacijskih struktura. Istovremeno, neophodno je karakterizirati vitalne sigurnosne i sigurnosne administrativne sisteme kako bi se zadovoljile administrativne potrebe, ali i kako bi se osigurali tajni podaci prilikom prijenosa.

Funkcije za podršku odlučivanju obezbjeđuju organizaciji kapacitet da brzo i produktivno analizira i dešifruje informacije. Ovaj pristup osnažuje skrivene dizajne i obrasce koji se mogu pronaći, vraćajući obrazovanom donošenju izbora. Korištenje podrške za izbor može ponuditi pomoć pri donošenju superiornih izbora na osnovu istina i informacija. U svakom slučaju, oni zahtijevaju napredno rukovanje informacijama i mjerljive istražne procedure.

Tu je i izazov prevođenja i nadgledanja velikih količina informacija. Stoga je od suštinskog značaja kontinuiran napredak u oblasti rudarenja i istraživanja informacija.

Istraživanja informacionih sistema fokusira se na zahtjeve i pristupe dizajnu kao okvir za razvoj sistema i njihovu procjenu. Postojeća istraživanja već su ponudila rješenja za informacione

sisteme, kao što autori Turoff *et al.*, 2004 navodi da su predstavljeni funkcionalni zahtjevi za dizajn informacijskim sistema za upravljanje hitnim odgovorima, kao i okvir konceptualnog dizajna (Alles *et al.*, 2004).

Autori Jain i Mclean, 2006 predstavljaju okvir koji olakšava primjenu modeliranja i simulacije u upravljanju incidentima i shodno tome temelji se na tri ključne dimenzije – incidentu, domeni i fazi životnog ciklusa (Jain i McLean, 2006).

Model dizajna sistema za podršku odlučivanja (DSS) namijenjenog kritičnim situacijama, autori Klashner i Sabet, 2007 ističu potrebu za širom i integriranijom primjenom različitih teorija i pristupa prilikom dizajniranja DSS-a za složene domene (Klashner i Sabet, 2007).

Autori Chen *et al.*, 2007 predstavljaju set principa dizajna za razvoj informacijskih sistema za upravljanje incidentima, koji se samim tim temelje na konceptima upravljanja vanrednim situacijama i uvidima prvih ispitanika. Tehnologije integracije podataka je oblast koja je ostvarila značajan napredak u razvoju algoritma i alata. Pristup se fokusira na analizu domena, razvoj reprezentativnih skupova podataka za procjenu informacijskih tehnologija i strategija (Chen *et al.*, 2007).

Visok nivo mogućnosti integracije i nadzora se prati za specifične sekundarne incidente. Kroz historiju, funkcionalne podrške odlučivanju za upravljanje incidentima su razvile niz odgovarajućih rješenja za identifikaciju prevelikih (abnormalnih) situacija korišćenjem tehnika Data mininga.

Sistemi mogu generisati alternative i pomoći donosiocu odluka da bolje razumije problem i izabere najbolju alternativu za dobrobit kompanije. Smatra se da strategije skladištenja informacija i pripreme ekspozitorija u realnom vremenu (OLAP) omogućavaju brzo ispitivanje vanrednih okolnosti i daju smjernice za podsticanje istraživanja o korištenju informacija (Peng *et al.*, 2011).

Donošenje odluka u okolnostima kao što su hitni slučajevi je od izuzetnog vremenskog značaja i može biti izazovan zadatak koji zahteva brzu i isplativu reakciju odabranih proizvođača ispod težine i u uslovima bespomoćnosti unutar kompanija.

Integracija podataka igra ključnu ulogu kao razvoj koji omogućava efikasno prikupljanje, dijeljenje, širenje, rukovanje i ispitivanje podataka o prilikama. Ove informacije su nevjerojatno imperativne da pomognu proizvođačima izbora u otvaranju i prilagođavanju izbora u vanrednim okolnostima. Upravljanje ogromnom količinom i složnošću podataka o događajima predstavlja izazov u integraciji informacija (Ludäscher *et al.*, n.d.).

2.2.2. Segmentacija dobivenih podataka

U industriji, funkcionalni koncept segmentacije podataka postaje sve važniji. Kompanije otkrivaju da je od vitalnog značaja da prenose informacije koristeći karakteristične kriterijume i da ih organizuju u određene kategorije kako bi podstakle njihovo korišćenje u aktivnostima predstavljanja i dogovora.

Podjela informacija osnažuje prikupljanje, sposobnost i integraciju informacija o klijentima, na taj način je od vitalnog značaja napraviti tačne profile savršenih klijenata koji su pretpostavka tehnika izlaganja.

U svakodnevnicima različitih sektora, informacije igraju ključnu ulogu u odabiru koga kontaktirati i kako navesti na kontakt. Poznavanje najvažnijih imperativnih klijenata, njihovog sučelja i omiljenih oblika komunikacije osnažuje predstavnike da bolje iskoriste svoje vrijeme. Nakon toga, nastanak podjele informacija može se koristiti za procjenu potencijalnih klijenata.

Segmentacija podataka ima i značajnu ulogu u *cyber* sigurnosti. Kroz ograničavanje pristupa podatkovnim skladištima samo određenim segmentima mreže, stvara se dodatna sigurnosna barijera. S manjim brojem segmenata koji imaju pristup podacima, hakerima će biti teže pristupiti i izvući vrijedne informacije. Ograničavanjem pristupa i implementiranjem lokalnih sigurnosnih politika, smanjuje se rizik gubitka i krađe podataka.

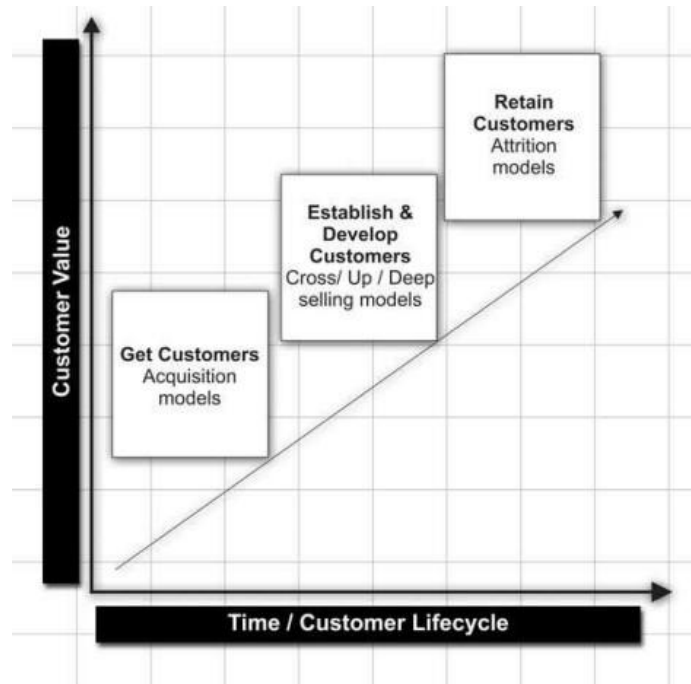
Važno je shvatiti da se sve u procesu segmentacije podataka svodi na određivanje koje vrste podataka u publici su zaista relevantne za određenu kompaniju i kako se namjeravaju koristiti. Shodno tome, pojedine kompanije imaju koristi od nekih ili svih pristupa segmentacije podataka, samim tim važno je temeljito razumjeti svrhu i posljedice svakog pristupa prije nego što se donese odluka za njegovu primjenu.

Podaci se kategoriziraju ne samo prema slučajevima upotrebe i vrstama informacija, nego i prema osjetljivosti informacija i razini ovlaštenih potrebnih za pristup. Cilj segmentacije podataka je kategorizirati podatke, izdvojiti najkritičnije podatke od ostatka i označiti ih kao površinu zaštite, te zatim implementirati dodatne sigurnosne mjere oko tih zaštita koje su se identificirale (*What Is Data Segmentation? The Complete Guide | Just Total Tech*, n.d.).

Postoje mnogo različitih tipova segmentacije temeljenih na specifičnim kriterijima ili atributima koji se koriste za segmentaciju. U segmentaciji ponašanja, kupci se grupiraju prema ponašanjima i korisničkim karakteristikama. Također, rudarenje podataka može stvoriti segmente ponašanja temeljene na podacima. Algoritmi grupiranja mogu analizirati ponašanje podataka, identificirati prirodne grupiranja kupaca i predložiti rješenje temeljeno na promatranim uzorcima podataka.

Samim tim, ako su modeli rudarenja podataka pravilno izgrađeni, mogu se otkriti grupe s različitim profilima i karakteristikama te dovesti do bogatih shema segmentacije s poslovnim značajem i vrijednošću. Rudarenje podataka se može koristiti za razvoj shema segmentacije temeljenih na trenutnoj ili očekivanoj vrijednosti kupaca. Ovi segmenti su nužni zbog prioritizacije obrade kupaca i marketinških intervencija prema važnosti svakog kupca / klijenta (Tsipstsis i Chorianopoulos, 2009).

Ilustracija 10 Rudarenje podataka i upravljanje životnim vijekom kupca



Izvor: Tsipsis i Chorianopoulos, 2009

Uloga modela rudarenja u marketingu je prilično nova. Iako se brzo širi, rudarenje podataka je još uvijek „nepoznato područje“ za marketinške stručnjake koji se oslanjaju na intuiciju i iskustvo u industriji. Njihove sheme segmentacije i liste marketinških akcija stvaraju se na temelju poslovnih pravila koja se temelji na njihovim poslovnim znanjima.

Različiti marketinški stručnjaci kampanje koordiniranih ponuda kako bi prenijeli poruku svojim klijentima putem pošte, weba, telemarketinga i drugih koordinatnih kanala kako bi predvidjeli odljev klijenata i osnažili moderne kupce da kupe dodatne artikle.

Napredak posebno ukazuje na privlačenje neiskorištenih i možda važnih klijenata od konkurenata. Napredovanja se sprovode u cilju ponude dodatnih artikala, više istih artikala, i ukazuju na predviđanje važnih klijenata od prekida odnosa sa kompanijama.

Na taj način, kako bi se izbjeglo gigantsko rasipanje imovine i iritiranje klijenata neželjenom komunikacijom, u tom trenutku dolazi do izražaja značaj informatičkog rudarenja i modela klasifikacije.

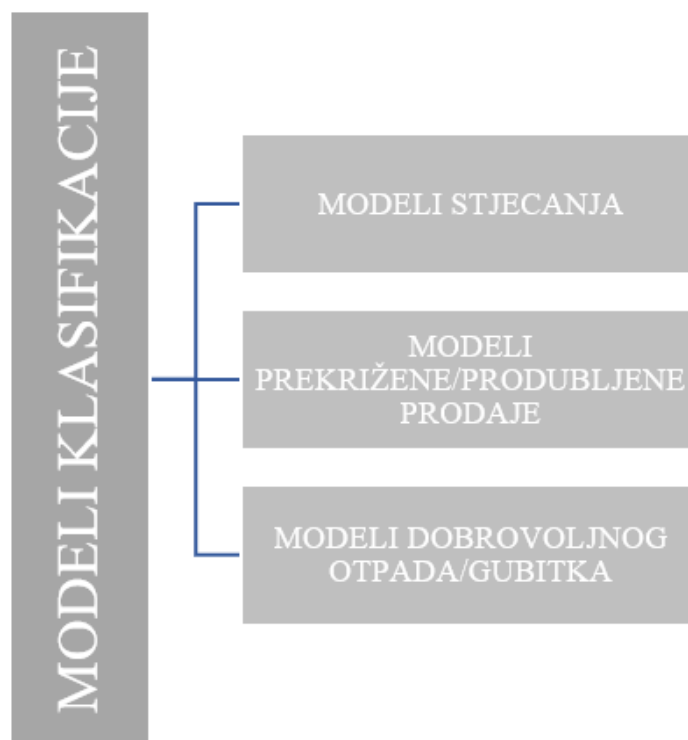
Oni osnažuju ispitivanje karakteristika kupaca i razlikovanje profila ciljnih klijenata, kao i optimizaciju marketinških akcija u cilju postizanja što veće stručnosti. Kroz informatičko rudarenje i primjenu modela klasifikacije, aktivnosti prikazivanja su na kraju bile fokusirane, personalizirane i uspješne, što na neki način dolazi do boljeg angažmana klijenata i ostvarivanja ciljeva promocije kompanije.

Sljedeći modeli klasifikacije koriste se za optimizaciju kasnijih marketinških akcija:

1. Modeli stjecanja – mogu se koristiti za prepoznavanje potencijalno profitabilnih kupaca na vanjskim listama kontakata.
2. Modeli prekrížene/produbljene prodaje – otkrivaju potencijal za kupovinu kod postojećih kupaca/klijenata.
3. Modeli dobrovoljnog otpada ili dobrovoljnog gubitka – identificiraju rane signale odljeva i prepoznaju kupce s povećanom vrijednošću da će dobrovoljno napustiti.

Također, ovi modeli mogu prepoznati prave kupce/klijente za kontaktiranje i rezultirati listama akcija s povećanom gustoćom ciljnih kupaca.

Ilustracija 11 Modeli klasifikacije



Izvor: Autor

2.2.3. Korisnost analize podataka za poslovanje kompanije

Rudarenje podataka predstavlja proces koji sadrži analizu ogromnih podataka zbog pronalaska korisnih znanja ili same informacije. Svakako da rudarenje podataka kroz različite tehnike i algoritme pruža mogućnost da se otkriju i skriveni uzorci i donošenja informiranih odluka u budućnosti (Zaki i Meira, n.d.).

U današnjici, neprestano se povećava rast podataka, a procesi rudarenja podataka su od važnosti kako bi se mogli razumjeti i klijenti i njihova ponašanja, te rudarenje podataka se smatra kao dinamičnim poljem koji se jako brzo razvija pružajući jake alate za daljnja istraživanja i donošenja bržih odluka (Han, Kamber, i Pei, 2012).

Poslovna inteligencija

Veoma je važno da preduzeća postignu veća razumijevanja u svom poslovnom okruženju, ali samim tim se to odnosi i na uključivanje i kupaca, tržišta, resursa, konkurencije i sl. Trenutne i prediktivne informacije o poslovanju preduzeća tehnologije poslovne inteligencije pruža mogućnost prikazivanja prošlih, sadašnjih i prediktivnih informacija kroz tehnologije poslovne inteligencije.

Navode se primjeri tehnologija, a to su analitička obrada u realnom vremenu, te kao i upravljanje rezultatima koji su dobiveni izvještaji, konkurentske informacije i sl.

Rudarenje podataka ima glavni i ključni zadatak u procesu jer pruža tu mogućnost efikasnosti analize tržišta, uporedba povratnih informacija o kupcima i o sličnim proizvodima, te otkrivanje prednosti i slabosti konkurencije, čuvanje bitnijih kupaca i donošenja odluka koja su bitne za budućnost. Shodno tome, Data mining je ključna komponenta poslovne inteligencije. Alati za analitičku obradu u stvarnom vremenu u poslovnoj inteligenciji temelje se na metodama Data mininga, a ključnu ulogu imaju metode klasifikacije i klasterizacije.

Web pretraživači

Web pretraživač je specijalizirani server namjenjen pretraživanju informacija na internetu. Rezultat pretraživanja se prikazuje kao lista koja sadrži različite vrste sadržaja, kao što su web stranice, slike i druge datoteke. Web pretraživači su kompleksne aplikacije za obradu podataka, a upotreba tehnika Data mininga je ključna u svim fazama pretraživanja. Samim tim, predstavljaju puno izazova, a kao prvo moraju se nositi s ogromnom i rastućom količinom podataka.

Pošto su pretraživači koristili računarske oblake sa puno čak i stotinama hiljada računara koji zajedničkim snagama obrađuju ogromne količine podataka, nije moguće da obrade samo na jednu ili nekoliko mašina ili računara.

Shodno tome, web pretraživač predstavlja da se koristi mrežnim podacima, te da bi se nosili sa tim, koriste se klasifikatori upita koji se kategoriziraju prema unaprijed definisanim kategorijama na temelju teme upita. Također, oni se susreću s izazovom rješavanja upita koji se postavljaju.

Drugim riječima, kada se korisniku postavi upit, pretraživač nastoji zaključiti kontekst tog upita koristeći korisnički profil i historiju pretraživanja kako bi se u kratkom vremenskom intervalu dobili prilagođavajući odgovori.

Korištenje Data mininga

Konkurencija je izuzetno jaka u današnjem poslovnom okruženju u mnogim sektorima i kompanijama, što samim tim se dovodi do promjena u načinu poslovanja s klijentima. Data mining se primjenjuje u nekoliko trendova, a to su:

Istraživanje aplikacija – Aplikacije za primjenu Data mininga ulažu značajne napore kako bi pomogli kompanijama da ostvare konkurentsku prednost. Istraživanje Data mininga za poslovne svrhe nastavlja se širiti jako brzo, posebno s porastom e-trgovine i e-marketinga koji su postali ključni u maloprodajnoj industriji. Data mining se sve više koristi i istražuje u drugim područjima kao što su analize weba i teksta, finansijske analize, industrije, nauke i dr.

Skalabilne i interaktivne metode Data mininga – Data mining se razlikuje od tradicionalnih metoda analize podataka jer mora efikasno i, po mogućnosti, interaktivno obraditi ogromne količine podataka. S obzirom na brzinu rasta količine prikupljenih podataka, postaju ključni skalabilni algoritmi za pojedinačne i integrirane funkcije Data mininga. Rudarstvo zasnovano na ograničenjima govori o jednom od ključnih pokretača koji ima za cilj postizanje iskoraka u opštoj produktivnosti rukovanja rudarenjem informacija, istovremeno osnažujući istaknutiju interakciju sa klijentima.

Integracija informacionog rudarenja sa motorima izgleda - U osnovi garantuje da rudarenje informacija ne igra ključnu ulogu u istraživanju informacija, već je kritična komponenta planiranja u okviru opsežnijeg okruženja za planiranje informacija. Također, ova integracija osnažuje interakciju između informacijskog rudarenja podataka i drugih napretka tehnologija i alata, te na taj način postiže visoku produktivnost i maksimalno iskorištenje dostupnih resursa. Rudarstvo zasnovano na ograničenjima govori o jednom od ključnih usmjerenja koji su usmjereni ka unapređenju opće vještine pripreme rudarenja informacija, istovremeno osnažujući značajniju interakciju s klijentom.

Rudarstvo društvenih i informacionih mreža – Osnovno je garantovati da rudarenje informacija ne igra ključnu ulogu u istraživanju informacija, već govori o kritičnoj komponenti planiranja u okviru opsežnijeg okruženja uređenja informacija. Štaviše, ova integracija jača interakciju između data mininga i drugih pokretača napretka i aparata, postizujući nakon toga visoku efikasnost i najekstremnije korištenje odgovarajuće imovine. Informaciono rudarenje može biti svestrano, tako da se na isti način može na odgovarajući način prilagoditi karakterističnim poslovnim potrebama i promjenama u kompanijama i u okruženju.

Mining društvenih sistema i sistema podataka - Govori o osnovnim zadacima u oblasti rudarenja informacija, koji su ključni zbog sve veće blizine i stalnog razvoja kompleksnosti ovakvih sistema. U današnjem kompjuterizovanom dobu, društveni sistemi i sistemi podataka progresivno pokazuju i uporno stvaraju ogromne količine informacija.

Intuitivno rudarenje podataka o medijima, supstanci i webu – U kasnijem dugom vremenu, informaciono rudarenje se usredsredilo na rudarenje različitih vrsta podataka, računajući društvene i informacione okvire. Danas, ovi okviri dinamično postaju sve složeniji u prisutnosti. Iskopavanje podataka iz toka akcije i analiza njegovih veza ima potencijal za otkrivanje važnih informacija, prepoznavanje planova i donošenje obrazovanih izbora.

Unatoč činjenici da je u ovoj regiji postignut veliki napredak, još uvijek postoje brojna otvorena pitanja koja zahtijevaju ispitivanje i pomoć u poboljšanju.

Data mining softverskim inženjeringom i sistemskim inženjeringom – Dimenzije softverskih programa i računarskih sistema neprestano rastu, postajući sve veći i složeniji. Samim tim, ova tendencija stvara sve veći izazov u osiguravanju robusnosti i pouzdanosti softvera. Shodno tome, očekuje se da će daljnji razvoj Data mining metoda za otklanjanje grešaka u softveru i sistemima poboljšati robusnost softvera i donijeti novo znanje u području softverskog i sistemskog inženjeringa (Tuffery, 2008).

Sama robusnost softvera podrazumijeva sposobnost softverskog sistema da se nosi s neočekivanim ili neispravnim ulazima, pogreškama ili promjenama u okolini, te da nastavi pravilno funkcionirati i pružiti očekivane rezultate. Odnosi se na stabilnost, otpornost i pouzdanost softvera u različitim situacijama.

Robusnost softvera je važna jer korisnici često koriste softvere na različitim platformama, s različitim ulaznim podacima i u različitim uvjetima. Softver treba biti dizajniran i implementiran na način koji minimizira mogućnost rušenja, pogrešnih rezultata. Robusnost se postiže provjerom, validacijom i testiranjem softvera na različitim scenarijima i ulazima, kako bi se osigurao da se softver može nositi s različitim situacijama i iznimkama.

Informaciono rudarenje za CRM - Informaciono rudarenje progresivno preovlađuje u brojnim kompanijama u organizovanju napredovanja administracije odnosa sa klijentima (CRM) i osnaživanju uređenja personalizovanih i prilagođenih administracija koje zadovoljavaju posebne potrebe svake osobe klijenta, ograničavajući pristup masovnog predstavljanja. Ispitivanjem dizajna pregledavanja i kupovine u web prodavnici, kompanije mogu prilagoditi reklame i promocije prema profilima klijenata, tako da je manja vjerovatnoća da će klijenti biti iritirani neželjenom poštom.

Na ovaj način aktivnosti mogu rezultirati značajnim investicionim fondovima u kompanije, smanjenjem troškova objavljivanja i povećanjem adekvatnosti promocije poduhvata. Koristeći rudarenje informacija, kompanija može izgraditi čvrste veze s klijentima, pružajući im personalizirane i relevantne podatke, što dovodi do značajnije privrženosti klijentima i proširene komercijalne realizacije.

Utjecaj istraživanja informacija osiguran je načinima na koje pojedinci koriste kompjuterske okvire, istražuju podatke i distribuiraju svoje zadatke. Za slučaj, najbrže podatke traži stranica Google (www.google.com), koja omogućava pristup podacima o milijardama snimljenih web stranica na njihovom serveru. Google je postao jedan od najpoznatijih i široko korištenih web preglednika, a njegovo korištenje za otkrivanje podataka postalo je način života za kompletnu populaciju (Han, Kamber, i Pei, 2012).

3. METODOLOGIJA ISTRAŽIVANJA

3.1. Identifikacija i selekcija izvorne literature

Korak identifikacije obuhvata pretragu istraživačkih radova koji su potencijalno relevantni za predefinisana istraživačka pitanja. Ovaj korak rezultira listom takvih radova (Booth, 2016). Ovdje se koristi pojam 'potencijalno' jer se konačna procjena odgovarajućeg sadržaja određenog istraživačkog rada za odgovor na pitanja pregleda vrši tek u fazi selekcije.

Najpopularnije baze podataka literature u polju se pretražuju, kao što su:

1. ACM Digital Library (dl.acm.org)
2. IEEE eXplore (ieeexplore.ieee.org)
3. ScienceDirect (sciencedirect.com)
4. Springer (springerlink.com)
5. Scopus (scopus.com)

Kreatori anketa mogu okarakterizirati kriterije bez sadržaja kako bi ograničili prepoznatljiv dokaz o eventualno značajnim istraživačkim radovima na određene vrste distribucije, određeni vremenski period osiguran revizijom ili način na koji se pronalazi upit o papirima. Redovne revizije se zasnivaju na izgledu elektronskih baza podataka kako bi se dobio sveobuhvatan revizorski test na organizovan i jednostavan način (Tranfield *et al.*, 2003).

Postoje prijedlozi u upitniku uprave i srodnim oblastima koje baze podataka odabrati za ove svrhe.

Autori iz drugih područja općenito predlažu minimalno korištenje dvije baze podataka kako bi se ublažili učinci različitih pokrivenosti među pojedinim bazama podataka (Green i Glasgow, 2006). Drugi autori zaključuju da istraživači često koriste premalo baza podataka, što može ugroziti generalizovanje i valjanost rezultata pregleda (Howard *et al.*, 2017).

Iako ovi uvidi mogu pružiti smjernice istraživačima u menadžmentu o izboru baza podataka u sistematskim pregledima, problem je što izbor baza podataka uveliko ovisi o proučavanoj temi. Stoga, nedostaje strukturirani pregled uobičajenih baza podataka koje se koriste u sistematskim pregledima istraživanja u području menadžmenta.

U slučaju da izgled baze podataka pokriva adekvatnu širinu i dubinu, izgledi zasnovani na ključnim riječima mogu propustiti istraživanje radova koji su važni za osobu koja postavlja pitanja o postojećim pitanjima. Na primjer, uprkos činjenici da se dnevnički članci obično lako prepoznaju kroz pregled baze podataka, drugi upiti o papirima kao što je sivo pisanje redovito i nisu prepoznatljivi (Adams *et al.*, 2017). Brojne postojeće uredne ankete su na ovaj način izbjegle tako sivo pisanje u svojim metodama utvrđivanja testova, uprkos činjenici da takvo izbjegavanje u osnovi dovodi do manjih anketnih testova i eventualno poziva na razmatranje sveobuhvatnosti tih anketa.

Bilo kako bilo, izgled zasnovan na ključnim riječima i dalje može propustiti moguće relevantne upite o papirima. Obrnuti pogledi mogu biti dopunjeni pogledima naprijed, koji reviziju vrše

kasnije istraživajući radove koji citiraju prvo priznate radove (Webster i Watson, 2002). U tu svrhu se također mogu koristiti elektroničke baze podataka poput Web of Science ili Google Scholar, što čini tehnike snježnog lopatanja koristima u radu na sveobuhvatnom pregledu.

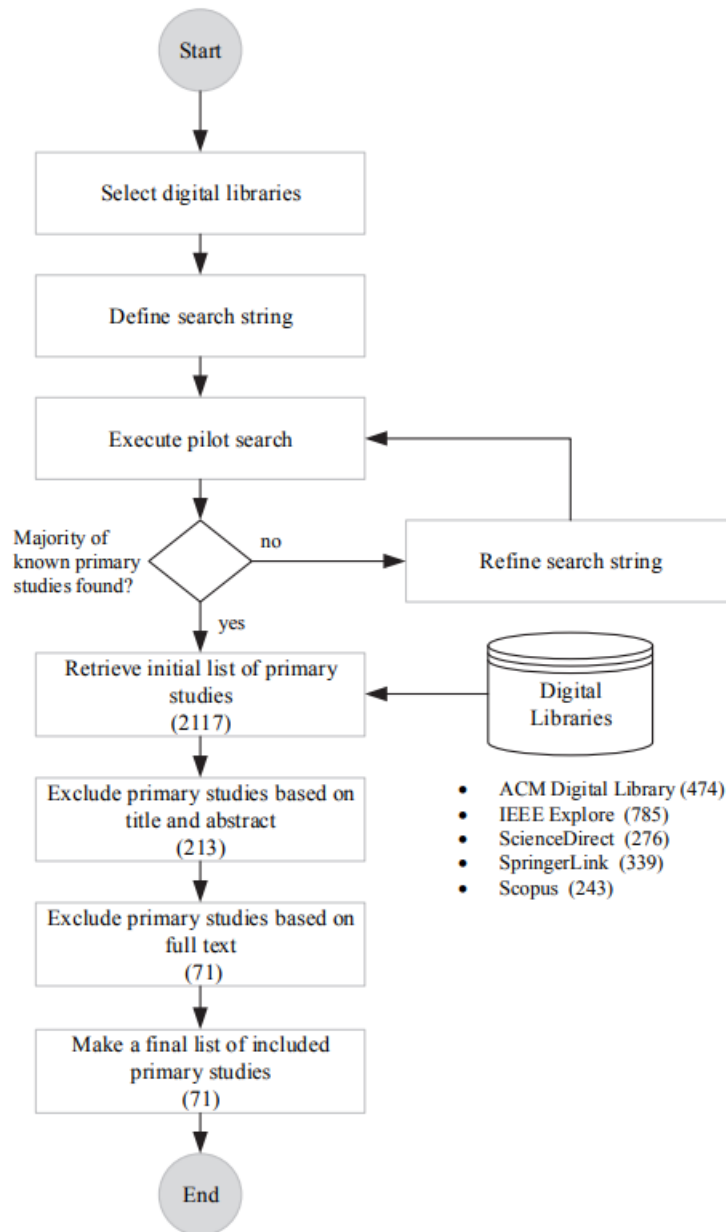
Unutar faze utvrđivanja, analizira se lista potencijalno važnih istraživačkih radova kako bi se odlučilo da li se upoređuje s unaprijed definiranim istraživanjem (Waddington *et al.*, 2012).

Kriterijumi koji se često povezuju u fazi izbora uključuju određivanje upita o papirima s obzirom na određene planove razmatranja ili upit o provedenim na određenim populacijama (Booth, 2016). Kriterijumi unaprijed, imperativnih određivanja su evaluacije kvaliteta (Briner i Denyer, 2012).

Slažući se sa Tranfieldom i dr., „osoba o kojoj razmišlja u urednoj reviziji se ispituje u odnosu na unaprijed određene kriterije i kontrolne liste“ kako bi se ocijenilo da li su odgovarajućeg kvaliteta da budu uključene u test revizije. Tranfield *et al.* prepoznali su da se u administrativnim upitima kvalitet uvodnika po pravilu ispituje na osnovu kvaliteta dnevnika u kojem je članak distribuiran. Da bi se izbjegle takve procjene kvaliteta zasnovane na reviziji dnevnika, Tranfield *et al.* dao je listu kriterija za ocjenu kvaliteta komentara, ali je zaključio da u urednim revizijama ispitivanja u administraciji ocjena kvaliteta ostaje "veliki izazov".

Uglavnom, budući da istražuje kvalitet može biti uobičajena tema o kojoj se govori i sporna tačka u administrativnoj istrazi, a najvećim dijelom priznati uvid u ono što čini ispitivanje kvaliteta još uvijek nije iznenađujući.

Ilustracija 12 Pretraga i selekcija literature



Izvor: Tranfield et al., 2003

Općenito, kriteriji koji su vezani za sadržaj i uključivanje ili isključivanje, kao što su procjene kvalitete, nose problem da se mogu primjenjivati različito među autorima sistematskih pregleda. Posebno tijekom faze selekcije, stoga može biti korisno da više recenzenta provodi analizu potencijalno relevantnih istraživačkih radova i da recenzenti raspravljaju o neslaganjima unutar šireg tima za pregled, što bi dodatno doprinijelo strukturiranom postupku odabira uzorka (Tranfield et al., 2003).

Unatoč takvim mjerama, različiti epistemološki normativi u istraživanjima menadžmenta i subjektivne procjene da li članak ispunjava ili ne ispunjava određene kriterije vezane za sadržaj ometaju reproducibilnost sistematskih pregleda istraživanja u menadžmentu.

Dakle, iako potpuna repliciranost možda nije ostvariva, zagovornici sistematskih pregleda u istraživanjima menadžmenta tvrde da izvještavanje o postupku pregleda treba biti što transparentnije moguće.

Kako bi se pohranili istraživački radovi, korišten je Mendeley. Proces odabira studija je proveden u dva koraka: isključivanje primarnih studija na osnovu naslova i sažetka, te isključivanje primarnih studija na osnovu punog teksta. Pregledne studije i ostale studije koje ne sadrže eksperimentalne rezultate, kao irelevantna istraživanja su isključene. Također je uzet u obzir stepen sličnosti studije sa predviđanjem grešaka u softveru kao kriterij za uključivanje studija.

Svaka systemska analiza literature treba da ima jasnu tačku fokusa na dokaze koji su objavljeni o određenoj temi ili pitanjima. Takvo jasno pitanje koje se fokusira na intervencije, mehanizme i ishode u određenom kontekstu ima prednosti za pregled, jer određuje jasne granice.

Kako bi pomogao u oblikovanju revizije, istraživač će možda htjeti napraviti istraživačku grupu ili, s druge strane, uključiti pomoć kolega ili savjetnika da usmjere obim ankete pisanja, pripremu revizije i triangulaciju ključnih izbora, ističu autori Briner i Denyer, 2012. Uprkos činjenici da efikasne revizije može da obavi osoba koja vrši ispitivanje, postoji opasnost da vitalne komponente (npr. izbor kriterijuma za uključivanje/isključivanje) ne budu predmet revizije, što može dovesti do predispozicije. Upit o grupi može ponuditi pomoć u prevazilaženju takvih prepreka, za ilustraciju, brojeći stručnjake za efikasne revizije, specijaliste za predmet/sadržaj i moguće stručnjake za metodologiju.

Trebalo bi izvršiti uvodni pregled kako bi se dobio pripremni dijagram trenutnog stanja istrage i zaista bi moglo biti dragocjeno za stručnjake u određenoj oblasti da naprave sistem za preciznu reviziju. Ovo istraživanje može uključiti široki izgled pisanja kako bi se procijenile odgovarajuće tehnike izgleda (npr. odgovarajuće baze podataka/izvori, vremenski periodi, termini/ključne riječi izgleda, dijalekatske prepreke) i dobili zajednički dijagram pisanja.

Istraživač bi trebao istražiti: Da li već postoje pregledi o specifičnom pitanju ili istraživačkom pitanju koje se istražuje?

U tom slučaju, repliciranje možda nije opravdano, ali proširenje ili ažuriranje bi moglo biti korisno. Da li ima mnogo objavljenih članaka o temi? Ako da, može biti plodotvornije usredotočiti se na literaturu objavljenu u posljednjim godinama (npr. od određenog datuma), u određenim časopisima ili u uže definiranom polju.

U svakom slučaju, u slučaju da je ovo područje razumno moderno, možda neće biti brojnih recenziranih naučnih dnevničkih članaka dostupnih za reviziju, ali mogu postojati referati sa konferencije, radovi unaprijed ili drugi oblici "mračnog" pisanja koji treba uzeti u obzir. Zbog toga se često postavlja pitanje da li je dostupno pisanje uglavnom distribuirano u jednoj oblasti, ili postoje fantastični upiti o pitanjima ili problemima i u drugim oblastima.

Analitičar će možda morati da razmotri intrigirane obaveze kako bi prevazišao opasnost od izgubljenih vitalnih obaveza izvan „standardnih“ dnevnika, smatraju autori Adams *et al.*, 2017. Posebna područja stavljaju poseban naglasak na distribuciju u knjigama, poglavljima u

knjigama ili zbornicima konferencija, tako da nekoliko ključnih radova možda neće biti distribuirano u recenziranim dnevnicima.

Još jedan kritičan korak je odabir kriterija za inkorporaciju i zabranu. Analitičar ne mora uključiti bilo kakve nepravilne informacije u preciznu reviziju; stoga je imperativ razmotriti koji će koraci biti poduzeti da se otkriju značajna razmišljanja o tome. Uobičajeni pristup je razlikovanje pisanja za razmatranje putem Booleovog pregleda unutar izgrađenih lista izgleda, kao što su Internet of Science ili Scopus faze.

Ova baza podataka omogućava da se istraži distribucija sa unapred definisanim frazama, za mala i mala slova, unutar naslova, sažetaka ili ključnih riječi. Bilo kako bilo, nedavno poduzimajući ovaj korak, analitičar bi trebao pažljivo odabrati metodologiju izgleda, izbor ili kombinaciju ključnih riječi i baze podataka, uključivanje/isključivanje radova iz drugih disciplina (brojne baze podataka dozvoljavaju analitičaru ograničiti izgled na poslovanje/menadžment, ali to bi moglo propustiti važne obaveze), kao i uključivanje/isključivanje referata sa konferencije, knjiga, poglavlja u knjigama, izvještaja i drugog "mutnog" pisanja. U osnovi ne postoji "ispravan" ili "pogrešan" način za formiranje tih izbora.

Da bi garantovali konzistentnost, analitičari mogu pregledati metodologije izgleda uporednih istraživanja. U ekspanziji, kako bi se održala strateška distanca od predispozicije, korake donošenja odluka treba recenzirati i prijaviti kako bi se povećala jednostavnost pripreme izgleda. Nakon identifikacije odgovarajućih studija, duplicirani zapisi trebaju biti uklonjeni iz analize. Osim toga, studije koje nisu relevantne trebaju biti odbačene.

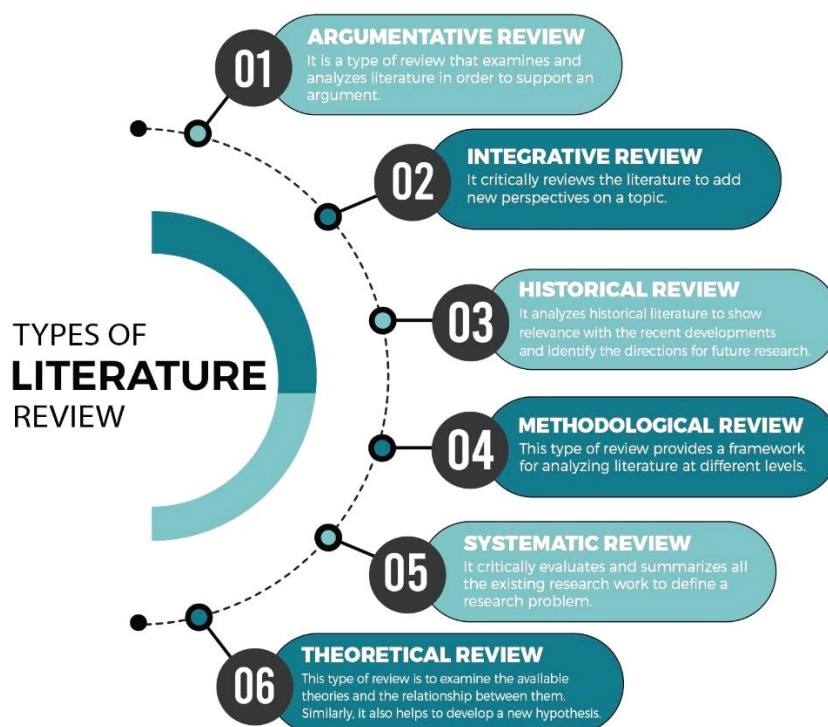
Ukoliko postoje rezultati koji ukazuju na 'lažne rezultate', odnosno uključuju moguće nesrodne teme, neophodno je isključiti ih. Umjesto toga, takve studije trebaju biti zadržane i pravilno kritizirane. Proces čišćenja podataka se sprovodi na način da dva ili više istraživača pregledaju literaturu dobivenu kroz pretraživanja te provjeravaju prikladnost, kroz provjeru naslova, sažetaka ili ključnih riječi u cilju razmatranja mogućih neslaganja.

Dodatno je moguće provjeriti citiranu literaturu odnosno reference ili drugu triangulaciju rezultata iz više baza podataka.

3.2. Bibliometrijska analiza podataka

U ovom istraživanju odabrani su pristupi efikasnog pisanja ankete i bibliometrijskog ispitivanja. Pisanje efikasnih anketa bi mogla biti preovlađujuća strategija koja uključuje oblike prepoznavanja, evaluacije i prevođenja svih dostupnih upita o dokazima za pružanje odgovora na određeni upit o pitanjima (Wahono, 2015). Sistematski pregled literature je korišten da bi se prepoznala važna razmatranja za ovu istragu i odlučila o budućem programu rada, kao i da se otkriju trenutne strategije i procedure u vezi sa ogromnim istraživanjem informacija i ispitivanjem u poslovanju (Lame, 2019).

Ilustracija 13 Vrste analize literature



Izvor: <https://myperfectwords.com/blog/research-paper-guide/how-to-write-a-literature-review>

Pored toga, koristila se bibliometrijska analiza kao tehnika koja omogućava razlikovanje dokaza istraživačkih trendova.

Bibliometrijsko istraživanje je korišćeno za razlikovanje značajnih istraživanja o radovima u oblastima ogromnog informacionog rudarenja i istraživanja u trgovini, njihovom međusobnom odnosu i ključnim parametrima za realizaciju istraživačke organizacije svih mentalnih asocijacija. Bibliometrijsko ispitivanje je završilo kao zaista dobro poznata strategija u posljednjoj deceniji budući da omogućava analizu supstanci ili citata unutar logičkog materijala i istraživanje raspona zaintrigiranih na detaljniji način. Njegov primarni cilj je opskrba znanjem iz različitih gledišta logičke generacije i komunikacije, kao i prepoznatljiv dokaz ključnih kreatora, dnevnika i raspitivanja o temama.

Na osnovu istih, koriste se ekspresivni uvidi koji omogućavaju saznanje preovlađujućih stvaralaca i dnevnika koji imaju imperativnu ulogu u određenoj oblasti istraživanja. Na osnovu istih, koriste se jasna mjerenja koja omogućavaju znanje u nadmoćnim stvaraocima i dnevnicima koji igraju kritičnu ulogu u određenom polju.

Na osnovu toga, cilj se može ostvariti na osnovu dva pristupa, kao što je subjektivni, koji se zasnivaju na rasvjetljavanju istražne zone od strane specijalista, do čega dolazi u subjektivnoj istrazi, za ilustraciju, preciznom pismenom revizijom ističu autori Tranfield *et al.*, 2003, te objektivni pristup zasnovan na bibliografskim i kvantitativnim strategijama koje su osnažene određenim programima (Di Stefano *et al.*, 2009; Pal i Gauri, 2018). Kombinacija ova dva pristupa se obično koristi da bi se dobila struktura i specifičnost istraživanja o regionu.

Efikasan pregled literature igra ključnu ulogu u istraživanju poduhvata, osnažujući analitičare da dobiju trenutno stanje informacija o određenoj tački i prepoznaju rupe u istraživanju. Potom se obraća ključnoj jedinici koja daje snažan format za podsticanje raspitivanja (Torraco, 2005).

Efikasnosti doprinosi činjenica da istraživač mora pažljivo ispitati kvalitet i relevantnost pisanja, natjerati ga da se raspita o strategiji i potencijalnim sklonostima, i fundamentalno ispitati nastanak prošlih razmatranja.

Dosadašnja praksa u mnogim pregledima literature u polju menadžmenta i srodnih oblasti je da analitičari daju plitki pregled pisanja i priče, usredsređujući se na izbor članaka koji se smatraju "visokim kvalitetom" koji se slažu sa određenim kriterijumima. Ovaj pristup može dovesti do izbora specifičnog testa pisanja, bez ubrajanja šireg niza dokaza i istražnih radova koji će biti relevantni za upit o istraživačkom pitanju.

Samim tim, takvi površni pregledi često nisu dovoljno reprezentativni za stanje postojećeg znanja. Odabir samo nekih studija, a isključivanje drugih, može dovesti do pristranosti uzorka i nepravednog isključivanja relevantnih informacija. Ovo se može nazvati sklonošću određivanju testova i može uticati na legitimnost zaključaka i dolazi do generalizacije istraživanja (Tranfield *et al.*, 2003).

Kako bi proširili informacije o određenoj tački, analitičari bi trebali biti prepoznatljivi po značajnom pisanju u tom rasponu, pa im stoga istraživanje literature omogućava da pakupe detaljnije znanje o prošlim istraživanjima, hipotetičkim konceptima, tehnikama i nastalim . Ovo osnažuje osobe da prepoznaju obrasce, rasprave, tekstore i kontraste unutar pisanja, nakon čega stvaraju opsežniju postavku za svoje tvrdnje da se raspitaju.

Citati omogućavaju provjeru kako se informacije stvaraju i šire unutar određenog učenja, a kroz bibliometrijska ispitivanja, analitičari mogu pokupiti razumijevanje u obrasce, prepoznati ključne analitičare i časopise, te stoga provjeriti utjecaj njihovog rada na nauku.

Kada analitičar doprinese postojećem istraživanju, ključna adresa postaje: „Šta je do sada poznato o temi ili pitanju koje se istražuje, i koji su produktivni otvori za prethodno ispitivanje?“

Redovno pisanje revizije redovno ima za cilj prepoznavanje osnovnih pitanja o tokovima i naslovima za buduću istragu.

Drugi slučaj podrazumijeva kada istraživač analizira neiskorištenu liniju istraživanja ili stvara novu hipotezu. U ovom slučaju, ključno pitanje je: "Šta je neiskorišteno pitanje o polju/teoriji i kako se ono odnosi na postojeća područja ili hipoteze?" U ovim okolnostima, sastavljanje studije može biti nevjerojatno korisno za istraživače koji smatraju da postoji potreba da se raspitaju o određenoj tematici i/ili području.

Unatoč činjenici da može postojati takoreći mnogo relevantnih razmišljanja, pisana anketa može poslužiti kao osnova za ispitivanje informacijskih pukotina i kako ekstremni analitičari doprinose njihovom popunjavanju (Linnenluecke, 2017).

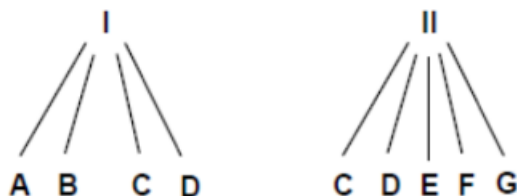
Sama bibliometrija se zasniva na ideji da nauka uključuje strukturu koja se može eksperimentalno okarakterisati kroz stepen zasnovan na odnosu između arhiva, kreatora ili dnevnika. Izvođenje bibliometrijskog pregleda uključuje nekoliko diskretnih aktivnosti, koje se mogu grupisati u pet glavnih koraka:

- Korak 1: Dizajn istraživanja - odabir istraživačkog pitanja(i) i bibliometrijske metode (npr. citiranje, kocitiranje, bibliografsko povezivanje).

U ovom koraku, istraživači odabiru istraživačko pitanje ili pitanja koja žele istražiti putem bibliometrijskog pregleda. Također, odabiru odgovarajuću bibliometrijsku metodu koja će im omogućiti analizu veza između dokumenata, autora ili časopisa. Na primjer, mogu odabrati analizu citata kako bi istražili utjecajnost i interakciju između radova.

Kocitati su oblik povezivanja dvaju dokumenata čija se jačina povezanost mjeri učestalošću kojom se ta dva dokumenta zajedno citiraju. U slučaju da imamo dokument I koji citira radove A, B, C, D i dokument II koji citira radove C, D, E, F i G, tada su kocitati citirani radovi C i D, u dokumentima I i II (Jokić M, 2005).

Ilustracija 14 Kocitirani radovi



Izvor: Jokić M, 2005

- Korak 2: Izdvajanje potrebnih podataka: odabir baze podataka (npr. SSCI, Scopus, Ebsco) i izvoz podataka (npr. upit, odabir časopisa, prag citiranosti).

U ovom koraku, istraživači odabiru relevantnu bazu podataka koja sadrži relevantne publikacije za njihovu studiju. Mogu koristiti popularne baze podataka poput SSCI (eng. *Social Sciences Citation Index*), Scopus ili Ebsco. Nakon odabira baze podataka, vrše upit za izdvajanje relevantnih radova i izvoze podatke koji su im potrebni za analizu.

- Korak 3: Analiza podataka - odabir bibliometrijskog softvera (npr. Bibexcel, Sitkis, VosViewer, CiteSpace II), generiranje matrice sličnosti (npr. korelacija, jačina veze), identifikacija podgrupe s odabranom tehnikom (npr. eksplorativna faktorska analiza, multidimenzionalno skaliranje, analiza klastera, analiza mreže).

Ovaj korak uključuje odabir odgovarajućeg bibliometrijskog softvera za analizu prikupljenih podataka. Primjeri takvog softvera uključuju Bibexcel, Sitkis, VosViewer, CiteSpace II i druge. Istraživači mogu koristiti ovaj softver kako bi generirali matricu sličnosti između radova, autora ili časopisa, koristeći metrike kao što su korelacija, su-pojava, kosinus ili jačina veze. Također, koriste različite tehnike poput eksplorativne faktorske analize, multidimenzionalnog skaliranja, analize klastera ili analize mreže kako bi identificirali podgrupe i obrasce u podacima.

- Korak 4: Vizualizacija - odabir željene metode vizualizacije (npr. multidimenzionalno skaliranje, analiza mreže, VOS).

U ovom koraku, analitičari biraju odgovarajuću strategiju vizualizacije koja će im pomoći da prikažu rezultate svoje istrage na jasan i instinktivan način. Oni mogu koristiti strategije kao što je multidimenzionalno skaliranje (MDS), organizirati istraživanje ili VOS (Vizualizirajući prinos nauke) kako bi ukazali na asocijacije između radova, kreatora ili dnevnika.

- Korak 5: Interpretacija - objašnjenje i kritička interpretacija rezultata.

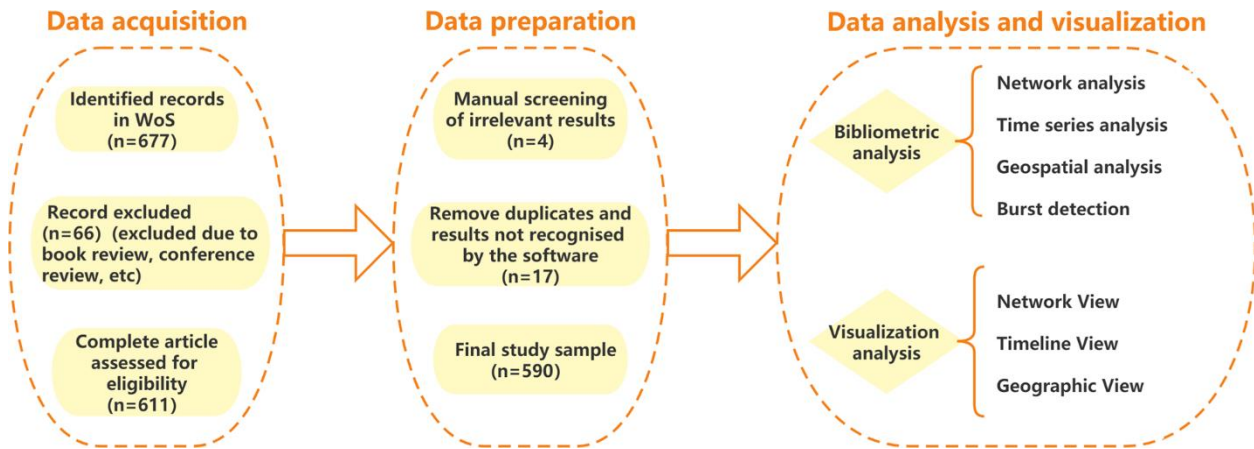
Završni korak uključuje pojašnjenje i osnovno razjašnjavanje rezultata dobijenih bibliometrijskim ispitivanjem. Analitičari moraju analizirati i dešifrovati šta se dešava unutar sistema njihovog istraživanja o regionu. Ovo uključuje prepoznavanje osnovnih dizajna, prepoznavanje ključnih kreatora ili radova, prepoznavanje i uočavanje potencijalnih problema u komponovanju.

Ovi koraci omogućavaju analitičarima da sistematiziraju bibliometrijske informacije, prepoznaju veze i dizajne i vizualiziraju ono što se događa na način koji podstiče razjašnjenje. Analitičari mogu pokupiti dublje znanje o napretku logičkog polja, prepoznati ključne kreatore i dnevnike i razlikovati obrasce i pukotine unutar pisanja.

Ključno je da prijevod bibliometrije dođe do zaključka o osnovnom razmatranju i kontekstualizaciji unutar upitnika o zoni.

Činilo se da ti koraci daju okvir za provođenje bibliometrijskog istraživanja, ali istraživači bi trebali primijeniti svoje stručne informacije i razumijevanje da prevedu ono što se dešava na osnovu specifičnosti svojih razmatranja.

Ilustracija 15 Dijagram toka istraživanja



Izvor: <https://www.nature.com/articles/s41599-023-01784-x>

U okviru oblasti upravljanja, gdje dosljedno kontrolirana ispitivanja nisu rutinski dostupna, izvodljiva sastavljena revizija drugih prenesenih kontemplacija govori o 'vođem okviru' dokaza i zahtijeva prikupljanje, temeljnu procjenu i objedinjavanje svih važnih kontemplacija koje upravljaju specifičnim pitanjima.

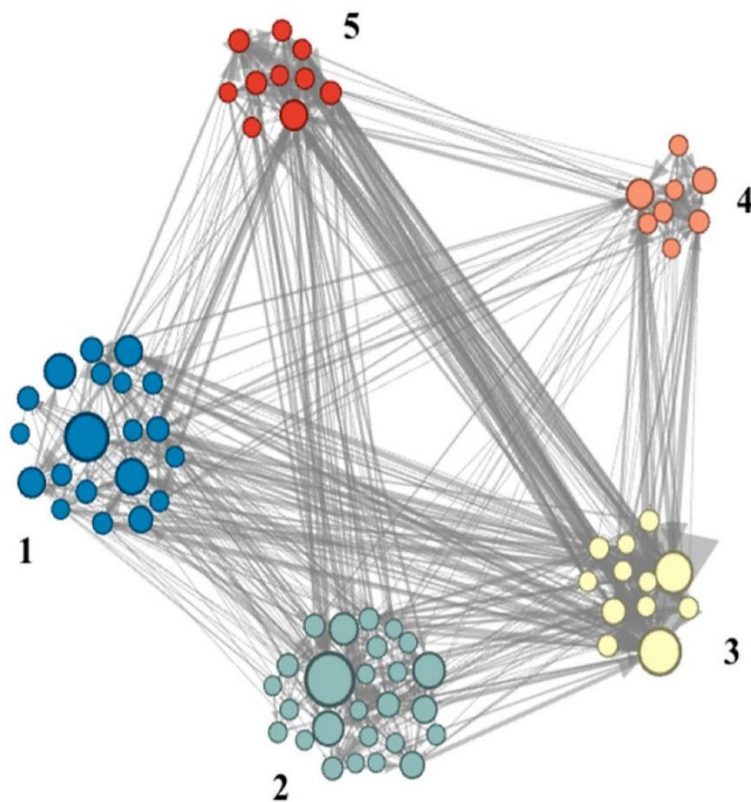
Za razliku od restaurativnih nauka, administracija ili društvenih nauka općenito, nisu stvoreni standardi i usklađene konvencije za prikupljanje dokaza o najboljem usavršavanju; razmatranja unutar oblasti administracije, zaista na istoj ili komparativnoj tački, mogu se razlikovati u smislu istraživačkih pitanja, vremenskih okvira, utvrđivanja testova i povezanih strategija.

Kako bi prevladali takva odvrćanja, ispitivači administracije progresivno shvaćaju metaanalitičke pristupe kako bi napravili različite "super-uzorke" koji omogućavaju da se informacije sumiraju i izvode iz kontemplacija (Combs *et al.*, 2019).

U svakom slučaju, uprkos dostupnosti ovakvih pristupa, postoji skoro najbolje izoštrena kontradikcija, a analitičari administracije nastavljaju da se suočavaju sa izazovima u istraživanju tako podjeljenih dokaza. Dodatno se ispituje značaj ili uticaj pojedinih stvaralaca ili pojedinih logičkih dijela.

Bibliometrijska analiza citiranosti predstavlja relativno novi oblik meta-analitičkog istraživanja ili "meta-pregleda" literature (Fetscherin *et al.*, 2010). Ova tehnika uzima citate kao osnovnu jedinicu analize i istražuje odnos među člancima u određenom istraživačkom području (J. Kim i McMillan, 2008). Također je korištena analiza ključnih pojmova kako bi se identificirali najrelevantniji segmenti u identificiranom okviru. Naknadno je provedena kvalitativna analiza sadržaja kako bi se dobili odgovori na postavljena istraživačka pitanja.

Ilustracija 16 Definisiranje 5 klastera literature



Izvor: <https://www.mdpi.com/2071-1050/10/3/595>

3.3. Interpretacija rezultata i diskusija

3.3.1. Najznačajniji autori

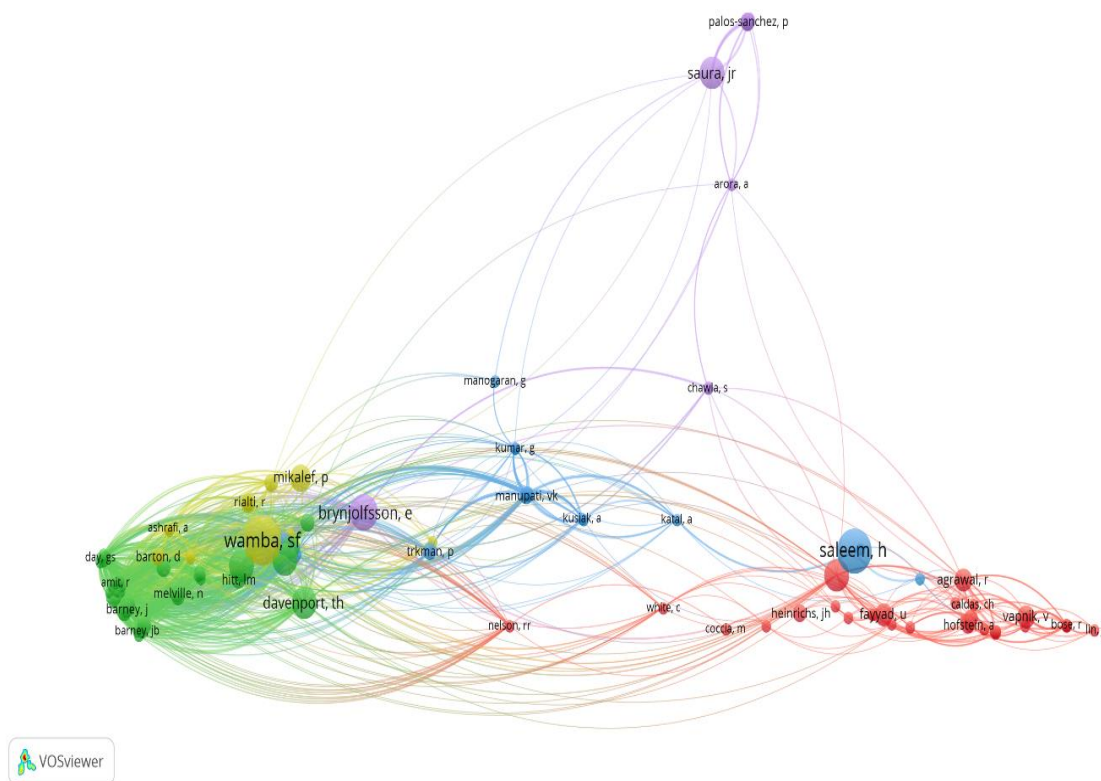
Tabela 2 Najznačajniji autori odabrane literature

AUTORI	UKUPAN BROJ OBJAVLJENIH RADOVA	UKUPAN BROJ CITATA
Ren, Steven Ji-fan; Wamba, Samuel Fosso; Akter, Shahriar; Dubey, Rameshwar; Childe, Stephen J.	491	22.215
Hao, M; Dayal, U; Hsu, M	487	7.861
Papadimitriou, Georgios; Komminos, Andreas; Garofalakis, John	316	1.842
Gopal, P. R. C.; Rana, Nripendra P.; Krishna, Thota Vamsi; Ramkumar, M.	274	9.174
Liu, Bin; Cao, Shu Gui; He, Wu	273	7.743
Liu, Song; Chen, Yukun; Xu, Kunpei; Lin, Jiaxin	269	3.070
Olson, David L.	266	3.759
Ramon Saura, Jose; Palos-Sanchez, Pedro; Grilo, Antonio	210	2.940
Wang, Hai; Wang, Shouhong	204	1.515
Choi, Jeongsub; Kim, Byunghoon; Hahn, Hyuk; Park, Hun; Jeong, Yongil; Yoo, Jaeyoung; Jeong, Myong Kee	157	4.053

Izvor: Autor

U korištenoj literaturi za sistemsku analizu navedene tematike, sprovedena je dodatna analiza grupe autora svih dijela. Usporedbom podataka dobivenih iz Web of Science baze, može se zaključiti da autori Steven Ji-fan Ren *et al.*, su najznačajniji autori među datom literaturom sa ukupnim brojem objavljenih radova 491 i ukupnim brojem citata 22.215 zajedno. Kako bi se postigao detaljniji uvid, deset (grupa) autora sa najvećim brojem ukupnih objavljenih radova i citata predstavljeni su u tabeli. Najznačajniji individualni autor jeste David L. Olson, koji individualno posjeduje 266 objavljenih radova sa 3.759 citata.

Ilustracija 17 Najznačajniji autori odabrane literature



Izvor: VOSviewer

3.3.2. Najcitiraniji radovi

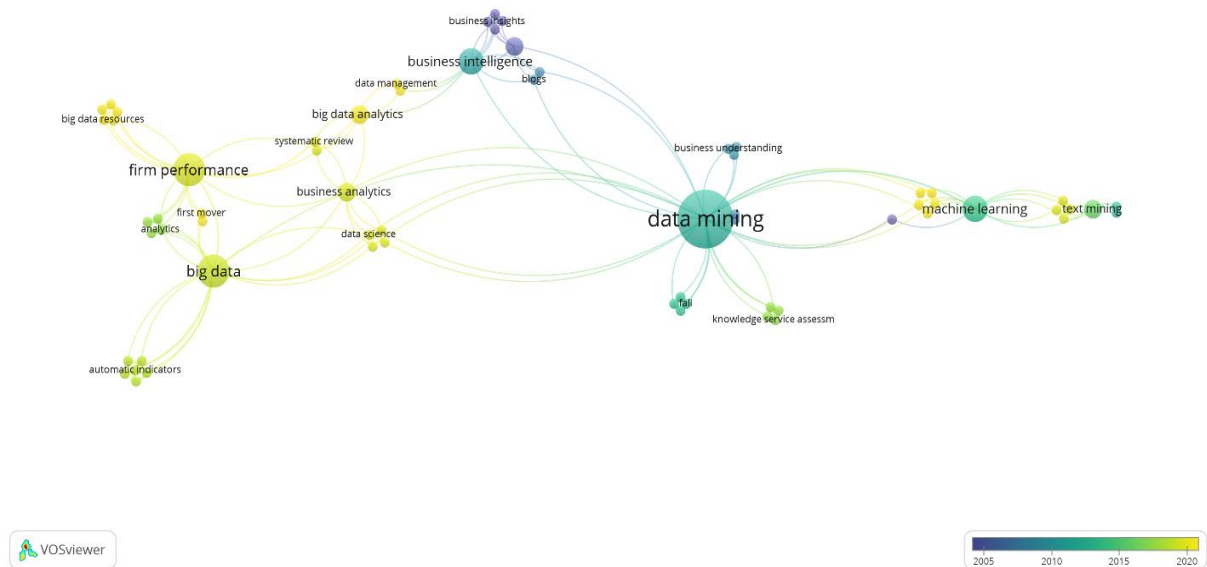
Ilustracija 18 Najcitiraniji radovi odabrane literature

NAZIV DJELA	BROJ CITATA
<i>A data mining approach to study the impact of the methodology followed in chemistry lab classes on the weight attributed by the students to the lab work on learning and motivation</i>	153
<i>Big data and firm performance: The roles of market-directed capabilities and business strategy</i>	121
<i>Modelling quality dynamics, business value and firm performance in a big data analytics environment</i>	90
<i>Big Data Analytics and Firm Performance: A Systematic Review</i>	80
<i>Distributed data mining for e-business</i>	64
<i>Business data mining - a machine learning perspective</i>	64
<i>Web data mining for monitoring business export orientation</i>	64
<i>Business environmental analysis for textual data using data mining and sentence-level classification</i>	58
<i>Detecting Indicators for Startup Business Success: Sentiment Analysis Using Text Data Mining</i>	53
<i>A knowledge management approach to data mining process for business intelligence</i>	47

Izvor: Autor

Najcitiraniji rad jeste *A data mining approach to study the impact of the methodology followed in chemistry lab classes on the weight attributed by the students to the lab work on learning and motivation* od autora Figueiredo et al. objavljen 2016., a posjeduje 153 citata. Zatim, na drugom mjestu je rad *Big data and firm performance: The roles of market-directed capabilities and business strategy* autora Suoniemi et al. objavljen 2020. sa 121 citatom.

Ilustracija 21 Ključni pojmovi istraživanja - vremenski prikaz

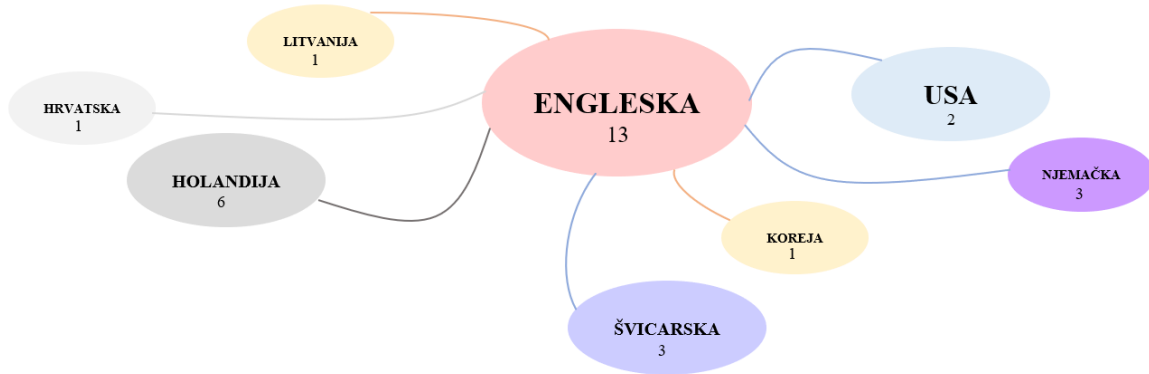


Izvor: VOSviewer

3.3.4. Geografska područja istraživačkih radova

Budući da je u ovoj sistemskoj analizi literature korišteno 30 istraživačkih radova na tematiku rudarenja i analize velikih podataka u poslovanju, važno je inkorporirati varijablu koja se odnosi na to gdje su istraživanja nastala. Države iz kojih autorska djela u navedenoj domeni potiču su: Njemačka, SAD, Holandija, Litvanija, Engleska, Švicarska, Hrvatska i Koreja. Od navedenih zemalja, najzastupljenije istraživanje jeste u Engleskoj odakle potiče 13 radova prema WoS citatnoj bazi. Zatim slijede Holandija (6), Švicarska (3), Njemačka (3) i ostale.

Ilustracija 22 Pregled po državama odabrane literature



Izvor: VOSViewer

3.3.5. Diskusija rezultata

Detaljna analiza istraživačkih radova podrazumijeva sistematičnu podjelu, gdje su u svim radovima pojedinačno analizirane hipoteze/ciljevi istraživanja, zaključci i eventualne preporuke ukoliko su autori diskutovali o istim.

Zaključci svih 30 radova usmjereni su na činjenicu da rudarenje podataka pozitivno utječe na poslovanje, te da u suštini predstavlja vrlo važan segment koji iz dana u dan dobija na značaju. Prvenstveno, smatrano je da svaki značajan proces koji podrazumijeva upravljanje podacima u poslovnoj inteligenciji mora ujedno i podrazumijevati upravljanje podacima nastalo rudarenjem podataka.

Sama poslovna inteligencija i analitika podataka su neposredno povezane sa velikim podacima, a samim time oba procesa doprinose donošenju odluka u organizacijama. Smatra se da je u današnjici, rudarenje podataka razvijeno u koristan analitički alat u svim znanstvenim i poslovnim oblastima, pružajući način za brzo sticanje novog razumijevanja na temelju analize podataka velikih razmjera. Dokazano je da poslovni subjekti koji primjenjuju rudarenje podataka zabilježavaju uspješne faktore za učinkovite performanse dovodeći se do uspjeha.

U isto vrijeme, kompanije koje imaju ogromne informacije sa zadovoljavajućim uređajima mogu napraviti napredak u kvaliteti kompletnog okvira i podataka, koji utiču na napredak. Upit o tome je bio predmet sistemske istrage, te je pokazao da upotreba ogromnih informacija pozitivno utiče na budžetsko poslovanje i uvažavanje izloga.

Sa napretkom tehnologije podataka, napravljen je napredak u primjeni i kontroli ogromnih informacija, pa je istraživanje pokazalo da razvoji dostupni putem računarstva u oblaku osnažuju kompanije da imaju naprednije pripreme za donošenje odluka, predviđanje obrazaca i pravednost.

Kao što vjerodostojnost predviđanja budućih okolnosti čini razliku kompanijama da planiraju na vrijeme za zamislive varijable koje bi uticale na njihovu trgovinu, ali je od vitalnog značaja da se izmeni proračun instrumenata za rudarenje informacija u razumijevanju sa različitim radoznalostima. Dalja istraživanja autora su dovela do zaključaka da samo rudarenje podataka ima pozitivan utjecaj na poslovnu inteligenciju, a samim time i poslovanje kompanije. Rezultati istraživanja dodatno utvrđuju da usvajanje tehnologija poput rudarenja podataka i mašinskog učenja doprinosi različitim sektorima kompanije sa ciljem veće isplativosti i postizanja konkurentske prednosti.

Ono što mnogi autori ističu jeste da se rudarenje podataka rapidno mijenja uz razvoj tehnologija, te da je vrlo važno da kompanije ozbiljno shvate suštinu procesa kako bi se iskoristio u punom kapacitetu, te doveo do bitnih promjena u organizaciji, preuzimajući veliki broj funkcija u kratkom periodu.

Osim poslovanja, neka istraživanja su dokazala da se napredak u razvoju nastave poslovnog engleskog jezika realizuje kroz velike količine podataka dobivene rudarenjem podataka, gdje su korišteni podaci dokazani kao validni, uz snažnu stabilnost i kvalitet.

Također, korištenjem inovativne integracije tehnika rudarenja tekstualnih podataka, moguće je postići optimizaciju performansi klasifikatora, te primjenom hibridnih metoda dolazi do preciznije klasifikacije tekstualnih podataka u odnosu na tradicionalne metode.

3.4. Ograničenja i preporuke

3.4.1. Ograničenja pregleda literature

Tradicionalni pristupi pregledu literature mogu biti podložni pristranosti i rezultirati netačnim odlukama. To je posebno zabrinjavajuće kada pregledi obuhvataju pitanja relevantna za politiku i praksu. Sistematski pregledi su uvedeni kao rigorozniji pristup sintezi dokaza iz različitih studija; oni se oslanjaju na skup metoda zasnovanih na dokazima, s ciljem maksimiziranja rigoroznosti i minimiziranja podložnosti pristranosti. Uprkos većoj popularnosti sistematskih pregleda, strategije zasnovane na sintezi dokaza i dalje se neefikasno primenjuju, što rezultira objavljivanjem sinteza koje su vrlo podložne sklonostima (Haddaway *et al.*, 2020).

Prepoznajući ograničenja s kojima se analitičari povremeno susreću prilikom dogovaranja, provođenja i distribucije temeljnih i sveobuhvatnih anketa dokaza, cilj je prepoznati osnovna pitanja u provođenju i detaljiziranju efikasnih istraživanja. Preuzimajući ulogu "osnovnog pratioca" u podršci potencijalnim preciznim anketama i izbjegavanju odgovora osoba na korištenje naziva "efikasne revizije", nastavlja se sa razlikovanjem metodoloških aranžmana za ublažavanje ovih zamki. Ističu se postojeća rješenja dostupna kako bi se izbjegli ovi problemi (*What Is a Literature Review? - Literature Reviews - LibGuides at University of Texas at Austin*, n.d.).

Prvenstveno, pregled literature pruža samo trenutnu sliku stanja znanja o određenoj temi. Kao takav, on ima početak, sredinu i kraj, ali ne može uključiti potencijalne buduće razvijanja koja mogu utjecati na relevantnost istraživanja.

Novi znanstveni radovi, otkrića ili promjene u području mogu se pojaviti nakon što je pregled literature završen, čime se može smanjiti njegova relevantnost.

Ključne definicije, strategije pretrage i kriteriji za uključivanje ili procjenu mogu se mijenjati tokom vremena ili se razlikovati između recenzenata. Kao rezultat toga, skup članaka neće biti reprezentativan za relevantnu osnovu dokaza, a važne studije mogu biti propuštene.

Kao rezultat toga, pregled može biti vrlo netačan i zavaravajući, te neće biti moguće ponoviti istraživanje. Prethodno definirani protokoli smanjuju pristranost, omogućavaju konstruktivne povratne informacije prije nego što dođe do grešaka u metodologiji pregleda, omogućavaju čitateljima provjeru metoda i izvještavanja, te služe kao unutar-grupni plan metoda tokom provođenja pregleda. Pregledi koji nemaju protokole ograničavaju ovu jasnoću i provjerljivost. Ovo je slično "prethodno registriranju" primarnih istraživanja u nekim poljima, gdje se metodološki planovi objavljuju, datiraju, verzioniraju i ne mogu se mijenjati.

Pregled literature također može imati određeni fokus na specifične studije zbog kojih dolazi do potencijalnih nedostataka relevantnosti, pri čemu nije moguća generalizacija niti iskazivanje jasnog zaključka.

Bitno je istaći da zaključak nekog istraživačkog rada može imati značajan utjecaj na akademsko polje.

To može poboljšati efikasnost, učinkovitost i utjecaj pregleda. Ujedno, može doći do korištenja definicija ključnih pojmova i termina pretrage koji nisu općeprihvaćeni ili prikladni, što ograničava prihvaćanje i usvajanje rezultata pregleda ili dovodi do netačnog ili pristranog odabira literature. Ovo može proizlaziti iz nedostatka koherentnosti unutar samih zajednica zainteresovanih strana.

Ukoliko kritička procjena manjka, moguće je da neka primarna istraživanja potencijalno imaju niži stepen pouzdanosti, koji dovode do pristranih rezultata.

Pregledi koji neadekvatno procjenjuju i uzimaju u obzir pouzdanost uključenih studija podložni su perpetuiranju ovih problema kroz sintezu, rezultirajući netačnim i pristranim rezultatima. Primarna istraživanja mogu imati poteškoće vezane uz "unutrašnju valjanost" (tj. tačnost metoda) koju uzrokuju, na primjer, konfounding varijable, nedostatak slijepe procjene, nepoštivanje prisutnosti konfounding varijabli i nedostatak randomizacije. Pregledi također mogu imati probleme s vanjskom valjanosti, gdje se primarna istraživanja razlikuju u relevantnosti za pitanje pregleda (na primjer, provedena na različitim prostornim mjerilima), ali to nije uzeto u obzir u sintezi.

Konačno, zaključci pregleda mogu biti zavodljivi ako se studije odabiru za meta-analizu na temelju kriterija koji nisu pravilno povezani s istraživačkim pitanjem.

Postizanje potpune cjelovitosti u pregledu literature može biti izazovno. Postoji veliki broj pisanja dostupnih o brojnim temama, tako da je neshvatljivo garantovati da je pristupačna informacija zaštićena i detaljno razmotrena. Stoga je od vitalnog značaja za istraživače da imaju na umu da će stalno postojati određena količina pisanja koja će biti isključena iz ankete.

Također, pregled literature može biti podložna predispoziciji u slučaju da se analitičari suzdržavaju od nevažnih dnevnika ili izvora.

Bitno je uzeti u obzir da se izvanredna istraživanja mogu objaviti i na drugim mjestima, kao što su publikacije s otvorenim pristupom ili predstavljanje na konferencijama. Pri tome, bitno je uključiti ograničavajuće stavove, kako bi se osigurala prilagodba i objektivnost pisane literature. Sklonost izboru se dešava kada članovi uključeni u reviziju nisu agenti opšte baze dokaza. Spajanje i zaključci zasnovani na ovim dokazima izuzetno su vjerovatno jednostrani ili pogrešni.

Uobičajeno, sklonost utvrđivanja može se desiti u revizijama zbog potrebe razmišljanja da se raspita o distribuciji i informacijama koje najavljuju sklonosti u distribuiranim ponderima, kao i zbog nedostatka strategija revizije koje utiču na odabir misli koje će biti uključene u pregled. U svakom slučaju, posebno u vezi sa tehnikom izgleda, sklonost odlučnosti utječe na miješanje kroz tehnike nepristojnog izgleda; za slučaj, odabir osobe o kojoj razmišlja za inkorporaciju, odabir pristrasnih/nereprezentativnih bibliografskih baza podataka ili korištenje nedovoljnih procedura pregleda za datu temu.

Mogućnost ponavljanja metoda pregleda tačno je ključni princip znanstvene metode, a metode koje se koriste za kreiranje revizija treba da budu direktno prikazane sa adekvatnim detaljima kako bi se omogućila ponovljivost pregleda. U slučaju da korisnik ne može shvatiti kako su razmišljanja razlikovana, birana i sintetizirana, a koja su zabranjena, opasnost od predispozicije se ne može ispitati, a magloviti subjektivni izbori mogu utjecati na nepokolebljiv kvalitet.

Neponovljivi pregledi se ne mogu potpuno pouzdano vjerovati, jer tokom izvođenja mogu biti napravljene greške. Osim toga, neponovljivi pregledi imaju ograničen uticaj, jer ih nije moguće ažurirati ili nadograditi, a razlike u rezultatima između više pregleda na istu temu ne mogu se uskladiti.

Konačno, neponovljivi pregledi umanjuju povjerenje u sintezu dokaza kao disciplinu, stvarajući prepreku za donošenje odluka na temelju dokaza. Slično, nedostatak transparentnosti u izvještavanju o rezultatima (tj. sirovi podaci studija, sažete statistike i analitički kod) sprječava analitičku replikaciju i verifikaciju.

Iskrivljenost objavljivanja podrazumijeva isključivanje sive literature i nedostatak testiranja na prisustvo iskrivljenosti objavljivanja. Pozitivni i statistički značajni rezultati istraživanja vjerovatnije će biti objavljeni od negativnih i nesignifikantnih rezultata.

Sinteze nastale na tradicionalnim, komercijalno objavljenim akademskim istraživanjima bit će iskrivljeni kao i same početne informacije. Istraživanja koja nisu pod kontrolom komercijalnih izdavača nazivaju se "siva literatura" te su istraživanja koja su namijenjena objavljivanju u akademskim medijima, ali iz nekog razloga nisu bila objavljena ili izvještaji kompanija nisu namijenjene akademskoj publici.

Izostavljanje relevantnih studija ovog tipa iz pregleda rezultira nepotpunom osnovom dokaza.

Testovi koji dovode do jakog sumnjanja u prisustvo iskrivljenosti objavljivanja i/ili kvantifikaciju njenog potencijalnog utjecaja predstavljaju važan element visokokvalitetne kvantitativne sinteze.

Pisanje literature imaju za cilj stvaranje novih saznanja kroz prikupljanje dokaza. U kvantitativnim revizijama, meta-analiza se postiže, kombinujući nastajanja i fluktuacije svih za koje se smatra da se proizvode mjerači uticaja sa sigurnim periodima.

Uprkos činjenici da meta-analiza nije obavezna strategija za sve efikasne ankete, revizije koje se razlikuju kao "meta- analiza" moraju ispuniti određene potrebe, računajući proračune uticaja za osobe koje razmišljaju i objedinjenu podatke.

Odabir i dizajn metoda sinteze predstavljaju bitnu stavku u sistemskom pregledu literature, uključujući formulaciju pitanja, pretraživanje, provjeru i izvlačenje podataka. Neodgovarajuća sinteza može narušiti valjanost cijelog pregleda.

Istraživač koji samostalno obavlja različite zadatke sistematskog pregleda može drugačije tumačiti definicije, koncepte i granice sistema u odnosu na druge. Ta varijabilnost je strukturalni dio ljudske prirode, ali u pregledu literature to može rezultirati uključivanjem ili isključivanjem različitog skupa studija, ovisno o individualnoj interpretaciji. Rad pojedinca, bez izazova drugih istraživača, ne omogućuje sigurnost u ispravno tumačenje protokola.

Također, rad pojedinca može dovesti do većeg broja grešaka (a posebno za pregledima, neprihvatljivo visoke stope lažno negativnih grešaka, odnosno pogrešno isključivanje relevantnih studija) u usporedbi s radom u suradnji s drugim istraživačem.

Sistematski pregledi su važni za donošenje odluka u politici i praksi u različitim disciplinama. Međutim, nedostaje svijesti o metodama koje osiguravaju pouzdanost i slobodu od pristranosti u tim pregledima. Autori pregleda bi trebali provoditi rigoroznije preglede, urednici i recenzenti trebaju strože vršiti selekciju, a metodolozi trebaju pružiti bolju podršku istraživačkoj zajednici. Bitno je ne uključivati pristrane rezultate koji sa sobom nose nizak stepen kvalitete, bez iskorištavanja resursa (Haddaway *et al.*, 2020).

Generalno, pregled literature je u suštini analiza bitnih i validnih istraživanja nekog područja ili tematike, pri čemu se pojašnjava ili dodatno podstiče istraživanje na određena postavljena pitanja (*Advantages and Disadvantages of Literature Review - Www.Howandwhat.Net*, n.d.).

Nakon sprovedene systemske analize literature, ograničenja koja je bitno spomenuti podrazumijevaju ograničen broj dostupne literature koja ispunjava kriterije tematike rudarenja i analize velikih podataka.

Naime, u ovom radu izvršena je detaljna analiza literature 30 istraživačkih radova, pri čemu nije moguće u potpunosti generalizovati zaključak o samoj temi.

Iako su autori radova dokazali pozitivan utjecaj rudarenja podataka na performanse poslovanja kompanija, ipak za potpune rezultate je potreban veći opseg istraživačkih radova. Također, dodatno ograničenje predstavlja činjenica da su dati radovi fokusirani na samo određeni broj kompanija, različitih zemalja.

3.4.2. Prijedlozi za buduća istraživanja

Budući istraživači, u području poslovanja i menadžmenta, bi trebali provesti više empirijskih studija o datoj temi, sa osvrtom na efekte raznih faktora, a koji podrazumijevaju menadžment podataka. Na primjer, mnoga istraživanja na temu velikih podataka i analitike istih vežu za velike kompanije ili kompanije u visokoprihodnim zemljama, bez osvrtnja na manje kompanije ili zemlje.

U tom slučaju, istraživači bi svoje studije uznapredovali šireći svoj fokus na druga, do sada ne tako istražena polja, kao što su SME preduzeća u zemljama srednjeg prihoda (Wang i Wang, 2008).

Pisanje istraživačkih radova ukazuje na stvaranje neiskorištenih informacija kroz prikupljanje dokaza. U kvantitativnim revizijama, meta-analiza se postiže, kombinujući nastajanja i fluktuacije svih za koje se smatra da se proizvode mjerači uticaja sa sigurnim periodima. Uprkos činjenici da metaanaliza nije obavezna strategija za sve efikasne ankete, revizije koje se razlikuju kao "meta-analiza" moraju ispuniti određene potrebe, računajući proračune uticaja za osobe koje razmišljaju i objedinjene.

Samim time, sistemski pregled literature bi bio relevantniji kada bi se proširio spektar istraživanja, te ukoliko bi ista inkorporirala mjere za poboljšanje dobivenih rezultata. Pregled takve literature, dovodi do adekvatnih zaključaka, a ujedno utječe na olakšanu i bolju integraciju prikupljenih podataka.

Za buduća istraživanja, preporučuje se istraživanje opasnosti povezanih s korištenjem rudarenja informacija i načinima na koje se rizici mogu ublažiti ili kontrolisati. Rudarenje podacima ili metoda pronalaženja vrijednih podataka iz ogromnih skupova informacija, nosi sa sobom određene izazove i potencijalne opasnosti.

Ključno je da te opasnosti i način na koji mogu uticati na organizaciju, sigurnost informacija, sigurnost i druge perspektive. Buduće ispitivanje može se usredsrediti na prepoznavanje određenih opasnosti, kao što su zloupotreba informacija, potreba za jasnoćom ili bezbijednosne opasnosti.

Dodatno je kritično istražiti načine na koje se ove opasnosti mogu ublažiti ili kontrolisati, kao što je stvaranje zadovoljavajućih pristupa, mehaničkih aranžmana ili podučavanje klijenata o sigurnoj upotrebi rudarenja informacija.

Razumijevanje i eliminisanje tih opasnosti je osnovno za efikasno aktualiziranje rudarenja informacija i garantiranje sigurnosti i povjerenja unutar rukovanja (Khder *et al.*, 2021).

U budućnosti je važno proširiti i integrirati model sa vremenskim napretkom u distribuciji tekstualnog rudarenja kako bismo mogli otkriti još više skrivenih znanja i obrazaca. Modeli rudarenja informacija zasnovani na tekstu mogu pružiti profitabilna iskustva iz ekspanzivnih skupova informacija o sadržaju, ali poboljšanje strategija i inovacija u oblasti analitike sadržaja kontinuirano napreduje.

Kombinovanje postojećih modela sa naprednim strategijama vremena omogućava nam da pratimo promene tokom vremena, razlikujemo obrasce i prepoznamo napredak pokrivenih informacija i dizajna. Integracija vremena unaprijed nam omogućava da bolje dobijemo energetske informacije i pokažemo prilagođavanje promjenljivim uslovima.

Na ovaj način možemo postići dublje razumijevanje štampanih informacija i iskoristiti njihov puni potencijal za kreiranje obrazovanih izbora. Proširivanje i integracija modela s vremenskim napretkom u distribuciji tekstualnog rudarenja omogućuje nam da budemo korak ispred u otkrivanju novih saznanja i razumijevanju složenih uzoraka u tekstualnim podacima (Liu *et al.*, 2011).

Buduće istraživanje može istražiti značaj SWOT varijanti kroz korištenje naprednijih proračuna mašinskog učenja i naprednijih procedura rudarenja informacija. Ovo može uključivati aplikacije kao što su sumiranje i prijedloge kako bi se dobilo superiorno i iskoristilo poštovanje SWOT varijable. U ekspanziji, analitičari mogu primijeniti druge napredne strategije objašnjenja kako bi potvrdili prikladnost SWOT istrage i unaprijedili njenu primjenu u stvarnim poslovnim situacijama. Ovo će dati dublje razumijevanje utjecaja SWOT varijable na poslovanje i dati bitan smjer za donošenje odluka u stvarnom svijetu (Y. S. Kim *et al.*, 2019).

Potrebno je nastaviti istraživanja kako bi se napredovale metodologije za predloženu tehniku, posebno kroz prezentaciju svih modernih strategija za modele mrežnog predstavljanja. Ova istraživanja će osnažiti poboljšanje naprednijih procedura koje će utjecati na preciznost i produktivnost klasifikacije štampanih informacija.

Primjenom naprednih procedura i proračuna, analitičari će moći bolje doći do njih i analizirati književne informacije, što može rezultirati preciznijim i vještijim klasifikacijama. Ovo će biti osnovno u brojnim rasponima kao što su ispitivanje pretpostavki, potvrda dizajna i donošenje obrazovanih izbora na osnovu štampanih informacija (Ur-Rahman i Harding, 2012).

Prijedlog za buduće istraživanje, odnosno sistemsku analizu literature, jeste da se prikupe istraživanja na određeni sektor kompanija ili kompanija zemalja koje je moguće porediti, te na taj način u potpunosti utvrditi zaključak za efikasnost rudarenja u poslovanju.

4. ZAKLJUČAK

Inovativni pokretači i najnovija kompjuterizacija svih gledišta trgovine doveli su do trajne baze podataka u kojoj su pohranjene različite informacije koje služe svim članovima oglašavanja. Koristeći određene kompjuterske programe za rudarenje podataka, mogu otkriti obrasce ponašanja u ekspanzivnim bazama podataka.

Sa takvim podacima kompanije mogu na odgovarajući način pozicionirati svoje procedure i smanjiti trgovinske troškove. Informaciono rudarenje može biti aranžman za pripremu vitalnih informacija iz različitih skupova podataka i baza podataka, i govori o velikoj prekretnici u stvaranju konkurentske prednosti za kompanije koje treba da maksimiziraju svoj potencijal u budućnost. Danas, koristeći internet, brojni pojedinci nesvjesno skidaju informacije koje imaju uvažavanje kod određenog dijela oglašivača i nose potrebne informacije. Ovo odražava neke od ključnih fokusnih tačaka koje informaciono rudarenje donosi u današnju trgovinu, i vjerujemo da će u bliskoj budućnosti biti fundamentalno koristiti rudarenje informacija unutar kompanije kako bi se nadživjelo oglašavanje.

Funkcijske podrške odlučivanju za upravljanje incidentima su kroz historiju razvijale niz odgovarajućih rješenja da identificiraju abnormalne situacije koristeći metode Data mininga. Sistemi mogu generisati alternative i pomoći donosiocu odluke u razumijevanju problema prilikom odlučivanja i odabiru alternative.

Metode Data mininga pružaju mogućnosti pronalaska skrivenog znanja u velikim količinama podataka što utiče na lakše donošenje odluka. Trenutno su se metode Data mininga razvile i ne koriste se samo u kriznim situacijama već kroz cijelu industriju. Sam proces donošenja odluka uslijed mnogih neizvjesnosti nije lagan zadatak upravo zbog toga je bitan Data mining kako bi obezbjedio sistem za efikasno prikupljanje, razmjenu, širenje, eksploataciju i analizu informacija, što može veoma pozitivno uticati na proces donošenja odluka u budućnosti ukoliko ga koristimo pravilno.

Precizno istraživanje upita o radovima na temu informatičkog rudarenja ukazalo je na razvojni zamah kada je i u pitanju primjena ogromnih informacija ili rudarenje istih u kompanijama, te uticaj daljeg poslovanja kompanije. Nastala istraživanja mogu se iskoristiti za buduće radove na uporedivu temu, s obzirom na istinu da su u istraživanju korišteni izvanredno odabrani radovi sa najkritičnijim stvaraocima u ovoj oblasti.

Bitno je naglasiti da se rudarenje podataka mijenja iz dana u dan, a dostiže sve veću popularnost, te je jako bitan i koristan za sve sektore na tržištu.

REFERENCE

1. Adams, R. J., Smart, P., i Huff, A. S. (2017). Shades of Grey: Guidelines for Working with the Grey Literature in Systematic Reviews for Management and Organizational Studies. *International Journal of Management Reviews*, 19(4), 432–454. <https://doi.org/10.1111/ijmr.12102>
2. *Advantages and disadvantages of literature review - www.howandwhat.net.* (n.d.). Datum pristupanja: May 13, 2023, from <https://www.howandwhat.net/advantages-disadvantages-literature-review/>
3. Akcetin, E., Kilic, A., Yurtay, N., Yurtay, O. Y., Ozturk, E., i Sahin, O. A. (2015). *Data mining: usage and applications in tourism industry.* <https://www.researchgate.net/publication/294580583>
4. Alles, M., Kogan, A., Vasarhelyi, M., Hiltz, R., i Turoff, M. (2004). *ASSURING HOMELAND SECURITY: CONTINUOUS MONITORING, CONTROL AND ASSURANCE OF EMERGENCY PREPAREDNESS.* www.optimizemag.com/issue/020/law.htm,
5. Apte, C., Liu, B., Pednault, E. P. D., i Smyth, P. (2002, August). *Business applications of data mining.* 49–53.
6. Ardito, L., Scuotto, V., Del Giudice, M., i Petruzzelli, A. M. (2019). A bibliometric analysis of research on Big Data analytics for business and management. *Management Decision*, 57(8), 1993–2009. <https://doi.org/10.1108/MD-07-2018-0754>
7. Bach, M. P., Krstič, Ž., Seljan, S., i Turulja, L. (2019). Text mining for big data analysis in financial sector: A literature review. *Sustainability (Switzerland)*, 11(5). <https://doi.org/10.3390/su11051277>
8. Bean, R., i Kiron, D. (2013). *Organizational Alignment is Key to Big Data Success.* <https://sloanreview.mit.edu/article/organizational-alignment-is-key-to-big-data-success/>
9. Booth, A. (2016). Searching for qualitative research for inclusion in systematic reviews: A structured methodological review. *Systematic Reviews*, 5(1). <https://doi.org/10.1186/s13643-016-0249-x>
10. Briner, R. B., i Denyer, D. (2012). Systematic Review and Evidence Synthesis as a Practice and Scholarship Tool. In *The Oxford Handbook of Evidence-Based Management.* Oxford University Press. <https://doi.org/10.1093/oxfordhob/9780199763986.013.0007>
11. Brynjolfsson, E., i Hitt, L. M. (1998). Beyond the productivity paradox. *Communications of the ACM*, 41(8), 49–55. <https://doi.org/10.1145/280324.280332>
12. Chen, R., Sharman, R., Rao, R. H., i Upadhyaya, S. (2007). Design principles for critical incident response systems. *Information Systems and E-Business Management*, 5(3), 201–227. <https://doi.org/10.1007/s10257-007-0046-0>
13. Chung, H. M., i Gray, P. (1999). Special Section: Data Mining. *Journal of Management Information Systems*, 16(1), 11–16. <https://doi.org/10.1080/07421222.1999.11518231>
14. Combs, J. G., Crook, T. R., i Rauch, A. (2019). Meta-Analytic Research in Management: Contemporary Approaches, Unresolved Controversies, and Rising Standards. *Journal of Management Studies*, 56(1), 1–18. <https://doi.org/10.1111/joms.12427>
15. Cypher, A., i Halbert, D. C. (1993). Watch what I Do: Programming by Demonstration - Google Books. In *Massachusetts Institute of Technology.* <https://books.google.ba/books?hl=enilr=iid=Ggzjo0->

- W1y0Cioi=ndipg=PA1idq=A.+Cypher,+D.+C.+Halbert+(1993),+Watch+What+I+Do:
+Programming+by+Demonstration,+MIT+Press,+Cambridge,+MA.iots=W752-
f9GSjisig=nTHOD8yKBOwvkYESK1SdzUYnOesiredir_esc=y#v=onpageiq=A.%20
Cypher%2C%20D.%20C.%20Halbert%20(1993)%2C%20Watch%20What%20I%20D
o%3A%20Programming%20by%20Demonstration%2C%20MIT%20Press%2C%20C
ambridge%2C%20MA.if=false
16. Di Stefano, G., Peteraf, M., i Verona, G. (2009). *DYNAMIC CAPABILITIES DECONSTRUCTED: A BIBLIOGRAPHIC INVESTIGATION INTO THE ORIGINS, DEVELOPMENT, AND FUTURE DIRECTIONS OF THE RESEARCH DOMAIN DYNAMIC CAPABILITIES DECONSTRUCTED*: A BIBLIOGRAPHIC INVESTIGATION INTO THE ORIGINS, DEVELOPMENT, AND FUTURE DIRECTIONS OF THE RESEARCH DOMAIN.*
 17. *Difference Between Data Mining and Data Analysis - Javatpoint.* (n.d.). Datum pristupanja: June 11, 2023, from <https://www.javatpoint.com/data-mining-vs-data-analysis>
 18. Fayyad, U., Piatetsky-Shapiro, G., i Smyth, P. (1996). From Data Mining to Knowledge Discovery in Database. *AI MAGAZINE*, 17. www.fly.com/
 19. Fetscherin, M., Voss, H., i Gugler, P. (2010). *30 Years of foreign direct investment to China: An interdisciplinary literature review.* http://scholarship.rollins.edu/as_facpub
 20. Gao, J., Wang, H., i Shen, H. (2022). Task Failure Prediction in Cloud Data Centers Using Deep Learning. *IEEE Transactions on Services Computing*, 15(3), 1411–1422. <https://doi.org/10.1109/TSC.2020.2993728>
 21. Garača, Ž., i Jadrić, M. (2011). *Rudarenje podataka: različiti aspekti informacijskog društva.* Ekonomski fakultet u Splitu.
 22. Gazzawe, F., i Alturki, R. (2022). Data Mining and Soft Computing in Business Model for Decision Support System. *Scientific Programming*, 2022. <https://doi.org/10.1155/2022/9147444>
 23. Ghapanchi, A. H., i Aurum, A. (2011). Antecedents to IT personnel's intentions to leave: A systematic literature review. *Journal of Systems and Software*, 84(2), 238–249. <https://doi.org/10.1016/j.jss.2010.09.022>
 24. Grbavac, V. (2018). *Rudarenje podataka kao metoda upravljanja znanjem.* Sveučilište u Zagrebu, Fakultet organizacije i informatike.
 25. Green, L. W., i Glasgow, R. E. (2006). Evaluating the relevance, generalization, and applicability of research: Issues in external validation and translation methodology. In *Evaluation and the Health Professions* (Vol. 29, Issue 1, pp. 126–153). <https://doi.org/10.1177/0163278705284445>
 26. Haddaway, N. R., Bethel, A., Dicks, L. V., Koricheva, J., Macura, B., Petrokofsky, G., Pullin, A. S., Savilaakso, S., i Stewart, G. B. (2020). Eight problems with literature reviews and how to fix them. In *Nature Ecology and Evolution* (Vol. 4, Issue 12, pp. 1582–1589). Nature Research. <https://doi.org/10.1038/s41559-020-01295-x>
 27. Han, J., Kamber, M., i Pei, J. (2012). *Data Mining Concepts and Techniques* (3rd ed.). Elsevier.
 28. Han, J., Kamber, M., Pei, J., i Kaufmann, M. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Elsevier.
 29. Hao, M., Dayal, U., i Hsu, M. (2000). *Visual Data Mining for Business Intelligence Applications.*

30. Hassani, H., Beneki, C., Unger, S., Mazinani, M. T., i Yeganegi, M. R. (2020). Text mining in big data analytics. *Big Data and Cognitive Computing*, 4(1), 1–34. <https://doi.org/10.3390/bdcc4010001>
31. Hearst, M. A. (2012). *Text Data Mining* (R. Mitkov, Ed.; Vol. 1). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199276349.013.0034>
32. Hill, T. (2012). *The Big Data Revolution And How to Extract Value from Big Data*. www.statsoft.com
33. Howard, B., Scott, J. T., Blubaugh, M., Roepke, B., Scheckel, C., i Vassar, M. (2017). Systematic review: Outcome reporting bias is a problem in high impact factor neurology journals. *PLoS ONE*, 12(7). <https://doi.org/10.1371/journal.pone.0180986>
34. Huang, C. K., Wang, T., i Huang, T. Y. (2020). Initial Evidence on the Impact of Big Data Implementation on Firm Performance. *Information Systems Frontiers*, 22(2), 475–487. <https://doi.org/10.1007/s10796-018-9872-5>
35. Ishwarappa, i Anuradha, J. (2015). A brief introduction on big data 5Vs characteristics and hadoop technology. *Procedia Computer Science*, 48(C), 319–324. <https://doi.org/10.1016/j.procs.2015.04.188>
36. Jain, S., i McLean, C. R. (2006). An Integrating Framework for Modeling and Simulation for Incident Management. *Journal of Homeland Security and Emergency Management*, 3(1). <https://doi.org/10.2202/1547-7355.1194>
37. Ji-fan Ren, S., Fosso Wamba, S., Akter, S., Dubey, R., i Childe, S. J. (2017). Modelling quality dynamics, business value and firm performance in a big data analytics environment. *International Journal of Production Research*, 55(17), 5011–5026. <https://doi.org/10.1080/00207543.2016.1154209>
38. Jokić M. (2005). *Bibliometrijski aspekti vrednovanja znanstvenog rada*.
39. K. Ojo, A., i B. Adeyemo, A. (2017). Characterisation of Academic Journal Publications Using Text Mining Techniques. *Journal of Computer Sciences and Applications*, 5(2), 42–49. <https://doi.org/10.12691/jcsa-5-2-1>
40. Kauffmann, E., Peral, J., Gil, D., Ferrández, A., Sellers, R., i Mora, H. (2019). Managing marketing decision-making with sentiment analysis: An evaluation of the main product features using text data mining. *Sustainability (Switzerland)*, 11(15). <https://doi.org/10.3390/su11154235>
41. Khder, M. A., Abdullrahman, J., i Bahar, Y. (2021). THE IMPACT OF IMPLEMENTING DATA MINING IN BUSINESS INTELLIGENCE. *International Journal of Entrepreneurship*, 25(2), 1–7.
42. Kim, J., i McMillan, S. J. (2008). Evaluation of internet advertising research: A bibliometric analysis of citations from key sources. *Journal of Advertising*, 37(1), 99–112. <https://doi.org/10.2753/JOA0091-3367370108>
43. Kim, Y. S., Rim, H. C., i Lee, D. G. (2019). Business environmental analysis for textual data using data mining and sentence-level classification. *Industrial Management and Data Systems*, 119(1), 69–88. <https://doi.org/10.1108/IMDS-07-2017-0317>
44. Klashner, R., i Sabet, S. (2007). A DSS Design Model for complex problems: Lessons from mission critical infrastructure. *Decision Support Systems*, 43(3), 990–1013. <https://doi.org/10.1016/j.dss.2005.05.027>
45. Kohavi, R., i Provost, F. (2001). *Applications of Data Mining to Electronic Commerce* (R. Kohavi i F. Provost, Eds.). Springer US. <https://doi.org/10.1007/978-1-4615-1627-9>

46. Lame, G. (2019). Systematic literature reviews: An introduction. *Proceedings of the International Conference on Engineering Design, ICED, 2019-August*, 1633–1642. <https://doi.org/10.1017/dsi.2019.169>
47. Lau, J., Antman, E. M., Jimenez-Silva, J., Kupelnick, B., Mosteller, F., i Chalmers, T. C. (1992). Cumulative Meta-Analysis of Therapeutic Trials for Myocardial Infarction. *New England Journal of Medicine*, 327(4), 248–254. <https://doi.org/10.1056/NEJM199207233270406>
48. Linnenluecke, M. (2017). Resilience in Business and Management Research: A Review of Influential Publications and a Research Agenda. In *International Journal of Management Reviews* (Vol. 19).
49. Liu, B., Cao, S. G., i He, W. (2011). Distributed data mining for e-business. In *Information Technology and Management* (Vol. 12, Issue 2, pp. 67–79). <https://doi.org/10.1007/s10799-011-0091-8>
50. Llave, M. R. (2018). Data lakes in business intelligence: Reporting from the trenches. *Procedia Computer Science*, 138, 516–524. <https://doi.org/10.1016/j.procs.2018.10.071>
51. Ludäscher, B., Lin, K., Bowers, S., Jaeger-Frank, E., Brodaric, B., i Baru, C. (n.d.). *Managing Scientific Data: From Data Integration to Scientific Workflows*.
52. Maroufkhani, P., Wagner, R., Wan Ismail, W. K., Baroto, M. B., i Nourani, M. (2019). Big data analytics and firm performance: A systematic review. In *Information (Switzerland)* (Vol. 10, Issue 7). MDPI AG. <https://doi.org/10.3390/INFO10070226>
53. McKinsey Global Institute. (2011). *Big data: The next frontier for innovation, competition, and productivity*. www.mckinsey.com/mgi.
54. Mengist, W., Soromessa, T., i Legese, G. (2020). Method for conducting systematic literature review and meta-analysis for environmental science research. *MethodsX*, 7. <https://doi.org/10.1016/j.mex.2019.100777>
55. Mikalef, P., Pappas, I. O., Krogstie, J., i Pavlou, P. A. (2020). *Big data and business analytics: A research agenda for realizing business value*.
56. Niu, Y., Ying, L., Yang, J., Bao, M., i Sivaparthipan, C. B. (2021). Organizational business intelligence and decision making using big data analytics. *Information Processing and Management*, 58(6). <https://doi.org/10.1016/j.ipm.2021.102725>
57. Olson, D. L. (2007). Data mining in business services. In *Service Business* (Vol. 1, Issue 3, pp. 181–193). <https://doi.org/10.1007/s11628-006-0014-7>
58. Pal, S., i Gauri, S. K. (2018). A desirability functions-based approach for simultaneous optimization of quantitative and ordinal response variables in industrial processes. *International Journal of Engineering, Science and Technology*, 10(1), 76–87. <https://doi.org/10.4314/ijest.v10i1.6>
59. Peng, Y., Zhang, Y., Tang, Y., i Li, S. (2011). An incident information management framework based on data integration, data mining, and multi-criteria decision making. *Decision Support Systems*, 51(2), 316–327. <https://doi.org/10.1016/j.dss.2010.11.025>
60. Pivk, A., Vasilecas, O., Kalibatiene, D., i Rupnik, R. (2013). On approach for the implementation of data mining to business process optimisation in commercial companies. *Technological and Economic Development of Economy*, 19(2), 237–256. <https://doi.org/10.3846/20294913.2013.796501>
61. Ramageri, B. M. (2014). Data mining techniques and applications. In *Indian Journal of Computer Science and Engineering* (Vol. 1).

62. Rialti, R., Marzi, G., Ciappei, C., i Busso, D. (2019). Big data and dynamic capabilities: a bibliometric analysis and systematic literature review. *Management Decision*, 57(8), 2052–2068. <https://doi.org/10.1108/MD-07-2018-0821>
63. Rizvić, A. (2016). *CJELOŽIVOTNO UČENJE I UTICAJ KORIŠTENJA INFORMACIONIH TEHNOLOGIJA NA INDIVIDUALNE PERFORMANSE: ANALIZA NA PRIMJERU BOSNE I HERCEGOVINE*. Univerzitet u Sarajevu Ekonomski fakultet.
64. Rodríguez del Águila, M. M., i Benítez-Parejo, N. (2011). Simple linear and multivariate regression models. *Allergologia et Immunopathologia*, 39(3), 159–173. <https://doi.org/10.1016/j.aller.2011.02.001>
65. Sabherwal, R., i Jeyaraj, A. (2015, December). *Information Technology Impacts on Firm Performance on JSTOR*. <https://www.jstor.org/stable/26628653>
66. Saha, D. (2020, May 20). *How The World Became Data-Driven, And What's Next*. <https://www.forbes.com/sites/googlecloud/2020/05/20/how-the-world-became-data-driven-and-whats-next/?sh=50e692357fc5>
67. Seng, J. L., i Chen, T. C. (2010). An analytic approach to select data mining for business decision. *Expert Systems with Applications*, 37(12), 8042–8057. <https://doi.org/10.1016/j.eswa.2010.05.083>
68. Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., i Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ*, 349(jan02 1), g7647–g7647. <https://doi.org/10.1136/bmj.g7647>
69. Shamsi, J. A., i Khojaye, M. A. (2018). Understanding Privacy Violations in Big Data Systems. *IT Professional*, 20(3), 73–81. <https://doi.org/10.1109/MITP.2018.032501750>
70. Shmueli, G., Bruce, P. C., i Patel, N. R. (2016). *Data mining for business analytics*.
71. Smith, B. (2018). Doing a Literature Review: Releasing the Research Imagination. *Journal of Perioperative Practice*, 28(12), 318–318. <https://doi.org/10.1177/1750458918810149>
72. Srića, V. (2018). *Otkrivanje Znanja Iz Podataka*. Ekonomski Fakultet u Zagrebu. <https://www.scribd.com/document/378101618/05-Otkrivanje-Znanja-Iz-Podataka#>
73. Suoniemi, S., Meyer-Waarden, L., Munzel, A., Zablah, A. R., i Straub, D. (2020). Big data and firm performance: The roles of market-directed capabilities and business strategy. *Information and Management*, 57(7). <https://doi.org/10.1016/j.im.2020.103365>
74. *Systematic literature reviews - Literature Review - Library Guides at Charles Sturt University*. (n.d.). Datum pristupanja: May 16, 2023, from <https://libguides.csu.edu.au/review/Systematic>
75. Torraco, R. J. (2005). Writing Integrative Literature Reviews: Guidelines and Examples. *Human Resource Development Review*, 4(3), 356–367. <https://doi.org/10.1177/1534484305278283>
76. Tranfield, D., Denyer, D., i Smart, P. (2003a). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management*, 14(3), 207–222. <https://doi.org/10.1111/1467-8551.00375>
77. Tranfield, D., Denyer, D., i Smart, P. (2003b). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review*. In *British Journal of Management* (Vol. 14).
78. Tsiptsis, K., i Chorianopoulos, A. (2009). *Data Mining Techniques in CRM: Inside Customer Segmentation*.

79. Tuffery, S. (2008). *Data mining and statistics for decision making*. Editions Technip. www.wiley.com.
80. Ur-Rahman, N., i Harding, J. A. (2012). Textual data mining for industrial knowledge management and text classification: A business oriented approach. *Expert Systems with Applications*, 39(5), 4729–4739. <https://doi.org/10.1016/j.eswa.2011.09.124>
81. Vera-Baquero, A., Colomo-Palacios, R., i Molloy, O. (2013). *Business Process Analytics Using a Big Data Approach*. <http://hadoop.apache.org>.
82. Vidal-García, J., Vidal, M., i Barros, R. H. (2019). Computational Business Intelligence, Big Data, and Their Role in Business Decisions in the Age of the Internet of Things. In *Web Services* (pp. 1048–1067). IGI Global. <https://doi.org/10.4018/978-1-5225-7501-6.ch055>
83. Waddington, H., White, H., Snilstveit, B., Hombrados, J. G., Vojtkova, M., Davies, P., Bhavsar, A., Eyers, J., Koehlmoos, T. P., Peticrew, M., Valentine, J. C., i Tugwell, P. (2012). How to do a good systematic review of effects in international development: A tool kit. *Journal of Development Effectiveness*, 4(3), 359–387. <https://doi.org/10.1080/19439342.2012.711765>
84. Wahono, R. S. (2015). A Systematic Literature Review of Software Defect Prediction: Research Trends, Datasets, Methods and Frameworks. *Journal of Software Engineering*, 1(1). <http://journal.ilmukomputer.org>
85. Wang, H., i Wang, S. (2008). A knowledge management approach to data mining process for business intelligence. *Industrial Management and Data Systems*, 108(5), 622–634. <https://doi.org/10.1108/02635570810876750>
86. Webster, J., i Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. In *Source: MIS Quarterly* (Vol. 26, Issue 2).
87. *What is a literature review? - Literature Reviews - LibGuides at University of Texas at Austin*. (n.d.). Datum pristupanja: May 13, 2023, from <https://guides.lib.utexas.edu/literaturereviews>
88. *What Is Data Segmentation? The Complete Guide | Just Total Tech*. (n.d.). Datum pristupanja: June 11, 2023, from <https://justtotaltech.com/what-is-data-segmentation/>
89. Wilson, T. D. (1999). Models in information behaviour research. *Journal of Documentation*, 55(3), 249–270. <https://doi.org/10.1108/EUM0000000007145>
90. Zaki, M. J., i Meira, W. (n.d.). *Data mining and analysis : fundamental concepts and algorithms*.
91. Zhang, H., Zang, Z., Zhu, H., Uddin, M. I., i Amin, M. A. (2022). Big data-assisted social media analytics for business model for business decision making system competitive analysis. *Information Processing i Management*, 59(1), 102762. <https://doi.org/10.1016/j.ipm.2021.102762>
92. Zong, K., Yuan, Y., Montenegro-Marin, C. E., i Kadry, S. N. (2021). Or-based intelligent decision support system for e-commerce. *Journal of Theoretical and Applied Electronic Commerce Research*, 16(4), 1150–1164. <https://doi.org/10.3390/JTAER16040065>