

PREDIKTIVNA ANALITIKA U PREDVIĐANJU TRENDOVA I DONOŠENJU INVESTICIJSKIH ODLUKA

Bogdanović, Doris

Undergraduate thesis / Završni rad

2023

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Josip Juraj Strossmayer University of Osijek, Faculty of Economics in Osijek / Sveučilište Josipa Jurja Strossmayera u Osijeku, Ekonomski fakultet u Osijeku**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:145:279580>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-12-23**



Repository / Repozitorij:

[EFOS REPOSITORY - Repository of the Faculty of Economics in Osijek](#)



Sveučilište Josipa Jurja Strossmayera u Osijeku

Ekonomski fakultet u Osijeku

Sveučilišni prijediplomski studij, Financijski menadžment

Doris Bogdanović

**PREDIKTIVNA ANALITIKA U PREDVIĐANJU TRENDOVA I
DONOŠENJU INVESTICIJSKIH ODLUKA**

Završni rad

Osijek, 2023.

Sveučilište Josipa Jurja Strossmayera u Osijeku

Ekonomski fakultet u Osijeku

Sveučilišni prijediplomski studij, Financijski menadžment

Doris Bogdanović

**PREDIKTIVNA ANALITIKA U PREDVIĐANJU TRENDOVA I
DONOŠENJU INVESTICIJSKIH ODLUKA**

Završni rad

Kolegij: Financijski menadžment

JMBAG: 0010234607

e-mail: dbogdanovic1@efos.hr

Mentor: prof.dr.sc., Domagoj Sajter

Osijek, 2023.

Josip Juraj Strossmayer University of Osijek
Faculty of Economics and Business in Osijek
University Undergraduate Study; Financial management


Doris Bogdanović

**PREDICTIVE ANALYTICS IN FORECASTING TRENDS AND
MAKING INVESTMENT DECISIONS**

Final paper

Osijek, 2023.

IZJAVA
O AKADEMSKOJ ČESTITOSTI,
PRAVU PRIJENOSA INTELEKTUALNOG VLASNIŠTVA,
SUGLASNOSTI ZA OBJAVU U INSTITUCIJSKIM REPOZITORIJIMA
I ISTOVJETNOSTI DIGITALNE I TISKANE VERZIJE RADA

1. Kojom izjavljujem i svojim potpisom potvrđujem da je završni
(navesti vrstu rada: završni / diplomski / specijalistički / doktorski) rad isključivo rezultat osobnoga rada koji se temelji na mojim istraživanjima i oslanja se na objavljenu literaturu. Potvrđujem poštivanje nepovredivosti autorstva te točno citiranje radova drugih autora i referiranje na njih.
2. Kojom izjavljujem da je Ekonomski fakultet u Osijeku, bez naknade u vremenski i teritorijalno neograničenom opsegu, nositelj svih prava intelektualnoga vlasništva u odnosu na navedeni rad pod licencom *Creative Commons Imenovanje – Nekomercijalno – Dijeli pod istim uvjetima 3.0 Hrvatska*. 
3. Kojom izjavljujem da sam suglasan/suglasna da se trajno pohrani i objavi moj rad u institucijskom digitalnom repozitoriju Ekonomskoga fakulteta u Osijeku, repozitoriju Sveučilišta Josipa Jurja Strossmayera u Osijeku te javno dostupnom repozitoriju Nacionalne i sveučilišne knjižnice u Zagrebu (u skladu s odredbama Zakona o visokom obrazovanju i znanstvenoj djelatnosti, NN 119/2022).
4. izjavljujem da sam autor/autorica predanog rada i da je sadržaj predane elektroničke datoteke u potpunosti istovjetan sa dovršenom tiskanom verzijom rada predanom u svrhu obrane istog.

Ime i prezime studenta/studentice: Doris Bogdanović

JMBAG: 0010234607

OIB: 16828149895

e-mail za kontakt: doris.bogdanovic42@gmail.com

Naziv studija: Sveučilišni prijediplomski studij, Financijski menadžment

Naslov rada: Prediktivna analitika u predviđanju trendova i donošenju investicijskih odluka

Mentor/mentorica rada: prof.dr.sc., Domagoj Sajter

U Osijeku, _____ 5.9.2023. _____ godine

Potpis Doris Bogdanović

Prediktivna analitika u predviđanju trendova i donošenju investicijskih odluka

SAŽETAK

Prediktivna analitika je proces korištenja analitičkih alata za predviđanje budućih pojava. Sastoji se od definiranja ciljeva, prikupljanja podataka, primjene tehnika modeliranja i implementacije plana u stvarnoj situaciji. Prediktivna analitika služi kao sredstvo za poboljšanje donošenja odluka, kvalitetnije predviđanje, povećanje efikasnosti i smanjenje rizika. Zahtijeva kompleksne modele, kvalitetne podatke i odgovarajuće resurse. Primjenjuje se na brojnim područjima, posebice za predviđanje trendova i donošenje investicijskih odluka. Budućnost prediktivne analitike leži u automatizaciji postupaka i primjeni uloge umjetne inteligencije. Cilj je ovog rada istražiti prediktivnu analitiku kao alat za predviđanje trendova i donošenje investicijskih odluka. Rezultati bi mogli pomoći investitorima u donošenju kvalitetnijih investicijskih odluka, a također će doprinijeti i napretku financijske tržišne strukture. U ovom radu obrađuju se razni alati strojnog učenja za prikupljanje, modeliranje i analiziranje podataka o trendovima i investiranju. Korišteni alati mogu uključivati algoritme strojnog učenja kao što su regresija, stablo odlučivanja i umjetna neuronska mreža. Spomenuti algoritmi mogu biti korisni investitorima, te zahvaljujući njima moguće je s većom pouzdanošću i točnošću predviđati buduće trendove i rizike što može dovesti do pouzdanijih investicijskih odluka.

Ključne riječi: Prediktivna analitika, modeliranje, analitika podataka, strojno učenje

Predictive analytics in forecasting trends and making investment decisions

ABSTRACT

Predictive analytics is the process of using analytical tools to predict future events. It involves defining goals, collecting data, applying modeling techniques, and implementing a plan in a real situation. Predictive analytics serves as a means to improve decision-making, enhance prediction accuracy, increase efficiency, and reduce risk. It requires complex models, quality data, and appropriate resources. It is applied in various fields, particularly for predicting trends and making investment decisions. The future of predictive analytics lies in automating procedures and incorporating artificial intelligence. The aim of this study is to explore predictive analytics as a tool for predicting trends and making investment decisions. The results could assist investors in making better investment decisions and contribute to the advancement of the financial market structure. This study discusses various machine learning tools for data collection, modeling, and analysis of trends and investments. The utilized tools may include machine learning algorithms such as regression, decision trees, and artificial neural networks. These mentioned algorithms can be beneficial to investors, as they enable more reliable and accurate predictions of future trends and risks, leading to more trustworthy investment decisions.

Keywords: Predictive analytics, modeling, data analysis, machine learning

SADRŽAJ

1. Uvod	1
2. Proces prediktivne analitike	4
2.1. Proces modeliranja	6
2.2. Tehnike.....	9
2.2.1. Stablo odlučivanja	9
2.2.2. Regresijski model	10
2.2.3. Umjetna neuronska mreža.....	11
2.2.4. Bayesova statistika	12
2.2.5. Učenje u ansamblu	13
2.2.6. Model povećanja gradijenta	14
2.2.7. Potporni vektorski stroj.....	14
2.2.8. Analiza vremenskih nizova.....	15
2.2.9. Analiza glavnih komponenti.....	16
2.3. Implementacija	17
3. Izazovi prediktivne analitike	19
4. Prednosti prediktivne analitike	22
5. Praktični primjeri predviđanja trendova i donošenja investicijskih odluka	24
6. Zaključak	27
Literatura	28
Popis slika	30

1. Uvod

Svrha je ovog rada ispitati prediktivnu analitiku kao alat u donošenju investicijskih odluka. Prediktivna analitika poprilično je udaljila proces donošenja odluka, posebice strateških poslovnih odluka od „intuicije“ i „osjećaja“ prema donošenju odluka na temelju činjenica i dokaza (Ogunleye, 2014).

U današnje vrijeme, prediktivna analitika i vizualizacija podataka moderne su teme zbog toga što je danas dostupno mnogo internih i eksternih podataka. Tehnologija za korištenje tih podataka je evoluirala te je opće prihvaćeno da u podacima leži mnogo vrijednih saznanja koja se mogu dobiti analizom tih podataka (Pagans, 2015).

Prediktivna analitika može se promatrati još od 650. g. pr. Kr. kada su Babilonci pokušavali predvidjeti promjenu vremena na temelju oblika oblaka i optičkih pojava u atmosferi u obliku svjetlećega prstena oko Sunca ili Mjeseca, koja nastaje kada se Sunce ili Mjesec naziru kroz tanak sloj oblaka (Winters, 2017). U medicini je također postajala potreba za klasificiranjem bolesti. Iz toga je proizašlo da je babilonski kralj Adad-apla-iddina naredio sakupljanje zdravstvenih zapisa kako bi formirali knjigu dijagnostičkih postupaka. U tom se korpusu nalaze predviđanja liječenja temeljena na broju dana koliko je pacijent bolestan i na njihovom pulsu. U kasnijim vremenima, prediktivna se analitika razvijala u začecima industrije osiguranja gdje se analitika upotrebljavala za predviđanje rizika povezanih uz osiguranje plovila. U to vrijeme društva za životno osiguranje počela su predviđati dob koju će osoba doživjeti kako bi odredili najprikladnije premijske stope. Iako se ideja predviđanja činila duboko ukorijenjenom u ljudsku potrebu za razumijevanjem i klasifikacijom, počela se doista snažno razvijati tek u 20. stoljeću, s nastankom modernog računarstva.

Alan Turing koji je pomogao američkoj vladi 1940-ih da dešifrira tajne poruke, također je radio i na prvim računalnim algoritmima u kojem bi računalo igralo šah protiv čovjeka. Tijekom Projekta Manhattan, koji je tajni program američke vlade za stvaranje prvih atomskih bombi za potrebe Drugog svjetskog rata, glavna računala računala su vjerojatnost nuklearnih napada koristeći klasu algoritama koje su danas poznate pod nazivom Monte Carlo metode. Teorija operacijskih istraživanja razvila se u 1950-ima pomoću koje je moguće optimizirati pronalaženje najkraće udaljenosti između dvaju točaka. U današnje vrijeme tvrtke poput UPS-a i Amazona koriste te iste tehnike u logistici. Operacijsko istraživanje koristili su i stručnjaci izvan područja matematike. U 70-tim godinama prošlog stoljeća, kardiolog Lee Goldman, koji

je bio zaposlen na podmornici, razvijao je stablo odlučivanja koje je pomoglo osoblju podmornice utvrditi treba li podmornica ponovno izroniti kako bi pomogla osobama koje pate od bolova u prsima. Svim je tim primjerima zajedničko što su korisnici prvo zabilježili zapažanja o nastalim događajima, a zatim su koristili te informacije da bi generalizirali i donosili odluke o tome što bi se moglo dogoditi u budućnosti. Uz predviđanja došlo je do daljnjeg razumijevanja uzroka i posljedica te povezanosti različitih dijelova problema međusobno. Otkrića i uvidi nastali su putem metodologije i poštivanjem znanstvene metode (Winters, 2017).

Voditelji poslovanja danas primjenjuju prediktivnu analitiku u svrhu predviđanja budućih tržišnih uvjeta kako bi saznali što kupci žele i trebaju (Wessler, 2014). Donositelji odluka moraju napraviti pretpostavke o budućnosti na temelju prethodnih događaja, te ono što voditelji poslovanja traže su točni podaci koji vode predviđanje budućnosti. Prediktivna analitika pruža donositeljima odluka i analitičarima mogućnost da precizno predviđaju buduće događaje na temelju složenih statističkih algoritama primijenjenih na relevantne podatke (Wessler, 2014).

Brojna istraživanja donose različita definiranja prediktivne analitike. Moderne organizacije skupljaju ogromne količine podataka, te da bi ti podaci bili korisni organizaciji, moraju ih analizirati kako bi iz njih izvukli spoznaje koje se mogu koristiti za donošenje boljih odluka. Nadalje, prediktivna se analitika može definirati kao umijeće korištenja modela koja mogu napraviti predviđanja temeljena na uzorcima izvučenih iz povijesnih podataka (Kelleher i D'Arcy, 2015).

S druge strane, prediktivna se analitika može promatrati kao proces otkrivanja zanimljivih i značajnih uzoraka u podacima. Oslanja se na nekoliko međusobno povezanih disciplina, koje su korištene više od stotinu godina za otkrivanje uzoraka među podacima. Neke od spomenutih disciplina su sljedeće: prepoznavanje uzoraka, statistika, strojno učenje, umjetna inteligencija te rudarenje podataka (Abbott, 2014).

Osim prethodno opisanog pogleda, prediktivna se analitika može objasniti i kao širok pojam koji opisuje različite statističke i analitičke tehnike koje služe za razvijanje modela čija je svrha predviđanje budućih događaja ili ponašanja. Njihov oblik varira s obzirom na ponašanje ili događaj koji predviđaju. Većina takvih modela generiraju kriterij prema kojemu veći rezultat pokazuje veću vjerojatnost da će se neki događaj ili ponašanje dogoditi (Nyce, 2007).

Rudarenje podataka smatra se dijelom prediktivne analitike u kojem se analiziraju podaci za otkrivanje trendova, uzoraka ili odnosa među podacima. Ta informacija se onda može koristiti

za izradu prediktivnog modela. Prediktivna analitika, zajedno s većinom prediktivnih modela i tehnika rudarenja podataka, oslanja se na sve sofisticiranije statističke metode, uključujući i multivarijacijske analize kao što su napredne regresije ili vremenski serijski modeli. Takve tehnike omogućuju organizacijama da utvrde trendove i odnose koji možda neće biti lako vidljivi, ali još uvijek omogućuju bolje predviđanje budućih događaja ili ponašanja (Nyce, 2007).

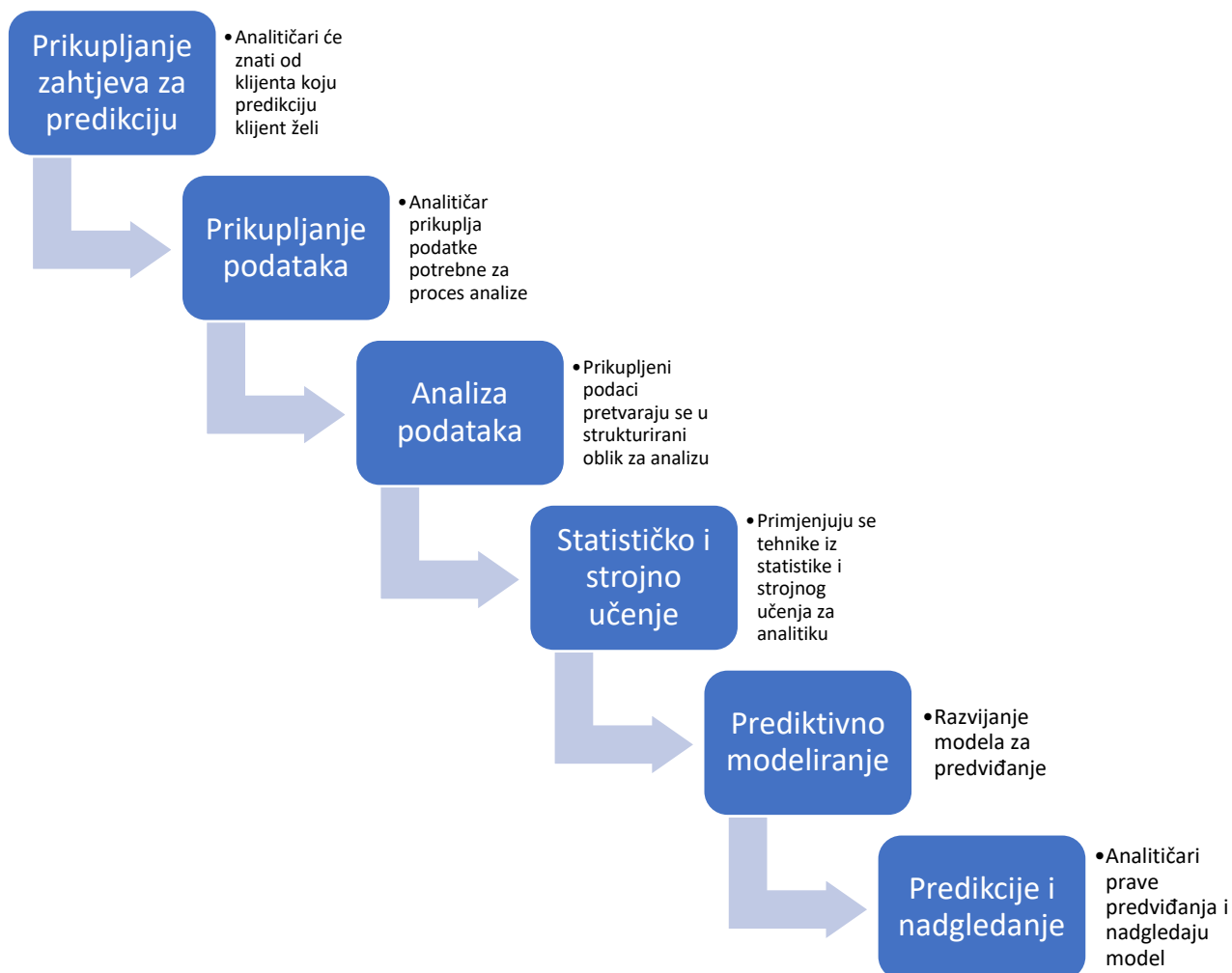
Nadalje, prediktivna analitika opisuje se kao primjena vještina, stručnosti i sposobnosti softvera za izdvajanje, ispitivanje, analizu i pretvaranje podataka u jasne, razumljive oblike koji se mogu koristiti u planiranju organizacije ili donošenju odluka. Ona kombinira ljudske vještine i stručnost s tehnologijom poput strojnog učenja i pronađene uzorke u trenutnim i povijesnim podacima i primjenom algoritama, ne samo za identificiranje uzoraka u podacima već i za prognoziranje buduće vjerojatnosti rezultata tih uzoraka (Ogunleye, 2014).

Konačno, dvije su spoznaje jasne iz ovakvih definicija prediktivne analitike. Prva je da su ljudi, alati i algoritmi ključni u projektu prediktivne analitike, a druga da je prediktivna analitika usmjerena prema budućnosti i kreira prognoze koje se temelje na podacima.

Stoga se na temelju spomenutog u ovom radu generalno istražuje navedena tematika počevši od uvoda koji govori o početcima prediktivne analitike. Zatim se u drugom poglavlju objašnjava proces prediktivne analitike i proces modeliranja te tehnike korisne u prediktivnom modeliranju. Uz to, spomenuta su i područja implementacije prediktivne analitike primjerice, u zdravstvu i osiguranju. U trećem poglavlju navode se prednosti koje ona ima u pogledu donošenja odluka i prognoziranja kao što su otkrivanje prijevara i smanjenje rizika. Nakon toga se u četvrtom poglavlju spominju izazovi s kojima se moguće susresti prilikom korištenja prediktivne analitike, neke od kojih su kompleksnost modela, problemi s kvalitetom podataka i potrebe za resursima. U petom poglavlju izneseni su praktični primjeri predviđanja trendova i donošenja investicijskih odluka u kojima poznate organizacije kao što su PepsiCo, Rolls-Royce i Netflix koriste prediktivnu analitiku kao alat u donošenju investicijskih odluka. Na temelju istraživanog, za kraj je iznesen zaključak.

2. Proces prediktivne analitike

Prediktivna analitika sadrži nekoliko koraka kroz koje analitičar može predvidjeti budućnost na temelju trenutnih i povijesnih podataka. Ovaj proces prikazan je na slici 1.



Slika 1: Proces prediktivne analitike, Izvor: Kumar, V., & Garg, M. L. (2018). *Predictive analytics: a review of trends and techniques*. *International Journal of Computer Applications*, 182(1), 32. str.

a) Prikupljanje zahtjeva za predikciju

Kako bi se razvio prediktivni model, mora se jasno definirati cilj predviđanja. Kroz predviđanje, potrebno je definirati vrstu znanja koja će se steći. Primjerice, farmaceutska tvrtka želi znati prognozu o prodaji lijeka na specifičnom području kako bi izbjegla istek tih lijekova. Analitičari potom komuniciraju s klijentima da bi saznali njihove zahtjeve za razvoj prediktivnog modela

i da im prenesu moguće koristi od tih predviđanja. Zatim će identificirati podatke koji su im potrebni od klijenta kako bi razvili model.

b) Prikupljanje podataka

Nakon spoznaje zahtjeva klijenta, analitičari će prikupiti skupove podataka, koji mogu biti iz različitih izvora, potrebne za razvoj prediktivnog modela. Ti podaci mogu biti u strukturiranom ili u nestrukturiranom obliku. Analitičar provjerava podatke koje je klijent prikupio.

c) Analiza podataka

Nakon prikupljanja podataka, analitičari analiziraju prikupljene podatke i pripremaju ih za analizu i za upotrebu u modelu. U tom koraku nestrukturirani podaci pretvaraju se u strukturirani oblik. Kada svi podaci budu dostupni u strukturiranom obliku, provjerava se njihova kvaliteta. Postoje mogućnosti grešaka u glavnom skupu podataka ili izostanaka nekih vrijednosti. Učinkovitost prediktivnog modela potpuno je ovisna o kvaliteti podataka. Faza analize podataka ponekad se naziva i obradom podataka što znači pretvoriti „sirove“ podatke u format koji se koristi za analitiku.

d) Statističko i strojno učenje

Proces prediktivne analitike koristi mnoge statističke tehnike i tehnike strojnog učenja. Teorija vjerojatnosti i regresijska analiza neke su od važnijih alata koje se široko koriste. Također, neuronske mreže, stabla odlučivanja, stroj za podršku vektora alati su strojnog učenja koji se široko koriste u mnogim zadacima prediktivne analitike. Svi modeli prediktivne analitike temelje se na statističkim i/ili strojnim tehnikama učenja. Stoga analitičari primjenjuju koncepte statistike i strojnog učenja kako bi razvili prediktivne modele. Strojne tehnike učenja imaju prednost u odnosu na konvencionalne statističke tehnike, ali statistika je neizostavno uključena u razvoj bilo kojeg prediktivnog modela.

e) Prediktivno modeliranje

U toj se fazi razvija model temeljen na statističkim i strojnim učenjima te na primjeru skupa podataka. Nakon razvoja, model se testira na testnom skupu podataka koji je dio glavnog skupa podataka kako bi se provjerila valjanost modela, a u slučaju uspjeha model se smatra valjanim. Nakon toga, model može točno izvršavati predviđanja na temelju novih podataka unesenih u sustav.

f) Predviđanja i nadgledanje

Nakon uspješnih testova u predviđanjima, model je postavljen u klijentovo područje za svakodnevna predviđanja i proces donošenja odluka. Model se kontinuirano nadgleda radi sigurnosti točnosti rezultata i napravljenih predviđanja. Može se zaključiti kako prediktivna analitika nije jedan korak u predviđanju budućnosti. To je proces koji se kreće postupno, od prikupljanja zahtjeva za predikciju do same predikcije i njezinog praćenja za učinkovito korištenje sustava kako bi napravili koristan sustav u procesu donošenja odluka.

2.1. Proces modeliranja

Općenito govoreći, model se objašnjava kao prikaz nečega. To može biti statički model ili dinamički model koji predstavlja neki proces. Primjer statičkog modela može biti žena koja hoda po modnoj pisti u novim dizajnerskim odjevnim predmetima, a ona predstavlja kako bi osoba izgledala u tim odjevnim predmetima. Primjeri dinamičkih modela mogu biti skupovi jednadžbi koji opisuju dinamiku fluida, promet koji se odvija u gradovima, vrijednosti dionica koje osciliraju s vremenom, i sl. (Wu i Coggeshall, 2012).

Modeli obično čine jednostavnije reprezentacije složenijih sustava. Modeli se grade kako bi nam pomogli da shvatimo što se događa, kako se stvari međusobno odnose ili da predvidimo što se može dogoditi, kako se stvari mijenjaju ili evoluiraju.

Modeli mogu biti fizički objekti ili algoritmi. Kod algoritamskih modela, oni mogu biti algoritmi prvih principa ili statistike. Za prvi princip modela ispituje se sustav i piše skup pravila ili jednadžbi koje opisuju esenciju sustava, ignorirajući detalje koji su manje važni. Prvi principi

modela mogu biti diferencijalne jednačbe koje opisuju dinamički sustav ili mogu biti skup evolucijskih pravila koji se pišu na temelju stručnog znanja o procesima sustava. Važno je razumjeti ključne elemente sustava, koje značajke karakteriziraju sustav, što se mijenja, koje su moguće ovisnosti i koje su važne značajke koje nas zanimaju. To promatranje vodi u izgradnju modela, bilo prvim principima ili gradnjom statističkog modela. Ponekad je postupak toliko složen ili nepoznat da ne možemo početi pisati jednačbe iz prvih principa, ali naša promatranja su dovoljna da imamo mnoge primjere podataka sustava koji se modelira. U tim slučajevima možemo izgraditi statistički model koji se definira kao skup jednačbi s prilagodljivim parametrima koje „treniramo“ upotrebom mnogih podataka. Postoji predodređeni sustav gdje imamo mnogo više primjera podataka nego prilagodljivih parametara i želimo „najbolji“ set parametara koji omogućuju da naše jednačbe što bolje odgovaraju podacima. Struktura jednačbi može biti vrlo jednostavna ili vrlo složena. To može biti jednostavna linearna kombinacija parametara i svojstava sustava ili može biti vrlo složena nelinearna kombinacija svojstava i parametara. Zadatak modeliranja sadrži nekoliko komponenti:

- a. skup jednačbi ili pravila s prilagodljivim parametrima,
- b. skup podataka koji uključuje primjere sustava koje pokušavamo modelirati,
- c. koncept dobrog uklapanja u podatke i
- d. skup pravila koji nam govori kako prilagoditi parametre da povećamo dobro uklapanje modela.

Proces izgradnje modela je mješavina znanosti i vještine za koju postoji jednostavna rutinska procedura koju slijedimo, a prilikom svakog koraka vodi nas intuicija i iskustvo. Osnovni koraci izgradnje statističkog modela prema Wu i Coggeshallu (2012) su sljedeći:

1. Definiranje ciljeva

Što želimo postići? Kakav ishod želimo predvidjeti? Pod kojim uvjetima će model biti korišten i u koje svrhe?

2. Sakupljanje podataka

Koji podaci su dostupni, u kojem obliku i kakve su kvalitete? Koliko zapisa ćemo imati? Koji modeli inputa mogu biti dostupni? Koliko daleko u prošlost idu podaci i koliko su relevantni stariji podaci?

3. Odabiranje strukture modela

Trebamo li raditi linearnu regresiju, logističku regresiju ili nelinearni model? Koje vremenske okvire i/ili populacije treba koristiti za razvoj i validaciju modela? Koji su inputi, outputi? Dobar pristup bi bio pokušati napraviti jednostavni linearni model kao bazu, a zatim pokušati različite nelinearne modele kako bi vidjeli hoće li jedan moći bolje. Odluke o strukturi modela zahtijevaju iskustvo i duboko znanje o prednostima i slabostima svakog tehničkog postupka. Važne karakteristike koje je potrebno uzeti u obzir prilikom odabira tehnike uključuju izlaz kontinuiranog ili kategoričkog (klasifikacijskog) outputa, broj zapisa i iznos nevažnih podataka.

4. Priprema podataka

Potrebno je sastaviti podatke u odgovarajući oblik za model, kodiranjem podataka u inpute, koristeći stručna znanja. Na odgovarajući način nužno je normalizirati numeričke podatke i kodirati kategorijska polja podataka. Zatim slijedi ispitivanje distribucije gdje treba pripaziti na ekstreme. Nakon toga, potrebno je odvojiti podatke u željene setove za obuku, testiranje i validacijske setove.

5. Izbor i eliminacija varijabli

U ovom koraku, koji bi mogao biti povezan ili ne s potencijalnim modelima, varijable se ispituju za važnost modela gdje se odabiru ili eliminiraju. Neke metode odabira varijabli su samostalne a neke su integrirane s određenom metodologijom modela. Uglavnom se u ovom koraku odlučuje lista potencijalno dobrih varijabli, rangiranih po važnosti.

6. Izgradnja modela

Počinju se graditi modeli i procjenjivati kvaliteta modela. Kreće se s linearnim modelima kao bazom, a zatim se pokušava poboljšati korištenjem složenijih nelinearnih modela. Prilikom ove i ostalih faza važno je imati na umu okruženje u kojem će se model implementirati.

7. Zaključivanje modela

Odabire se među kandidatima najprikladniji model za implementaciju. Model se dokumentira prema potrebi.

8. Implementacija i praćenje

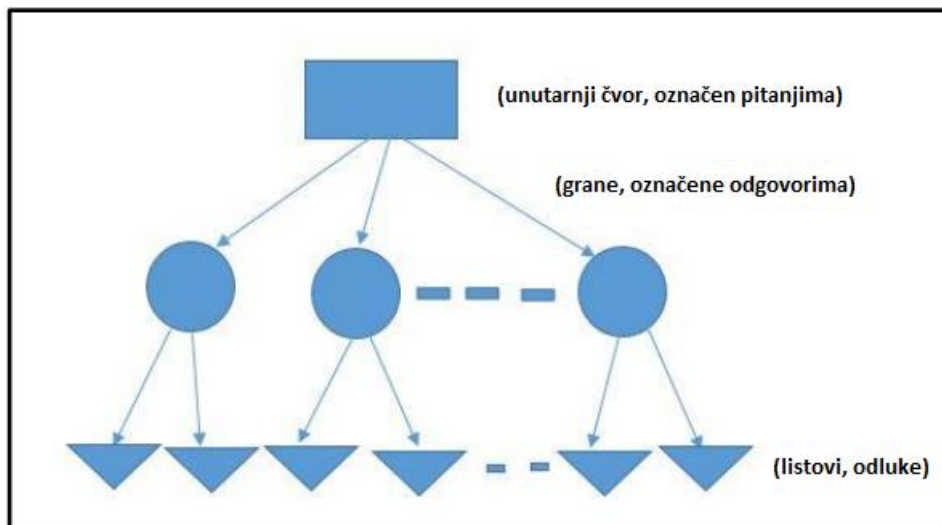
U ovom se koraku model ugrađuje u odgovarajuće procese sustava, a izgrađuju se postupci praćenja za ispitivanje performansi modela tijekom redovite uporabe.

2.2. Tehnike

Prema Kumaru i Gargu (2018), svi modeli prediktivne analitike raspoređeni su u klasifikacijske modele i regresijske modele. Klasifikacijski modeli predviđaju pripadnost vrijednosti određenoj klasi, dok regresijski modeli predviđaju broj. Spomenuti autori izdvojili su važne tehnike koje se popularno koriste u razvoju prediktivnih modela te su one ukratko predstavljene.

2.2.1. Stablo odlučivanja

Stablo odlučivanja je model klasificiranja, ali ga se može koristiti i u regresiji. To je model nalik stablu koji povezuje odluke i njihove moguće posljedice. Posljedice mogu biti ishod događaja, troškovi resursa ili učinkovitost. U svojoj strukturi svaka grana predstavlja izbor između nekoliko alternativa, a svaki list predstavlja odluku. Na temelju kategorija ulaznih varijabli, podatke dijeli u podskupove. Jednostavnost razumijevanja i interpretacije čine stablo odlučivanja popularnim za korištenje. Tipičan model stabla odlučivanja predstavljen je na slici 2.



Slika 2: Stablo odlučivanja, Izvor: Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. International Journal of Computer Applications, 182(1), 34. str.

Unutarnji čvor definiran je pitanjima vezanima za odluku. Sva grananja predstavljaju moguće odgovore na određeno pitanje. Vanjski elementi stabla, poznati kao listovi, predstavljaju odluku o problemu. Taj model ima svojstvo obrađivanja podataka koji nedostaju te je također koristan u odabiru preliminarnih varijabli. Ono koristi većinu podataka u skupu podataka i minimizira razinu pitanja. Uz spomenuta svojstva, stabla odlučivanja imaju i prednosti i mane. Novi

mogući scenariji mogu se dodati modelu koji otkriva njegovu fleksibilnost i prilagodljivost. Oni se mogu integrirati s drugim modelima odlučivanja po potrebi. Imaju ograničenu mogućnost prilagođavanja promjenama. Mala promjena u podacima dovodi do velikih promjena u strukturi. Zaostaju u točnosti predviđanja u usporedbi s drugim prediktivnim modelima. U slučaju korištenja nepouzdanih podataka, izračun postaje otežan i složen.

2.2.2. Regresijski model

Regresija se smatra jednom od najpopularnijih statističkih tehnika koja predviđa odnos između varijabli. Modelira odnos između zavisne varijable i jedne ili više nezavisnih varijabli pomoću funkcije. Analizira kako se vrijednost zavisne varijable mijenja promjenom vrijednosti nezavisnih varijabli u modeliranom odnosu. Taj modelirani odnos između zavisne i nezavisne varijable prikazan je na slici 3.



Slika 3: Regresijski model, Izvor: Kumar, V., & Garg, M. L. (2018). *Predictive analytics: a review of trends and techniques. International Journal of Computer Applications*, 182(1), 34. str.

U kontekstu kontinuiranih podataka, za koje se pretpostavlja da imaju normalnu distribuciju, regresijski model pronalazi ključne uzorke u velikim skupovima podataka. U regresiji se vrijednost zavisne varijable predviđa temeljem prediktivne varijable. U tom se slučaju regresijska funkcija koristi svim nezavisnim varijablama i povezuje ih odgovarajućim zavisnim varijablama. Koristeći distribuciju vjerojatnosti, varijacija zavisne varijable okarakterizirana je predviđanjima regresijske funkcije.

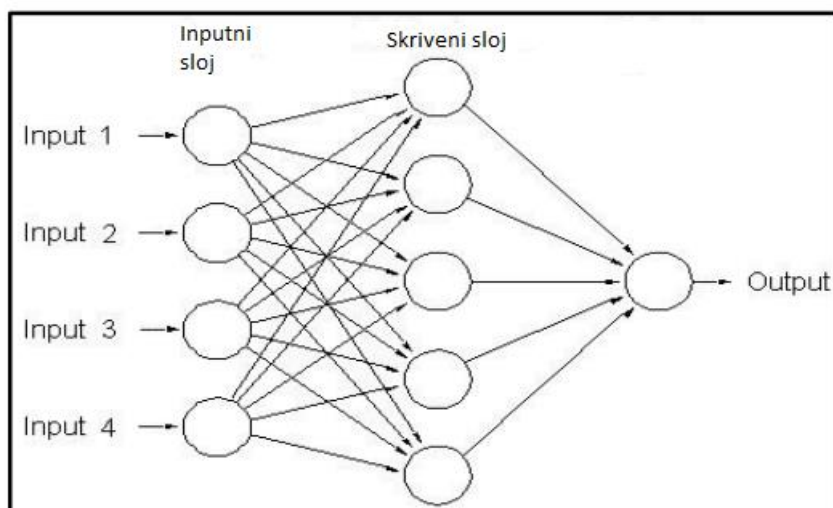
Uobičajeno se koriste dvije vrste regresijskih modela u prediktivnoj analitici za predviđanje ili prognoziranje: linearni regresijski model i logistički regresijski model. Linearni regresijski model primjenjuje se za modeliranje linearne veze između zavisnih i nezavisnih varijabli.

Linearna regresija je statistički pristup koji pomaže organizacijama u razumijevanju ponašanja kupaca, poslovnih operacija i profitabilnosti te može procijeniti trendove i generirati procjene ili prognoze u poslovanju (ProjectPro, 2023). S druge strane, logistička regresija koristi se kada postoje kategorije zavisnih varijabli. Time se nepoznate vrijednosti diskretnih varijabli predviđaju na temelju poznatih vrijednosti nezavisnih varijabli. U marketingu, algoritam logističke regresije bavi se stvaranjem modela vjerojatnosti koji prognoziraju vjerojatnost da će kupac obaviti kupnju koristeći podatke o kupcima (ProjectPro, 2023).

2.2.3. Umjetna neuronska mreža

Umjetna neuronska mreža na temelju bioloških neurona simulira sposobnosti ljudskog živčanog sustava za obradu ulaznih signala proizvodeći izlazne vrijednosti. To je sofisticirani model koji je sposoban modelirati izuzetno složene odnose. Umjetna neuronska mreža koristi se u prediktivnoj analitici kao moćno sredstvo za učenje na primjerima podataka i donošenje predviđanja na novim podacima. Kroz ulazni sloj mreže primjenjuje se uzorak obuke podataka za obradu i prosljeđuje se skrivenom sloju koji sadrži vektor neurona. Tipovi aktivacijskih funkcija koriste se na neuronima ovisno o zahtjevu outputa (izlaza). Output jednog neurona prenosi se na neurone sljedećeg sloja.

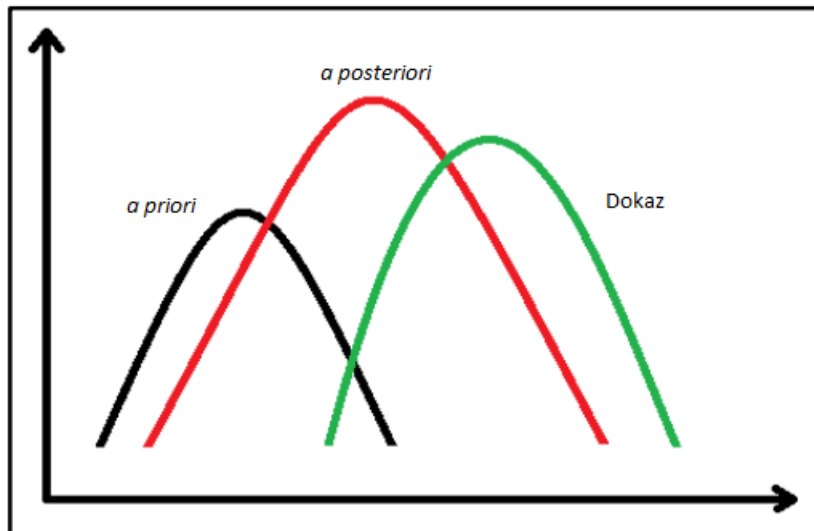
Postoje različiti modeli umjetne neuronske mreže i svaki model koristi drugi algoritam. Neki primjeri su propagacija unazad („Backpropagation“), algoritam koji je dominantan u mnogim problemima nenadgledanog učenja. U radu s nelinearnim podacima koristi se i tehnika grupiranja. U takvim problemima se neuronske mreže mogu koristiti za rad s nelinearnim odnosima u podacima. Neuronske mreže mogu se koristiti i u procjenjivanju rezultata regresijskih modela kao i stabala odlučivanja. Zbog mogućnosti prepoznavanja uzoraka, koriste se i u problemima prepoznavanja slika. Model je ilustriran na slici 4.



Slika 4: Umjetna neuronska mreža, Izvor: Kumar, V., & Garg, M. L. (2018). *Predictive analytics: a review of trends and techniques*. *International Journal of Computer Applications*, 182(1), 34 str.

2.2.4. Bayesova statistika

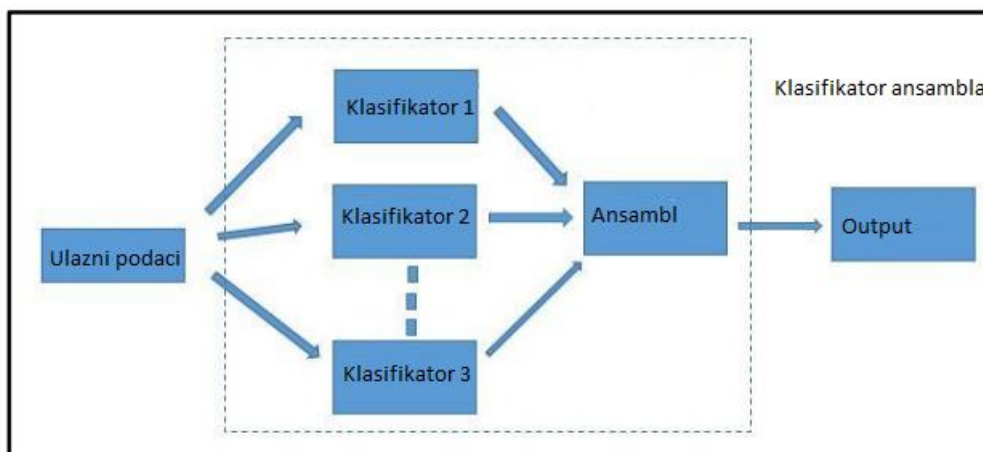
Bayesova statistika temelji se na Bayesovom teoremu kojim se opisuju događaji kao *a priori* i *a posteriori*. U uvjetnoj (kondicionalnoj) vjerojatnosti cilj je pronaći vjerojatnost naknadnog događaja s obzirom na to da se prethodni već dogodio. S druge strane, Bayesov teorem pronalazi vjerojatnost događaja *a priori* onda kada se zna da se događaj *a posteriori* već dogodio. Ono koristi vjerojatnosni model nazvan Bayesova mreža koja predstavlja uvjetnu ovisnost među slučajnim varijablama. Taj koncept može se primijeniti da bi se pronašli uzroci s rezultatima tih uzročnika. Na primjer, on se može primijeniti u pronalaženju bolesti temeljenoj na simptomima. Bayesova statistika prikazana je na slici 5.



Slika 5: Bayesova statistika, Izvor: Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. International Journal of Computer Applications, 182(1), 35 str.

2.2.5. Učenje u ansamblu

Učenje u ansamblu pripada kategoriji nadziranih algoritama u grani strojnog učenja. Ti modeli razvijeni su treniranjem nekoliko modela sličnih tipova i na kraju spajanjem njihovih rezultata u predviđanje. Na taj način se poboljšava točnost modela. Razvoj na taj način smanjuje pristranost i smanjuje varijancu modela. To pomaže u identificiranju najboljeg modela za korištenje novim podacima. Primjer modela je na slici 6.



Slika 6: Klasifikator ansambla, Izvor: Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. International Journal of Computer Applications, 182(1), 35 str.

2.2.6. Model povećanja gradijenta

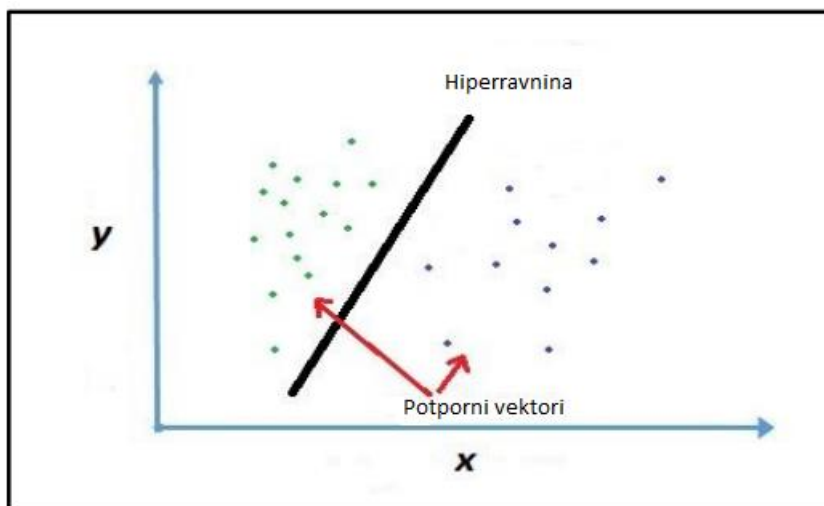
Koristi se u prediktivnoj analitici kao tehnika strojnog učenja. Najčešće se koristi u klasifikaciji i u primjenama u regresiji. On spaja slabije prediktivne modele stabala odlučivanja. Radi se o pristupu pojačavanja koji više puta uzima uzorke iz danog skupa podataka i generira rezultate kao težinski prosjek skupova iz kojih su ponovno uzorkovani podaci. Prednost tog modela je da je manje sklon „overfittanju“ tj., pretjeranoj manipulaciji modela da bude što više nalik danim podacima. Korištenje stabala odlučivanja u tom modelu pomaže u objektivnijem modeliranju podataka. Taj model prikazan je na slici 7.



Slika 7: Model povećanja gradijenta, Izvor: Kumar, V., & Garg, M. L. (2018). *Predictive analytics: a review of trends and techniques*. *International Journal of Computer Applications*, 182(1), 35 str.

2.2.7. Potporni vektorski stroj

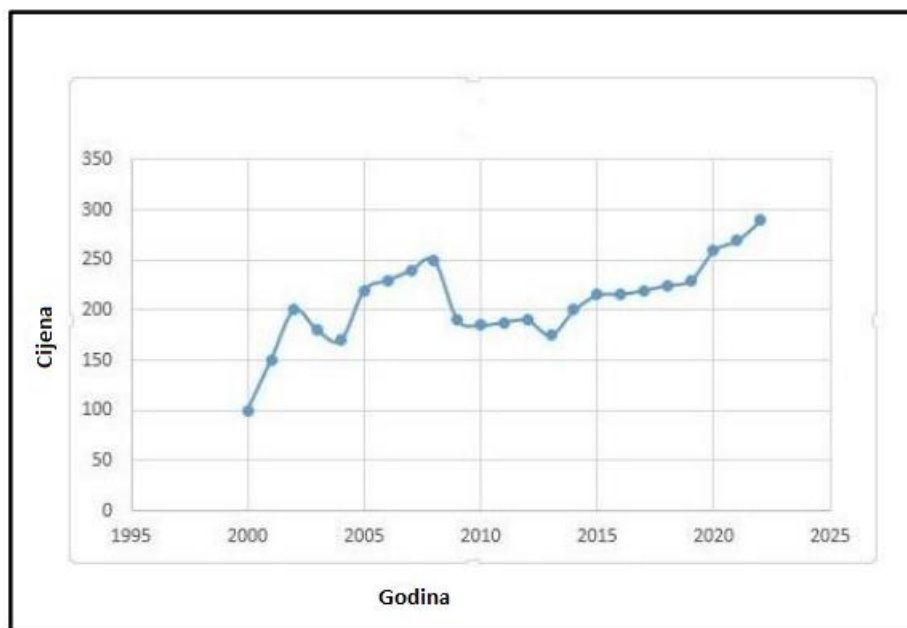
Prema Kumaru i Gargu (2018), potporni vektorski stroj je nadzirani tip strojnog učenja popularno korišten u prediktivnoj analitici. Primjenjujući algoritme povezanog učenja, analizira podatke za klasifikaciju i regresiju. Međutim, najčešće se koristi u zadacima klasifikacije. To je diskriminantni klasifikator koji je definiran hiperravninom za klasificiranje primjera u kategorije. To je predstavljanje primjera na ravnini tako da su primjeri razdijeljeni u kategorije s jasnom razlikom. Novi primjeri se onda predviđaju da pripadaju određenoj klasi prema tome koji razmak one zauzimaju. Primjer odvajanja pomoću potpornog vektorskog stroja prikazan je na slici 8.



Slika 8: Potporni vektorski stroj. Izvor: Kumar, V., & Garg, M. L. (2018). *Predictive analytics: a review of trends and techniques*. *International Journal of Computer Applications*, 182(1), 35 str.

2.2.8. Analiza vremenskih nizova

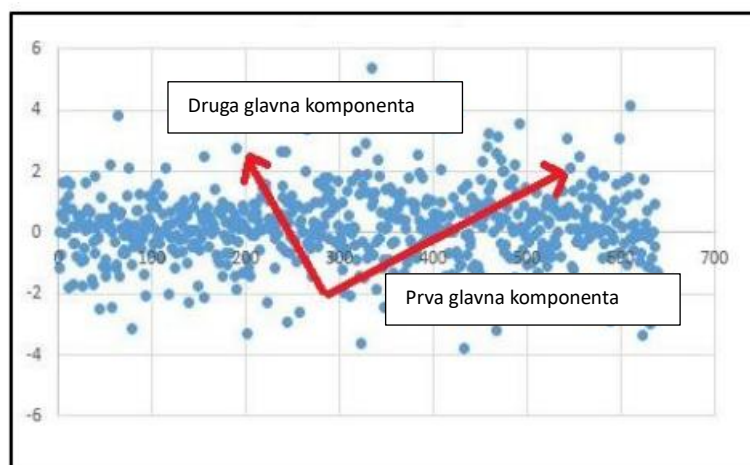
Statistička tehnika koja koristi podatke vremenskih nizova koji se prikupljaju tijekom određenog vremenskog razdoblja s određenom frekvencijom kombinira tradicionalne tehnike istraživanja podataka i predviđanja. Analiza vremenskih nizova podijeljena je u dvije kategorije, odnosno domene frekvencije i vremenske domene. Ona predviđa budućnost varijable u budućim vremenskim intervalima na temelju analize vrijednosti u prošlim vremenskim intervalima. Koristi se u predviđanju tržišta, a često i u vremenskoj prognozi. Primjer varijacije cijene nekog proizvoda tijekom razdoblja i predviđanje njegovih trendova u budućim godinama prikazan je na slici 9.



Slika 9: Analiza vremenskih serija, Izvor: Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. International Journal of Computer Applications, 182(1), 36 str.

2.2.9. Analiza glavnih komponenti

To je statistička procedura najčešće korištena u prediktivnim modelima za istraživačku analizu podataka. Ona je blisko povezana s faktorskom analizom koja se koristi za rješavanje svojstvenih vektora matrice. Također se koristi za opisivanje varijacije u skupu podataka. Primjer principa glavne komponente u skupu podataka prikazan je na slici 10.



Slika 10: Analiza glavnih komponenti, Izvor: Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. International Journal of Computer Applications, 182(1), 36. str

2.3. Implementacija

Prediktivna analitika može se koristiti u brojne svrhe. Neke koje su donijele pozitivan utjecaj će se u nastavku ukratko objasniti.

a) Sustav podrške odlučivanju u medicini

Stručnjaci koriste analizu predviđanja u zdravstvu prvenstveno da bi odredili kojim pacijentima prijete određeni uvjeti poput dijabetesa, astme, srčanih oboljenja i drugih trajnih bolesti.

b) Detekcija prijevara

Prijevara je široko rasprostranjena u ekonomiji. Slučajevi prijevara pojavljuju se u različitim poljima, poput aktivacije kreditnih kartica, računa, poreznih prijava, online aktivnosti, zahtjeva za osiguranjem i aktivnosti telefonskih poziva. Svi ti sektori zanimaju se za otkrivanje prijevara i dovođenje odgovornih pred lice pravde te sprječavanjem i praćenjem prijevara. Analiza predviđanja može im pomoći da postignu te ciljeve. Također mogu se koristiti i za otkrivanje prijevara u financijskim izvještajima tvrtke.

c) Osiguranje

Slično prijevarama, neočekivano visoki iznosi štete osiguranika najveći je problem za osiguravajuća društva koja bi željela izbjeći plaćanje takvih naknada. Iako je cilj jasan, analiza predviđanja uspjela je tek djelomično eliminirati izvor spomenutih visokih gubitaka za osiguravajuća društva čime je to obećavajuće područje za daljnja istraživanja.

d) Zdravstvo

Iako su sistematične primjene analize predviđanja u zdravstvu relativno nove, temeljna primjena je slična onoj u drugim područjima. Sve u svemu, smanjivanje rizika bolesnika je cilj. U zdravstvu to je rizik ponovne hospitalizacije, koji se može smanjiti identificiranjem pacijenata s visokim rizikom te njihovim praćenjem.

e) Financijsko predviđanje

Analiza predviđanja korisna je u financijskim predviđanjima.

f) Zadržavanje kupaca

Učestalim pregledom za usluge koje je ranije koristio, njegovim performansama, potrošnjom i drugim uzorima ponašanja, modeli predviđanja mogu odrediti vjerojatnost da će kupac u bliskoj budućnosti prekinuti uslugu.

g) Analitičko upravljanje odnosima s kupcima

Analitičko upravljanje odnosima s kupcima česta je komercijalna primjena analize predviđanja. Koristi se u primjenama marketinških kampanja, u prodaji i uslužnim djelatnostima (Yun, Shun, Junta, i Browndi, 2022).

3. Izazovi prediktivne analitike

Prediktivna analitika može generirati značajna poboljšanja u učinkovitosti, donošenju odluka i povratu ulaganja. Budući da prediktivna analitika nije uvijek uspješna, velika je vjerojatnost da većinu modela prediktivne analitike nije moguće iskoristiti operativno. Neki od najčešćih razloga zašto prediktivni modeli ne uspijevaju, grupiraju se u četiri kategorije: prepreke u menadžmentu, prepreke s podacima, prepreke u modeliranju i prepreke u implementaciji (Abbott, 2014).

a) Prepreke u menadžmentu

Da bi bili korisni, prediktivni modeli moraju biti implementirani. Često, sama implementacija zahtijeva značajnu promjenu resursa u organizaciji, stoga projekt često treba podršku menadžmenta kako bi se prešlo sa istraživanja i razvoja na operativno rješenje. Ako menadžment ne zagovara korištenje prediktivnog modeliranja, zadovoljavajući modeli ostat će neiskorišteni zbog nedostatka resursa i slabog financiranja iz političkog okruženja. Na primjer, pretpostavimo da organizacija gradi model za otkrivanje prijevara kako bi identificirala transakcije koje izgledaju sumnjivo i zahtijevaju daljnju istragu. Nadalje, pretpostavimo da organizacija može identificirati 1.000 transakcija mjesečno koje bi trebale biti dodatno istražene. Potrebno je raspodijeliti slučajeve istražiteljima, a model za otkrivanje prijevara mora biti dovoljno pouzdan da bi istražitelji mogli provesti istrage. Ako menadžment ne podržava korištenje prediktivnih modela, slučajevi se mogu podijeliti istražiteljima, no od njih neće biti ništa zbog smanjenog korištenja resursa.

b) Prepreke s podacima

Prediktivni modeli zahtijevaju podatke u obliku jedne tablice ili nekog drugog formata baze podataka. Ako su podaci spremljeni u transakcijskim bazama podataka, moraju se identificirati ključevi kako bi se spojili podaci iz izvora podataka i formirala jedinstvena tablica. Projekti mogu propasti prije nego što uopće započnu ako ključevi ne postoje u tablicama potrebnim za izgradnju podataka. Čak i ako se podaci mogu spojiti u jednu tablicu, ako primarni inputi ili outputi nisu dovoljno ili dosljedno popunjeni, podaci su besmisleni. Primjer može biti model privlačenja kupaca. Prediktivni modeli trebaju primjere kupaca koji su kontaktirani, ali nisu odgovorili, kao i onih koji su kontaktirani, ali su odgovorili. Ako se aktivni kupci pohranjuju u jednoj tablici, a marketinški kontakti (potencijalni kupci) u zasebnoj tablici, nekoliko se

problema može pojaviti pri modeliranju. Prvo, osim ako tablice kupaca ne uključuju kampanju iz koje su stekli kupca, možda će biti nemoguće rekonstruirati popis potencijalnih kupaca u kampanji zajedno s oznakom jesu li odgovorili ili nisu odgovorili na kontakt.

Drugo, ako se podaci kupaca, uključujući demografiju (dob, dohodak, poštanski broj), prepisuju kako bi se održali ažurnima, a demografija u vrijeme kada su stekli kupca nije pohranjena, tablica koja sadrži potencijalne kupce kako su se pojavili u vrijeme marketinške kampanje nikada ne može biti rekonstruirana. Kao jednostavan primjer, pretpostavimo da se telefonski brojevi dobivaju tek nakon što potencijalni kupac postane stvarni kupac. Odličan prediktor postajanja kupca tada bi bio ima li potencijalni kupac telefonski broj. Kako telefonski broj nije unesen, model nema na temelju čega predvidjeti da će pojedinac postati kupac što ilustrira problem naknadnog dodavanja podataka u model.

c) Prepreke u modeliranju

Jedna od najvećih prepreka u izgradnji prediktivnih modela s analitičke perspektive je tzv. „overfitting“, odnosno stvaranje modela koji pretjerano nalikuje podacima na kojima je model izgrađen. Time nastaje kompleksan model koji u suštini kopira podatke na kojima je treniran umjesto da ih interpretira. Učinak „overfittinga“ je dvostruk: model loše izvodi nove podatke i interpretacija modela je nepouzdana. Ako se ne pazi na eksperimentalni dizajn prediktivnih modela, opseg prekomjernog prilagođavanja modela nije poznat sve dok se model ne implementira.

Druga prepreka u izgradnji prediktivnih modela javlja se kada analitičari postanu previše ambiciozni u vrsti modela koji se može izgraditi s dostupnim podacima i u dodijeljenom vremenskom okviru. Ako pokušaju "postići veliki uspjeh" i ne mogu dovršiti model u vremenskom okviru, model se uopće neće implementirati. Često je bolja strategija prvo izgraditi jednostavnije modele kako bi se osigurao model neke vrijednosti za implementaciju. Modeli se mogu nadopuniti i poboljšati kasnije ako vrijeme dopušta.

Na primjer, može se razmotriti model zadržavanja kupaca. Ambiciozni analitičar može identificirati tisuće potencijalnih inputa u model zadržavanja kupaca, a u nastojanju da izgradi najbolji mogući model, može biti usporen čistom kombinatorikom koja se odnosi na pripremu podataka i odabir varijabli prije i tijekom modeliranja. Međutim, iz analitičarovog iskustva, on može identificirati 100 varijabli koje su povijesno bile dobri prediktori. Iako analitičar sumnja

da bi se bolji model mogao izgraditi s više potencijalnih inputa, prvi model se može izgraditi od 100 varijabli u puno kraćem vremenskom okviru.

d) Prepreke u implementaciji

Projekti prediktivnog modeliranja mogu propasti zbog prepreka u fazi implementacije modela. Sami modeli obično nisu vrlo složeni, zahtijevajući samo desetke, stotine, tisuće ili desetke tisuća množenja i zbrajanja, što se lako obrađuje današnjim procesorima.

Međutim, na fundamentalnoj razini, modeli moraju biti sposobni biti ispitani operativnim sustavom i izdavati predviđanja u skladu s tim sustavom. U transakcijskim sustavima to obično znači da se model mora kodirati u programskom jeziku kojega može pozvati sustav, poput SQL-a, C++-a, Jave ili nekog drugog programskog jezika. Ako se model ne može prevesti ili se pogrešno prevede, model je operativno beskoristan.

Ponekad prepreku predstavlja priprema podataka u obliku potrebnom za implementaciju. Ako modeliranje podataka zahtijeva spajanje nekoliko tablica kako bi se stvorila jedinstvena modelirajuća tablica, implementacija mora replicirati iste korake spajanja kako bi se izgradili podaci koje modeli trebaju za ocjenjivanje. U nekim transakcijskim sustavima s različitim podacima koji tvore tablicu modeliranja, složena spajanja možda neće biti moguća u potrebnom vremenskom okviru. Primjer može biti model koji preporučuje sadržaj koji će se prikazati na web stranici. Ako taj model treba podatke iz povijesnih uzoraka ponašanja pregledavanja za posjetitelja i stranica se treba prikazati u manje od jedne sekunde, sva dohvaćanja podataka i transformacije moraju zadovoljiti taj vremenski okvir (Abbott, 2014).

4. Prednosti prediktivne analitike

Iako postoji duga povijest rada s prediktivnom analitikom i primjenjuje se široko u mnogim područjima već desetljećima, ona je u današnjici na svojem vrhuncu zbog napretka tehnologije i ovisnosti o podacima. Mnoge organizacije teže prediktivnoj analitici kako bi povećale svoj profit, a postoji nekoliko razloga zašto je to tako:

1. Rast količine i vrsta podataka razlog je za korištenje prediktivne analitike kako bi se pronašli uvidi iz velikih podataka.
2. Brža, jeftinija i korisnički prijateljska računala su dostupna za obradu.
3. Dostupan je raznovrstan softver i razvijaju se novi softvereri koji su jednostavni za korištenje za korisnike.
4. Konkurentno okruženje rasta organizacije s profitom i ekonomskim uvjetima organizacije navodi ih na korištenje prediktivne analitike.

S razvojem interaktivnog softvera jednostavnog za korištenje i njegove dostupnosti, prediktivna analitika nije ograničena samo na statističare i matematičare. U punom zamahu koriste je poslovni analitičari i menadžeri u procesu odlučivanja. Neke od najčešćih prednosti u području prediktivne analitike navedene su u nastavku.

a) Otkrivanje prijevare

Otkrivanje i sprečavanje uzoraka kriminalnog ponašanja mogu se poboljšati kombiniranjem više metoda analize. Rast kibernetičke nesigurnosti postaje zabrinjavajući. Bihevioralna analitika može se primijeniti za nadzor nad radnjama na mreži u stvarnom vremenu. Može otkriti nepravilne aktivnosti koje mogu dovesti do prijevare. Prijetnje se također mogu otkriti primjenom tog koncepta.

b) Smanjenje rizika

Vjerojatnost neplaćanja kupca ili korisnika usluge može se unaprijed procijeniti prema kreditnom rejtingu primjenom prediktivne analitike. Kreditni rejting generira prediktivni model koristeći sve podatke koji se odnose na kreditnu sposobnost osobe. To se primjenjuje kod izdavanja kreditnih kartica i osiguravajućih društava kako bi se identificirali nepošteni klijenti.

c) Optimizacija marketinške kampanje

Kupnja određenog proizvoda može se predvidjeti primjenom prediktivne analitike. Također se može koristiti za promociju dodatnih usluga ili proizvoda kupcu. Pomaže tvrtkama da privuku i zadrže najprofitabilnije klijente.

d) Poboljšanje rada

Prognoziranje inventara i upravljanje resursima mogu se postići primjenom prediktivnih modela. Za postavljanje cijena karata aviokompanije mogu koristiti prediktivnu analitiku. Za maksimiziranje popunjenosti i povećanje prihoda, hoteli mogu koristiti prediktivne modele za predviđanje broja gostiju u određenoj noći. Organizacija može funkcionirati učinkovitije primjenom prediktivne analitike.

e) Sustav podrške odlučivanju u kliničkoj praksi

Ekspertni sustavi temeljeni na prediktivnim modelima mogu se koristiti za dijagnosticiranje pacijenata. Mogu se koristiti i u razvoju lijekova (Kumar i Garg, 2018).

5. Praktični primjeri predviđanja trendova i donošenja investicijskih odluka

Prema Nwajiaku C. (2023), donošenje odluka o investicijama složen je proces koji uključuje pronalaženje prilika za investiranje, procjenu rizika i prinosa te odabiru investicije koje odgovaraju ciljevima investitora. Analitika velikih podataka transformira taj proces pružajući investitorima pristup velikoj količini podataka koji se mogu koristiti za donošenje informiranijih investicijskih odluka.

Područje na kojem je analitika velike količine podataka („big data analysis“) imala značajan utjecaj je analiza financijskih podataka. Financijski podaci uključuju informacije o cijenama dionica, financijskim izvješćima tvrtki, ekonomskim indikatorima i drugim čimbenicima koji mogu utjecati na odluke o investiranju. Velika količina podataka omogućuje investitorima analizu tih podataka u stvarnom vremenu, otkrivajući trendove i uzorke koji se mogu koristiti za donošenje informiranijih odluka o investiranju. Na primjer, analitika podataka može se koristiti za obradu društvenih medija kako bi se identificirala osjetljivost prema određenoj dionici ili tvrtki. Analizom podataka društvenih medija investitori mogu dobiti uvid u javno mišljenje o tvrtki, što može utjecati na cijenu dionice. Također, može se koristiti za analizu novinskih članaka, priopćenja za medije i drugih izvora informacija kako bi se pronašli trendovi i uzorci koji se mogu koristiti za predviđanje kretanja cijena dionica.

Osim financijskih podataka, analitika velike količine podataka također se koristi za analizu nefinancijskih podataka, kao što su demografski podaci, podaci o ponašanju potrošača i podaci o okolišu. Ti podaci mogu pružiti uvid u potrošačke preferencije i ponašanje, što se može koristiti za identifikaciju prilika za investiranje u industrije poput zdravstva, maloprodaje i tehnologije.

Prema izvješću istraživanja koje je provela tvrtka Microstrategy (2020), 53% tvrtki koristi analitiku podataka radi smanjenja troškova i optimizacije procesa poslovanja te stvaranje strategije poslovanja. Anketa je također pokazala da tvrtke koje koriste analitiku podataka češće nadmašuju svoje konkurente u pogledu performansi i konkurentske prednosti.

Vanjski utjecaji, poput cijene sirovina, ekonomskih promjena i burze, otežali su predviđanje performansi imovine. Korištenjem tehnologija velikih podataka poput strukturnog modeliranja i prediktivne analize, tehnologija može pomoći investitorima da odrede kako će se imovina ponašati u svjetlu nadolazećih promjena na tržištu.

Prema internet izvoru Analytics Insight (2023), analitika podataka može pomoći investitorima na mnoge načine u pitanju donošenja investicijskih odluka. Navedeno je da analitikom podataka investitori mogu uočiti trendove na financijskim tržištima što im pomaže donijeti odluku o tome kada kupiti, prodati ili zadržati investiciju. Također, moguće je identificirati i kvantificirati potencijalne rizike povezane s investiranjem što omogućava investitorima da donesu bolje i informiranije odluke. Zatim, koristeći se analitikom podataka za optimizaciju portfelja, investitori si mogu osigurati da im ulaganja budu raznolika i usklađena sa svojim financijskim ciljevima. Uz to, analizom financijskih izvješća tržišnih trendova i drugih relevantnih podataka mogu pomoći investitorima u određivanju potencijala prilike za investiranje, omogućujući im donošenje utemeljenih odluka.

Organizacije u svim granama industrije koriste prediktivnu analitiku kako bi svoje usluge učinile učinkovitijima, optimizirale održavanje, pronašle potencijalne prijetnje, pa čak i spasile živote. Ovo su nekoliko primjera.

a) Rolls-Royce optimizira rasporede održavanja i smanjuje ugljični otisak

Rolls-Royce, jedan od najvećih svjetskih proizvođača zrakoplovnih motora, primijenio je prediktivnu analitiku kako bi pomogao u dramatičnom smanjenju količine ugljika u svojim motorima, dok je također optimizirao održavanje kako bi klijentima pomogao da zadrže svoje zrakoplove dulje u zraku.

b) DC Water smanjuje gubitak vode

Uprava za vodu i kanalizaciju District of Columbia (DC Water) koristi prediktivnu analitiku kako bi smanjila gubitak vode u svom sustavu. Njegov vodeći alat, Pipe Sleuth, koristi napredni model neuronske mreže dubokog učenja za analizu slike kanalizacijskih cijevi malog promjera, njihovo klasificiranje, a zatim stvaranje izvješća o procjeni stanja.

c) PepsiCo se bavi lancem opskrbe prediktivnom analitikom

PepsiCo transformira svoje prodajne timove za e-trgovinu i terensku prodaju s prediktivnom analitikom koja će mu pomoći da zna kada će trgovina na malo biti rasprodana. Tvrtka je stvorila platformu za obavještanje o prodaji (eng. Sales Intelligence Platform), koja kombinira podatke trgovaca s podacima o lancu opskrbe PepsiCo kako bi predvidjela nestašicu zaliha i upozorila korisnike da ponovno naručuju robu (Olavsrud, 2023).

S druge strane, prediktivnu analizu može koristiti gotovo svaka tvrtka koja želi unaprijediti poslovanje i pospješiti rad za postizanje boljih rezultata, bez obzira u kojoj industriji posluje.

Primjerice prodavaonice e-trgovine kako koriste prediktivnu analitiku za bolje reklamiranje svojih zaliha i povećanje prodaje. Zatim najveća svjetska internetska trgovina Amazon koja analizira tone unosa velikih podataka kako bi proizvela najbolje moguće preporuke za proizvode. Internetske trgovine ne samo da nas mogu natjerati da kupujemo više zahvaljujući toj tehnologiji, već pružaju i iskustvo koje je najviše usmjereno na kupca, što je ključno za pobjedu s konkurencijom. Nadalje, marketinški timovi koriste prediktivnu analitiku za ciljanje oglasa na temelju mrežnog (eng. „online“) ponašanja svojih potencijalnih kupaca i klijenata. Na taj način mogu potrošiti manje novca, ali stvoriti više potencijalnih kupaca. Tehnologija također omogućuje marketinškim stručnjacima stvaranje personaliziranih iskustava za kupce, pomažući im da dobiju proizvode i usluge koji točnije odgovaraju njihovim potrebama. Nadalje, Netflix, internetski prijenos videa na pretplatu koji članovima omogućuje gledanje serija i filmova na uređaju, kao jedan od izvrsnih primjera kada je riječ o marketingu temeljenom na povijesnim podacima. Oni ne samo da daju vrlo personalizirane preporuke sadržaja već također pokušavaju predvidjeti popularnost svojih filmova i emisija, kao i optimizirati proizvodnju sadržaja i druge interne poslovne procese (Trachim, 2023).

6. Zaključak

Prediktivna analitika predstavlja proces korištenja podataka, statističkih algoritama i strojnog učenja kako bi se predvidjeli budući događaji i trendovi. U posljednjih nekoliko godina prediktivna analitika postala je izuzetno korisna u donošenju investicijskih odluka i predviđanju trendova na tržištima. Korištenjem naprednih algoritama strojnog učenja i velikih količina podataka, investitori mogu dobiti vrijedne uvide u buduće trendove i rizike, što im omogućuje donošenje informiranih odluka o ulaganjima. Ti algoritmi omogućuju analizu velikih količina podataka i stvaranje preciznih predviđanja.

Na primjer, mogu se koristiti za predviđanje kretanja cijena dionica ili kretanja tržišta u cjelini. Međutim, važno je imati na umu da prediktivna analitika nije savršena i da se temelji na pretpostavkama i procjenama. Istaknuto je da se prediktivna analitika, zajedno s većinom prediktivnih modela i tehnika rudarenja podataka oslanja na sve sofisticiranije statističke metode, uključujući i složenije analize kao što su napredne regresije ili vremenski serijski modeli čime naglašava svoju kompleksnost.

Također se susreće s mnogim izazovima kao što su prepreke u menadžmentu, prepreke s podacima, prepreke u modeliranju i prepreke u implementaciji. S druge strane, ona može biti izuzetno korisna u predviđanju trendova i donošenju investicijskih odluka pružajući investorima pristup velikoj količini podataka koji se mogu koristiti za donošenje informiranijih investicijskih odluka. Omogućuje im veću efikasnost u investiranju te identifikaciju i kvantifikaciju potencijalnih rizika povezane s investiranjem kao i uočavanje prilika.

Međutim, važno je koristiti ju kao jedan od alata u procesu donošenja odluka, a ne kao jedini faktor. Investitori bi trebali uzeti u obzir i druge faktore koji utječu na tržište, te biti spremni uložiti vrijeme i resurse u učenje i razumijevanje tog procesa kako bi ga pravilno primijenili.

Literatura

1. Abbott, D. (2014). *Applied predictive analytics: Principles and techniques for the professional data analyst*. John Wiley & Sons.
2. Analytics Insight (2023). Leveraging Data Analytics in Investment Decisions: How Seeking Alpha Can Help. Dostupno na: <https://www.analyticsinsight.net/leveraging-data-analytics-in-investment-decisions-how-seeking-alpha-can-help/> [pristupljeno 22.6.2023.]
3. Kelleher, D. J., Namee, M. B., D'Arcy, A. (2015). *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*. Cambridge, Massachusetts London, England: The MIT Press
4. Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. *International Journal of Computer Applications*, 182(1), 31-37.
5. Nwajiaku C. (2023). *The Role Of Big Data Analytics In Investment Decision-Making*. Dostupno na: <https://www.smashnegativity.com/the-role-of-big-data-analytics-in-investment-decision-making/#penci-Role-of-Big-Data-Analytics-in-Investment-Decision-Making> [pristupljeno 20.6.2023.]
6. Nyce, C., & Cpcu, A. (2007). Predictive analytics white paper. *American Institute for CPCU. Insurance Institute of America*, 1-15.
7. Ogunleye, J. (2014). The concepts of predictive analytics. *International Journal of Developments in Big Data and Analytics*, 1(1), 82-90.
8. Olavsrud, T., Edwards, J. (2023), *What is predictive analytics? Transforming data into future insights*. Dostupno na: <https://www.cio.com/article/228901/what-is-predictive-analytics-transforming-data-into-future-insights.html> [pristupljeno: 12.7.2023.]
9. Pagans, F. G. (2015). *Predictive Analytics Using Rattle and Qlik Sense*. Packt Publishing Ltd.
10. ProjectPro (2023). Predictive Modeling Techniques- A Comprehensive Guide. Dostupno na: https://www.projectpro.io/article/predictive-modelling-techniques/598#mctoc_1g5gnldk23t [pristupljeno: 25.8.2023.]
11. Trachim, A. (2023.), *8 industry examples of predictive analytics*, dostupno na: <https://indatalabs.com/blog/predictive-analytics-examples> [pristupljeno 12.7.2023.]
12. Wessler, M. (2014), *Predictive analytics for dummies: Wiley, Alteryx Special Edition*, Hoboken, NJ: John Wiley & Sons, Inc
13. Winters, R. (2017). *Practical predictive analytics*. Packt Publishing Ltd.

14. Wu, J., Coggeshall, S. (2012). *Foundations of predictive analytics*. Boca Raton, FL: CRC Press, Chapman & Hall/CRC data mining and knowledge discovery series
15. Yun, C., Shun, M., Junta, U., & Browndi, I. (2022). Predictive Analytics: A Survey, Trends, Applications, Opportunities' and Challenges for Smart City Planning. *International Journal of Computer Science and Information Technology*, 23(56), 226-231.
16. 2020 Global State of Enterprise Analytics: Minding the Data-driven Gap (2020), MicroStrategy. Dostupno na: <https://www.microstrategy.com/en/go/always-on-campaign/product-track/federated-analytics/the-global-state-of-enterprise-analytics-2020> [pristupljeno: 27.8.2023.]

Popis slika

Slika 1: Proces prediktivne analitike, Izvor: Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. International Journal of Computer Applications, 182(1), 32. str.	4
Slika 2: Stablo odlučivanja, Izvor: Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. International Journal of Computer Applications, 182(1), 34. str.	9
Slika 3: Regresijski model, Izvor: Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. International Journal of Computer Applications, 182(1), 34. str.	10
Slika 4: Umjetna neuronska mreža, Izvor: Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. International Journal of Computer Applications, 182(1), 34 str.	12
Slika 5: Bayesova statistika, Izvor: Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. International Journal of Computer Applications, 182(1), 35 str.	13
Slika 6: Klasifikator ansambla, Izvor: Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. International Journal of Computer Applications, 182(1), 35 str.	13
Slika 7: Model povećanja gradijenta, Izvor: Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. International Journal of Computer Applications, 182(1), 35 str.	14
Slika 8: Potporni vektorski stroj, Izvor: Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. International Journal of Computer Applications, 182(1), 35 str.	15
Slika 9: Analiza vremenskih serija, Izvor: Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. International Journal of Computer Applications, 182(1), 36 str.	16
Slika 10: Analiza glavnih komponenti, Izvor: Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. International Journal of Computer Applications, 182(1), 36. str.	16