

Prediktivno modeliranje ponašanja korisnika društvenih mreža

Kristijan, Perković

Master's thesis / Diplomski rad

2020

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Organization and Informatics / Sveučilište u Zagrebu, Fakultet organizacije i informatike**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:211:373507>

Rights / Prava: [Attribution-NonCommercial-NoDerivs 3.0 Unported / Imenovanje-Nekomercijalno-Bez prerađivanja 3.0](#)

Download date / Datum preuzimanja: **2024-07-06**



Repository / Repozitorij:

[Faculty of Organization and Informatics - Digital Repository](#)



**SVEUČILIŠTE U ZAGREBU
FAKULTET ORGANIZACIJE I INFORMATIKE
VARAŽDIN**

Kristijan Perković

**PREDIKTIVNO MODELIRANJE
PONAŠANJA KORISNIKA DRUŠTVENIH
MREŽA**

DIPLOMSKI RAD

Varaždin, 2020.

SVEUČILIŠTE U ZAGREBU
FAKULTET ORGANIZACIJE I INFORMATIKE
V A R A Ž D I N

Kristijan Perković

Matični broj:

Studij: Informacijsko i programsko inženjerstvo

PREDIKTIVNO MODELIRANJE PONAŠANJA KORISNIKA
DRUŠTVENIH MREŽA

DIPLOMSKI RAD

Mentor/Mentorica:

Doc. dr. sc. Oreški Dijana

Varaždin, svibanj 2020.

Kristijan Perković

Izjava o izvornosti

Izjavljujem da je moj završni/diplomski rad izvorni rezultat mojeg rada te da se u izradi istoga nisam koristio drugim izvorima osim onima koji su u njemu navedeni. Za izradu rada su korištene etički prikladne i prihvatljive metode i tehnike rada.

Autor/Autorica potvrdio/potvrdila prihvaćanjem odredbi u sustavu FOI-radovi

Sažetak

U ovome radu provest će se predviđanje zadovoljstva korisnika društvenih mreža. Smatra se kako korisnici društvenih mreža provode jako puno vremena na njihovo korištenje te je iz toga razloga važno da su oni zadovoljni istima. Za provođenje predviđanja koristit će se skup podataka koji se dobio popunjavanjem ankete koja je prethodno kreirana i podijeljena na društvenim mrežama kako bi korisnici mogli izraziti svoje mišljenje i dati odgovore na pitanja koja se odnose na njihovu aktivnost na društvenim mrežama. Pitanja su kreirana na način da kreiraju profil osobe, odnosno da se dobije informacija o tome koliko osoba provodi vremena na društvenim mrežama, koliko prijatelja/pratitelja ima, koliko lajkova (sviđanja) dobiva po objavi, komentira li objave ostalih osoba i slično. Anketu je popunilo 390 osoba. Predviđanje će se provesti kroz tri metode koje ćemo potkrijepiti teorijom te će se na temelju predviđanja donijeti zaključak o karakteristikama pojedinih osoba za dodjeljivanje ocjene zadovoljstva.

Ključne riječi: društvene mreže, neuronska mreža, Bayesova mreža, stablo odlučivanja, BigML, Netica

Sadržaj

1. Uvod	1
2. Prethodna istraživanja	2
2.1. Priča o dvije društvene mreže: Twitter protiv Facebook-a i prediktori osobnosti društvenih mreža	2
2.2. Modeliranje upotrebe Facebook-a među studentima na Tajlandu: uloga emocionalne privrženosti na prošireni model prihvaćanja tehnologije	4
2.3. Ispitivanje analize korisničkog ponašanja na društvenim mrežama	6
2.4. Razumijevanje ponašanja korisnika na društvenim mrežama: Istraživanje.....	10
2.5. Istraživanje korisničkog ponašanja na društvenim mrežama	12
2.6. Računala u ponašanju ljudi	14
3. Opis podataka.....	16
3.1. Pitanja s ankete	16
3.2. Prikaz dobivenih rezultata	17
4. Stablo odlučivanja.....	21
4.1. Izrada stabla odlučivanja u BigML-u.....	24
4.2. Prikaz dobivenih rezultata	27
4.3. Izrada ekspertnog sustava u XRKB-u	29
4.4. Testiranje sustava.....	32
5. Neuronske mreže	37
5.1. Kreiranje neuronske mreže	39
5.2. Primjena neuronskih mreža.....	40
5.2.1. Mreža širenja unatrag.....	41
5.2.2. Vjerojatnostna mreža	41
5.2.3. Mreže s radijalno zasnovanom funkcijom	42
5.3. Primjer zadatka neuronskih mreža:	42
5.4. Izrada neuronske mreže pomoću alata BigML	44
6. Bayesove mreže	47
6.1. Bayesova formula	47
6.2. Primjer primjene Bayesove formule.....	47
6.3. Teorija grafova.....	48
6.4. Definicija Bayesove mreže	49
6.4.1. Postupak kreiranja Bayesove Mreže	49
6.4.1.1. Primjer kreiranja Bayesove mreže	49
6.4.1.2. Propagacija unaprijed i unazad	51
6.5. Izrada Bayesove mreže korištenjem alata Netica	52
6.6. Interpretacija dobivenih rezultata.....	54

7. Usporedba rezultata dobivenih kroz tri metode predviđanja	56
8. Zaključak	58
9. Literatura	59
Popis slika	61
Popis tablica	63

1. Uvod

U današnje vrijeme društvene mreže se koriste svakodnevno, nerijetko izazivaju ovisnost. Korisnici društvenih mreža u velikom broju slučajeva koriste društvene mreže za komunikaciju sa sebi važnim osobama, ponekad su to osobe koje žive u njihovoj blizini, a ponekad se radi o osobama koje žive daleko od njih. Na društvenim mrežama se nerijetko nalazi veliki broj reklama, kao i profili koji sadrže poveznice na informativne web stranice kao što je na primjer Dnevnik.hr. Iz toga razloga korisnici društvenih mreža sve manje kupuju novine, gledaju televiziju te sve manje koriste tradicionalne izvore informacija. Mlađi korisnici društvenih mreža kreiraju razredne grupe unutar kojih dijele materijale vezane za nastavu, sportski klubovi kreiraju grupe unutar kojih obavještavaju svoje članove o rasporedu treninga, bliski prijatelji kreiraju grupe unutar kojih dijele sadržaj koji žele zadržati unutar manjeg kruga ljudi. Možemo primijetiti da se društvene mreže koriste u različite svrhe te da njihovi korisnici provode jako puno vremena na njima. Zanima nas koliko su korisnici društvenih mreža zadovoljni samim društvenim mrežama koje koriste jer jedan veliki dio svojega vremena izdvajaju na korištenje istih. Iz toga razloga u ovome radu će se provesti predviđanje koliko su korisnici društvenih mreža zadovoljni društvenim mrežama koje koriste.

Prvo će se prikazati podaci o ranijim istraživanjima koja su povezana s društvenim mrežama, nakon toga ćemo prikazati prikupljene podatke putem ankete. Predviđanje ćemo provesti pomoću tri metode, kao skup podataka za učenje koristit će se podaci prikupljeni putem ankete. Predviđanje će se prvo provesti metodom stabla odlučivanja, zatim metodom neuronske mreže te kao posljednju metodu predviđanja koristit ćemo Bayesovu mrežu. Prije samoga provođenja predviđanja pomoću pojedine metode navest ćemo teoriju koja se odnosi na tu metodu. Pod pojmom teorije podrazumijevamo način na koji se dolazi do samih rezultata koji se dobivaju pomoću metoda predviđanja, ukratko to znači da ćemo objasniti način kreiranja stabla odlučivanja, način kreiranja neuronske mreže te način kreiranja Bayesove mreže. Nakon navedene teorija provest ćemo predviđanje pojedinom metodom te prokomentirati dobivene rezultate, nakon primjene sve tri metode predviđanja napraviti ćemo usporedbu rezultata te donijeti zaključak.

2. Prethodna istraživanja

U ovome poglavlju obrađena su šest prethodna istraživanja koja su povezana s društvenim mrežama. Svako istraživanje se baziralo na različitom skupu podataka te na različitoj skupini korisnika. Pojedina istraživanja su istraživala samo skupinu studenata, što je jako povezano sa skupom podataka koji se koristi u ovome istraživanju jer su studenti u najvećem broju popunili anketu koja predstavlja skup podataka u ovome istraživanju.

2.1. Priča o dvije društvene mreže: Twitter protiv Facebook-a i prediktori osobnosti društvenih mreža

Stranice društvenih mreža sve brže postaju jedan od najpopularnijih alata za društvenu interakciju i razmjenu informacija. Prošla istraživanja su pokazala vezu između korisnikove osobnosti i korištenja društvenih mreža. Koristeći opći uzorak populacije (N=300), ovo istraživanje nastavlja istraživanja ispitujući uzajamne veze ličnosti (neuroticizam, ekstraverzija, otvorenost, ugodnost, savjesnost, društvenost i potrebu za spoznajom) društvene i informacijske potrebe dviju najvećih društvenih mreža, Facebook-a i Twitter-a. [2]

Facebook-ova popularnost se povećala eksponencijalno u zadnjih nekoliko godina, s 5,5 milijuna aktivnih korisnika u 2005. godini do oko 500 milijuna aktivnih korisnika u 2011. godini. Facebook omogućuje korisnicima da objavljuju osobne informacije o sebi, gdje rade, što im je hobi ili zanimacija, koje su religije te kakve političke poglede imaju, što vole slušati te koje filmove i serije vole. Putem „Messenger-a“ mogu komunicirati s drugim ljudima koji su im prvotno „postali prijatelji“ na Facebook društvenoj mreži. Također na svom „zidu“ mogu pisati objave koje onda vide njihovi prijatelji i ostali ljudi (ukoliko su to u postavkama dopustili) (Huberman, Romero, & Wu, 2009). [2]

Iako je Facebook najveća društvena mreža, postoje i druge. Sve društvene mreže olakšavaju online, društvenu interakciju, ali ne pružaju jednaku uslugu ili nemaju fokus na istim stvarima kao Facebook. Najzanimljivija društvena mreža poslije Facebook-a je Twitter, čiji je fokus i svrha objavljivanje i dijeljenje mišljenja i informacija za razliku od interakcije s drugim ljudima (Huberman, Romero, & Wu, 2009). [2]

Twitter omogućuje korisnicima da svoj račun ažuriraju kratkim rečenicama, koje se nazivaju „tweets“ te su limitirane na 140 znakova. Drugi korisnici mogu pratiti ažuriranja. Twitter strelovito raste te prema podacima iz Siječnja 2010. godine ima 73,5 milijuna jedinstvenih korisnika i gledatelja objava (Twitter, 2011). [2]

Tijekom istraživanja osobnosti u korištenju interneta, istraživači su koristili „Five-Factor-Model“ ili „Big-Five“. Ti se modeli sastoje od 5 osobina ličnosti, a to su: neuroticizam,

ekstraverzija, otvorenost, ugodnost te savjesnost. Iako teoretske i metodološke podloge modela mogu biti upitne, smatra se kako su to neki od bitnih aspekata ličnosti (McCrae & Costa, 1999). [2]

Neuroticizam se definira kao mjera utjecaja i emocionalne kontrole s niskim razinama koja sugerira dobru kontrolu emocija i stabilnost gdje pojedinci s visokim razinama možda mogu biti senzitivne i nervozne na neki način sa sklonosti za brigu. Ranija istraživanja su pokazivala kako osobe s visokom razinom neuroticizma vrlo vjerojatno izbjegavaju Internet, međutim empirijsko ispitivanje nije uspjelo dokazati tu tezu. Sada se zaključuje kako te osobe zapravo često koriste Internet, upravo kako bi izbjegle usamljenost, a to se posebno u istraživanju pokazalo za osobe ženskog spola. [2]

Ekstroverti su tipično pustolovan, društven i pričljiv tip osobe dok su introverti tihi i sramežljivi. Ekstraverzija se pokazala da je u korelaciji s korištenjem „messenger“ poruka i društvenih mreža. Ekstroverti su se pokazali kao članovi više grupa, imali su više prijatelja na društvenim mrežama iako se čini kako ta prijateljstva nisu započeta online. Upravo zbog svoje naravi ekstroverti uspijevaju skupiti puno poznanika ili prijatelja u pravom životu pa ovakav rezultat zapravo i ne čudi. [2]

Pojedinci koji demonstriraju visoku otvorenost prema iskustvu imaju široke interese i traže novosti i nove stvari za isprobati. Otvorenost je u korelaciji s pisanjem poruka online, tj. Online dopisivanjem te s korištenjem većeg broja opcija i značajki društvenih mreža. Također, otvorenost je u relaciji s traženjem informacija. [2]

Ugodnost je zapravo mjera o tome koliko su ljudi srdačni i prijateljski nastrojeni, što znači da osobe s visokom razinom ugodnosti su one koje su ljubazne, simpatične i tople. Zbog toga se vjerovalo kako pojedinci koji imaju nisku razinu ugodnosti imaju više online prijatelja jer im je teško pronaći i zadržati prijatelje u stvarnom životu. Međutim, najnovija istraživanja su pokazala kako zapravo razina ugodnosti i broj prijatelja online i općenito društvene mreže nisu u nekoj relaciji jer su rezultati bili miješani. [2]

Savjesnost se odnosi na radnu etiku osobe, urednost i ustrajnost. Vjerovalo se da osobe koje su savjesne izbjegavaju društvene mreže jer one promoviraju oklijevanje i služe kao distrakcija od stvarnosti i bitnih zadataka. Bilo kako bilo, ova se teza nije uspjela dokazati nijednim istraživanjem. Ali, dokazalo se kako će ljudi koji su više savjesni koristiti društvene mreže u svrhu skupljanja informacija i istraživanja vezanih za njihov posao i za obavljanje nekih bitnih stvari, dok oni koji su manje savjesni će stvarno koristiti društvene mreže za traćenje vremena. [2]

Kratki osvrt na istraživanja:

- Neuroticizam će biti pozitivan u korelaciji s društvenim korištenjem Facebook-a i Twitter-a
- Ekstraverzija će biti pozitivno u korelaciji s korištenjem Facebook-a te u negativnoj korelaciji s korištenjem Twitter-a
- Otvorenost će biti u korelaciji s društvenim i informacijskim korištenjem Facebook-a i Twitter-a
- Ugodnost će biti potpuno ne povezana s korištenjem društvenih mreža
- Savjesnost je u pozitivnoj korelaciji s informacijskim korištenjem društvenih mreža

Dakle, Internet i društvene mreže su promijenile način traženja informacija i način komunikacije s drugim osobama te brzo postaju jedan od najdominantnijih načina za društvenu interakciju i širenje informacija. Sa sve više pojedinaca koji počinju koristiti društvene mreže, važno je razumjeti tko koristi koje društvene mreže te iz kojih razloga. [2]

U istraživanju se koristilo nekoliko metoda i analiza istraživanja. Prva metoda je bila korištenje društvenih mreža, koja je pokazala kako sudionici provode od 0,25 do 25h tjedno na društvene mreže. Sljedeća je bila analiza istraživačkog faktora (eng. Exploratory factor analysis) u kojoj je provedeno 12 Facebook i Twitter predmeta koristeći ponderirane najmanje kvadrata sredstva i različitosti metode. Još su se koristile: analiza potvrdnog faktora, model mjerenja, modeliranje strukturalnih jednadžbi, korelacijska analiza te razlike osobnosti prema preferencijama društvenih mreža. [2]

Rezultati ovog velikog istraživanja su pokazali kako su osobine ličnosti iz „Big-Five“ modela povezane s društvenosti i razmjenom informacija u online okruženju društvenih mreža. Pokazalo se kako je osobnost u relaciji, kako te korelacije nisu izravne ili jednako utjecajne kao neka prijašnja istraživanja. Također, pokazala se diferencijalna veza između ponašanja na Facebook-u i Twitter-u, sugerirajući da različiti ljudi koriste iste društvene mreže za različite svrhe.[2]

2.2. Modeliranje upotrebe Facebook-a među studentima na Tajlandu: uloga emocionalne privrženosti na prošireni model prihvaćanja tehnologije

Cilj ovog istraživanja je ispitati faktore koji utječu na korištenje Facebook-a kod studenata. Korištenjem proširenog modela tehnološkog prihvaćanja (eng. technology acceptance model) s emocionalnom vezanosti kao vanjskom varijablom, uzorak od 498 studenata iz državnog

sveučilišta u Tajlandu je ispitano na temelju njihovih odgovora na 5 pretpostavljenih varijabli kako bi se predvidjela njihova stvarna aktivnost na Facebook-u. Podaci su analizirani korištenjem modela strukturalnih jednadžbi te je rezultat pokazao kako su uočena korisnost, stav prema korištenju tehnologije te emocionalna vezanost imale direktan i veliki utjecaj na stvarno korištenje Facebook-a. [3]

Tehnološka prihvaćenost odnosi se na želju osobe za korištenjem tehnologije za zadatke za koje je ta tehnologija i služi. Kako bi razumjeli zašto studenti žele koristiti tehnologiju, istraživači su se koristili teorijama i modelima informatičke znanosti i psihologije. Tako je za objašnjenje korištenja tehnologije kod studenata široko korišten model tehnološkog prihvaćanja. Taj model ima svoje korijene u teoriji obrazloženog djelovanja koja govori kako su namjere ponašanja pojedinca pokrenute njegovim stavovima i subjektivnim normama. Također, u ovom modelu uočena korisnost te uočena jednostavnost uporabe se ponašaju kao prethodnici stavovima prema tehnologiji. To ukazuje kako korisnici razvijaju pozitivne stavove prema korištenju tehnologije jer opažaju kako je tehnologija korisna i jednostavna za korištenje. [3]

U ovom istraživanju, emocionalna povezanost je dodana kao vanjska varijabla. Dok su psihološki i društveni konstrukti široko korišteni u literaturi tehnološkog prihvaćanja, nekoliko istraživača je istražilo utjecaj emocionalne povezanosti s okvirom tehnološke prihvaćenosti. Paušalna istraživanja koja ispituju emocionalnu povezanost kao varijablu više se fokusiraju na radnu okolinu i prostor osobe nego na samu osobu. Povezanost je tu definirana kao stupanj privrženosti drugim sustavima, baziranim na efektu društvenog utjecaja na korisnika. Usprkos vezama između emocionalne povezanosti i tehnološke prihvaćenosti, nekoliko istraživanja je koristilo model tehnološke prihvaćenosti s emocionalnom povezanosti kao vanjskom varijablom. [3]

Kao što je već navedeno, u ispitivanju je sudjelovalo 498 ispitanika iz državnog sveučilišta u Tajlandu. Bilo je 313 ženskih studenata što je 62,9% ispitanika te 85,9% ukupnog broja ispitanika su stari između 25 i 34 godine. Također, odgovori su se prikupljali online te su bili potpuno anonimni. Pitanja na koja su morali odgovarati ispitanici su bili na temelju skale, točnije imali su skalu od 1 do 7 gdje je za svako pitanje 1 značilo da se ne slažu uopće s navedenom tvrdnjom dok je 7 značilo da se slažu u potpunosti te je ukupno bilo 19 tvrdnji na koje su morali dati svoj odgovor. [3]

Rezultat istraživanja je da srednja vrijednost svih 19 tvrdnji je iznad srednje točke od 4.0, u rasponu od 4,41 do 5,96, što zapravo znači kako su ispitanici generalno davali pozitivne odgovore na pitanja. Standardna devijacija je u rasponu od 0,91 do 1,60, što znači kako su se odgovori ravnopravno podijelili oko srednje vrijednosti. Koristila se i potvrдна analiza faktora u 2 koraka. U prvom se koraku provjeravalo koliko dobro su promatrani pokazatelji mjerili

nezamijećene varijable. U drugom koraku, strukturalni dio analize koji specificira veze između latentnih varijabli je bio procijenjen. [3]

Zaključak istraživanja je da studentske preference za iskustvo s digitalnim alatima se mijenja tijekom vremena i s njihovom životnom situacijom. Kako bi optimizirali tehnološke inovacije za učenje, potrebno je razumjeti kako i zašto studenti koriste Facebook. Suprotno učenju Facebook nije primarno smatran edukacijskim alatom. Nadalje, ne postoji neka institucijalna kontrola kako studenti troše svoje vrijeme na Facebook-u. S obzirom na trenutno istraživanje o digitalnim domaćinima u visokom obrazovanju, dolazi se do zaključka kako studenti očekuju da će se angažirati s razumnim standardom tehnologije za učenje i interakciju na njihovom mjestu učenja. Takvo angažiranje uključivalo bi vezu na Internet, učenje i čitanje izvora te lako dostupnu korisničku podršku bilo kada i bilo gdje. [3]

2.3. Ispitivanje analize korisničkog ponašanja na društvenim mrežama

Postoje istraživanja o ponašanju korisnika na društvenim mrežama u literaturi. Ipak, potrebno je više istraživanja kako bi se mogli opaziti razni izazovi u točnim analizama ponašanja korisnika. Većina društvenih mreža je po prirodi dinamično te to predstavlja ogromne izazove u karakterizaciji točnog ponašanja. U ovom istraživanju se pruža strukturirani i sveobuhvatni pregled raznih aspekata analize ponašanja korisnika na društvenim mrežama. Kategoriziraju se postojeće analize metoda i sustava društvenih mreža baziranih na temeljnim računalnim tehnikama koje korisnik koristi. Detaljno se opisuje i uspoređuje veliki broj analiza korisničkog ponašanja na društvenim mrežama s posebnim isticanjem karakterizacije ponašanja, prepoznavanja ponašanja te predviđanje ponašanja u dvije široke perspektive – analiza postojanog korisničkog ponašanja i nepostojanog korisničkog ponašanja. Dodatno, izdvojeni su neki važni pravci u analizi korisničkog ponašanja na društvenim mrežama. Također, raspravlja se o raznim tipovima skupova podataka koji mogu biti korišteni od strane istraživača kod tih analiza. [4]

Društvena mreža generira ogromne podatke dinamično na brži način sa sudjelovanjem milijuna korisnika širom svijeta u raznim oblicima društvenih mreža kao što su: mreža sljedbenika, mreža praćenja, mreža citiranja, wikipedia, twitter, itd. Ovi razni tipovi mreža kategoriziraju korisnike u drugačije grupe kako bi ih analizirali, odnosno saznali popularne korisnike za statistiku prodaje proizvoda za ažuriranje mrežnih performansi, dizajna i ostalog. Ova kategorizacija korisnika u razne grupe bazirana na različitim kriterijima grupiranja, pravilima odluka, skrivenim tokom i napokon to utječe na obrasce aktivnosti i mrežne obrasce nekog generalnog korisnika. Ti obrasci su ponekad ovisni o samoj društvenoj mreži,

fleksibilnosti i bogatim svojstvima društvenih mreža i napokon to stvara oblik određenog korisničkog ponašanja. Neka od tih ponašanja su direktno podložna računanju iz podataka društvenih mreža javno dostupnih koji su trajno ponašanje u ovoj konceptualizaciji i neki koji možda neće biti objavljeni iz baze podataka direktno. Oni se smatraju nestalnim korisničkim ponašanjem na društvenim mrežama. To ponašanje može biti indirektno izračunato sa skrivenih sadržaja kao što su zapisnici pristupa web poslužitelju, skriveni tijek rada i slično. Cijelo ovo istraživanje je podijeljeno u gore dvije navedene kategorije koristeći razne metodologije kao što su kopanje podataka, uključujući klasteriranje, strojno učenje, pristup vođen podacima i drugi. [4]

Kao što je već navedeno, analiza korisničkog ponašanja je jako popularna te postoji mnogo istraživanja na tu temu. Reza Motamedi je proveo detaljno istraživanje mjerenja kako bi kategorizirao i usporedio ponašanje korisnika na Facebook-u, Twitter-u te Google-u na razini određenih grupa. Koristile su se metrike, a to su: aktivnost korisnika, reakcije korisnika te povezivanje korisnika. Rezultat takvog istraživanja se čak koristio i za pronalazak terorista i terorističkih grupa. Postoje neke doktorske teze gdje su istraživači pokušali karakterizirati, prepoznati i predvidjeti korisničko ponašanje na društvenim mrežama. [4]

Istraživanje analize korisničkog ponašanja je važno područje pretraživanja u računalnoj društvenoj mreži iz raznih razloga kao što su: identifikacija kriminalnih aktivnosti, detekcija anomalija, širenje informacija u kratkim vremenima, istraživanje osjećaja većih društvenih skupina, procjena popularnosti i slično. Ovdje su podaci prikupljeni s društvenih mreža kako bi se analiziralo ponašanje korisnika kvantitativno i kvalitativno. [4]

Ovo istraživanje razlikuje se od postojećih po sljedećem: raspravlja se o izvorima, uzrocima i aspektima korisničkog ponašanja na društvenim mrežama te je također uključena detaljna diskusija analize ponašanja. Pokušava se pružiti klasifikacija raznih analiza ponašanja korisnika na društvenim mrežama, sustava te softverskih alata za klasifikaciju društvenih mreža i njihovih karakteristika. No, unatoč svemu postoje određeni problemi s analizom ponašanja korisnika. Kako bi se što jasnije objasnilo ponašanje korisnika, potrebno je pronaći uzorke u komunikaciji i interakciji između korisnika društvenih mreža ovisno o društvenoj mreži koju koriste. Točnije, vrlo važno je pronaći najutjecajniji uzorak za pojedinu društvenu mrežu. Neki od glavnih doprinosa ovog istraživanja su navedeni u nastavku:

- Korisničko ponašanje u mrežnim društvenim mrežama je dobro proučeno, ali korelacija ovog ponašanja između ovih mreža još nije jasno riješena. Pokušalo se povezati analizu ponašanja korisnika društvene mreže na jednostavan način s primjerima.
- Analiza ponašanja korisnika svrstana je u dvije široke perspektive: postojana i nepostojeća. Ovo je novi način gledanja na analizu ponašanja korisnika na društvenim mrežama.

- Većina postojećih anketa ne pokriva sve pristupe do danas za analizu ponašanja korisnika društvenih mreža, ali pokriven je širok raspon pristupa koji su dostupni u literaturi do 2016. godine.
- Većina postojećih istraživanja izbjegava karakterizaciju i prikaz ponašanja korisnika društvenih mreža, koji su presudni u zadatku analize ponašanja. Predstavljeno je nekoliko tehnika za karakterizaciju ponašanja iz dostupnih skupova podataka te se uspoređuju

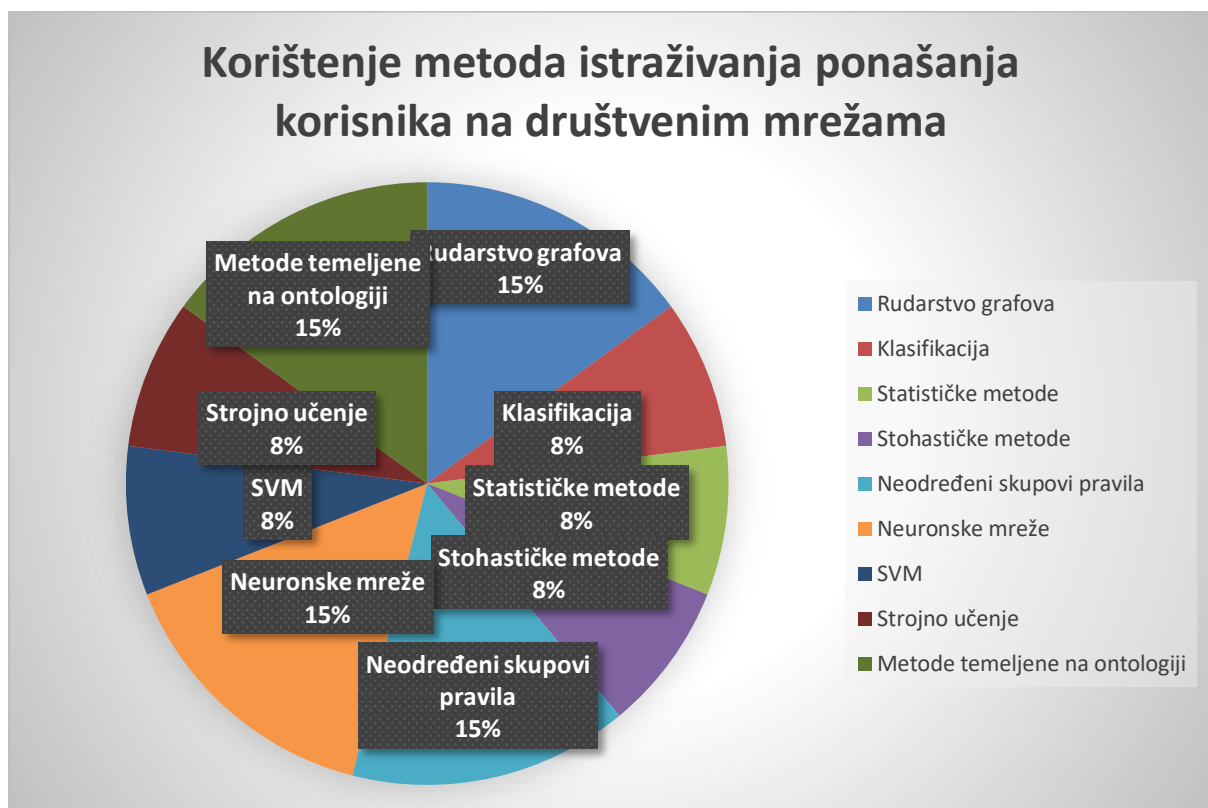
Karakterizacija ponašanja korisnika društvenih mreža je definirati ponašanje korisnika matematički na odgovarajući način. Matematička reprezentacija ponašanja će pomoći u pronalaženju odgovarajućeg obrasca ponašanja iz ogromnih podataka o društvenim mrežama. Ovo je vrlo složeno predstavljanje jer se ljudsko ponašanje uvelike razlikuje od ponašanja korisnika od jedne do druge društvene mreže. Isti korisnik na drugim drugačijim društvenim mrežama ostavlja različite uzorke aktivnosti koji opet variraju od osobe do osobe što formira generalno ponašanje korisnika. [4]

U zadnjem desetljeću su mnogi autori proveli neku vrstu istraživanja na temu karakterizacije ponašanja. Tako je Maria Kihl 2010. godine objavila članak na temu „Analiza prometa i karakterizacije ponašanja korisnika na internetu“. Fokusirala se na korištenje interneta na temelju uzoraka, opsege i primjenu tih uzoraka. Nadalje, 2012. godine je Erheng Zhong istraživao složeno ponašanje korisnika, što znači da je istraživao ponašanje jednog korisnika na različitim platformama društvenih mreža. Francis T. Odonovan je u 2013. godini istraživao karakterizaciju korisničkog ponašanja i informatičko razmnožavanje na društvenim mrežama. Jedna od aktivnosti koja je bila dio istraživanja je objavljivanje različitih multimedijским sadržavaj, dijeljenje pjesama, video-isječaka i ostalih sadržaja s interneta na društvene mreže. Uz njih postoji još nekoliko desetaka članaka i istraživanja koji su pomnije raspravljali o nekom području bitnom za proučavanje korisničkog ponašanja na društvenim mrežama s različitim varijablama i rezultatima. [4]

U ovom istraživanju, koristile su se tri podkategorije za analizu ponašanja korisnika na društvenim mrežama, kao što su karakterizacija ponašanja, prepoznavanje ponašanja te predviđanje ponašanja. Zaključilo se kako unatoč relativnoj slučajnosti i manjoj strukturnoj posvećenosti, odnosi na internetskim forumima, pridruživanje korisnika grupama prema vrsti njegovog ponašanja prikazuje snažnu pravilnost. Tako se zaključilo da ponašanje korisnika ovisi o platformi na kojoj su korisnici. Ne postoji karakterizacija ponašanja korisnika za sve vrste društvenih foruma. Nekoliko je skupova podataka dostupno za analizu ponašanja korisnika na društvenim mrežama. Najčešće korišteni skupovi podataka su:

- Osnovni podaci – podaci koje korisnik postavlja o sebi tijekom registracije na društvenu mrežu.
- Objavljeni sadržaji – to su podaci, odnosno sadržaji koje je korisnik objavljivao na svojem profilu, a mogu biti tekstualni sadržaji, slike, video isječci, glazba ili poveznice na vanjske sadržaje.
- Izvedeni podaci – to su podaci o pretraživanju korisnika, vrijeme u koje je nešto na društvenoj mreži korisnik učinio i slično.

Iznad navedeni skupovi podataka se analiziraju kroz nekoliko metoda, točnije njih 9, a one su: rudarstvo grafova, klasifikacija, statističke metode, stohastičke metode, neodređeni skupovi pravila, neuronske mreže, SVM, strojno učenje, metode temeljene na ontologiji. U grafu ispod se nalaze podaci o korištenju određenih metoda za istraživanje ponašanja korisnika na društvenim mrežama od 2000. do 2015. godine. [4]



Slika 1. Korištenje metoda istraživanja ponašanja korisnika na društvenim mrežama (Prema: Krishna Das i Smriti Kumar Sinha, 2016)

Najveći izazov u ovom istraživanju je točna karakterizacija ponašanja korisnika. Iako se pokušavalo mnogo puta i s raznim metodama karakterizirati ponašanja korisnika na društvenim mrežama, nikada se nije u potpunosti uspjelo dobiti rješenje. [4]

U ovom se istraživanju ispitivalo ponašanje korisnika na društvenim mrežama u tri perspektive, karakterizacija ponašanja, prepoznavanje ponašanja te predviđanje ponašanja.

Istraživanje je razmatralo korisnike kroz razne aktivnosti na platformama društvenih mreža, kao što su veze i interakcije, aktivnosti prometa, pozicioniranje korisnika u mreži itd. Prikazale su se postojeće sheme te su pruženi potencijalni budući putevi. Također, fokusiralo se na 2 različita pristupa: postojano i nepostojano ponašanje na društvenim mrežama. Ovo će istraživanje pomoći u razumijevanju kronološkog napretka u istraživanju i analizama korisničkog ponašanja. [4]

2.4. Razumijevanje ponašanja korisnika na društvenim mrežama: Istraživanje

Trenutno, društvene mreže kao što su Facebook, Twitter, Google+, LinkedIn i druge postale su ekstremno popularne u svijetu i imaju veliku ulogu u svakodnevnim životima ljudi. Ljudi pristupaju društvenim mrežama koristeći tradicionalna stolna računala te u novije vrijeme mobilne telefone. S više od milijardu korisnika u cijelome svijetu, društvene mreže su novo mjesto inovacija s puno izazovnih istraživačkih problema. U ovom istraživanju, cilj je dati sveobuhvatni pregled istraživanja povezanih s ponašanjem korisnika na društvenim mrežama iz različitih perspektiva. Prvo se raspravlja o društvenoj povezanosti i interakciji između korisnika. Također, istražuje se promet aktivnosti iz perspektive društvene mreže. Nadalje, kako su mobilni uređaji postali uobičajen alat za povezivanje na društvene mreže, istražuje se i ponašanje u mobilnom okruženju. I za kraj, recenzira se zlonamjerno ponašanje korisnika društvenih mreža te se raspravlja o nekoliko mogućih rješenja za detekciju korisnika koji se nedolično ponašaju. [5]

Društveni graf je vrlo učinkovit i često korišteni matematički alat koji prikazuje veze između korisnika na društvenim mrežama, koja koristi analizu društvenih interakcija i karakterizacije korisničkog ponašanja. Obično, društvene mreže mogu biti modelirane kao neusmjereni grafovi ili usmjereni grafovi ovisno o svojstvima društvenih mreža. Neki od tipova društvenih grafova su:

- Graf prijateljstva – prikazuje prijateljstvo između korisnika
- Graf interakcija – prikazuje vidljive interakcije, kao npr. Objavljivanje na Facebook „zidu“
- Latentan graf – latentna interakcija, kao npr. Pretraživanje i pronalazak drugih korisnika na društvenoj mreži
- Graf pratitelja – pretplatite se kako bi dobili i vidjeli sve obavijesti i poruke

Bazirano na tim grafovima, proučava se povezanost i interakcija između korisnika društvenih mreža. Nadalje, ogromna količina društvenih grafova izaziva efektivnost analiza.

Neusmjereni model grafa – za graf prijateljstva, svaki korisnik je zapravo prikazan kao čvor, a prijateljstvo između korisnika je prikazano bridom. Pokušalo se saznati jesu li društvene veze važeći indikatori korisničke interakcije. Objave na profilu te komentare na slikama definiraju kao interakciju. Na temelju podataka prikupljenih s Facebook-a došlo se do zaključka kako su korisnici skloni interakciji s malom grupom ljudi koji su im Facebook prijatelji, dok često s više od pola prijatelja nemaju nikakvu interakciju ili komunikaciju. Stoga, prijateljstvo na društvenim mrežama se ne može deklarirati jednako kao i pravo prijateljstvu u stvarnom životu.

Usmjereni model grafa – latentne interakcije su pasivne akcije korisnika društvenih mreža koje se ne mogu promatrati tradicionalnim tehnikama mjerenja. U istraživanju latentne interakcije koristili su se prikupljeni podaci s društvene mreže Renren, koja je najveća mreža u Kini. Renren prati i prikazuje zadnjih 9 posjetitelja na svakom profilu, stvarajući mjerenje latentne interakcije mogućom. U usmjerenom latentnom grafu, usmjereni vrh od A do B govori kako je čvor A posjetio profil čvora B. Ovime se došlo do 3 velika zaključka:

- Prvi zaključak – latentna interakcija je značajno veća i učestalija od vidljive interakcije
- Drugi zaključak – latentna interakcija je ne-recipročna po svojoj prirodi
- Treći zaključak – popularnost profila nije u korelaciji s učestalošću ažuriranja sadržaja ili brojem prijatelja za veoma popularne korisnike

Različiti načini društvenih grafova mogu otkriti kako se korisnici povezuju i komuniciraju jedni s drugima. Međutim, zbog ograničenih informacija koje graf prikazuje, razni tipovi korisničkih aktivnosti ne mogu biti karakterizirani. Promatranje od strane mrežnih operatera može pratiti takve informacije jednostavno te interpretirati kako korisnici bolje koriste društvene mreže. [5]

U ovom istraživanju, istraživalo se korisničko ponašanje na društvenim mrežama iz četiri različite perspektive: povezanost i interakcija, aktivnost prometa, mobilno društveno ponašanje te zlonamjerno ponašanje. Pregledavale su se postojeće reprezentativne sheme te su se također pružili potencijalni budući putevi. Zamišlja se kako će se ovo istraživanje povećati korisničko iskustvo iz različitih aspekata i zadovoljiti različite igrače, uključujući davatelje infrastrukture, davatelje usluga te krajnje korisnike. Također, vjeruje se kako će buduća istraživanja korisničkog ponašanja na društvenim mrežama generirati više zanimljivih istraživačkih problema i uzbudljivih rješenja u tom području. [5]

2.5. Istraživanje korisničkog ponašanja na društvenim mrežama

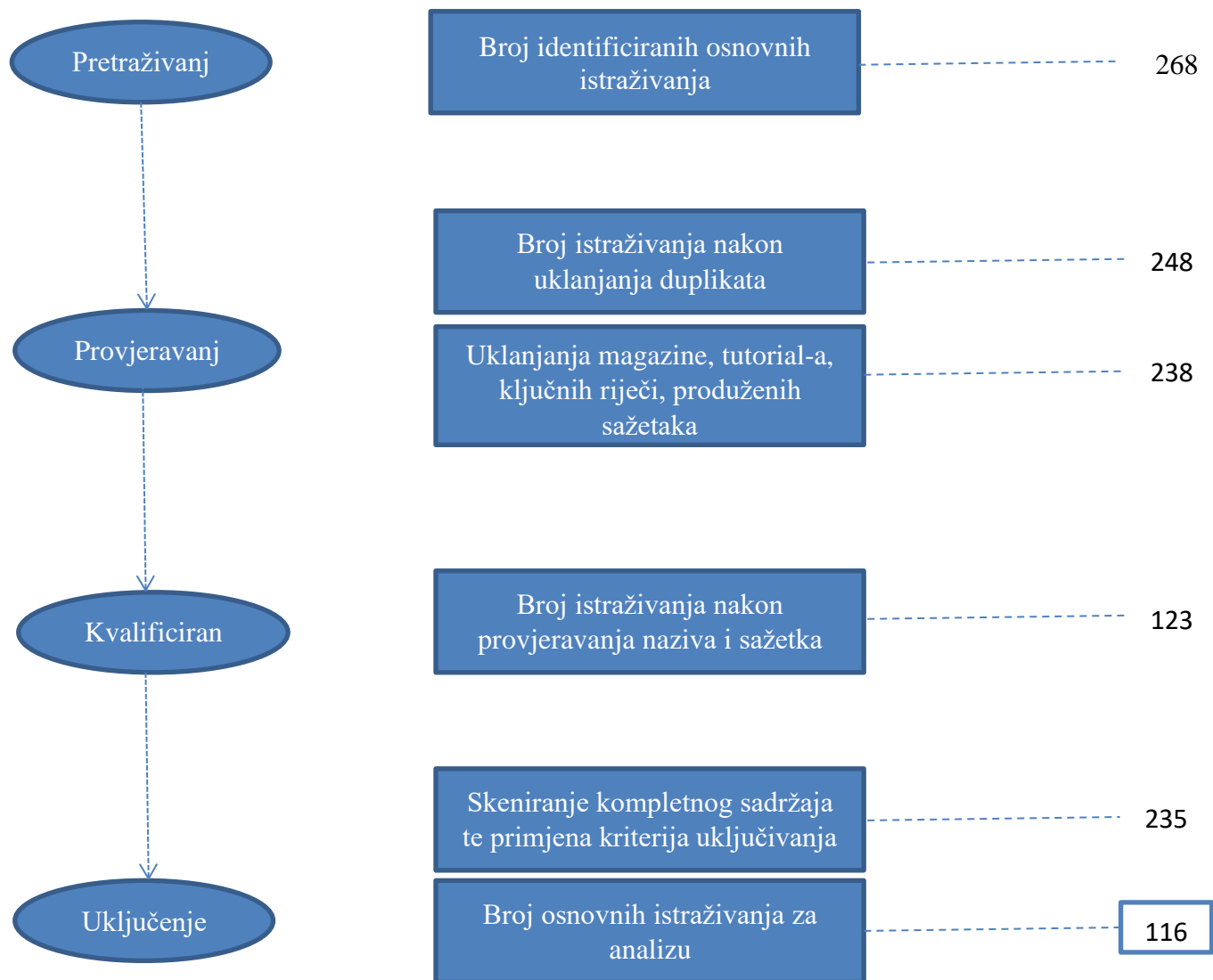
Društvene mreže su korištene za društvenu i profesionalnu interakciju s ljudima. Njihova popularnost je ohrabrila istraživače da analiziraju veze između aktivnosti formirane na društvenim mrežama s ponašanjem korisnika. Sami pojam „ponašanje korisnika“ je korišten dvosmisleno s različitim interpretacijama, što otežava identifikaciju istraživanja ponašanja korisnika u relaciji s društvenim mrežama. Ovaj fenomen je potaknuo ovo temeljito istraživanje na karakteristike korisničkog ponašanja o kojem se raspravlja u literaturi. Zbog toga, u ovom se istraživanju pokušava identificirati, analizirati i klasificirati karakteristike povezane s ponašanjem korisnika kao odgovor na istraživačka pitanja dizajnirana za provođenje ovog istraživanja. Identificirano je 116 osnovnih istraživanja za potpunu analizu te je ovo istraživanje pronašlo 7 karakteristika povezanih s ponašanjem koje imaju direktan utjecaj na korištenje društvenih mreža te 9 faktora koji imaju indirektan utjecaj. Rezultati ovog istraživanja mogu biti korišteni od strane vježbača da ocjenjuju svoje platforme društvenih mreža te razvijaju aplikacije više usmjerene prema korisniku, a također istraživanje može pomoći i organizacijama da bolje razumiju potrebe svojih zaposlenika. [6]

Ovo istraživanje je koristilo istraživanje mapiranja koji se također zove i ispitivanje opsega te je tip sistemskog pregleda. Ispitivanje opsega je više otvoreni obrazac sistemskog pregleda literature koja se provodi na proširenoj temi kako bi pružila opsežan i iscrpan pregled na područje za ocjenu kvalitete i količine dostupnih dokaza o toj temi. Osnovna istraživanja tražena iz odabranih baza podataka. Istraživanja su identificirana primjenom kriterija uključenosti i isključenosti. Izvršeno je vađenje podataka i provedena je sinteza za klasifikaciju ponašanja korisnika. Konačno, pruženi su odgovori na istraživačka pitanja. Glavno pitanje ovog istraživanja je glasilo „Koje su karakteristike ponašanja korisnika bile raspravljane od strane istraživačke zajednice društvenih mreža“? Povezana pitanja koja su povezana s glavnim pitanjem su:

- Kakav je trend u provođenju istraživanja na web stranicama društvenih mreža?
- Gdje su objavljena osnovna istraživanja?
- Koje zemlje su odabrane za istraživanje?
- Koju vrstu web lokacija društvenih mreža polaznici koriste?
- Koju metodu istraživanja u velikoj mjeri usvajaju studije?
- U kojem se kontekstu provode ponašanja korisnika društvenih mreža?

Za strategiju istraživanja vrlo važne su bile sljedeće četiri stavke: pretraživački niz, vremenski period, izbor elektroničke baze podataka te izbor osnovnih istraživanja. Kako bi se osnovna istraživanja pronašla potrebna je različita kombinacija pretraživačkih nizova. Bilo je

potrebno odabrati dobre nizove koji bi dali određeni rezultat vezan uz korištenje i ponašanje na društvenim mrežama. Sljedeće što je vrlo bitno je vremenski period. U ovom slučaju period je od 2005. do 2013. godine te ne postoji neko slično istraživanje prije 2005. godine. Kako bi se ovim svime pronašla kvalitetna osnovna istraživanja, korištene su tri baze podataka koje su vrlo pouzdane i pružaju kvalitetan i vrlo recenziran sadržaj. U nastavku slijedi prikaz 4 koraka kako se došlo do osnovnih istraživanja potrebnih za provođenje ovog istraživanja. [6]



Slika 2. Postupak dobivanja osnovnih istraživanja potrebnih za provođenje istraživanja (Prema: Hajra Waheed, Maria Anjum, Mariam Rehman, Amina Khawaja, 2016)

Cilj ovog istraživanja bio je istražiti karakteristike ponašanja korisnika društvenih mreža. Kako bi se pronašao dokaz koji pokazuje opseg rasprave o ponašanju korisnika u postojećoj literaturi, dizajnirala su se istraživačka pitanja na koje je bilo potrebno pronaći kvalitetne odgovore. Kako bi se na ta pitanja moglo odgovoriti korištena je tehnika mapiranja istraživanja koja je tip sistematičnog pregleda. Kako karakteristike ponašanja korisnika nisu točno

definirane u istraživanjima, tematska analiza je korištena za identifikaciju, analizu i klasificiranje značajki. Karakteristike ponašanja su identificirale učestalost korištenja, kontrolu informacija, društvenu pripadnost, samostalnu orijentaciju, društvenu smjelost. Kontekst korišten u provođenju istraživanja uključuje informacije o privatnosti, godinama, kulturi, dijeljenju informacija, spolu i slično. [6]

U današnje vrijeme, korištenje društvenih mreža se jako povećalo u odnosu na godine koje su prošle. Jedan od mnogih razloga je što postoji jednostavan pristup do mreža te široka dostupnost određenih društvenih mreža. Zbog toga je ovo vrijeme bilo odlično za provođenje istraživanja o društvenim mrežama. Ovo istraživanje je koristilo niz istraživačkih metoda za ispitivanje ponašanja korisnika u određenim područjima kao što su: privatnost, dijeljenje informacija, povjerenje uzimajući u obzir različite tipove korisnika, koji nisu limitirani samo na studente, odrasle, tinejdžere ili bilo koju drugu grupu. [6]

Provođenje istraživanja mapiranja, identificiralo se da su istraživanja na temu društvenih mreža jedva stara desetak godina te da ne postoji dokaz o postojanju istraživanja ranijih od 2005. godine. Čak i ona objavljena 2005. godine nisu relevantna i pouzdana. Od 116 istraživanja identificiranih za punu analizu, 70 od njih je objavljeno na konferencijama, a 40 u stručnim časopisima. Korisnici koji su sudjelovali u istraživanjima su s područja Amerike, Europe, Kine, Meksika, Singapura, Malezije i drugih zemalja. Najveći broj istraživanja je ipak napravljen u Americi. Najpopularnija društvena mreža među ispitanicima je bio Facebook, iako je, primjerice, u Kini vrlo popularna njihova društvena mreža Renren. Pregledom korištenih metoda u osnovnim istraživanjima dolazi se do zaključka kako su korišteni intervjui te obrasci s pitanjima na koje korisnici trebaju odgovoriti u većini tih istraživanja kako bi se saznalo njihovo ponašanje na društvenim mrežama. Postoje i istraživanja gdje se koristilo i više različitih metoda, no njihov broj je vrlo malen. Kada se sagledaju sva istraživanja, ne postoje dokazi da se ponašanje korisnika potpuno razmatralo u jednom istraživanju. [6]

Ovo istraživanje doprinosi teoriji sistematičnom analizom dokaza iz literature pružajući integrirani prikaz ponašanja korisnika. Također, istraživanje prikazuje klasifikaciju ponašanja korisnika koje može pomoći istraživačima da provedu detaljno istraživanje o vezama tih ponašanja te njihovim utjecajima na korisnike društvenih mreža. [6]

2.6. Računala u ponašanju ljudi

Osim pružanja pristupa do informacija, Internet je omogućio ljudima stvaranje i očuvanje veza s drugima. Društvene mreže, kao Facebook, su jedna od metoda kako se ljudi povezuju s drugima na internetu (Bonds-Raacke & Raacke, 2010). Facebook je dominantna društvena mreža s preko 500 milijuna aktivnih korisnika (Facebook website, 2011). Tipični korisnici će

potrošiti od 10 minuta do više od 2 sata dnevno na Facebook (Christofides, Muise, & Desmarais, 2009; Orr et al., 2009; Pempek, Yermolaveya, & Calvert, 2009; Raacke & Bonds-Raacke, 2008; Ross et al., 2009). Jedna od grupa koja puno koristi društvene mreže, posebno Facebook, su studenti. Nedavna istraživanja su pokazala kako preko 90% studenata ima korisnički račun na Facebook-u (Cheung, Chiu, & Lee, 2010; Pempek et al., 2009). Iako Facebook može olakšati učenje, studenti ga najčešće koriste kako bi komunicirali s prijateljima, gledali slike i korisničke profile drugih osoba te kako bi trošili vrijeme, izbjegavali dosadu i odugovlačili (Pempek et al., 2009). [1]

U istraživanju je sudjelovalo 548 studenata prve godine studija psihologije iz sveučilišta u Melbourne-u. 94% tih ispitanika ima Facebook račun. Podaci ispitanika nisu korišteni u analizi ukoliko je nedostajalo više od 10% podataka. Finalni uzorak je sadržavao 300 studentica te 93 studenata s prosjekom godina od 20,59. Svi sudionici finalnog uzorka su imali Facebook račun. Skale koje su dobivene u trenutnom istraživanju uključuju mjerenje korištenja Facebook-a te „Big Five“ osobina ličnosti. Odgovori su dobiveni putem interneta korištenjem „Opinio“ alata za izradu anketa za ispitivanje korisnika. Rezultati su analizirani korištenjem „SPSS“ alata, verzije 19.0 te „Mplus“ alata verzije 6.0. Prije analize, provjeravali su se nedostajući podaci te nevažeci podaci. Učestalosti i deskriptivne statistike su generirane za svaku varijablu te za nasumični uzorak od 10% ispitanika je provjerena greška kod unosa podataka. Srednja vrijednost korištenja Facebook društvene mreže je iznosila 66,15 minuta dok je tipično vrijeme korištenja Facebooka na dnevnoj bazi iznosilo u rasponu od 20 do 90 minuta. Jedan ispitanik koristi Facebook izuzetno puno na dnevnoj bazi, čak 500 minuta. Prosječan broj prijatelja na Facebook-u iznosi 349,97, a pripadnost grupi iznosi 66,10. Funkcija Facebook-a koju korisnici najviše koriste je tzv. „Facebook zid“ koji je prikupio 35,1% glasova, a druga najzastupljenija funkcija su poruke s 18,3%. Najčešći odgovor na pitanje: „Zašto koristite Facebook?“ je bio „To je način na koji komuniciram s prijateljima.“ Te je takav odgovor stavilo više od 50% ispitanika. [1]

Nakon ispitivanja istodobnih učinaka nekoliko psiholoških varijabli pomoću regresijskog modela posebno prilagođenog za podatke o upotrebi Facebook-a – model negativne binomne regresije, čini se kako su ostali psihološki faktori u dodatku s osobinama ličnosti povezane s korištenjem Facebook društvene mreže. Ovo istraživanje je pokazalo da studenti s otvorenim ličnostima koriste Facebook da se povežu s ostalim korisnicima kako bi međusobno razgovarali o njihovom širokom rasponu interesa, gdje usamljeni studenti koriste Facebook da potisnu svoj manjak veza u stvarnom životu. Buduća istraživanja trebaju uzeti u obzir prikupljanje nekoliko različitih skupova podataka kako bi testirali uzročan smjer veza između psiholoških varijabli te Facebook mjerenja. [1]

3. Opis podataka

Kako bi se prikupili podaci za predviđanje ponašanja korisnika društvenih mreža provedena je anketa koja sadrži 29 pitanja. Na svako pitanje je potrebno unijeti odgovor kako kasnije u obradi ne bi nedostajali pojedini podatci. Anketa se provodila putem Google forms-a, navedenoj anketi se moglo pristupiti putem linka koji je bio objavljen na društvenim mrežama.

3.1. Pitanja s ankete

Kao što je navedeno anketa se sastoji od 29 pitanja. Jedan dio pitanja se odnosi na osnovne podatke o osobi kao što je dob ili spol osobe, dok se drugi dio pitanja odnosi isključivo na aktivnost osobe na društvenim mrežama. Zadnje pitanje koje se odnosi na zadovoljstvo osobe na društvenim mrežama će biti temelj daljnjeg predviđanja. Nećemo previše objašnjavati pitanja, iz njih je lako vidljivo koji se podatak traži. Pitanja koja se nalaze u anketi su navedena ispod:

1. Spol osobe
2. Starost osobe
3. Ljubavni status osobe
4. Stručna sprema osobe
5. Hobi osobe
6. Bavite li se sportom?
7. Koji je vaš radni status?
8. Koliko društvenih mreža koristite?
9. Koliko vremena u prosjeku u danu provedete na društvenim mrežama?
10. Koji je vaš razlog posjete društvenih mreža?
11. Koliko prijatelja/pratitelja imate na društvenim mrežama?
12. Koje vrste objava najviše objavljujete?
13. Koja je najčešća tema vaše objave?
14. Smatrate li da bi bez društvenih mreža bili učinkovitiji u obavljanju ostalih dnevnih aktivnosti?
15. Smatrate li se ovisnima o društvenim mrežama?
16. Koji ste horoskopski znak?
17. Jeste li bili pod utjecajem nasilja putem društvenih mreža ili sudjelovali u njemu?
18. Utječu li društvene mreže na vaš stil života?
19. Poznajete li sve osobe koje pratite ili imate za prijatelje na društvenim mrežama?

20. Koliko prosječno dobivate lajkova po objavi?
21. Koristite li emotikone?
22. U kojoj mjeri koristite emotikone?
23. Koristite li interpunkcijske znakove u komunikaciji putem društvenih mreža?
24. Utječe li način pisanja na društvenim mrežama na vašu pismenost?
25. Označavate li druge osobe na objavama?
26. Komentirate li objave drugih osoba?
27. Označavate li tuđe objave sa sviđa mi se?
28. Koristite li više grupne razgovore ili pojedinačne?
29. Koliko ste zadovoljni korištenjem društvenih mreža?

3.2. Prikaz dobivenih rezultata

Anketu je popunilo 390 osoba. Od njih se 222 izjasnilo da su ženskog spola, 167 da su muškog spola, dok se samo jedna osoba izjasnila da spada u skupinu ostalo. Starost osoba se kreće u intervalu od 17 do 9999, što nam govori da su neki ispitanici pogrešno popunili polje. Od svih ispitanika tri osobe su se izjasnile da imaju 17 godina, sedam osoba se izjasnilo da ima 18 godina, pedeset i četiri osobe su se izjasnile da su starosti 19 godina, pedeset i šest osoba se izjasnilo da su starosti 20 godina, četrdeset i šest osoba se izjasnilo da su starosti 21 godina, njih četrdeset i četiri se izjasnilo da imaju 22 godine, sedamdeset i pet osoba se izjasnilo da imaju 23 godine, dvadeset i dvije osobe se izjasnilo da imaju 24 godine, osamnaest osoba se je izjasnilo da imaju 25 godina, jedanaest osoba se izjasnilo da su starosti 26 godina, u ostale dobne skupine su raspoređeni ispitanici u intervalu od 1 do 3.

Na pitanje o ljubavnom statusu njih 169 se izjasnilo da su u vezi, 26 da su u bračnoj zajednici, dok ih je 195 reklo da su slobodni. Prema stručnoj spremi ispitanici su se izjasnili da njih petero ima završenu samo osnovnu stručnu spremu, njih 217 se izjasnilo da imaju srednju stručnu spremu, visoku stručnu spremu ima njih 117 te njih 51 se izjasnilo da imaju višu stručnu spremu. Ispitanici na pitanje o njihovom hobiju su se izjasnili na sljedeći način:

1. Igranje video igara (63 osobe)
2. Bavljenje sportom (98 osoba)
3. Sviranje glazbenog instrumenta (19 osoba)
4. Crtanje (12 osoba)
5. Čitanje knjiga (58 osoba)
6. Nešto drugo (99 osoba)
7. Nemaju hobi (41 osoba)

Na pitanje o bavljenju sportom njih 199 se izjasnilo da se ne bave sportom, dok se njih 191 izjasnilo da se bave sportom. Ovaj podatak je jako zanimljiv iz razloga što se je anketa pretežno provodila na području grada Varaždina u krugu studenata koji studiraju informatičke znanosti. Prema radnome statusu ispitanici su se izjasnili da njih četvero su još uvijek učenici (srednja ili osnovna škola), njih 289 se izjasnilo da su studenti, petnaest ih je zaposleno dok ih je dvadeset i šestero nezaposleno. Ispitanici su na pitanje o broju društvenih mreža koje koriste se izjasnili na sljedeći način. Njih sto dvanaest je izjavilo kako koristi više od tri društvene mreže, sto i pet osoba je navelo kako koristi tri društvene mreže, sto četrdeset i dvije osobe je reklo kako koristi samo dvije društvene mreže i njih trideset je reklo kako koristi jednu društvenu mrežu. Zanimljivo je istaknuti kako ni jedna osoba nije navela da ne koristi ni jednu društvenu mrežu.

Ispitanici su odgovarali na pitanje „Koliko vremena u prosjeku u danu provedete na društvenim mrežama?“. Njih trideset i dvoje je istaknulo kako koristi društvene mreže od 0-1 sat, devedeset i četvero se izjasnilo da koristi društvene mreže 1-2 sata, sto osamnaest ih se izjasnilo da koriste društvene mreže u prosjeku 2-3 sata, njih šezdeset i dvoje je reklo da na društvenim mrežama provode 3-4 sata, dok je preostalih osamdeset i četvero navelo da provode više od 4 sata na dan na društvenim mrežama. Nakon što su ispitanici odgovorili na prethodno pitanje, trebali su se izjasniti o razlogu korištenja društvenih mreža. Osmero od ukupnog broja ispitanika se izjasnilo kako koriste društvene mreže za učenje. Njih sto trideset i osam se izjasnilo kako koriste društvene mreže za zabavu. Društvene mreže na druženje s prijateljima koristi njih devedeset i jedan, njih sto je navelo kako društvene mreže koriste isključivo kako bi potrošili slobodno vrijeme, dok je preostalih pedeset i troje navelo da društvene mreže koriste kao izvor vijesti.

Na pitanje o broju prijatelja odnosno pratitelja ispitanici su se izjasnili na sljedeći način:

1. Broj prijatelja u intervalu 0-100 (33 osobe)
2. Broj prijatelja u intervalu 100-200 (69 osoba)
3. Broj prijatelja u intervalu 200-300 (88 osoba)
4. Broj prijatelja u intervalu 300-400 (71 osoba)
5. Broj prijatelja veći od 400 (129 osobe)

Na pitanje o vrsti objave koju koriste njih tristo trideset i sedam je navelo kako pretežno objavljuju slike, video objavljuje njih dvadeset i četvero, dok preostalih dvadeset i devet pretežno objavljuje tekstualne objave. Kao temu objave njih devetnaest je navelo kako pretežno objavljuju sadržaj vezan uz sport, trinaest je navelo kako objavljuju sadržaj vezan uz modu, dvadeset i devetero je navelo kako su to neki citati, pedeset i osmero kako su to neke smiješne stvari, dok je njih dvjesto sedamdeset i jedan navelo da su to objave vezane uz

njihove aktivnosti. Zanimljivo je da se je dvjesto dvadeset i devetero ispitanika izjasnilo kako bi bili učinkovitiji u obavljanju svojih dnevnih aktivnosti, dok je njih sto šezdeset i jedan izjavilo kako se ništa ne bi promijenilo, da bi bili jednako uspješni. Njih čak dvjesto trideset i devetero se izjasnilo kako su ovisni o društvenim mrežama, dok je ostatak njih sto pedeset i jedan navelo kako nisu ovisni o društvenim mrežama. Dvjesto tri ispitanika navelo je kako smatra da društvene mreže ne utječu na njihov stil života, dok je njih sto osamdeset i sedmero navelo kako društvene mreže utječu na njihov stil života. Od ukupnog broja ispitanika njih dvjesto dvadeset i petero se izjasnilo kako ne poznaju sve prijatelje ili pratitelje na društvenim mrežama, a njih sto šezdeset i petero kako poznaju sve svoje prijatelje na društvenim mrežama. Na pitanje o dobivanju lajkova po objavi ispitanici su se izjasnili na sljedeći način:

1. Broj lajkova po objavi u intervalu 0-100 (250 osoba)
2. Broj lajkova po objavi u intervalu 100-200 (114 osoba)
3. Broj lajkova po objavi u intervalu 200-300 (19 osoba)
4. Broj lajkova po objavi u intervalu 300-400 (3 osobe)
5. Broj lajkova po objavi veći od 400 (4 osobe)

Na pitanja vezana uz emotikone ispitanici su se izjasnili da njih tristo šezdeset i troje ih koristi, dok njih dvadeset i sedam ih ne koristi u svojoj komunikaciji putem društvenih mreža. Prema učestalosti korištenja emotikona ispitanici su se izjasnili da njih dvanaestero nikada ne koristi emotikone, njih devedeset ih koristi rijetko, čak dvjesto dvadeset i jedna osoba često koristi emotikone, dok njih šezdeset i sedmero ih koristi stalno.

Pomoću sljedećih pitanja sam htio provjeriti pismenost ispitanika. Na pitanje o korištenju interpunkcijskih znakova na društvenim mrežama njih tristo i dvoje se izjasnilo kako ih koristi, dok ih je osamdeset i osmero navelo kako ih ne koristi. Dvjesto šezdeset i četiri ih je izjavilo kako pisanje na društvenim mrežama ne utječe na njihovu pismenost, dok ih je sto dvadeset i šestero izjavilo da pisanje na društvenim mrežama utječe na njihovu pismenost.

Kako bi se provjerila aktivnost korisnika na društvenim mrežama postavljena su sljedeća pitanja. Na pitanje „Označavate li druge osobe na objavama?“ tristo dvadeset ispitanika je odgovorilo potvrdno, dok su preostali ispitanici odgovorili negativno. Na pitanje „Komentirate li objave drugih osoba?“ ispitanici su odgovorili na sljedeći način:

1. Nikada (36 osoba)
2. Ponekad (321 osoba)
3. Često (27 osoba)
4. Stalno (6 osoba)

Na pitanje „Označavate li tuđe objave sa sviđa mi se?“ ispitanici su odgovorili na sljedeći način:

1. Nikada (7 osoba)

2. Ponekad (145 osoba)
3. Često (158 osoba)
4. Stalno (80)

Od ukupnog broja ispitanika njih tristo sedamnaest je navelo kako pretežno koristi pojedinačne razgovore, dok je preostalih sedamdeset i troje navelo kako pretežno koristi grupne razgovore. Kroz anketu se je postavilo pitanje „Jeste li bili pod utjecajem nasilja putem društvenih mreža ili sudjelovali u njemu?“. Zanimljiv je podatak da je čak šezdeset i jedna osoba na to pitanje odgovorila potvrdno. Ne znamo radi li se o tome da su ispitanici nasilnici ili su žrtve nasilja, no u svakom slučaju 15% sudionika je sudjelovalo u nekom obliku nasilja. Posljednje pitanje koje se postavilo na anketi je zadovoljstvo korisnika društvenim mrežama. Četvero ispitanika je odgovorilo sa ocjenom 1, dvanaestero je dalo ocjenu 2, sto dvadeset i dvoje ocjenu 3, sto devedeset i osmero ocjenu 4 te preostalih pedeset i četvero je dalo maksimalnu ocjenu.

4. Stablo odlučivanja

Stablo odlučivanja je jedna od najčešće korištenih metoda prilikom klasifikacije, predviđanja, za procjenu vrijednosti, grupiranje, opisivanje podataka i vizualizaciju. [7]

Stablo odlučivanja je struktura koja je slična dijagramu tijeka u kojem svaki unutarnji čvor predstavlja "testiranje" atributa, svaka grana predstavlja ishod testiranja, a svaki čvor listova predstavlja klasu (odluka – rezultat koji se dobiva nakon računanja svih atributa). [8]

Na slici ispod prikazano je jedno jednostavno stablo odlučivanja, koje se sastoji od dva čvora i četiri grane. U našem slučaju korijen stabla je visina igrača. Putanje od korijena do listova predstavljaju klasifikacijska pravila. Iz toga razloga možemo kreirati jedno jednostavno pravilo. Ukoliko je igrač visok te primi udarac u nogu, tada će igrač zaraditi prijelom noge. Ovo je jedno jednostavno stablo odlučivanja koje nije najpreciznije, mogli smo dodati posjedovanje kostobrana i slične sportske opreme kako bismo ga dodatno razgranali.



Slika 3. Prikaz jednostavnog stabla odlučivanja (Prema: B. Kliček, 2020)

Stablo odlučivanja se stvara pomoću formule entropije. Formula entropije predstavlja mjerenje ukupnog nereda ili nehomogenosti iz baze podataka te ona glasi:

$$E = \sum_b \left[\left(\frac{n_b}{n_t} \right) \cdot \left(\sum_c \left(-\frac{n_{bc}}{n_b} \right) \cdot \log_2 \left(\frac{n_{bc}}{n_b} \right) \right) \right]$$

Značajne formule:

1. E - entropija (mjera nereda)
2. n_t – ukupan broj primjera u svim granama
3. n_b – broj primjera u grani b
4. n_{bc} – ukupan broj primjera u grani b klase c [8]

Na sljedećem primjeru ćemo prikazati na koji način se kreira stablo odlučivanja pomoću entropije. U tablici ćemo prikazati podatke koji nisu istiniti, već samo služe za primjer.

Redni broj	Liga	Visina	Jačina	Ozljeda
1.	1	Visok	Slab	Da
2.	1	Visok	Jak	Ne
3.	1	Nizak	Slab	Ne
4.	1	Nizak	Jak	Ne
5.	2	Visok	Slab	Da
6.	2	Visok	Jak	Ne
7.	2	Nizak	Slab	Da
8.	2	Nizak	Jak	Ne
9.	3	Visok	Slab	Da
10.	3	Visok	Jak	Da
11.	3	Nizak	Slab	Da
12.	3	Nizak	Jak	Da

Tablica 1. Podaci za izradu stabla odlučivanja (Prema: B. Kliček i D. Oreški, 2020)

Prvo ćemo računati entropiju za atribut liga. Primjenom ranije navedene formule dobivamo sljedeće:

$$E = \frac{4}{12} \cdot \left(-\frac{3}{4} \cdot \log_2 \frac{3}{4} - \frac{1}{4} \cdot \log_2 \frac{1}{4} \right) + \frac{4}{12} \cdot \left(-\frac{2}{4} \cdot \log_2 \frac{2}{4} - \frac{2}{4} \cdot \log_2 \frac{2}{4} \right) + \frac{4}{12} \cdot \left(-\frac{4}{4} \cdot \log_2 \frac{4}{4} \right)$$

$$E = 0,6038$$

Sada ćemo računati entropiju za visinu.

$$E = \frac{6}{12} \cdot \left(-\frac{4}{6} \cdot \log_2 \frac{4}{6} - \frac{2}{6} \cdot \log_2 \frac{2}{6} \right) + \frac{6}{12} \cdot \left(-\frac{3}{6} \cdot \log_2 \frac{3}{6} - \frac{3}{6} \cdot \log_2 \frac{3}{6} \right)$$

$$E = 0,9591$$

Računamo entropiju za posljednji atribut, jačinu.

$$E = \frac{6}{12} \cdot \left(-\frac{5}{6} \cdot \log_2 \frac{5}{6} - \frac{1}{6} \cdot \log_2 \frac{1}{6} \right) + \frac{6}{12} \cdot \left(-\frac{4}{6} \cdot \log_2 \frac{4}{6} - \frac{2}{6} \cdot \log_2 \frac{2}{6} \right)$$

$$E = 0,7842$$

Prema navedenom izračunu za korijen stabla uzimamo ligu jer entropija za nju je najmanja.

Sada dalje računamo entropije za sve attribute unutar lige.

Prva liga - visina:

$$E = \frac{2}{4} \cdot \left(-\frac{1}{2} \cdot \log_2 \frac{1}{2} - \frac{1}{2} \cdot \log_2 \frac{1}{2} \right) + \frac{2}{4} \cdot \left(-\frac{2}{2} \cdot \log_2 \frac{2}{2} \right)$$

$$E = 0,5$$

Prva liga – jačina:

$$E = \frac{2}{4} \cdot \left(-\frac{1}{2} \cdot \log_2 \frac{1}{2} - \frac{1}{2} \cdot \log_2 \frac{1}{2} \right) + \frac{2}{4} \cdot \left(-\frac{2}{2} \cdot \log_2 \frac{2}{2} \right)$$

$$E = 0,5$$

Dobili smo da su entropije jednake, što znači da možemo uzeti bilo koji od ta dva čvora kao sljedeći čvor za grananje.

Druga liga – visina:

$$E = \frac{2}{4} \cdot \left(-\frac{1}{2} \cdot \log_2 \frac{1}{2} - \frac{1}{2} \cdot \log_2 \frac{1}{2} \right) + \frac{2}{4} \cdot \left(-\frac{1}{2} \cdot \log_2 \frac{1}{2} - \frac{1}{2} \cdot \log_2 \frac{1}{2} \right)$$

$$E = 1$$

Druga liga – jačina:

$$E = \frac{2}{4} \cdot \left(-\frac{2}{2} \cdot \log_2 \frac{2}{2} \right) + \frac{2}{4} \cdot \left(-\frac{2}{2} \cdot \log_2 \frac{2}{2} \right)$$

$$E = 0$$

U ovome primjeru dobivamo da je entropija za jačinu manja od entropije za visinu te za čvor grananja odabiremo jačinu.

Treća liga – visina:

$$E = \frac{2}{4} \cdot \left(-\frac{2}{2} \cdot \log_2 \frac{2}{2} \right) + \frac{2}{4} \cdot \left(-\frac{2}{2} \cdot \log_2 \frac{2}{2} \right)$$

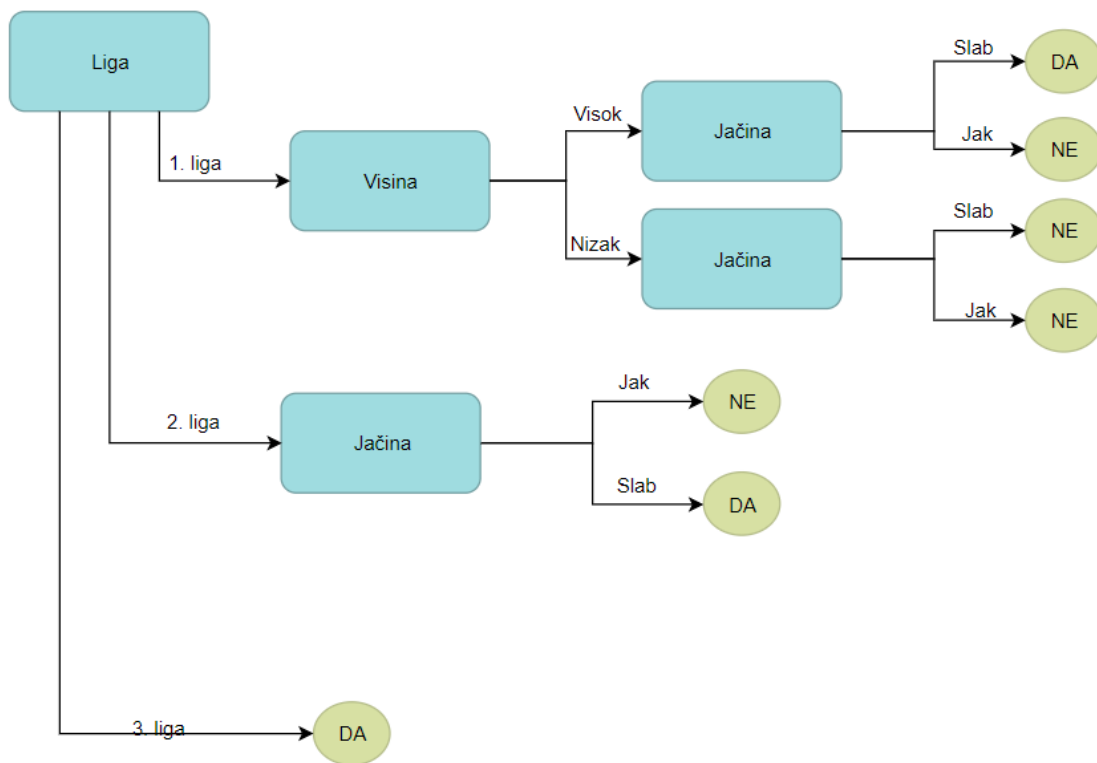
$$E = 0$$

Treća liga – jačina:

$$E = \frac{2}{4} \cdot \left(-\frac{2}{2} \cdot \log_2 \frac{2}{2} \right) + \frac{2}{4} \cdot \left(-\frac{2}{2} \cdot \log_2 \frac{2}{2} \right)$$

$$E = 0$$

Kao i u prvoj ligi možemo odabrati bilo koji atribut za grananje jer su entropije jednake. Možemo malo promotriti tablicu te uvidjeti da možemo skratiti stablo iz razloga jer svaki slučaj daje ishod DA.

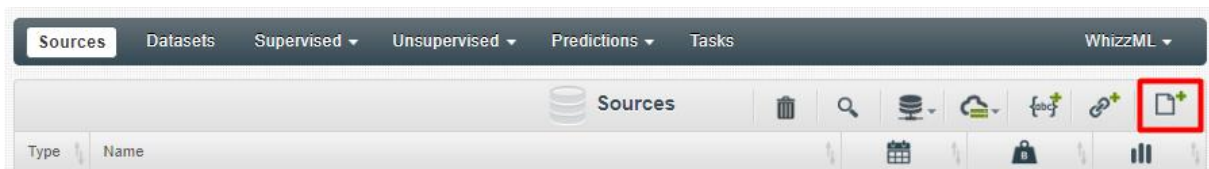


Slika 4. Prikaz stabla odlučivanja dobivenog računanjem entropije

Računanje entropije te izrada stabla odlučivanja izrađeno je prema [11].

4.1. Izrada stabla odlučivanja u BigML-u

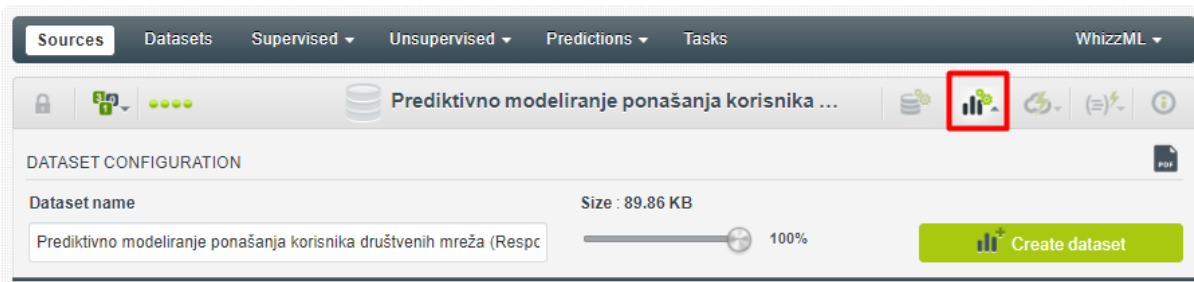
Kako bi se izgradilo stablo odlučivanja u BigML-u prvo je potrebno učitati podatke na temelju kojih će se isto generirati. BigML podržava skupove podataka u raznim formatima, neki od njih su csv, txt, json ili u ovome slučaju to je xlsx. Kako bi se učitali podaci potrebno je odabrat Source te kliknuti na područje označeno crvenim pravokutnikom kao što je prikazano na slici ispod.



Slika 5. Ikona Source BigML (Prema: B. Kliček i D. Oreški, 2020)

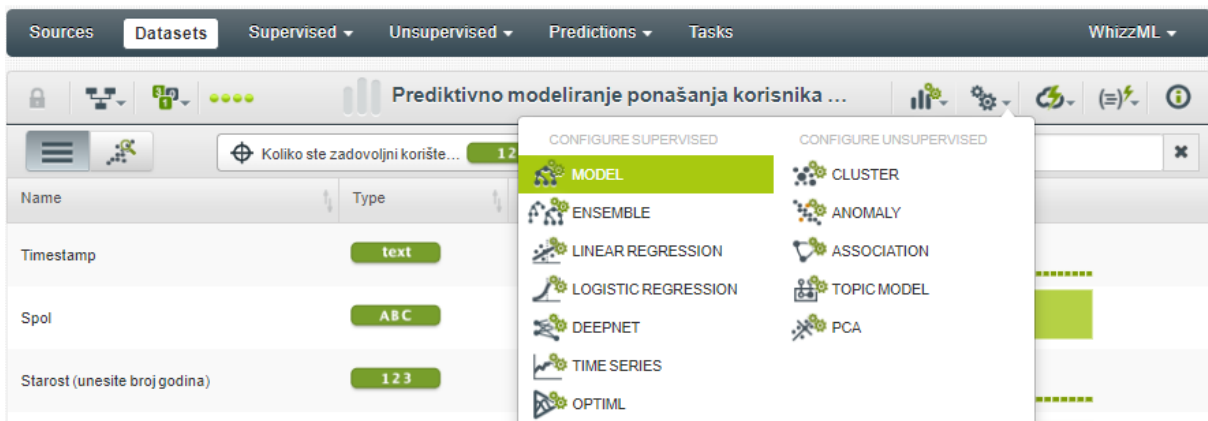
Otvora se File Explorer u kojemu se odabere skup podataka koji se želi učitati. Na temelju podataka koji su se učitali kreira se „Dataset“ kako bi se mogli koristiti podaci za kreiranje stabla.

Odabiru se učitani podaci te se klikne mišem na polje označeno crvenim pravokutnikom kao što je prikazano na slici ispod.



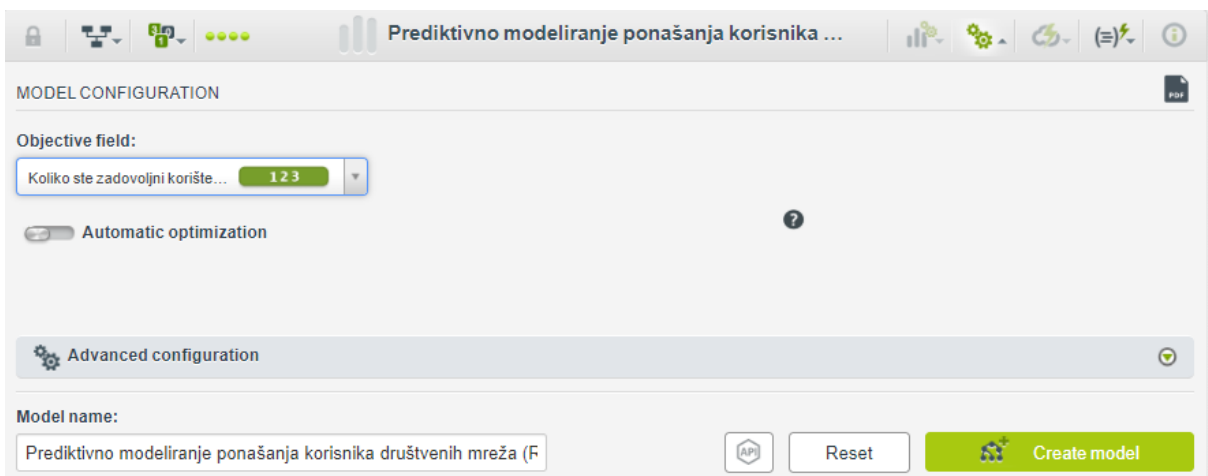
Slika 6. Ikona Dataset BigML (Prema: B. Kliček i D. Oreški, 2020)

Odabire se „Create dataset“ za kreiranje skupa podataka na temelju kojega će se kreirati stablo odlučivanja. Otvara se prikaz učitanoog skupa podataka. Kao što je prikazano na slici ispod, odabire se opcija „MODEL“ koja kreira stablo odlučivanja.



Slika 7. Ikona Model BigML (Prema: B. Kliček i D. Oreški, 2020)

Stablo odlučivanja služi za predviđanje te iz toga razloga mora se odabrati koji će se atribut predviđati.



Slika 8. Odabir atributa za predviđanje (Prema: B. Kliček i D. Oreški, 2020)

U ovome slučaju to je atribut koji govori koliko su korisnici društvenih mreža zapravo zadovoljni samim društvenim mrežama. Nakon odabira akcije „Create model“ dobiva se stablo koje je prikazano na slici ispod.



Slika 9. Stablo odlučivanja na temelju učitanih podataka (Prema: B. Kliček i D. Oreški, 2020)

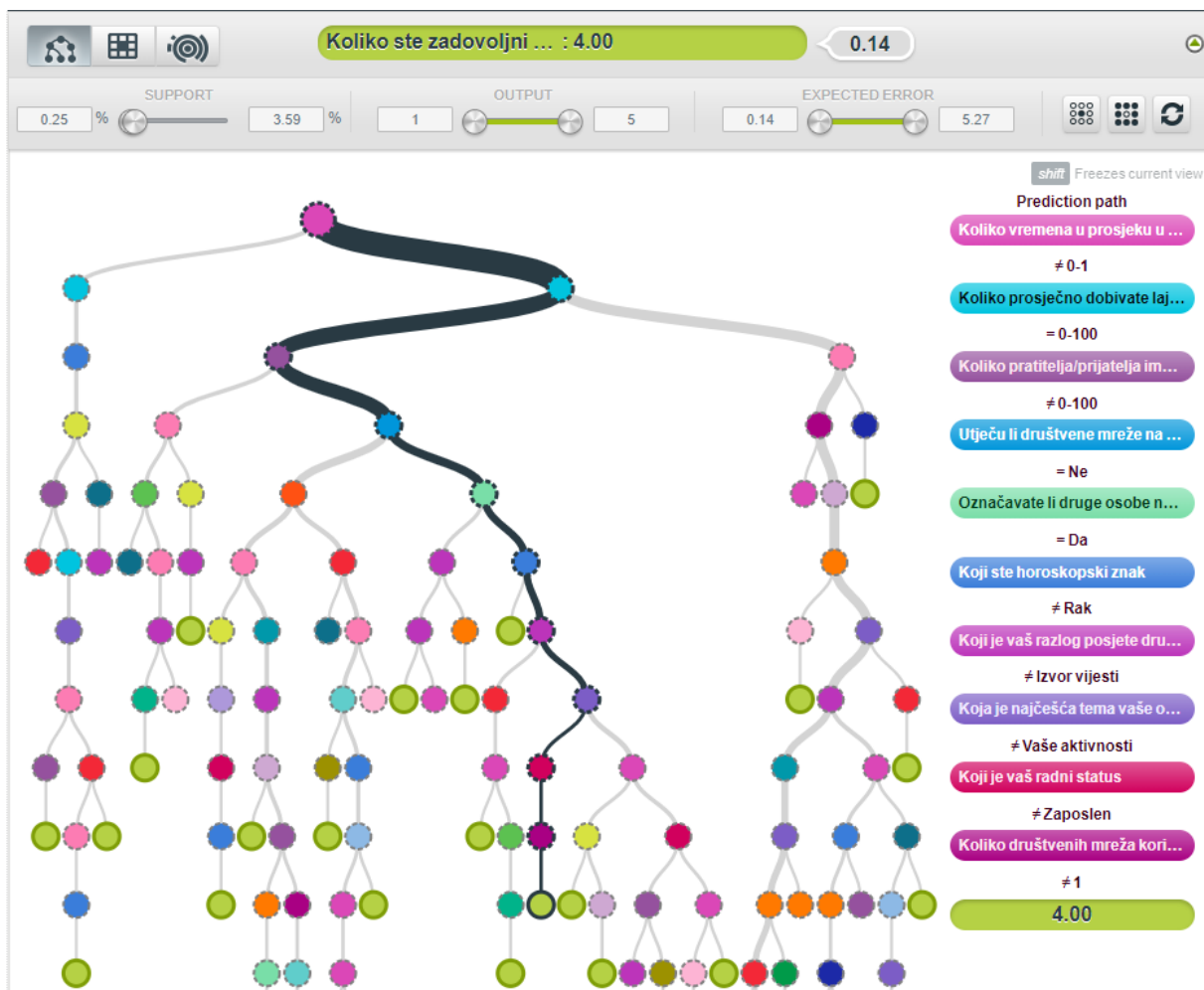
4.2. Prikaz dobivenih rezultata

Odabirom akcije „Model Summary Report“ može se pogledati o kojim atributima najviše ovisi promatrani atribut. Ovisnost atributa zadovoljstvo korisnika društvenim mrežama:

1. Koliko vremena u prosjeku u danu provedete na društvenim mrežama? (u satima):
10.45%
2. Hobi: 8.74%
3. Koji ste horoskopski znak?: 8.54%
4. Koliko prosječno dobivate lajkova po objavi?: 7.93%
5. Označavate li tuđe objave sa sviđa mi se?: 6.68%
6. Koji je vaš razlog posjete društvenih mreža?: 6.60%
7. Koliko pratitelja/prijatelja imate na društvenim mrežama?: 5.60%
8. Stručna sprema: 4.45%
9. Koja je najčešća tema vaše objave?: 4.39%
10. Starost (unesite broj godina): 3.77%
11. Poznajete li sve osobe koje pratite ili imate za prijatelje na društvenim mrežama?:
3.60%
12. Komentirate li objave drugih osoba?: 3.58%
13. Koliko društvenih mreža koristite?: 2.92%
14. Utječu li društvene mreže na vaš stil života?: 2.40%
15. Koji je vaš radni status?: 2.21%
16. U kojoj mjeri koristite emotikone?: 2.12%
17. Bavite li se sportom?: 2.10%
18. Koje vrsta objava najviše objavljujete?: 2.08%
19. Koristite li emotikone?: 2.02%
20. Označavate li druge osobe na objavama?: 1.96%
21. Koristite li više grupne razgovore ili pojedinačne?: 1.88%
22. Smatrate li da bi bez društvenih mreža bili učinkovitiji u obavljanju ostalih dnevnih
aktivnosti?: 1.37%
23. Koristite li interpunkcijske znakove u komunikaciji putem društvenih mreža?: 1.17%
24. Jeste li bili pod utjecajem nasilja putem društvenih mreža ili sudjelovali u njemu?:
1.14%
25. Ljubavni status: 1.11%
26. Utječe li način pisanja na društvenim mrežama na vašu pismenost?: 0.64%
27. Smatrate li se ovisnima o društvenim mrežama?: 0.55%

Prema navedenim podacima vidi se kako zadovoljstvo korisnika najviše ovisi o vremenu koje provode na njemu. Takav podatak može se protumačiti na dva načina, korištenje društvenih mreža jako malo ili korištenje društvenih mreža jako puno. Ukoliko osoba jako malo vremena izdvoji na korištenje društvenih mreža onda postoji mogućnost da je ona zadovoljna društvenim mrežama iz razloga što jednostavno ne primjećuje njihove nedostatke, drugim riječima ne provede dovoljno vremena na društvenim mrežama kako bi mogla naići na neki njihov nedostatak. Promatra se drugo gledište, da osoba provodi jako puno vremena na društvenim mrežama. Ukoliko osoba provodi jako puno vremena na društvenim mrežama tada ona može uočiti jako puno prednosti i nedostataka društvenih mreža te njezino zadovoljstvo ovisi o omjeru prednosti i nedostataka. Ukoliko osoba uvidi da društvene mreže imaju više prednosti tada je osoba zadovoljna društvenim mrežama, ukoliko osoba pronađe veći broj nedostataka tada je ona nezadovoljna društvenim mrežama.

Kako bi se pronašlo pravilo unutar stabla koje najpouzdanije predviđa zadovoljstvo korisnika društvenim mrežama potrebno je pronaći ono pravilo koje ima najmanji „Expected error“, odnosno najmanje odstupanje. Slika prikazuje pravilo koje daje najpouzdanije predviđanje. Pravilo se nalazi s desne strane od stabla odlučivanja.



Slika 10. Prikaz najpouzdanijeg pravila (Prema: B. Kliček i D. Oreški, 2020)

Maksimalna vrijednost koju su korisnici društvenih mreža u anketi birali za zadovoljstvo iznosi 5, što znači da ako je očekivano odstupanje 0.14, tada se pouzdanost pravila predviđanja izračunava na način da od maksimalne vrijednosti oduzmemo očekivano odstupanje te njihovu razliku podijelimo s maksimalnom vrijednosti. $(5-0.14)/5=0.972$. Pouzdanost pravila predviđanja iznosi 97.2%.

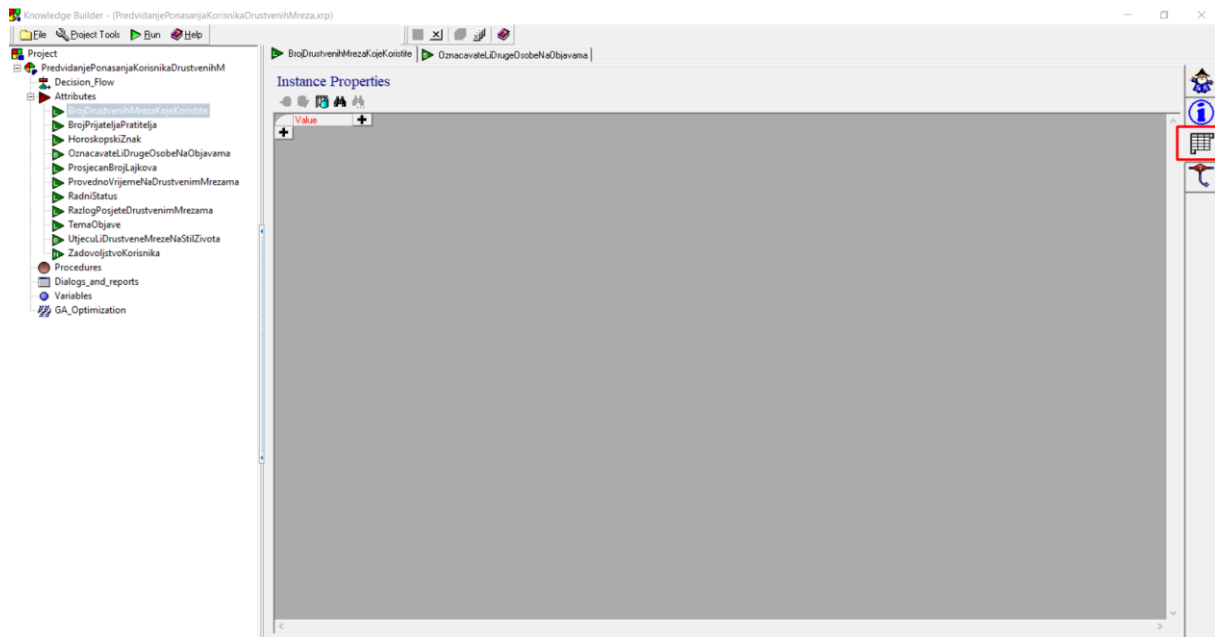
4.3. Izrada ekspertnog sustava u XRKB-u

XRKB je skraćena za Expert Knowledge Bulider. Navedeni program se koristi za kreiranje ekspertnog sustava. Ekspertni sustavi se grade na temelju stabla odlučivanja. U poglavlju 4.1. Izrada stabla odlučivanja u BigML-u kreirano je stablo odlučivanja koje će se koristiti kao temelj za ekspertni sustav. Mora se naglasiti kao se koristi demo verzija XRKB-a te se može izraditi samo 15 elemenata, što znači da će se izraditi samo pravilo za koje je u ranijem poglavlju određeno da je najpouzdanije.

Atributima su dodjeljena sljedeća imena:

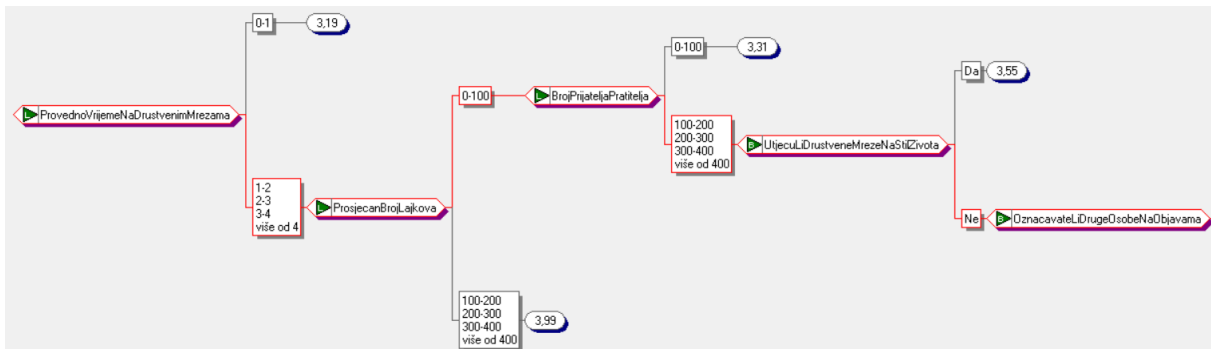
1. BrojDruštvenihMrežaKojeKoristite (List_Attribute, predstavlja odgovor na pitanje „Koliko društvenih mreža koristite?“)
2. BrojPrijateljaPratitelja (List_Attribute, predstavlja odgovor na pitanje „Koliko pratitelja/prijatelja imate na društvenim mrežama?“)
3. HoroskopskiZnak (List_Attribute, predstavlja odgovor na pitanje „Koji ste horoskopski znak?“)
4. OznacavateLiDrugeOsobeNaObjavama (Boolean_Attribute, predstavlja odgovor na pitanje „Označavate li druge osobe na objavama?“)
5. ProsjecanBrojLajkova (List_Attribute, predstavlja odgovor na pitanje „Koliko prosječno dobivate lajkova po objavi?“)
6. ProvednoVrijemeNaDruštvenimMrežama (List_Attribute, predstavlja odgovor na pitanje „Koliko vremena u prosjeku u danu provedete na društvenim mrežama? (u satima)“)
7. RadniStatus (List_Attribute, predstavlja odgovor na pitanje „Koji je vaš radni status?“)
8. RazlogPosjeteDruštvenimMrežama (List_Attribute, predstavlja odgovor na pitanje „Koji je vaš razlog posjete društvenih mreža?“)
9. TemaObjave (List_Attribute, predstavlja odgovor na pitanje „Koja je najčešća tema vaše objave?“)
10. UtjecuLiDruštveneMrežeNaStilZivota (Boolean_Attribute, predstavlja odgovor na pitanje „Utječu li društvene mreže na vaš stil života?“)
11. ZadovoljstvoKorisnika (Numeric_Attribute, predstavlja odgovor na pitanje „Koliko ste zadovoljni korištenjem društvenih mreža?“)

Za sve List_Attribute moraju se unijeti vrijednosti koje atributi mogu poprimiti. Kao primjer može se uzeti prvi atribut BrojDruštvenihMrežaKojeKoristite. Dvoklik na atribut te s desne strane se odabire „Instance Properties“. Na slici ispod označen crvenim pravokutnikom, ovdje se unose vrijednosti koje atribut može poprimiti.

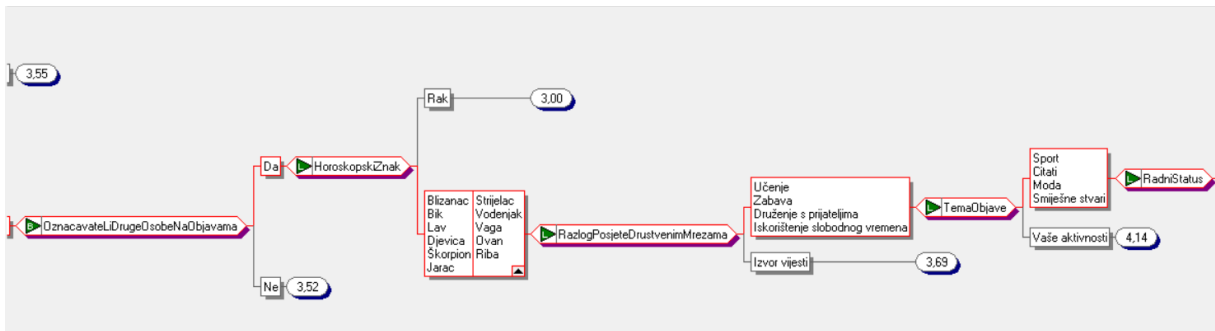


Slika 11. Akcija koja omogućava dodavanje vrijednosti atributima u alatu XRKB (Prema: B. Kliček i D. Oreški, 2020)

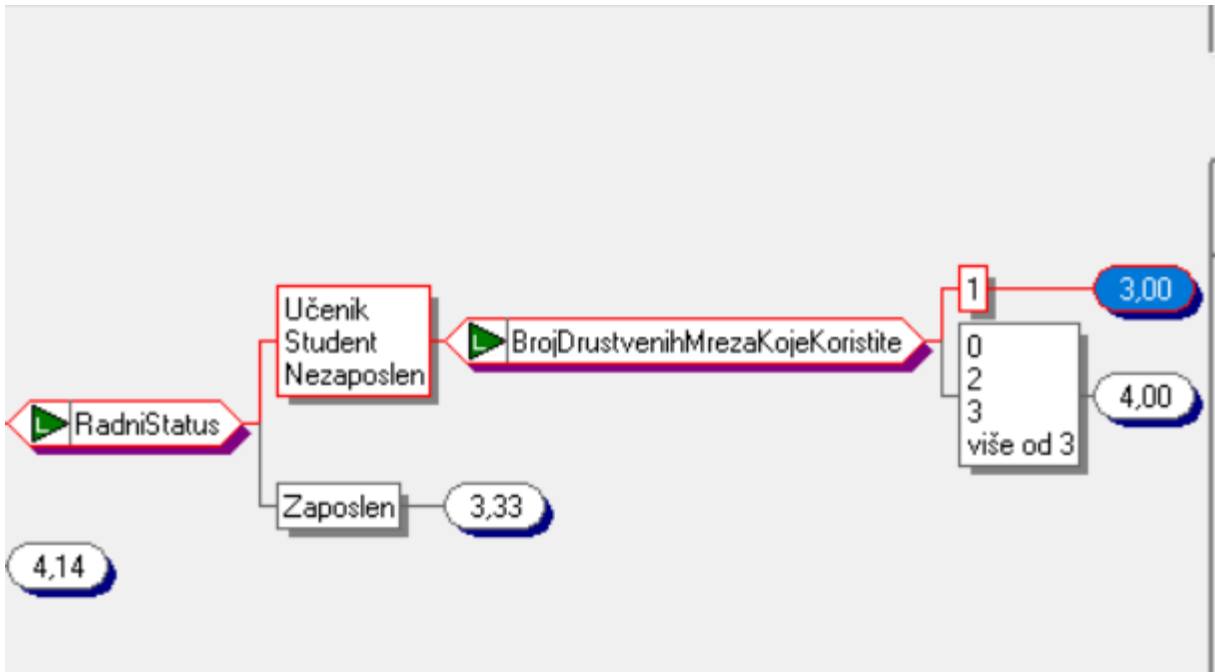
Nakon što se unesu vrijednosti za sve List_Attribute može se postaviti način prikaza znanja. Otvara se atribut ZadovoljstvoKorisnika i odabire se Object Wizards (ikona malog čarobnjaka). Odabire se „Knowledge Representation“ te je još samo bitno da se za odabir prikaza znanja odabere stablo. Navedeni korak ostvaruje se na način da se označi checkbox Tree. Sada se može unijeti pravilo koje se dobilo pomoću alata BigML. Atributi se dodaju povlačenjem iz lijevog prozora na prazno mjesto. Prikaz reprezentacije znanja:



Slika 12. Kreirano najpouzdanije pravilo (dio 1) (Prema: B. Kliček i D. Oreški, 2020)



Slika 13. Kreirano najpouzdanije pravilo (dio 2) (Prema: B. Kliček i D. Oreški, 2020)



Slika 14. Kreirano najpouzdanije pravilo (dio 3) (Prema: B. Kliček i D. Oreški, 2020)

4.4. Testiranje sustava

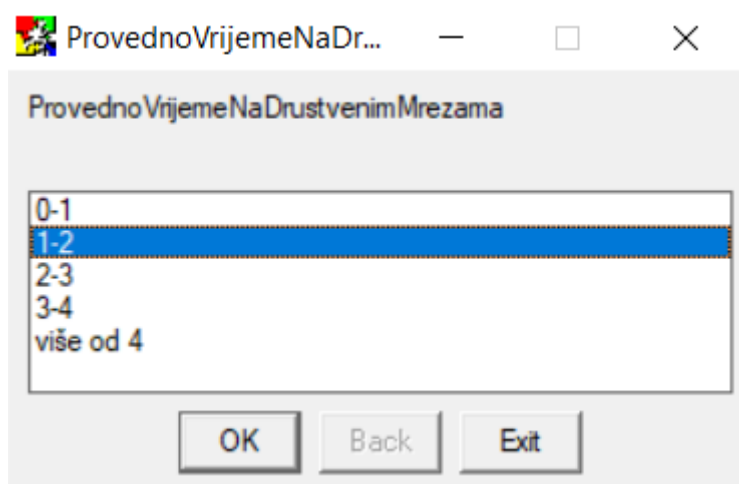
Kako bi se provelo predviđanje na temelju pravila potrebno je otvoriti „Knowledge Module“ koji je kreiran. U ovome slučaju je to PredviđanjePonasanjaKorisnikaDrustvenihMreza.

Ovdje se zalijepi tekste:

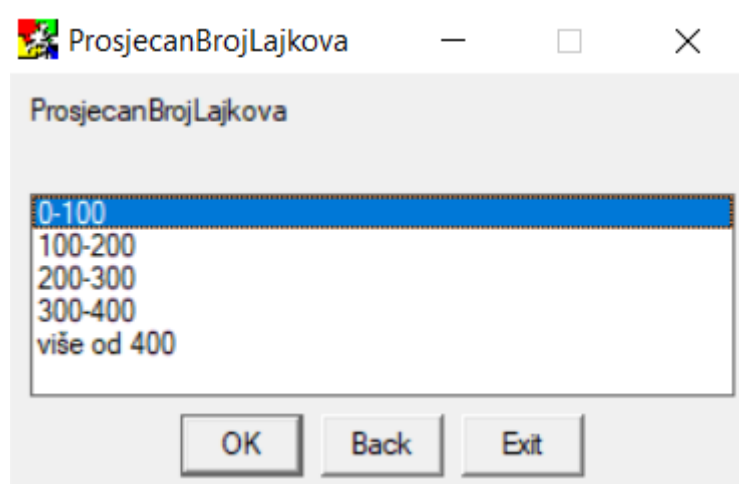
@Do ZadovoljstvoKorisnika

@Debug ZadovoljstvoKorisnika

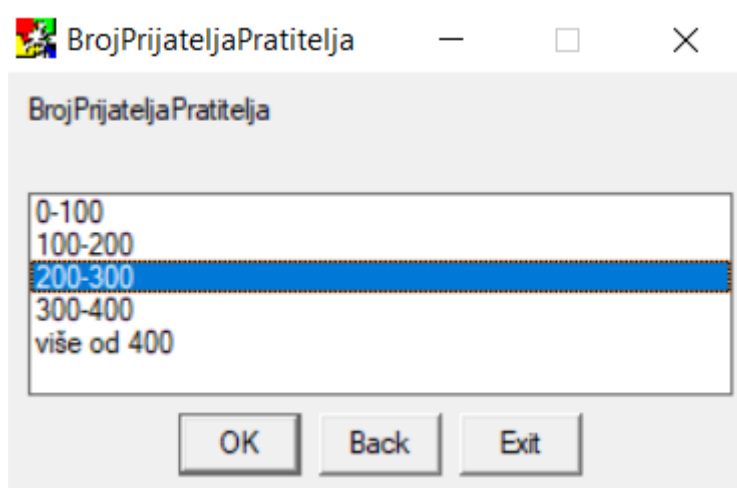
Kako bi se odredilo što će se pokrenuti kada se pokrene ekspertni sustav. Pomoću ekspertnog sustava će se izvršiti predviđanje za slučaj pravila.



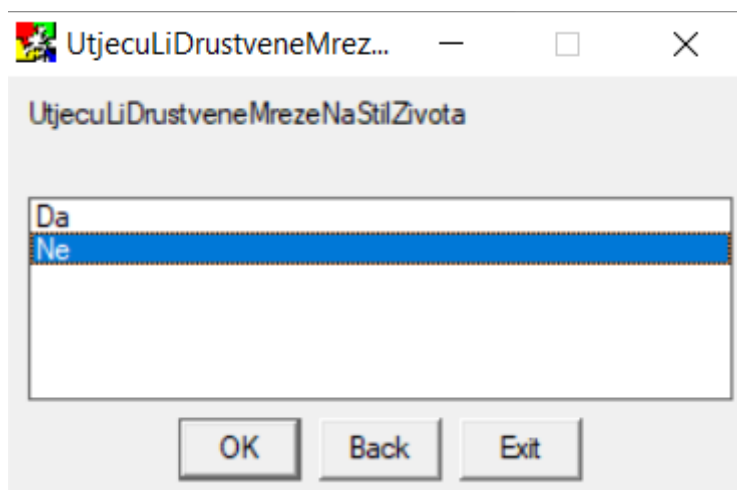
Slika 15. Provođenje predviđanja (Koliko vremena u prosjeku u danu provedete na društvenim mrežama? (u satima)) (Prema: B. Kliček i D. Oreški, 2020)



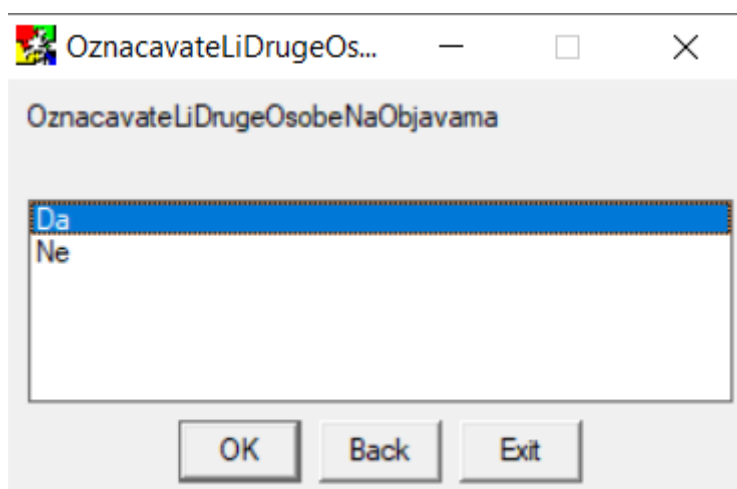
Slika 16. Provođenje predviđanja (Koliko prosječno dobivate lajkova po objavi?) (Prema: B. Kliček i D. Oreški, 2020)



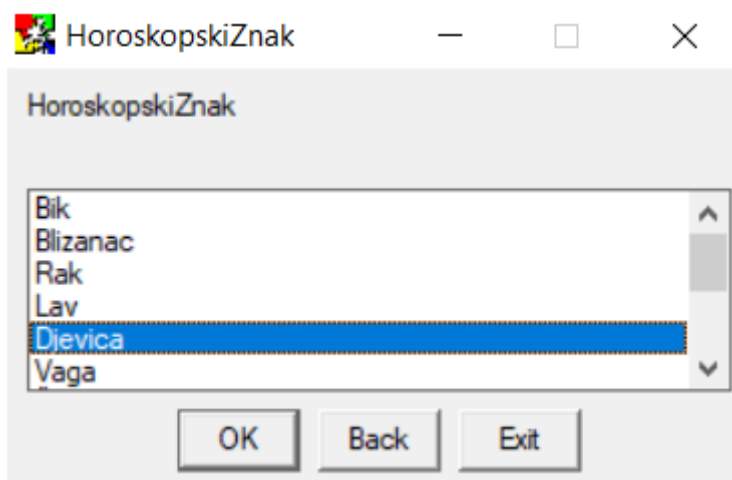
Slika 17. Provođenje predviđanja (Koliko pratitelja/prijatelja imate na društvenim mrežama?)
(Prema: B. Kliček i D. Oreški, 2020)



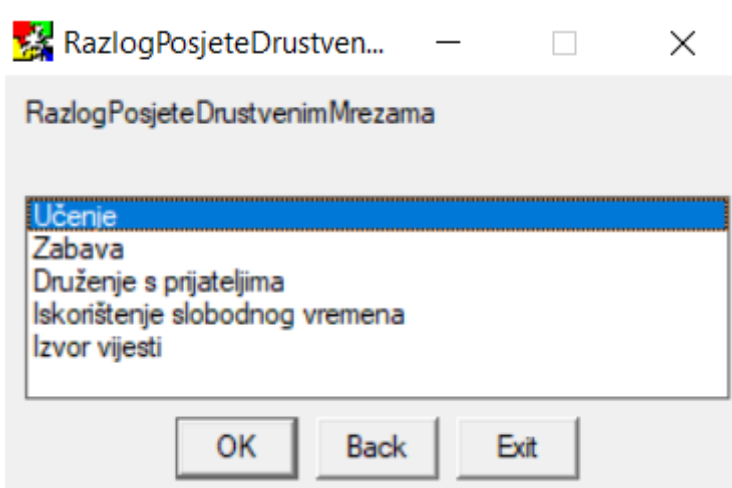
Slika 18. Provođenje predviđanja (Utječu li društvene mreže na vaš stil života?) [19]



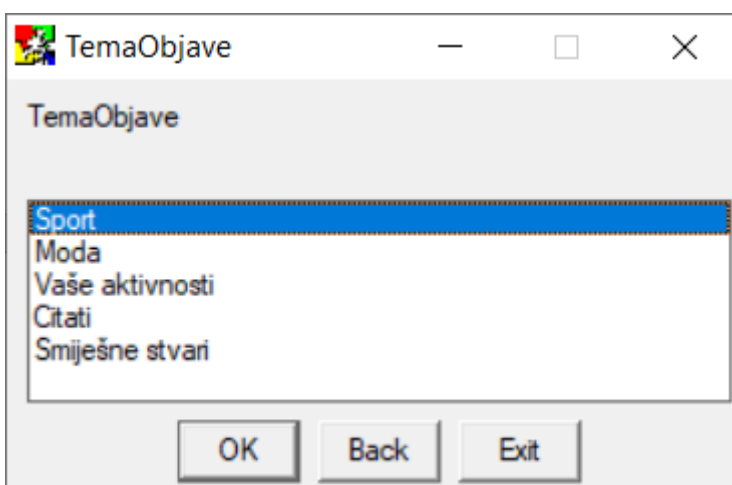
Slika 19. Provođenje predviđanja (Označavate li druge osobe na objavama?) (Prema: B. Kliček i D. Oreški, 2020)



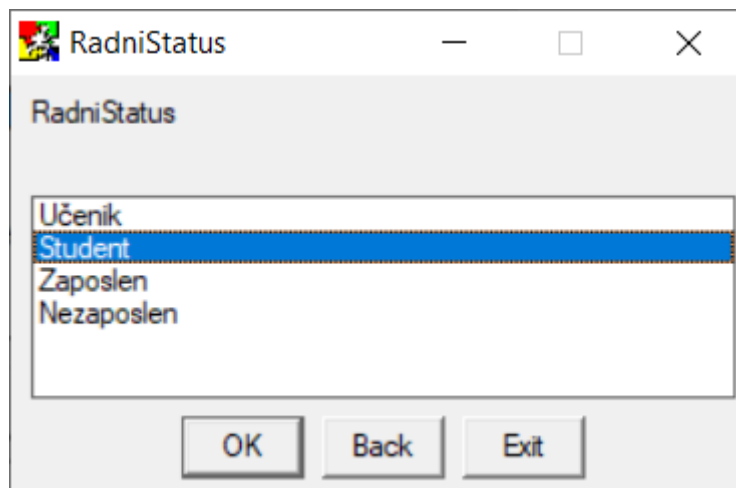
Slika 20. Provođenje predviđanja (Koji ste horoskopski znak?) (Prema: B. Kliček i D. Oreški, 2020)



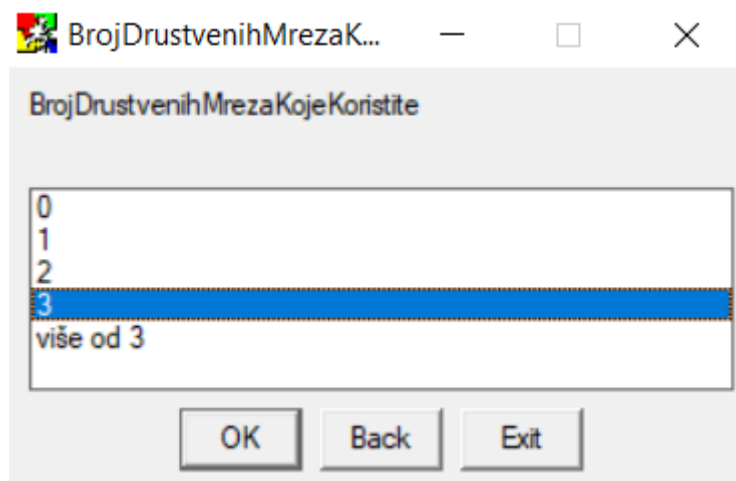
Slika 21. Provođenje predviđanja (Koji je vaš razlog posjete društvenih mreža?) (Prema: B. Kliček i D. Oreški, 2020)



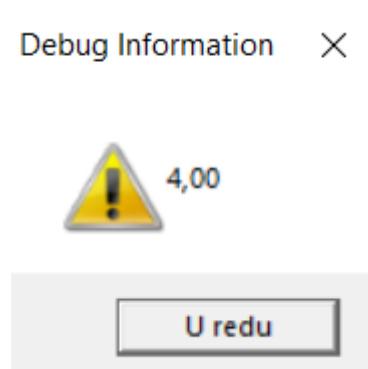
Slika 22. Provođenje predviđanja (Koja je najčešća tema vaše objave?) (Prema: B. Kliček i D. Oreški, 2020)



Slika 23. Provođenje predviđanja (Koji je vaš radni status?) (Prema: B. Kliček i D. Oreški, 2020)



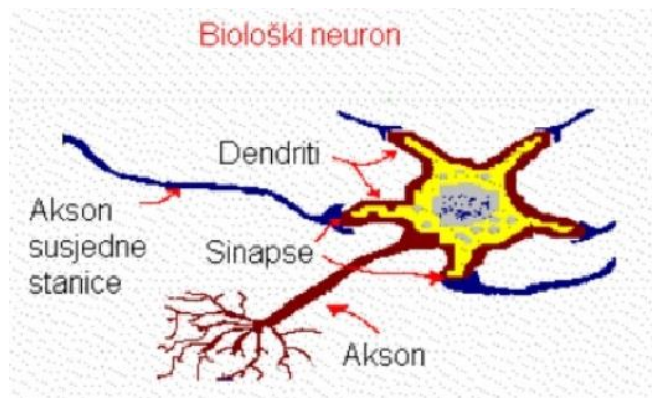
Slika 24. Provođenje predviđanja (Koliko društvenih mreža koristite?) (Prema: B. Kliček i D. Oreški, 2020)



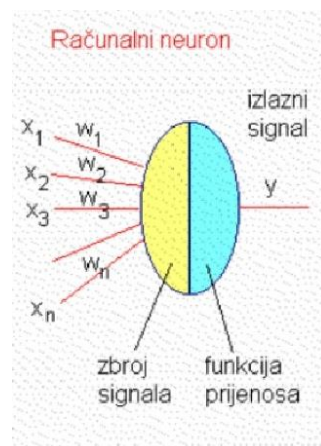
Slika 25. Rezultat provedenog predviđanja (Prema: B. Kliček i D. Oreški, 2020)

5. Neuronske mreže

Pojam „neuronska mreža“ može imati dva značenja. Jedno značenje se odnosi na tradicionalno značenje da se radi o mreži neurona koji se nalaze u živčanom sustavu, dok je drugo značenje da se radi o računalnoj neuronskoj mreži koja se također sastoji od neurona, ali računalnih. Njihov naziv dolazi od pravih prirodnih neuronskih mreža jer su modelirane prema biološkim procesima u živčanome sustavu. Neuron je osnovna jedinica živčanog sustava, a ujedno i najsloženija jedinica ljudskog organizma. Na slikama ispod su prikazani biološki neuron te računalni neuron. [9]



Slika 26. Biološki neuron [8]



Slika 27. Računalni neuron [8]

Neuroni se sastoje od dendrita, aksona, sinapsi te jezgre neurona. Svaki dio neurona ima svoju ulogu, tako na primjer dendriti primaju podražaj, odnosno njih se može zamisliti kao ulaze u neuron. Jezgru neurona obrađuje signal, što se može poistovjetiti sa zbrajanjem signala te funkcijom prijenosa. Računalni neuroni imaju izlazni signal što može predstavljati impulsom koji se preko aksona prenosi dalje do sinapsi koje podražavaju dendrite drugih neurona.

Na slici računalnog neurona prikazani su ulazi $X_1, X_2, X_3, \dots, X_n$ sa svojim pridruženim težinama. Drugi element računalnog neurona je zbroj signala koji obavlja sumiranje signala te poziva funkciju prijenosa. Funkcija prijenosa se bira prema vrsti problema koju mreža treba riješiti. [10]

Najčešće funkcije prijenosa s kojima se modelira da li je neuron aktivan ili ne na osnovu ulaznih signala su:

1. Linearna funkcija
2. Funkcija praga
3. Djelomična linearna funkcija
4. Sigmoidalna funkcija
5. Tangentna hiperbolička funkcija

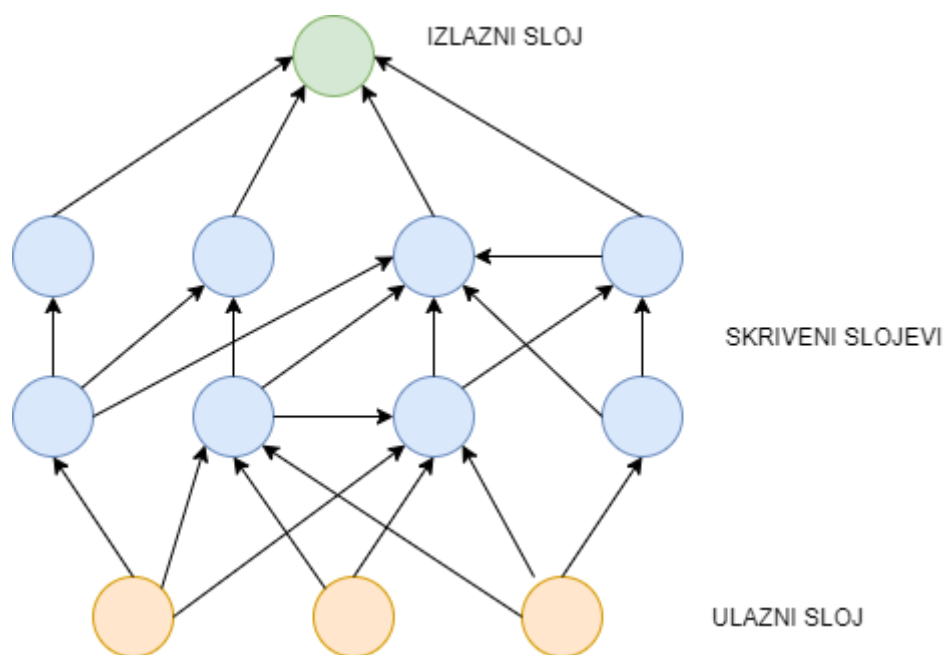
Od navedenih funkcija najkorištenije su funkcija praga, djelomična linearna funkcija te sigmoidalna funkcija. [10]

Funkcija praga vraća 1 ukoliko je suma ulaznih signala pozitivna te 0 ukoliko je suma ulaznih signala negativna. U slučaju da se radi o bipolarnom pragu tada funkcija vraća 1 ukoliko je suma ulaznih signala pozitivna, ukoliko je suma ulaznih signala negativna tada funkcija vraća -1. Djelomično linearna funkcija je nagnuta funkcija koja poprima sljedeće vrijednosti. Ukoliko je suma ulaznih signala pozitivna tada vraća 1, ukoliko je suma ulaznih signala negativna tada vraća -1 te ukoliko je suma signala nije ni pozitivna ni negativna tada vraća proporcionalnu vrijednost ulaznim vrijednostima u intervalu od -1 do 1. Sigmoidalna funkcija je zakrivljena funkcija te je to najčešći tip funkcije korištene za kreiranje neuronske mreže. Sigmoidalna funkcija prikazana je sljedećom formulom:

$$\text{Sigmoida}(I) = \frac{1}{1 + e^{-\lambda I}}$$

U formuli λ predstavlja nagib funkcije. Kao i sve ostale funkcije prijenosa, sigmoidalna funkcija poprima vrijednosti od 0 do 1. [10]

Neuronske mreže primaju ulazne parametre koje obrađuju i na temelju njih daju izlanu vrijednost te iz toga razloga ju dijelimo u slojeve. Ulazni sloj je sloj koji prima ulazne parametre, izlazni sloj koji vraća vrijednost te jedan ili više skrivenih slojeva koji transformiraju ulazne parametre u izlanu vrijednost. Svaki neuron može biti povezan iz jednog sloja može biti povezan sa svakim neuronom iz istog sloja ili s neuronom iz drugog sloja, što je prikazano na slici ispod.



Slika 28. Slojevi neuronske mreže (Prema: Goran Pintarić, 2013)

5.1. Kreiranje neuronske mreže

Faze rada umjetnih neuronskih mreža moguće je podijeliti na:

- Fazu učenja
- Fazu selekcije
- Fazu testiranja
- Operativnu fazu

Dvije najvažnije faze su faza učenja i faza testiranja, u fazi selekcije nastoji se optimizirati duljina treniranja te broj neurona u skrivenom sloju. Operativna faza se odnosi na korištenje neuronske mreže.[9]

Sve započinje odabirom ulaznih i izlaznih varijabli te pripremom podataka. Potrebo je podijeliti podatke na dva skupa, prvi skup je skup podataka za treniranje (učenje) mreže, dok je drugi skup namijenjen za testiranje neuronske mreže. Što se tiče podjele podataka ne postoji pisano pravilo koje govori o tome koliko bi se podataka trebalo dodijeliti kojoj fazi, preporuča se najveći dio podataka iskoristiti za treniranje mreže. Faza učenja je zapravo proces podešavanja težina u mreži koja se odvija kroz nekoliko tisuća ponavljanja. Kod samoga treniranje mreže potrebno je odrediti koliko se dugo mreža treba trenirati kako bi rezultat bio što bolji, odnosno kako bi greška bila što manja. Ukoliko se mreža pretrenira, može doći do toga da se mreža previše navikne na skup podataka za učenje. Nakon faze učenja dolazi faza

testiranja, u ovoj fazi mreža više ne uči, već se donosi ocjena mreže. Kroz fazu testiranja se daje odgovor na pitanje „Koliko točne rezultate mreža daje?“. U ovoj fazi koriste se podaci za testiranje, na ulaz mreže postavljaju se vrijednosti te mreža daje izlaznu vrijednost. Dobivenu izlaznu vrijednost uspoređujemo sa očekivanom vrijednošću te donosimo ocjenu. [9]

5.2. Primjena neuronskih mreža

Neuronske mreže primjenjuju se u raznim područjima, neka od njih su zdravstvo, obrazovanje, elektrotehnika, poslovanje, vojska, financije i brojne druge gospodarske grane. Primjeri primjene neuronskih mreža su:

- Otkrivanje eksploziva z prtljazi na aerodromu
- Određivanje problema kod rada avio-motora
- Određivanje tipova oblaka temeljem postojećih satelitskih snimki
- Obrada signala
- Prepoznavanje govora
- Pretvaranje teksta u govor
- Financije (kredit, burze)
- Marketing [9]

Kako s vremenom raste uporaba neuronskih mreža tako se i povećava broj algoritama za učenje neuronskih mreža. Neki od tih algoritama su prilagođeni za rješavanje samo jedne vrste problema, dok su neki univerzalni i moguće ih je koristiti za sve tipove problema. U nastavku slijedi podjela algoritama. [9]

Algoritmi se dijele na:

- Nadgledane – algoritmi za koje su potrebne poznate ulazne i izlazne vrijednosti
- Nenadgledane – algoritmi za koje su potrebne samo ulazne vrijednosti

U nadgledane algoritme ubrajamo:

1. Za probleme predviđanja
 - Mreža širenja unatrag
 - Modularna mreža
 - Mreža s radijalno zasnovanom funkcijom
 - Mreža opće regresije
2. Za probleme klasifikacije
 - Vjerojatnostna mreža
 - Mreža učeće vektorske kvantizacije

U nenadgledane algoritme ubrajaju se:

- Kohonenova mreža
- Mreža adaptivne rezonantne teorije

5.2.1. Mreža širenja unatrag

Algoritam širenja natrag je najpopularnije pravilo učenja. Način na koji algoritam radi je opisan u nastavku kroz pet koraka:

1. Prvo se postavljaju ulazne vrijednosti koje se preko skrivenog sloja prenose do izlaznog sloja i ostvaruju izlani vektor.
2. Za vrijeme prenošenja ulazno vektora računaju se ulazne i izlazne vrijednosti za svaki neuron u skrivenom sloju.
3. Za svaki neuron u izlaznom sloju računa se lokalna greška.
4. Za svaki sloj, počevši od predzadnjeg i sloja neposredno nakon ulaznog potrebno je izračunati skaliranu lokalnu grešku i delta težinu.
5. Obnoviti sve težine veza u mreži dodavanjem delta težina prijašnjim vrijednostima.

Prednost mreže širenja unatrag je što ima dodatne slojeve koji dopuštaju da se rezultat jednog sloja dodatno obrađuje, uređuje i stvara kompleksni sustav, kao nedostatak navodi se dugotrajno treniranje te osjetljivost na početne vrijednosti težina. [8]

5.2.2. Vjerojatnostna mreža

Vjerojatnosna mreža je statistički pristup koji koristi tako zvane Bayesove klasifikatore i Perzenove procjenjivače, koji računaju funkcije gustoće vjerojatnosti koje zahtjeva Bayesova teorija razmatrajući relativnu vjerojatnost događaja i koristeći informacije kako bi poboljšala predviđanje i minimizirala očekivanu grešku. Unutar mreže se svaki element tretira samoj jednom. Prednost ovoga algoritma je što faza učenja mreže je znatno brža jer se treniranje većinom provodi u jednom prolasku. Druga prednost je što mreža dozvoljava proširenje skupa za treniranje u bilo kojem trenutku te se ne mora cijela mreža ponovno trenirati. Nedostatak mreže je što mreža može postati jako velika i spora prilikom treniranja velikog uzorka pa time postane nepraktična za rješavanje problema klasifikacije. [8]

5.2.3. Mreže s radijalno zasnovanom funkcijom

Kao što samo ime kaže funkcije koje koristi ova mreža su radijalno simetrične, odnosno svaki neuron proizvodi identičan izlaz koji se nalazi na određenoj radijalnoj udaljenosti od centra.

Kao i sve ostale vrste neuronskih mreža tako i ova vrsta ima svojih prednosti i nedostataka. Kao prednosti i nedostatke se navodi sljedeće:

- Izvrsno aproksimiraju i mogu se istrenirati jednostavno i brzo, no daju slab odaziv u fazi ponavljanja zbog velikog broja neurona povezanih s drugim slojem.
- Linearne težine povezane s izlaznim slojem mogu se obrađivati odvojeno od skrivenog sloja neurona.
- Težine skrivenog sloja prilagođavaju se kroz nelinearnu optimizaciju, a težine izlaznog sloja kroz linearnu optimizaciju.
- Ne preporuča se korištenje mreža s radijalno zasnovanom funkcijom za pronalaženje rješenja ukoliko imamo mali skup podataka.
- Mreže s radijalno zasnovanom funkcijom su idealne za rješavanje klasifikacijskih problema. [8]

5.3. Primjer zadatka neuronskih mreža:

Izračunajte korekciju težine veze neurona br. 2 prema neuronu br. 4 kod neuronske mreže „širenja unatrag“. Težine veze i struktura prikazani su u tablici. Neuroni od 1 do 3 ostvaruju prijenos ulazne veličine bez ikakve izmjene. Neuroni od 4 do 6 imaju funkciju prijenosa sigmoide. Koeficijent učenja za skriveni sloj je 0,40. Ne koristi se momentni član niti druge korekcije funkcije u učenju. Traženi izlaz iz neurona br. 6 je 0,600.

Ulazi neurona su 1: 0,900; 2: 0,850; 3: -0,650

Neuron br. 4		Neuron br. 5		Neuron br. 6	
Veza s	Težina v.	Veza s	Težina v.	Veza s	Težina v.
1	0,5432	1	0,5678	1	0,8000
2	-0,7000	2	1,7000	4	-0,6234
3	1,0450	3	-0,6900	5	-0,8765

Tablica 2. Podaci o neuronima (Prema: B. Kliček i D. Oreški, 2020)

Izlaz za neuron 4 (Y_4):

$$I = \sum x_i \cdot w_i$$

I-predstavlja sumu ulaza u neuron 4

x-ulazi iz neurona

w-težina veze

$$I_4 = (0,900 \cdot 0,5432) + (0,850 \cdot (-0,7000)) + (-0,650 \cdot 1,0450) = -0,7854$$

$$\text{Sigmoida}(I) = \frac{1}{1 + e^{-I}}$$

$$Y_4 = \frac{1}{1 + e^{0,7854}} = 0,3132$$

Izlaz za neuron 5 (Y_5):

$$I_5 = (0,900 \cdot 0,5678) + (0,850 \cdot 1,7000) + (-0,650 \cdot (-0,6900)) = 2,4045$$

$$Y_5 = \frac{1}{1 + e^{-2,4045}} = 0,9172$$

Izlaz za neuron 6 (Y_6):

$$I_6 = (0,900 \cdot 0,8000) + (0,3132 \cdot (-0,6234)) + (0,9172 \cdot (-0,8765)) = -0,2791$$

$$Y_6 = \frac{1}{1 + e^{0,2791}} = 0,4307$$

Greška za izlazni neuron:

$$E = f'(I) \cdot (D - A)$$

$$f'(I) = f(I) \cdot (1 - f(I)) = 0,4307 \cdot (1 - 0,4307) = 0,2452$$

$$E_6 = 0,2452 \cdot (0,600 - 0,4307) = 0,0415$$

Greška za skriveni neuron:

$$E = f'(I) \cdot (E_i \cdot W_i)$$

$$f'(I) = f(I) \cdot (1 - f(I)) = 0,3132 \cdot (1 - 0,3132) = 0,2151$$

$$E_4 = 0,2151 \cdot 0,0415 \cdot (-0,6234) = -0,0056$$

Korekcija težine veze:

$$\text{delta}(w) = \text{LCoef} \cdot E_{j+1} \cdot x_{j-1}$$

$$\text{delta}(w) = 0,40 \cdot (-0,0056) \cdot 0,850 = -0,0019$$

(Prema: B. Kliček i D. Oreški, 2020)

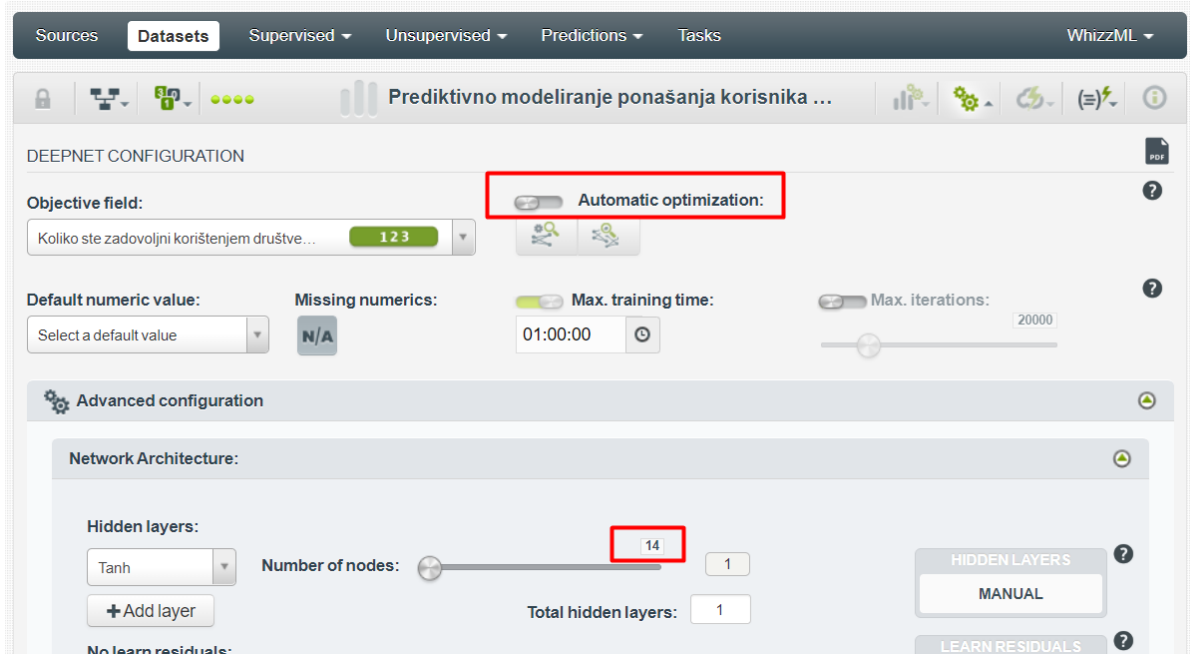
5.4. Izrada neuronske mreže pomoću alata BigML

Kako bi se kreirala neuronska mreža pomoću alata BigML mora se odabrati Dataset koji je kreiran prilikom kreiranja stabla odlučivanja. Umjesto „Model“ koji je odabran za izradu stabla odlučivanja, ovaj put odabire se „Deepnet“ kao što je prikazano na slici ispod.



Slika 29. Ikona Deepnet BigML (Prema: B. Kliček i D. Oreški, 2020)

Otvoraju se postavke za kreiranje neuronske mreže. U ovome koraku mora se odrediti koliko neurona treba biti u skrivenom sloju. Pošto skup podataka ima ukupno 29 atributa, a predviđa se jedan od njih, što znači da je on na izlazu. Ostaje još 28 atributa te se uzima polovica njih za skriveni sloj, njih 14. Za postavljanje broja neurona u skrivenom sloju potrebno je isključiti Automatic optimization, kao što je prikazano na slici ispod.



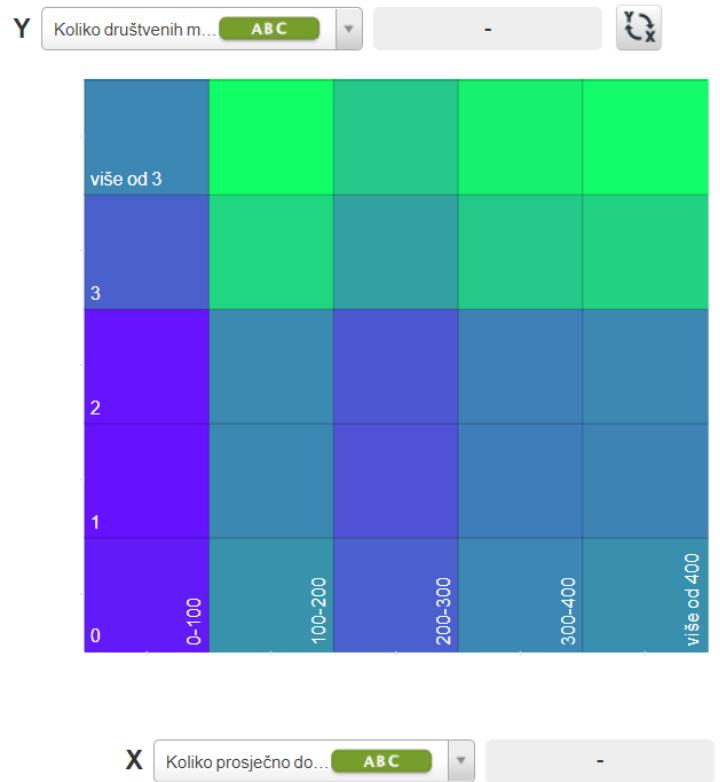
Slika 30. Postavljanje broja neurona u skrivenom sloju (Prema: B. Kliček i D. Oreški, 2020)

Kako bismo provjerili o kojim atributima najviše ovisi predviđanje zadovoljstva korisnika društvenim mrežama odabiremo Deepnet Summary Report.

Atributi o kojima najviše ovisi predviđanje:

1. Koliko prosječno dobivate lajkova po objavi? 18.13%
2. Koliko društvenih mreža koristite? 15.49%
3. Hobi 11.72%
4. Koji ste horoskopski znak? 10.72%
5. Označavate li tuđe objave sa sviđa mi se? 8.91%
6. Koja je najčešća tema vaše objave? 8.87%
7. Koliko vremena u prosjeku u danu provedete na društvenim mrežama? (u satima) 5.28%
8. Koji je vaš radni status? 3.94%
9. Komentirate li objave drugih osoba? 3.03%
10. Koje vrsta objava najviše objavljujete? 2.87%
11. Smatrate li da bi bez društvenih mreža bili učinkovitiji u obavljanju ostalih dnevnih aktivnosti? 1.94%
12. Koliko pratitelja/prijatelja imate na društvenim mrežama? 1.89%
13. Smatrate li se ovisnima o društvenim mrežama? 1.57%
14. Koristite li interpunkcijske znakove u komunikaciji putem društvenih mreža? 1.23%
15. Koristite li emotikone? 0.88%
16. U kojoj mjeri koristite emotikone? 0.71%
17. Poznajete li sve osobe koje pratite ili imate za prijatelje na društvenim mrežama? 0.57%
18. Utječe li način pisanja na društvenim mrežama na vašu pismenost? 0.38%
19. Stručna sprema 0.33%
20. Označavate li druge osobe na objavama? 0.3%
21. Koji je vaš razlog posjete društvenih mreža? 0.27%
22. Koristite li više grupne razgovore ili pojedinačne? 0.26%
23. Spol 0.19%
24. Starost 0.14%
25. Utječu li društvene mreže na vaš stil života? 0.13%
26. Bavite li se sportom? 0.09%
27. Ljubavni status 0.08%
28. Jeste li bili pod utjecajem nasilja putem društvenih mreža ili sudjelovali u njemu? 0.07%

Odabiru se na Y os atribut „Koliko društvenih mreža koristite?“ te na X os „Koliko prosječno dobivate lajkova po objavi?“ iz razloga jer su to dva atributa o kojima najviše ovisi predviđanje.



Slika 31. Prikaz kombinacije atributa "Koliko društvenih mreža koristite?" i "Koliko prosječno dobivate lajkova po objavi?" (Prema: B. Kliček i D. Oreški, 2020)

Kombiniranjem njihovih vrijednosti predviđa se ocjena zadovoljstva društvenih mreža u intervalu od 3.33 do 3.88. Kombinacijom atributa „Hobi“ i „Koliko prosječno dobivate lajkova po objavi?“ predviđa se ocjena zadovoljstva društvenim mrežama u intervalu od 3.40 do 3.97. Kao posljednju kombinaciju ćemo uzeti attribute „Hobi“ i „Koliko društvenih mreža koristite“. Prema njihovoj kombinaciji ocjena zadovoljstva korisnika društvenim mrežama kreće se u intervalu od 3.49 do 4.05.

6. Bayesove mreže

U ovome poglavlju prikazat će se izvod Bayesove formule, navest će se sve pojmove vezane uz teoriju grafova kako biste lakše razumjeli samu definiciju Bayesovih mreža.

6.1. Bayesova formula

Za izvod Bayesove formule potrebna nam je uvjetna vjerojatnost. Formulom će se definirati vjerojatnost događaja A uz uvjet događaja B te ona glasi:

$$p(A|B) = \frac{p(AB)}{p(B)}$$

Na isti način se definira vjerojatnost događaja B uz uvjet događaja A.

$$p(B|A) = \frac{p(AB)}{p(A)}$$

Ukoliko se iz formule $p(B|A) = \frac{p(AB)}{p(A)}$ izrazi vjerojatnost $p(AB)$ dobiva se izraz $p(AB) = p(B|A) \cdot p(A)$. Njegovim uvrštavanjem u formulu za izračun vjerojatnosti događaja A uz uvjet događaja B dobiva se najjednostavniji oblik Bayesove formule:

$$p(A|B) = \frac{p(B|A) \cdot p(A)}{p(B)}$$

Izvod Bayesove formule izrađen je prema [13].

Budući da je: $P(B) = P(A \cap B) + P(A^c \cap B) = P(B|A)P(A) + P(B|A^c)P(A^c)$ gdje je A^c komplement od A, iz toga razloga Bayesovo pravilo često se primjenjuje kao:

$$p(A|B) = \frac{p(B|A) \cdot p(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

Izrađeno prema [14].

6.2. Primjer primjene Bayesove formule

U jednome nogometnome kampu vrši se selekcija igrača. Pokazalo se da od ukupnog broja igrača koji krenu u školu nogometa njih 60% uspješno priđe kroz sve nogometne uzraste. Ukoliko su igrači završili barem jednu nogometnu školu njih 60% će biti primljeno na pripreme. Preostali igrači koji nisu prošli školu nogometa, a mogu proći selekciju iznosi 30%. Odredite koja je vjerojatnost da ako igrač ode na nogometni kamp će biti primljen na pripreme.

Za izračun će se koristiti Bayesovu formulu. Prvo je potrebno iz teksta izvući vrijednosti:

A- igrač je završio školu nogometa

B- igrač je primljen na pripreme

1. Vjerojatnost da će igrač biti primljen na pripreme ukoliko je prošao školu nogometa iznosi $p(B|A) = 0,60$
2. Vjerojatnost da je igrač prošao školu nogometa, a nije primljen na pripreme iznosi $p(B^c|A) = 0,40$
3. Vjerojatnost da igrač nije prošao školu nogometa, ali je primljen na pripreme iznosi $p(B|A^c) = 0,30$
4. Vjerojatnost da igrač nije prošao školu nogometa, ali nije ni primljen na pripreme iznosi $p(B^c|A^c) = 0,70$

Navedeni podaci se unose u Bayesovu formulu:

$$p(A|B) = \frac{p(B|A) \cdot p(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$
$$p(A|B) = \frac{0,60 \cdot 0,60}{0,60 * 0,60 + 0,30 * 0,40}$$
$$p(A|B) = 0,75$$

Zadatak je izrađen prema primjeru iz [15].

6.3. Teorija grafova

Grafovi su matematičke strukture koje se sastoje od skupa vrhova i skupa bridova koji ih povezuju. Također mnogi matematički problemi se rješavaju pomoću grafova. Jedan od najpoznatijih problema je Koninsberških mostova. U ovome poglavlju navest ćemo definicije koje su potrebne za razumijevanje svih pojmova koji će se naknadno spominjati.

1. (Neusmjereni) graf G je uređeni par $G = (V, E)$, gdje je $V = V(G) \neq \emptyset$ skup vrhova, a $E = E(G)$ skup bridova pri čemu je $E \cap V = \emptyset$, a svaki brid $e \in E$ spaja dva vrha $u, v \in V$ koji se zovu krajevi brida e . Kažemo još da su vrhovi u i v incidentni s bridom e , odnosno da su vrhovi u i v susjedni i pišemo $u \sim_e v$ ili $e = uv$, ili pak $e = \{u, v\}$.
2. Graf G je konačan ako su V i E konačni skupovi, u protivnom je beskonačan.
3. Veličina od G je broj bridova od G i označavamo ga s $\varepsilon(G)$.
4. Potpuni graf je jednostavni graf u kojemu su svaka dva vrha incidentna.
5. Šetnja u grafu G je netrivialni konačni niz $W = v_0 e_1 v_1 e_2 v_2 \dots e_k v_k$ čiji su članovi naizmjenice vrhovi v_i i bridovi e_i , tako da su krajevi od e_i vrhovi v_{i-1} i v_i za svaki $1 \ll i \leq k$.
6. Zatvorena šetnja je šetnja $W = v_0 e_1 v_1 e_2 v_2 \dots e_k v_k$ kod koje je $v_0 = v_k$.

7. Ako su bridovi e_1, e_2, \dots, e_k u šetnji $W = v_0e_1v_1e_2v_2 \dots e_kv_k$ međusobno različiti, onda se W zove staza. Ako su u stazi W i vrhovi međusobno različiti, onda se W zove put.
8. Ciklus je zatvorena šetnja kod koje su početak i unutrašnji vrhovi različiti.
9. Graf koji sadrži barem jedan ciklus je ciklički graf. U suprotnome kažemo da je graf aciklički (ne sadrži ciklus).
10. Dva vrha $u, v \in V(G)$ u grafu G su povezana ako postoji (u, v) -put u G . Graf je povezan ako između svaka dva vrha postoji put.

Sve navedene definicije su preuzete iz [16] i [17].

6.4. Definicija Bayesove mreže

Bayesova mreža je usmjereni aciklički graf u kojemu skup čvorova predstavlja varijable, dok bridovi predstavljaju povezanost između varijabli. Također je potrebno naglasiti da postoji uvjetna distribucija za svaki čvor, za dane njegove roditelje (kvantificira utjecaj roditelja na dijete). U diskretnom slučaju, uvjetna distribucija je reprezentirana tablicom uvjetne vjerojatnosti koja daje distribuciju preko svih varijabli X_i , za svaku kombinaciju vrijednosti roditelja.[15][8]

6.4.1. Postupak kreiranja Bayesove Mreže

Kako bi se kreirala Bayesovu mrežu potrebno je prvo definirati problem koji će se proučavati. Nakon što je problem definiran potrebno je definirati varijable (čvorove u mreži). Potrebno je također definirati ovisnosti između varijabli (bridovi). Za svaku varijablu potrebno je odrediti koliko utječu njezini roditelji na nju (raspodjela vjerojatnosti).

Nakon što je graf skiciran potrebno je dodijeliti a priori vrijednosti varijablama. A priori vrijednost je polazna vrijednost do koje se nije došlo pomoću mjerenja ili prikupljanja dokaza, već je ona rezultat odabira kreatora mreže. Kroz vrijeme se ta početna vrijednost mijenja ovisno o novim dokazima. Ukoliko varijabla ima roditelje, tada se a priori vrijednost može izračunati. Za kraj je potrebno dodijeliti tablicu uvjetne vjerojatnosti za svaku varijablu. Dimenzija tablice ovisi o broju roditeljskih čvorova varijable. [15]

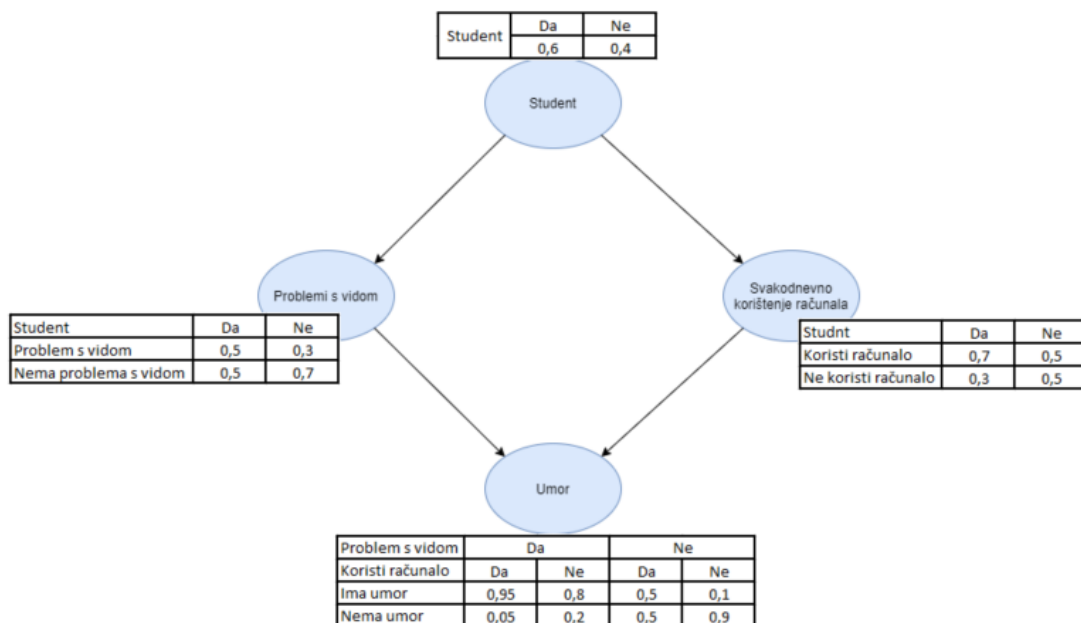
6.4.1.1. Primjer kreiranja Bayesove mreže

Velik broj ljudi osjeća umor, mnogi su razlozi koji do njega dovode. U ovome primjeru u obzir su uzeti korištenje računala te probleme s vidom. Korištenje računala samo po sebi

zahtjeva puno sjedenja što izaziva umor u donjem dijelu leđa, ali također utječe na zdravlje očiju. Prekomjerno korištenje računala, ali i u današnje doba mobilnih telefona dovodi do pogoršanja vida te iz toga razloga se gubi oštrina vida i oči se prekomjerno naprežu što dovodi do umora ili glavobolje.

U ovome primjeru se za korijenski čvor izabire „Studente“ jer oni često moraju koristiti računala kako bi mogli obavljati svoje fakultetske obveze (pisanje seminarskih radova, izrada projekata i sl.). Korijenski čvor također ima dva djeteta „Problemi s vidom“ i „Svakodnevno korištenje računala“, također navedena dva čvora su roditeljski čvorovi čvora „Umor“.

Navedene vrijednosti za uvjetne vjerojatnosti postavljene u tablicama nisu dobivene istraživanjem već su samo postavljene radi primjera.



Slika 32. Primjer Bayesove mreže (Prema: Marina Mijić, 2018)

Na temelju danih vrijednosti izračunava se a priori vrijednosti. Korijenski čvor „Student“ nema roditelja te je njezina vjerojatnost dana i koristi se u daljnjem računanju. Ostale tri vrijednosti će se računati. Prvo ćemo izračunati a priori vrijednost za varijablu „Problem s vidom“. Poznate vrijednosti su:

$$P(S) = 0,6$$

$$P(V|S) = 0,5$$

$$P(V|S^c) = 0,3$$

Navedene vrijednosti unosimo u jednadžbu:

$$P(V) = P(V|S)P(S) + P(V|S^c)P(S^c)$$

$$P(V) = 0,42$$

Na isti način izračunava se i a priori vrijednost za varijablu „Svakodnevno korištenje računala“.

$$P(S) = 0,6$$

$$P(R|S) = 0,7$$

$$P(R|S^c) = 0,5$$

$$P(R) = P(R|S)P(S) + P(R|S^c)P(S^c)$$

$$P(V) = 0,62$$

Ovim postupkom se računaju a priori vrijednosti sve dokle ne dođemo do čvora koji nema djecu. Primjer je izrađen na temelju primjera iz [15].

6.4.1.2. Propagacija unaprijed i unazad

U nekim situacija su neki podaci poznati. Tako na primjer ako je poznato de je osoba student, a priori vrijednost za svakodnevno korištenje računala više ne iznosi 0,62 jer vjerojatnost da je osoba student iznosi 1.

$$P(S) = 1$$

$$P(R|S) = 0,7$$

$$P(R|S^c) = 0,5$$

$$P(R) = P(R|S)P(S) + P(R|S^c)P(S^c)$$

$$P(V) = 0,7$$

Ovaj način računanja se naziva propagacija unaprijed jer se krećemo u smjeru strelice, odnosno od roditelja prema djetetu. Također može se računati i u suprotnom smjeru od djeteta prema roditelju, takav način se naziva propagacija unazad. Može se pretpostaviti da osoba svakodnevno koristi računalo te računamo koja je vjerojatnost da je osoba student.

$$P(S) = 0,6$$

$$P(S|R) = 0,7$$

$$P(R) = 0,62$$

$$P(R|S) = \frac{P(S|R) \cdot P(R)}{P(S)}$$

$$P(R|S) = \frac{0,7 \cdot 0,62}{0,6} = 0,72$$

Prema izračunatome, u većini slučajeva osoba koja koristi računalo svakodnevno je student. Kao što možete primijeti za računanje se koristio osnovni oblik Bayesove formule. Izrađeno prema primjeru iz [15].

6.5. Izrada Bayesove mreže korištenjem alata Netica

Za izradu Bayesove mreže može se koristiti alat Netica. Sučelje alata je uobičajenog izgleda te u ovome radu neće biti izdvajanja pojedinosti. Za kreiranje mreže potrebno je odabrati File → New → Network. Nadalje potrebno je kreirati čvorove mreže. Za kreiranje čvorova mreže potrebno je odabrati Modify → Add → Nature Node te kliknuti bilo gdje na bijelu površinu. Nakon što su izvršene prethodne akcije trebalo bi biti vidljivo sljedeće.

A	
state0	100

Slika 33. Izgled jednog dodanog Nature Noda (Prema: B. Kliček i D. Oreški, 2020)

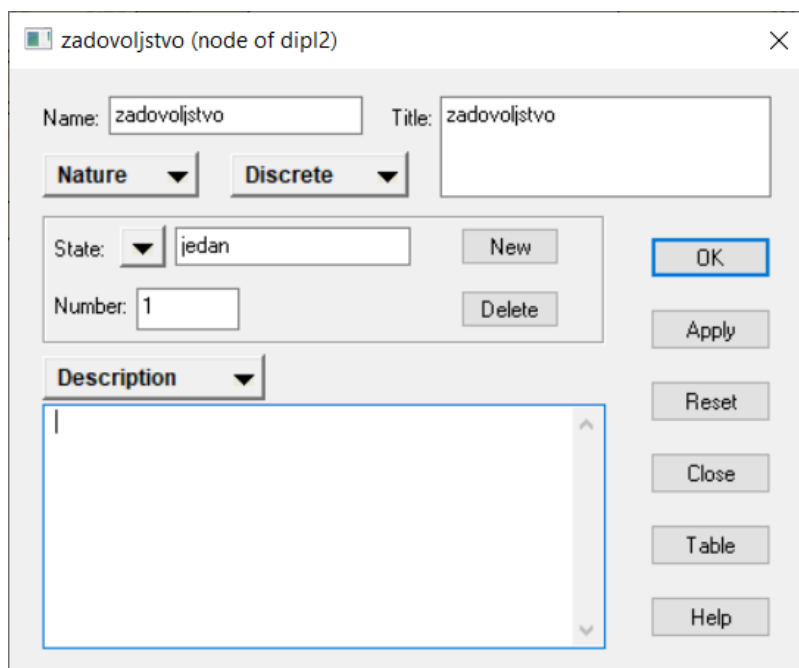
Svaki čvor ima svoja stanja koja je potrebno unijeti, kako bi se unijela stanja čvora potrebno je napraviti dvoklik na čvor.

The screenshot shows a dialog box titled "A (node of dipl2)". It contains the following fields and controls:

- Name: A
- Title: (empty)
- Nature: (dropdown menu)
- Discrete: (dropdown menu)
- State: (dropdown menu) and (input field)
- Number: (input field)
- Description: (dropdown menu) and (text area)
- Buttons: New, Delete, OK, Apply, Reset, Close, Table, Help

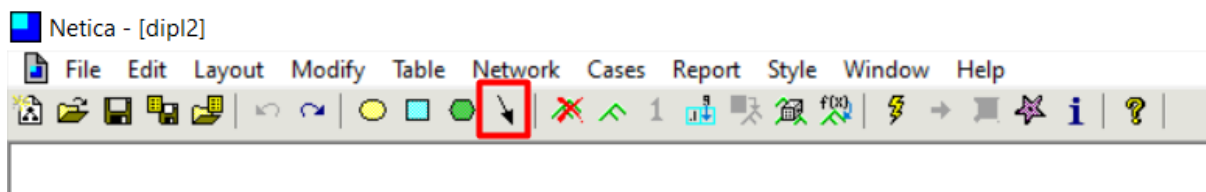
Slika 34. Forma koja se podiže na dvoklik na čvor (Prema: B. Kliček i D. Oreški, 2020)

U polje „Name“ unosi se ime čvora, „State“ jedno od stanja te u polje „Number“ unosi se vrijednost (jačinu stanja). Ukoliko čvor ima više stanja klikom na „New“ dobiva se mogućnost dodavanja stanja.



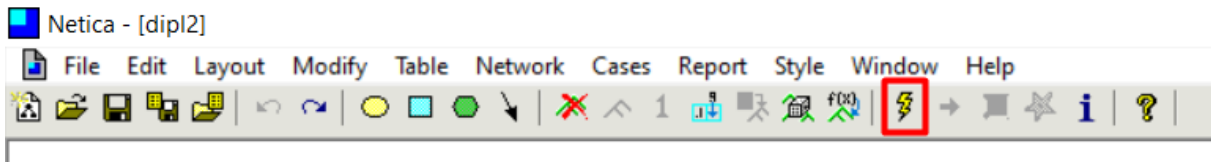
Slika 35. Dodavanje stanja čvora, imena čvora i naslova čvora (Prema: B. Kliček i D. Oreški, 2020)

Za svaki čvor mreže potrebno je ponoviti postupak. Nakon što su dodani svi čvorovi, potrebno ih je povezati u mrežu. To se radi na sljedeći način, odabire se strelica te povuče se mišem od čvora roditelja prema čvoru djetetu.



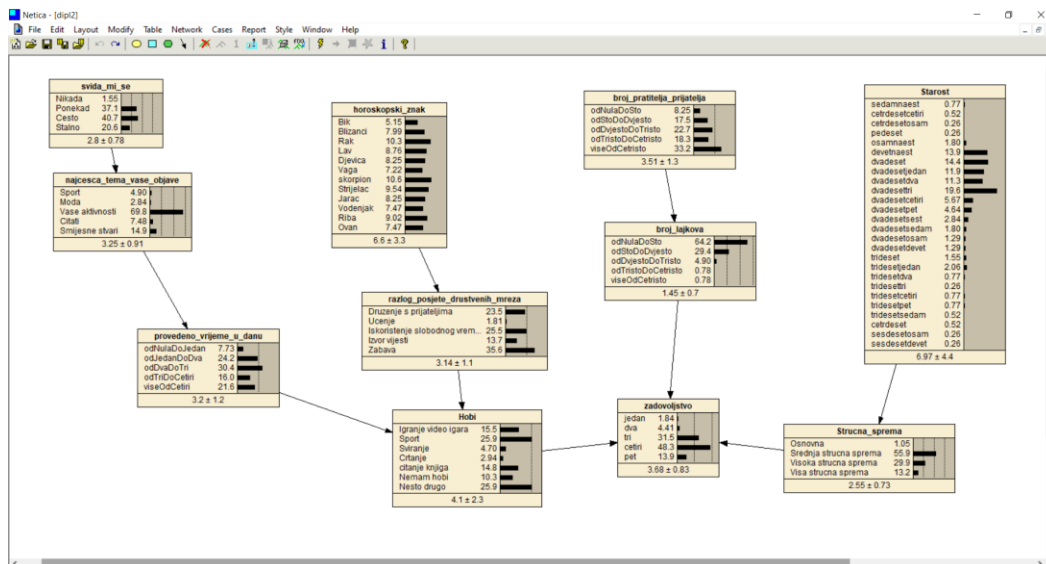
Slika 36. Ikona strelice u alatu Netica (Prema: B. Kliček i D. Oreški, 2020)

Nakon kreiranja cijele mreže, potrebno je dodijeliti skup podataka iz kojega će mreža učiti. Za učenje mreže koristit će se podaci dobiveni iz ankete. Navedeni podaci su spremljeni u Excel datoteci te ih je potrebno prilagoditi. U prvome retku se nalaze nazivi čvorova, što nije ispravno te se dodaje jedan prazan redak, kao prvi redak. U ćeliju A1 dodaje se tekst: “// ~->[CASE-1]->~“ . Nakon navedenog potrebno je spremiti datoteku u formatu Tekst (razdvojen tabularom)(* .txt). Nakon prilagodbe podataka može se dodati datoteku s podacima. Odabire se Cases → Learn → Incorporate Case File te prilagođena datoteku. Posljednji korak u kreiranju mreže je kompajliranje mreže koje se izvodi na način da se odabere gumb u obliku „munje“.



Slika 37. Ikona za kompajliranje u alatu Netica (Prema: B. Kliček i D. Oreški, 2020)

Krajnji rezultat kreiranja Bayesove mreže u alatu Netica prikazan je na slici.



Slika 38. Prikaz Bayesove mreže u alatu Netica (Prema: B. Kliček i D. Oreški, 2020)

6.6. Interpretacija dobivenih rezultata

Kako bi se istražili dobiveni podaci moraju se mijenjati slučajevi, što se izvodi na način da se mijenjaju vjerojatnosti. Klikom na neko stanje bilo kojeg čvora, njegova vjerojatnost iznosi 100%, drugi način mijenjanja vjerojatnosti je da na skali jednog stanja se pomiče mišem lijevo desno. Pogledom na vjerojatnosti u čvoru zadovoljstvo može se zaključiti kako je vjerojatnost da će osoba ocijeniti zadovoljstvo društvenim mrežama ocjenom 4 visokih 48,3%, što je moguće iščitati iz prethodne slike, također to je najveća vjerojatnost, što znači da prosječni korisnik svoje zadovoljstvo društvenim mrežama ocjenjuje ocjenom četiri. Ukoliko se odredi da se radi o osobi koja po svojim objavama ima broj lajkova („sviđanja“) u intervalu od 100 do 200 te da joj je hobi sviranje, tada je vjerojatnost da će osoba ocijeniti svoje zadovoljstvo ocjenom jedan 2,95%, ocjenom dva 2,95%, ocjenom tri 58,5%, ocjenom četiri 32,7% te ocjenom pet 2,95%. Prema ovome može se vidjeti da osobe ovoga profila ocjenjuju zadovoljstvo nižom ocjenom od prosječne osobe. Također, može se proučiti drugi slučaj u kojemu se radi o osobi koja ima broj lajkova („sviđanja“) veći od četiristo, radi se o osobi kojoj je hobi „nešto drugo“, odnosno nešto što nije navedeno u anketi, radi se o osobi koja ima samo srednju stručnu spremu, takve osobe svoje zadovoljstvo ocjenjuju ocjenom četiri s vjerojatnošću od 99%. Još

jedan slučaj je u kojemu se radi o osobi koja nema hobi, njezin broj lajkova po objavi kreće u intervalu od 100 do 200 te ima završenu srednju stručnu spremu, tada vjerojatnost da osoba ocjenjuje svoje zadovoljstvo s ocjenom četiri iznosi 50%, također ista vjerojatnost je da će ocijeniti zadovoljstvo ocjenom pet. Također je pronađen profil osobe koja će sa vjerojatnošću skoro 50% ocijeniti svoje zadovoljstvo ocjenom pet. Takva osoba nema hobi, njezin broj lajkova po objavi se kreće u intervalu od 100 do 200 te joj je stručna sprema srednja ili visoka.

7. Usporedba rezultata dobivenih kroz tri metode predviđanja

U ovome istraživanju provedene su tri metode predviđanja te će se usporediti dobiveni rezultati. Metodom stabla odlučivanja pronađeno je pravilo koje ima pouzdanost 97,2% da će ocjena zadovoljstva korisnika iznositi 4. Ova pouzdanost je jako visoka što znači da ovome pravilu možemo vjerovati. Primjenom metode neuronske mreže dobivena je informacija da atribut zadovoljstvo korisnika najviše ovisi o atributima „Koliko prosječno dobivate lajkova po objavi?“, „Koliko društvenih mreža koristite?“ i „Hobi“. Kombiniranjem tih atributa dobivaju se intervali unutar kojih se kreće zadovoljstvo korisnika. Kombinacija atributa „Koliko prosječno dobivate lajkova po objavi?“ i „Koliko društvenih mreža koristite?“ predviđa da će zadovoljstvo korisnika biti u intervalu od 3.33 do 3.88. Kombinacija atributa „Hobi“ i „Koliko prosječno dobivate lajkova po objavi?“ predviđa da će zadovoljstvo korisnika biti u intervalu od 3.40 do 3.97. Kombinacija atributa „Hobi“ i „Koliko društvenih mreža koristite?“ predviđa da će zadovoljstvo korisnika biti u intervalu od 3.49 do 4.05. Prema tome može se zaključiti da se predviđa kako će korisnici svoje zadovoljstvo najčešće ocijeniti ocjenama 3 ili 4, ali s prevagom na 4.

Mora se primijetiti da primjenom metode stablo odlučivanja atribut zadovoljstvo korisnika najviše ovisi o atributima „Koliko vremena u prosjeku u danu provedete na društvenim mrežama?“, „Hobi“, „Koji ste horoskopski znak?“ te „Koliko prosječno dobivate lajkova po objavi?“. Usporede li se atributi s atributima iz metode neuronske mreže može se primijetiti da se čak tri atributa poklapaju u prva četiri atributa o kojima najviše ovisi atribut zadovoljstvo korisnika. Usporede li se vrijednosti predviđanja obadvije metode, vidljivo je kako su vrijednosti podjednake. Metodom stabla odlučivanja dobivena je vrijednost 4, dok metodom neuronskih mreža je dobivena vrijednost 3-4.

Atributi koji se koriste u metodi Bayesove mreže su preuzeti iz metode stabla odlučivanja, prvih 10 atributa o kojima najviše ovisi atribut zadovoljstvo korisnika, iz toga razloga se neće komentirati attribute. Predviđanje metodom Bayesove mreže govori kako će vrijednost atributa zadovoljstvo korisnika iznositi 4 u 48,3% slučajeva. Kada se malo prouče korisnici i kreiraju njihovi profili dobivaju se zanimljivi rezultati. Jedna od zanimljivosti je da je uspješno pronađen profil osobe koja će u 99% slučajeva svoje zadovoljstvo ocijeniti ocjenom 4. Takva osoba u prosjeku ima broj lajkova veći od četiristo, njezin hobi je nešto drugo te ima samo srednju stručnu spremu. Kao i u prethodne dvije metode u najveći broj slučajeva se predviđa vrijednost 4 za atribut zadovoljstvo korisnika. Pomoću Bayesove mreže isprobano je

kreirati pravilo iz stabla odlučivanja koliko je to bilo moguće jer Bayesova mreža ne sadrži sve čvorove iz pravila. No, u svakom slučaju uspješno su pronađena četiri atributa koja se nalaze u pravilu te se primjenom jednog po jednog djela pravila vjerojatnost da će korisnik ocijeniti zadovoljstvo ocjenom četiri povećava.

Za predviđanje se koriste različite metode predviđanja, svaka od njih se koristi u različite svrhe, Oreški i Begičević Ređep su istražile koje bi metode bilo potrebno koristiti za predviđanje s obzirom na karakteristike skupa podataka. Navode kako skupovi podataka s visokim šumom trebaju koristiti stablo odlučivanja za klasifikaciju, također navode kako na skupovima podataka koji imaju mali šum te imaju puno nedostajućih vrijednosti trebala bi se koristiti neuronska mreža. Šum se odnosi na modifikaciju originalnih vrijednosti. [23]

Skupovi podataka se sastoje od dva tipa varijabli, a to su numeričke i kategoričke. Oreški i Hajdin navode kako Bayesove mreže najbolje rade sa kategoričkim varijablama, stablo odlučivanja radi jednako dobro s obadvije vrste dok neuronske mreže najbolje rade s numeričkim varijablama.[22]

Skup podataka koji se koristi u ovome istraživanju ima pretežno kategoričke varijable. Također, unutar njega ne postoje nedostajuće vrijednosti iz razloga jer su ispitanici morali popuniti sva polja kako bi mogli predati popunjenu anketu. Smatra se da ne postoji ni šum unutar podataka jer ne postoji način na koji bi se mogle modificirati originalne vrijednosti. Prema svemu navedenome može se zaključiti kako za provođenje predviđanja nad skupom podataka koji se koristi u ovome istraživanju bi bilo najbolje koristiti ili metodu stabla odlučivanja ili Bayesove mreže.

8. Zaključak

Nakon provedenog istraživanja može se reći da je jako teško skupiti veliki broj ljudi koji će popuniti anketu. Iako je uspješno skupljeno 390 odgovora na anketu, voljeli bismo da je taj broj bio puno veći. Veličina skupa podataka može utjecati na rezultate istraživanja te je najbolje koristiti što veći skup podataka kako bi došlo do što manje greške ili da bi se u najboljem slučaju greška izbjegla.

Istraživanje je provedeno pomoću tri metode predviđanja, kao što je već ranije rečeno. Svaka od navedenih metoda ima svoja svojstva prema kojima se uvelike razlikuje od drugih metoda. Unatoč svojim različitim načinima predviđanja, svaka metoda je dala podjednake rezultate. U svakoj od metoda smo uspjeli dobiti pravilo, profil osobe ili kombinaciju atributa koji daju pouzdanost ili vjerojatnost iznad 90% da će osoba svoje zadovoljstvo ocijeniti ocjenom četiri. Na temelju ovih podataka, kombinacijom pravila i profila osobe može se preciznije odrediti koja će osoba ocijeniti svoje zadovoljstvo ocjenom četiri. Ako se radi o osobi koja ima broj lajkova po objavi u intervalu od 0 do 100 ili veći od 400, hobi te osobe je „nešto drugo“ što znači da u anketi nije naveden njezin hobi, osoba ima završenu srednju stručnu spremu, takva osoba provodi svakodnevno više od jednog sata na društvenim mrežama, društvene mreže ne utječu na način života takve osobe, označavaju druge osobe na objavama, koristi više od jedne društvene mreže, njezin radni status je različit od „zaposlen“, njezine objave nisu vezane uz njezine aktivnosti te društvene mreže ne koriste kao izvor vijesti. Ukoliko pronađete osobu koja odgovara ovome profilu, tada sa velikom sigurnošću možete predvidjeti njezinu ocjenu zadovoljstva.

9. Literatura

- [1] Jason L. Skues, Ben Williams, Lisa Wise, „The effects of personality traits, self-esteem, loneliness, and narcissism on Facebook use among university students“, 2012.
- [2] David John Hughes, Moss Rowe, Mark Batey, Andrew Lee, „ A tale of two sites: Twitter vs. Facebook and the personality predictors of social media usage“, 2011.
- [3] Timothy Teo, „ Modelling Facebook usage among university students in Thailand: the role of emotional attachment in an extended technology acceptance model“, 2014.
- [4] Krishna Das, Smriti Kumar Sinha, „ A Survey on User Behaviour Analysis in Social Networks“, 2016.
- [5] Long Jin, Yang Chen, Tianyi Wang, Pan Hui, Athanasios V. Vassilakos, „Understanding User Behavior in Online Social Networks: A Survey“, 2013.
- [6] Hajra Waheed, Maria Anjum, Mariam Rehman, Amina Khawaja, „ Investigation of user behavior on social networking sites“, 2016.
- [7] Skladistenje.com, „Stabla odlučivanja“, 2002. [Na internetu]. Dostupno: <http://www.skladistenje.com/stabla-odlucivanja/> , [pristupano 01.07.2020.].
- [8] B. Kliček, "Inteligentni sustavi", nastavni materijali na predmetu Inteligentni sustavi [Moodle], Sveučilište u Zagrebu, Fakultet organizacije i informatike, Varaždin, 2020.
- [9] Suzana Dumančić, „ Neuronske mreže“, 2014. [Na internetu]. Dostupno: <http://www.mathos.unios.hr/~mdjumic/uploads/diplomski/dum05.pdf> ,[pristupio 01.07.2020.]
- [10] Goran Pintarić, „Umjetne neuronske mreže i mogućnosti njihove primjene u obrazovanju“, 2013. [Na internetu]. Dostupno: <http://www.mathos.unios.hr/~mdjumic/uploads/diplomski/PIN12.pdf> ,[pristupio: 01.07.2020.]
- [11] B. Kliček, D. Oreški, "Stablo odlučivanja", nastavni materijali na predmetu Inteligentni sustavi [Moodle], Sveučilište u Zagrebu, Fakultet organizacije i informatike, Varaždin, 2020.
- [12] B. Kliček, D. Oreški "Neuronska mreža zadatak", nastavni materijali na predmetu Inteligentni sustavi [Moodle], Sveučilište u Zagrebu, Fakultet organizacije i informatike, Varaždin, 2020.
- [13] S. Ribarić, B. Dalbelo Bašić, „5. Modeliranje neizvjesnosti u sustavima temeljenim na znanju“, 2002. [Na internetu]. Dostupno: http://www.zemris.fer.hr/predmeti/is/nastava/Modeliranje_neizvjesnosti.pdf , [pristupio: 01.07.2020]
- [14] „Vjerojatnost“, [Na internetu]. Dostupno: http://www.unizd.hr/portals/13/nastavni_materijali/01%20-%20vjerojatnost.pdf, [pristupio: 01.07.2020.]

- [15] Marina Mijić, „Bajezove mreže i njihova primena“, 2018. [Na internetu]. Dostupno: https://matematika.pmf.uns.ac.rs/wp-content/uploads/zavrzni-radovi/primenjena_matematika/MarinaMijic.pdf, [pristupio: 01.07.2020.]
- [16] B. Divjak, "Osnovni pojmovi iz teorije grafova", nastavni materijali na predmetu Diskretne strukture s teorijom grafova [Moodle], Sveučilište u Zagrebu, Fakultet organizacije i informatike, Varaždin, 2020.
- [17] B. Divjak, "Usmjereni grafovi", nastavni materijali na predmetu Diskretne strukture s teorijom grafova [Moodle], Sveučilište u Zagrebu, Fakultet organizacije i informatike, Varaždin, 2020.
- [18] B. Kliček, D. Oreški, "Stablo odlučivanja zadatak 1", nastavni materijali na predmetu Inteligentni sustavi [Moodle], Sveučilište u Zagrebu, Fakultet organizacije i informatike, Varaždin, 2020.
- [19] B. Kliček, D. Oreški, "Izrada ekspertnog sustava", nastavni materijali na predmetu Inteligentni sustavi [Moodle], Sveučilište u Zagrebu, Fakultet organizacije i informatike, Varaždin, 2020.
- [20] B. Kliček, D. Oreški, "Neuronska mreža primjer zadatka", nastavni materijali na predmetu Inteligentni sustavi [Moodle], Sveučilište u Zagrebu, Fakultet organizacije i informatike, Varaždin, 2020.
- [21] B. Kliček, D. Oreški, "Bayesove mreže", nastavni materijali na predmetu Inteligentni sustavi [Moodle], Sveučilište u Zagrebu, Fakultet organizacije i informatike, Varaždin, 2020.
- [22] D. Orešli, G.Hajdin, A Comparative Study of Machine Learning Approaches on Learning Management System Data, 2019.
- [23] D. Oreški, N. B. Ređep, Data-driven decision-making in classification algorithm selection, 2018.

Popis slika

Slika 1. Korištenje metoda istraživanja ponašanja korisnika na društvenim mrežama [4]	9
Slika 2. Postupak dobivanja osnovnih istraživanja potrebnih za provođenje istraživanja [6]..	13
Slika 3. Prikaz jednostavnog stabla odlučivanja [8].....	21
Slika 4. Prikaz stabla odlučivanja dobivenog računanjem entropije.....	24
Slika 5. Ikona Source BigML [18]	24
Slika 6. Ikona Dataset BigML [18]	25
Slika 7. Ikona Model BigML [18].....	25
Slika 8. Odabir atributa za predviđanje [18].....	25
Slika 9. Stablo odlučivanja na temelju učitanih podataka [18].....	26
Slika 10. Prikaz najpouzdanijeg pravila [18].....	29
Slika 11. Akcija koja omogućava dodavanje vrijednosti atributima u alatu XRKB [19]	31
Slika 12. Kreirano najpouzdanije pravilo (dio 1) [19]	31
Slika 13. Kreirano najpouzdanije pravilo (dio 2) [19]	32
Slika 14. Kreirano najpouzdanije pravilo (dio 3) [19]	32
Slika 15. Provođenje predviđanja (Koliko vremena u prosjeku u danu provedete na društvenim mrežama? (u satima)) [19].....	33
Slika 16. Provođenje predviđanja (Koliko prosječno dobivate lajkova po objavi?) [19].....	33
Slika 17. Provođenje predviđanja (Koliko pratitelja/prijatelja imate na društvenim mrežama?) [19]	34
Slika 18. Provođenje predviđanja (Utječu li društvene mreže na vaš stil života?) [19].....	34
Slika 19. Provođenje predviđanja (Označavate li druge osobe na objavama?) [19].....	34
Slika 20. Provođenje predviđanja (Koji ste horoskopski znak?) [19].....	35
Slika 21. Provođenje predviđanja (Koji je vaš razlog posjete društvenih mreža?) [19]	35
Slika 22. Provođenje predviđanja (Koja je najčešća tema vaše objave?) [19]	35
Slika 23. Provođenje predviđanja (Koji je vaš radni status?) [19].....	36
Slika 24. Provođenje predviđanja (Koliko društvenih mreža koristite?) [19]	36
Slika 25. Rezultat provedenog predviđanja [19].....	36
Slika 26. Biološki neuron [8].....	37
Slika 27. Računalni neuron [8].....	37
Slika 28. Slojevi neuronske mreže [10].....	39
Slika 29. Ikona Deepnet BigML [12].....	44
Slika 30. Postavljanje broja neurona u skrivenom sloju [12]	44
Slika 31. Prikaz kombinacije atributa "Koliko društvenih mreža koristite?" i "Koliko prosječno dobivate lajkova po objavi?" [12]	46
Slika 32. Primjer Bayesove mreže [15].....	50

Slika 33. Izgled jednog dodanog Nature Noda [21]	52
Slika 34. Forma koja se podiže na dvoklik na čvor [21]	52
Slika 35. Dodavanje stanja čvora, imena čvora i naslova čvora [21].....	53
Slika 36. Ikona strelice u alatu Netica [21].....	53
Slika 37. Ikona za kompajliranje u alatu Netica [21]	54
Slika 38. Prikaz Bayesove mreže u alatu Netica [21]	54

Popis tablica

Tablica 1. Podaci za izradu stabla odlučivanja [11]	22
Tablica 2. Podaci o neuronima [20]	42