

Kako profilirati birače u političkim kampanjama primjenom umjetne inteligencije?

How to Profile Voters in Political Campaigns by Means of Artificial Intelligence?

Sažetak

Umjetna inteligencija i strojno učenje velikom brzinom mijenjaju sve sfere društva, pa tako i način kako funkcioniraju političke kampanje. U ovom se radu ispituje primjena dva algoritma strojnog učenja, klasteriranja i neuronske mreže, u svrhu profiliranja birača u političkim kampanjama. Empirijsko istraživanje koristi podatke društvene mreže Facebook o aktivnostima političkih kandidata na lokalnim izborima u Francuskoj, s fokusom na sadržaj koji kandidat kreira i odjek koji to izaziva. Analizirani su podatci za kandidate na izborima u 41 gradu u Francuskoj. Podatci su uključivali varijable koje se odnose na broj dijeljenih događaja, fotografija, poveznica, videa i statusa, kao i komentare na sadržaj. Prikupljanje podataka društvenih mreža jednostavnije je i jeftinije od tradicionalnih načina prikupljanja javnog mnijenja. Primjenom umjetnih neuronskih mreža kreirani su prediktivni modeli ishoda izbora i identificirane su varijable koje su prediktori ishoda izbora. Primjena algoritma klasteriranja na varijablama koje su najznačajniji prediktori izbora omogućila je segmentaciju birača u homogenije skupine prema interesima. Hibridni pristup prediktivnog i deskriptivnog modela temeljem algoritama strojnog učenja omogućio je profiliranje birača. Ovo istraživanje primjenjuje strojno učenje i umjetnu inteligenciju u politici te daje uvide u kvalitete profiliranja birača koje se temelji na podatcima društvene mreže.

Ključne riječi: umjetna inteligencija, strojno učenje, prediktivni modeli, profiliranje

JEL klasifikacija: D83

Abstract

Artificial intelligence and machine learning are rapidly changing all spheres of society, including how political campaigns work. This paper explores the application of two machine learning algorithms, artificial neural networks and clustering, for voter profiling in political campaigns. The empirical research is based on data about candidate activity in local elections in France, analyzing interactions between candidates and voters through the social network Facebook. Data was analyzed for candidates in 41 France cities. The data included variables related to the number of shared events, photographs, links, videos, statuses, and comments on content. Collecting data from social networks is easier and cheaper than traditional methods of collecting public opinion. By applying artificial neural networks, predictive models of election outcomes were created and identified predictor variables of election results. The application of clustering algorithms using variables that are the most significant predictors of elections enabled the segmentation of voters into more homogeneous groups according to interests. The hybrid approach of predictive and descriptive models based on machine learning algorithms enabled voter profiling. This study applies artificial

dr.sc. Alen Kišić, univ.spec.oec.

Zona Sjever d.o.o.

E-mail: kisic.alen@gmail.com

Alen Kišić, PhD

Zona Sjever Ltd.

E-mail: kisic.alen@gmail.com

Kišić, A.

Kako profilirati birače u političkim kampanjama primjenom umjetne inteligencije?

intelligence and machine learning in politics for quality voter profiling based on social network data.

Key words: artificial intelligence, machine learning, predictive models, voters profiling

Jel classification: D83

1. Uvod

Političke kampanje ubrzano se prilagođavaju digitalnoj transformaciji, s obzirom na sve češću primjenu naprednih tehnologija, pri čemu umjetna inteligencija (UI) postaje ključan alat za analizu i profiliranje birača. Kroz povijest su političke stranke i politički kandidati nastojali razumjeti birače kako bi prilagodili svoje poruke njihovim interesima i stavovima. Tradicionalne metode istraživanja javnog mnijenja, poput anketa i fokus grupa, pokazale su određena ograničenja, uključujući visoke troškove, vremensku zahtjevnost i ograničenu reprezentativnost. Razvojem UI i povećanom dostupnošću velikih količina podataka, otvorene su nove mogućnosti za preciznije profiliranje birača uz istovremeno smanjenje troškova (npr. Kišić, 2018.).

Algoritmi strojnog učenja omogućuju kvalitetnu analizu velikih podataka, pružajući političkim kampanjama uvide u političke preferencije, ponašanje na društvenim mrežama, demografske karakteristike i psihološke osobine birača (Kišić i Kliček, 2021.). Korištenjem naprednih tehnika poput dubokog učenja, strojnog učenja i obrade prirodnog jezika, kampanje mogu prepoznati ključne teme koje su važne biračima te prema tome oblikovati poruke. Osim toga, sofisticirane metode, poput prediktivnog modeliranja temeljenog na algoritmima strojnog učenja i klasteriranja, omogućuju segmentaciju birača prema sličnostima u njihovom ponašanju i stavovima.

Danas se rade personalizirane kampanje koje koriste tehnologije umjetne inteligencije i algoritme strojnog učenja, i pomažu kandidatima u dubljem razumijevanju motivacije birača. Cilj ovog rada je istražiti profiliranje birača u političkim kampanjama pomoću algoritama strojnog učenja, posebno umjetnih neuronskih mreža. Rad se oslanja na istraživanja povezana s doktorskom disertacijom autora i usmjerava na implementaciju modela za profiliranje birača u političke kampanje

(Kišić, 2023.).

Empirijsko istraživanje temelji se na podatcima francuskih lokalnih izbora za 41 grad. Podatci o aktivnostima kandidata prikupljeni su na društvenoj mreži Facebook.

Rad je strukturiran kako slijedi. Nakon prvog, uvodnog poglavlja, drugo poglavlje daje pregled recentnih relevantnih istraživanja na ovu temu. Treće poglavlje opisuje metodologiju rada s naglaskom na opis korištenih podataka te prikazuje algoritme strojnog učenja koji su implementirani u empirijskom istraživanju: umjetne neuronske mreže i klasteriranje. Četvrto poglavlje prikazuje rezultate istraživanja i daje njihovu interpretaciju. U zaključku se naglašava znanstveni doprinos rada i daju se smjernice za buduća istraživanja.

2. Pregled literature

Algoritmi strojnog učenja pojavili su se kao moćni alati za predviđanje izbornih rezultata, koristeći različite izvore podataka i metodologije. Nedavna istraživanja ističu učinkovitost različitih modela, uključujući algoritme slučajnih šuma (engl. random forest), metodu potpornih vektora (engl. support vector machines), umjetne neuronske mreže (engl. artificial neural networks) i analize sentimenta (engl. sentiment analysis), u preciznom prognoziranju izbornih rezultata u različitim političkim kontekstima. Sljedeći tekst daje kratak pregled ključnih pristupa i spoznaja iz nedavnih relevantnih istraživanja na ovu temu. Pregled literature temeljen je na identifikaciji algoritama strojnog učenja koji su se koristili u ove svrhe, korištenih podataka i političkih konteksta iz kojih su prikupljeni te kvalitete dobivenih modela.

Autori Manhyan i Bose (2024.) u nedavnom su istraživanju demonstrirali iznimnu točnost (točnost predikcije od 99,89%) u predviđanju indijskih parlamentarnih izbora analizom povijesnih podataka o glasanju i socio-ekonomskih pokazatelja (Manhyan i Bose, 2024.). Njihovo

istraživanje je uključivalo algoritme slučajnih šuma (engl. random forest), gradijentnog pojačavanja (engl. gradient boosting) i stabla odlučivanja (engl. decision tree) za predviđanje izbornih rezultata u Indiji. Model slučajnih šuma postigao je najvišu točnost od 99,89%, slijedi gradijentno pojačavanje s 99,78% te stablo odlučivanja s 98,75%.

Alkhiary i suradnici (2024.) predviđaju rezultate egipatskih parlamentarnih izbora kroz analizu sentimenta podataka s društvenih mreža, prikazujući potencijal takvog načina mjerenja javnog mišljenja u izbornim prognozama (Alkhiary i sur., 2024.). Algoritam neuronske mreže višeslojni perceptron postigao je 96% točnosti u predviđanju predsjedničkih izbora SAD-a 2020. godine integracijom ocjena sentimenta s tradicionalnim ekonomskim indikatorima (Gu i sur., 2023.). Hibridni pristupi rješavanja ovog problema uključuju kombiniranje analize sentimenta s algoritmima strojnog učenja. Taj se pristup pokazao učinkovitim u rješavanju složenosti ponašanja birača i trendova raspoloženja (Gu i sur., 2023.).

Gowtham (2024.) koristi algoritme strojnog učenja, uključujući metodu potpornih vektora (SVM), naivni Bayesov klasifikator, stabla odlučivanja i slučajne šume za predviđanje rezultata egipatskih parlamentarnih izbora kroz analizu sentimenta javnog mišljenja prema političkim strankama i kandidatima na društvenim mrežama. Ovo istraživanje koristi algoritme strojnog učenja za izgradnju prediktivnih modela koristeći povijesne podatke o izborima, povećavajući preciznost predikcije. Ovaj pristup, kombiniran s analizom sentimenta s društvenih mreža i novinarskih podataka, nudi sveobuhvatno razumijevanje javnog mišljenja prema političkim entitetima. Rad Emeta i suradnika (2024.) raspravlja o upotrebi Bayesovih mreža (engl. Bayesian nets) i algoritama k-najbližih susjeda (engl. k-nearest neighbours) za predviđanje izbornih rezultata, koristeći kompleksne podatke o izborima s više od 120.000 glasačkih mjesta u 36 saveznih država Nigerije za analizu trendova glasača pod utjecajem političkih stranaka. Istraživanje Fenga i suradnika (2023.) koristi algoritme stabala gradijentnog pojačavanja i višeslojnih perceptrona za prognoziranje američkih predsjedničkih izbora, postižući 96% točnosti u predviđanju pobjede demokratskog

kandidata. Ova hibridna metoda kombinira analizu sentimenta s različitim nezavisnim varijablama za poboljšane predikcije na razini saveznih država (Feng i suradnici, 2023.). Elo i suradnici (2024.) primjenjuju algoritme strojnog učenja za izborne predikcije koji uključuju pojedinačne klasifikatore poput strojeva potporne vektore, slučajne šume, naivni Bayesov klasifikator, logističku regresiju i stabla odlučivanja, kao i ansambel klasifikatore, koji su pokazali bolju točnost, ali se rjeđe primjenjuju u ovom kontekstu. Mishra i Lee (2023.) su koristili modele neuronskih mreža s jednim slojem, koji su obrađivali ili tekst naslova vijesti ili numeričke podatke, te modele s dva sloja koji kombiniraju oba. Model s najboljim učinkom bio je dvoslojni model. Rad Guanija i suradnika (2023.) ocjenjuje različite modele strojnog učenja za predviđanje gubitaka vladajuće stranke na federalnim izborima, pri čemu je višeslojni perceptron identificiran kao najtočniji klasifikator, dok su Gaussovi modeli, poput Gaussovog klasifikatora procesa i Gaussovog naivnog Bayesa, pokazali najlošiju točnost klasifikacije. U istraživanju iz 2023. Myilvahan i suradnici (2023.) raspravljaju o nadziranoj tehnici strojnog učenja koja koristi podatke o indijskim izborima za predviđanje rezultata nacionalnih i lokalnih izbora, pokazujući poboljšanu točnost i otpornost na manipulaciju u usporedbi s drugim metodologijama, posebno u predviđanju indijskih izbora 2019. godine.

Iako je primjena strojnog učenja dala veliki napredak u prognozama izbornih rezultata, izazovi poput potrebe za sveobuhvatnim skupovima podataka i dalje ostaju važne smjernice za buduća istraživanja. Naime, mada su ova istraživanja dala kvalitetne modele u smislu točnosti predikcije, prikupljanje podataka potrebnih za razvoj takvih modela je dugotrajan i skup posao. Ovaj rad koristi podatke društvene mreže Facebook kojima je olakšan pristup i besplatni su, u usporedbi s drugim načinima prikupljanja javnog mnijenja.

3. Metodologija

3.1. Empirijski podatci

U fokusu istraživanja bili su podaci o aktivnostima koje su kandidati za funkciju gradonačelnika prikazivali na društvenoj mreži Facebook, i to na razini najvećih francuskih gradova, odnosno

Kišić, A.

Kako profilirati birače u političkim kampanjama primjenom umjetne inteligencije?

Kišić, A.

Kako profilirati birače u političkim kampanjama primjenom umjetne inteligencije?

Tablica 1. Opis varijabli korištenih u istraživanju i njihove vrijednosti

Naziv varijable	Opis
Grad	Ime francuskog grada u kojem se kandidira istraživani političar, ukupno 41 grad u kojem živi više od 100 tisuća stanovnika
Broj likeova stranice	Sveukupno dani likeovi stranice
Broj aktivnosti	Ukupan broj aktivnosti odnosno događaja koje je kandidat dijelio u samoj kampanji na svojoj stranici
Broj fotografija	Ukupni broj fotografija koje je kandidat podijelio u kampanji na svojoj stranici
Broj linkova	Ukupni broj linkova koje je kandidat podijelio u kampanji na svojoj stranici
Broj objavljenih videa	Ukupni broj videa koje je kandidat podijelio u kampanji na svojoj stranici
Broj Facebook statusa	Ukupni broj Facebook statusa koje je kandidat podijelio u kampanji na svojoj stranici
Broj likeova po događaju	Prosječni broj likeova za svaki pojedinačni događaj
Broj likeova po fotografiji	Prosječni broj likeova za svaku pojedinačnu fotografiju
Broj likeova po poveznici	Prosječni broj likeova za svaku pojedinačnu poveznicu
Broj likeova po videu	Prosječni broj likeova za svaki pojedinačni video
Broj likeova po objavljenom statusu	Prosječni broj likeova za svaki pojedinačni status
Broj komentara po događaju	Prosječni broj komentara za svaki pojedinačni događaj
Broj komentara po fotografiji	Prosječni broj komentara za svaku pojedinačnu fotografiju
Broj komentara po linku	Prosječni broj komentara po svakom linku
Broj komentara po videu	Prosječni broj komentara za svaki video
Broj komentara po statusu	Prosječni broj komentara po svakom statusu
Broj shareanja događaja	Prosječni broj shareanja svakog pojedinačnog objavljenog događaja
Broj shareanja fotografija	Prosječni broj shareanja po svakoj pojedinačnoj fotografiji
Broj shareanja linkova	Prosječni broj shareanja svakog linka
Broj shareanja videa	Prosječni broj shareanja svakog pojedinačnog videa
Broj shareanja statusa	Prosječni broj shareanja svakog pojedinačnog statusa
Pripadnost stranci s političkog francuskog spektra	Kojoj stranci pripada pojedini političar za svaku stranku čiji kandidat sudjeluje u izborima
Spol	Kandidatov spol: žena ili muškarac
Ostvaren brojačni rezultat na provedenim lokalnim izborima	Postotak osvojenih glasova kandidata

Izvor: autor.

izabrano je ukupno 41 gradova koje povezuje činjenica da svaki od njih ima stotinu tisuća stanovnika. Također, bio je zadan i dodatni kriterij u smislu činjenica da je pritom svaki grad morao imati minimalno dva kandidata koji su imali Facebook profile ili stranice jer u protivnom istraživanje za taj grad ne bi imalo smisla.

Sve navedene podatke je kreirao sam autor rada tijekom mjeseca lipnja 2020. godine i to na način da je ručno pojedinačno preuzimao podatke koje su kandidati na izborima postavljali na svoje Facebook profile, odnosno stranice.

Kod onih kandidata koji su javno istaknuti na stranicama povjerenstva za izbore te su imali otvorene profile ili stranice, pregledavala se i bilježila svaka aktivnost, odnosno svaki sadržaj koji je potom bio selekcioniran po raznim kategorijama kao što su događaj, status, poveznica, fotografija ili video. Za svaki od tih sadržaja se onda ekstrahirala brojka s podacima o tome koliko je ljudi stisnulo like za sadržaj, ili koji je broj komentara sadržaja, koliko

puta se dijelio sadržaj ili koliko je uopće sadržaja postavljeno odnosno objavljeno na stranicama ili profilima svakog od istraživanih kandidata.

U početnoj fazi istraživanja analizirano je 25 varijabli, uključujući 24 ulazne varijable koje obuhvaćaju:

- aktivnosti kandidata na svojoj stranici ili profilu te kako na to reagiraju oni koji prate stranicu
- varijable koje označavaju je li kandidat muškarac ili žena te kojoj stranci s francuskog političkog spektra pripada. U prediktivnom modelu se kao varijabla koja je označena kao izlazna pojavljuje ostvareni rezultat koji je kandidat ostvario na provedenim izborima, i to prikazan kroz aspekt postotka osvojenih brojeva glasova. Nakon provedene pripreme podataka i selekcije značajki, u modele su uključene samo one značajke koje su identificirane kao relevantne.

3.2. Algoritmi strojnog učenja

Umjetne neuronske mreže (ANN) predstavljaju algoritme nadziranog strojnog učenja inspiriranih strukturom i funkcijom ljudskog mozga, posebno u obradi informacija. Sastoje se od slojeva neurona, gdje svaki neuron obrađuje podatke i prenosi ih dalje kroz mrežu, omogućavajući složene transformacije podataka. ANN algoritmi osobito su korisni za rješavanje nelinearnih problema, prepoznavanje obrazaca i klasifikaciju podataka, jer mogu naučiti nelinearne ovisnosti unutar podataka kroz proces treniranja. Koristeći algoritme poput propagacije unatrag, neuronske mreže se prilagođavaju kako bi minimizirale pogreške i poboljšale točnost u prepoznavanju obrazaca. Primjena ANN-a u političkim kampanjama uključuje analiziranje aktivnosti kandidata i birača te prepoznavanje preferencija koje utječu na njihovo glasovanje.

Klasteriranje je metoda nenadziranog strojnog učenja koja grupira podatke prema sličnostima, omogućavajući identificiranje struktura unutar podataka. Algoritmi klasteriranja, poput k-srednjih vrijednosti, koriste se za segmentaciju velikih skupova podataka, gdje se instance unutar jednog klastera smatraju sličnima, dok se razlikuju od onih u drugim klasterima. U kontekstu političkih

kampanja, klasteriranje pomaže u segmentaciji birača na temelju njihovih karakteristika, interesa ili političkih stavova, omogućujući prilagođavanje poruka prema specifičnim skupinama. Primjena klasteriranja doprinosi boljem razumijevanju birača.

U nastavku se prikazuju rezultati primjene umjetnih neuronskih mreža i klasteriranja na skupu podataka društvene mreže Facebook.

4. Rezultati istraživanja i rasprava

4.1. Rezultati empirijskog istraživanja

Algoritam ReliefF korišten je u selekciji varijabli s ciljem utvrđivanja manjeg broja varijabli koje su relevantne za predikciju ishoda izbora. Rezultati su prikazani u tablici 2.

Sve varijable s pozitivnim vrijednostima mjera vrednovanja identificirane su kao relevantne za predviđanje rezultata te će služiti kao ulazni parametri za prediktivne modele. Na temelju selekcije atributa primjenom algoritma ReliefF, odabrano je 8 varijabli iz početnog skupa od 21 varijable povezanih s podatcima s društvene mreže Facebook, koje će se koristiti u procesu modeliranja. Sljedeći dio rada prikazuje rezultate razvoja prediktivnih modela.

Algoritam umjetne neuronske mreže sastoji se od

Kišić, A.

Kako profilirati birače u političkim kampanjama primjenom umjetne inteligencije?

Tablica 2. Odabir varijabli

Varijabla	Mjera vrednovanja
Broj likeova stranice	0.03
Broj događaja	0.03
Broj likeova po statusu	0.01
Broj fotografija	0.01
Broj poveznica	0.005
Broj likeova po fotografiji	0.002
Broj likeova po poveznici	0.002
Broj statusa	0.001
Broj likeova po događaju	-0.001
Broj komentara po događaju	-0.002
Broj likeova po objavljenom videu	-0.002
Broj dijeljenja fotografija	-0.004
Broj komentara po statusu	-0.004
Broj objavljenih videa	-0.004
Broj komentara po videu	-0.006
Broj dijeljenja poveznica	-0.007
Broj dijeljenja statusa	-0.007
Broj dijeljenja videa	-0.008
Broj dijeljenja događaja	-0.009
Broj komentara po poveznici	-0.011
Broj komentara po fotografiji	-0.012

Izvor: autor.

Kišić, A.

Kako profilirati birače u političkim kampanjama primjenom umjetne inteligencije?

Tablica 3. Odabir varijabli

Parametri	Vrijednosti
Validacija	5-struka validacija
Vrsta funkcije aktivacije	Hiperbolni tangens
Broj srednjih slojeva	1
Broj neurona u srednjem sloju	4

Izvor: autor.

Tablica 4. Vrednovanje modela dobivenog neuronskom mrežom

Mjera	Vrijednost
RSquare	0,81
RASE	5,98
Mean Abs Dev	3,74
LogLikelihood	141,11
SSE	1572,39

Izvor: autor.

tri sloja. Prvi je ulazni sloj koji sadrži neurone (sve ulazne varijable), drugi je skriveni sloj čija svrha je obrada podataka, a treći je izlazni sloj s čvorom (varijablom) rezultata na izborima definiranim kroz postotni udio glasova kandidata na izborima. Hornik i suradnici (1989.) sugerira jedan srednji sloj u arhitekturi neuronske mreže kao dovoljan za obradu i složenih problema (Hornik i sur., 1989.).

Broj skrivenih neurona jedan je od ključnih parametara modela neuronske mreže. Mali broj neurona može dovesti do podtreniranosti, dok

veliki broj povećava rizik od pretreniranosti. Optimalan broj neurona najčešće se određuje metodom pokušaja i pogrešaka. Neka ranija istraživanja predlažu korištenje aritmetičke sredine broja neurona na ulazu i izlazu kao smjernice za odabir broja neurona u skrivenom sloju.

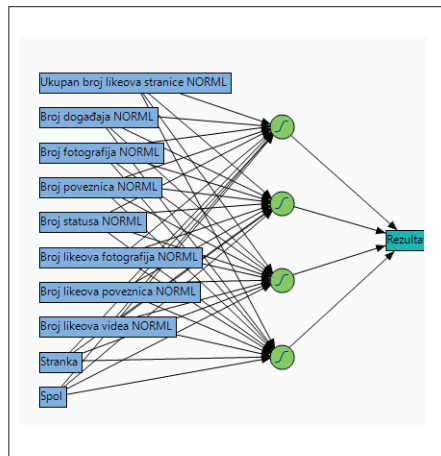
U ovom istraživanju razvijeno je desetak modela neuronske mreže s različitim brojem neurona u skrivenom sloju. Kao najbolja konfiguracija odabran je model s četiri neurona u skrivenom sloju. Tablica u nastavku prikazuje vrijednosti parametara kvalitete modela za ovu konfiguraciju.

Model ima pouzdanost 0,81 (mjera RSquare) što vodi zaključku da je razina generalizacije van skupa podataka francuskih izbora vrlo velika.

Na prvoj slici dan je model neuronske mreže s tri sloja. Prvi sloj koji ima 10 neurona, drugi sloj koji ima 4 neurona te izlazni sloj koji ima jedan neuron. Ovakav model umjetne neuronske mreže nije pretreniran, a pokazuje dobre kvalitete u smislu točnosti i pouzdanosti modela.

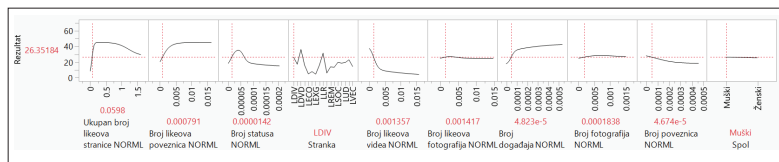
Slike 2. i 3. daju prikaz ovisnosti vrijednosti izlazne varijable o kretanjima ulaznih varijabli. Sve ulazne varijable prikazane su temeljem razine utjecaja na promjenu vrijednosti izlazne varijable.

Slika 1. Model umjetne neuronske mreže



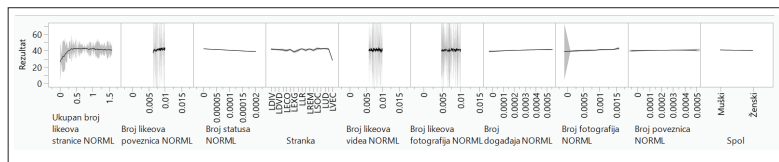
Izvor: autor.

Slika 2. Predviđanje neuronskom mrežom



Izvor: autor.

Slika 3. Granični modeli umjetne neuronske mreže



Izvor: autor.

Tablica 5. Najznačajniji prediktori rezultata u modelu neuronske mreže

Varijable	Utjecaj
Ukupan broj likeova po stranici	0.42
Broj likeova po poveznici	0.34
Broj statusa	0.32
Stranka	0.30
Broj likeova po videu	0.19
Broj likeova po fotografiji	0.15
Broj događaja	0.07
Broj fotografija	0.04
Broj poveznica	0.02
Spol	0.02

Izvor: autor.

Prikazi graničnih modela temelje se na predviđenim graničnim vrijednostima i odgovarajućim rezidualima. Varijable tamnije nijanse na grafu imaju značajniji utjecaj na promjenu izlazne vrijednosti u usporedbi s varijablama svjetlije nijanse.

Nad prediktivnim modelom umjetne neuronske mreže izvedena je analiza osjetljivosti. Rezultati prikazani u tablici 5. daju uvid u najveću prediktorsku snagu varijable Ukupan broj likeova stranice. S druge strane, varijabla Spol ima najmanju snagu predikcije.

4.2. Diskusija rezultata istraživanja

Primjenom neuronske mreže dobiveni su prediktivni

modeli koji služe kao alat za planiranje i upravljanje kampanjama te kao instrument za potporu odlučivanju kroz profiliranje (Klepac i sur., 2016.). Profiliranje će se provesti temeljem razvijenih prediktivnih modela dobivenih neuronskim mrežama iz razloga jer su ti modeli najkvalitetniji. Neuronske mreže igraju značajnu ulogu u prediktivnom modeliranju, pa tako i za profiliranje. Uzimanjem u obzir arhitekture neuronskih mreža i konačnog izlaza, modeliranje neuronskom mrežom i profiliranje iz prediktivnih modela svakako nije jednostavno (Larose, 2014.).

Analiza osjetljivosti radila se za prediktivni model za čiju je izradu korištena neuronska mreža. Utvrđeno je kako najslabije prediktorske snage ima

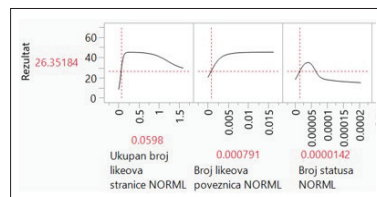
Kišić, A.

Kako profilirati birače u političkim kampanjama primjenom umjetne inteligencije?

Kišić, A.

Kako profilirati birače u političkim kampanjama primjenom umjetne inteligencije?

Slika 4. Odnos najznačajnijih varijabli modela



Izvor: autor.

korištena varijabla koja označava kandidatov spol, dok je najsnažniji prediktor varijabla koja pokazuje ukupan broj koliko je puta likeana stranica kandidata.

Profiliranje temeljeno na analizi osjetljivosti koristi se za podršku odlučivanju u političkim kampanjama, budući da podaci o profilima birača koji simpatiziraju svoga odabranoga kandidata služe kao osnova kako planirati, provoditi i upravljati političkom kampanjom. Identifikacija profila birača vrši se na bazi bitnosti varijabli za predviđanje rezultata izbora. Svakom pojedinačnom kandidatu kroz provedenu analizu osjetljivosti ciljna varijabla ishod izbora treba dati najveću moguću vjerojatnost da će taj kandidat biti odabran. Za svakog kandidata, ciljna varijabla ishod izbora treba postići maksimalnu vjerojatnost odabira kandidata kroz analizu osjetljivosti. Takav pristup ide prema tome kako odgovoriti na pitanje o tome koje su ključne karakteristike glasača koji odlučuju dati potporu nekom kandidatu.

Planiranje kampanje prema biračima je jako ovisno o tome kakvih su profila ti birači. Naime, jasno je kako nema svaki birač jednak način provođenja svojih aktivnosti na Facebooku, te je stoga iznimno važno što više se približiti i prepoznati profile birača kako bi kampanja kojom će se upravljati na Facebooku bila učinkovita i polučila kvalitetne rezultate. Upravo analiza relevantnosti svakog pojedinog atributa ima bitnu ulogu u tome kako shvatiti ključne čimbenike u vođenju aktivnosti prema biračima.

Analiza relevantnosti atributa kroz analizu osjetljivosti ima važnu ulogu u razumijevanju ključnih čimbenika u vezi s upravljanjem odnosima s biračima. Dobro shvaćena relevantnost za svaku pojedinu varijablu kao i razumijevanje

međuvisnosti varijabli su jako snažan alat vođenja uspješnih kampanja, s posebnim naglaskom na vođenje kampanje na društvenim mrežama, u ovom slučaju Facebooka.

Analiza osjetljivosti ima nekoliko bitnih uloga:

- Lociranje ključnih varijabli koje najviše utječu na zavisnu varijablu
- Shvaćanje međudnosa između važnih prediktora i zavisne varijable kao i shvaćanje međudnosa važnih prediktora iz perspektive zavisne varijable (Klepac, 2017). I jedna i druga funkcija je kompatibilna s prepoznavanjem profila birača (Klepac, 2015).
- Kako bi se bolje razumio utjecaj ključnih varijabli na ciljnu varijablu, provedena je analiza osjetljivosti. Analizirat će se na koji način najznačajnije varijable prediktivnog modela, dobivenog korištenjem algoritma neuronskih mreža (broj likeova stranice, broj statusa i broj likeova poveznica), utječu na ciljnu varijablu. U tu svrhu ispitan je odnos između tri najutjecajnije prediktivne varijable i ciljne varijable, a rezultati su vizualizirani na slici 4.

Rezultati dani na slici otkrivaju da određene varijable imaju profile s pozitivnim, a druge s negativnim nagibima. Na primjer, varijabla broj likeova poveznica pokazuje pozitivan nagib, što znači da veći broj likeova na poveznicama koje je kandidat ostvario na Facebooku povećava predviđenu srednju vrijednost postotka osvojenih glasova na izborima. Varijable ukupan broj likeova stranice i broj statusa pokazuju različite nagibe ovisno o intervalu vrijednosti – pri nižim vrijednostima nagib je pozitivan, dok je pri višim vrijednostima negativan. Ovakvi obrasci upućuju na nelinearnu povezanost između prediktivnih i ciljne varijable.

Na temelju tri najznačajnije varijable provedeno je profiliranje potencijalnih birača primjenom k-sredina klasteriranja. Rezultati prikazani u tablici 6. pokazuju da je optimalan broj klastera za definiranje profila birača četiri.

Za evaluaciju je korišten Cubic Cluster Criterion (CCC), pri čemu veće vrijednosti CCC ukazuju na kvalitetnije grupiranje birača. Testirani su različiti brojevi klastera (od 2 do 5), a model s četiri klastera postigao je najveću vrijednost CCC-a. Karakteristike

Tablica 6. Vrednovanje klastera

Broj klastera	CCC
2	-2.14
3	-4.40
4	1.19
5	1.13

Izvor: autor.

Tablica 5. Najznačajniji prediktori rezultata u modelu neuronske mreže

Klaster	Ukupan broj likeova po stranici	Broj statusa	Broj likeova po poveznici
1	0.036	0.103	0.053
2	0.768	0	0.309
3	0.038	0.571	0.072
4	0.077	0.041	0.564

Izvor: autor.

identificirane ovim modelom određene su na temelju srednjih vrijednosti varijabli. Rezultati su prikazani u tablici u nastavku.

Prva grupa obuhvaća profile s najmanjim prosječnim brojem likeova na stranicama i poveznicama, ali s umjerenim brojem statusa kandidata. Druga grupa uključuje kandidate s najvećim brojem likeova na stranicama, umjerenim brojem likeova na poveznicama i bez objavljenih statusa. Treća grupa ističe se najvećim brojem objavljenih statusa, dok četvrtu grupu karakterizira najveći broj likeova na poveznicama kandidata.

5. Zaključak

Kroz ovaj rad predstavilo se provedeno istraživanje o tome kako primjenjivati algoritme strojnog učenja i prikupljene podatke na društvenoj mreži Facebook u svrhu razvoja modela za predikciju ishoda rezultata političkih izbora i grupiranje kako bi se provela profilacija birača na političkim izborima. Znanstveni doprinos ovog istraživanja očituje se u prikazu kako se integriraju napredni algoritmi korišteni u strojnom učenju, kao što su to primjerice umjetne neuronske mreže ili klasteriranje po podacima s Facebooka, za profilaciju izbornih birača za političke kampanje, pri čemu je naglasak bio na provedenim francuskim lokalnim izborima 2020. godine. Doprinos rada iskazuje se i u tome kako primijeniti umjetnu inteligenciju za političke aktivnosti, te se daje dodatni prošireni pogled u to kako se podaci koji o političkim akterima, bilo da je riječ o političarima ili onima koji ih biraju, mogu

iskoristiti kako bi imali efikasniju i produktivniju političku kampanju, te kako se mogu na bolji način shvatiti preferencije birača. Prikupljeni rezultati daju vrlo korisne smjernice koje akteri političkih utakmica mogu primjenjivati u osmišljavanju osobnih poruka kandidata u izbornim kampanjama, pri čemu obuhvaća segmentirane interese i mišljenja raznih birača kroz njihovu komunikaciju s političkim kandidata putem društvenih mreža. Ovakav pristup ima i neka svoja ograničenja. Prije svega, ne treba generalizirati rezultate na druga okruženja van Francuske, te bi bilo potrebno prilagođavati modele prema drugim obilježjima nekog drugog geografskog prostora. Također, treba primijetiti i kako nisu svi birači jednako aktivni na društvenim mrežama, i to itekako može dati utjecaj na to koliko je profil točan, pa uzorak može biti „kontaminiran“ odnosno pristran. I na kraju, ali ništa manje važno, primjena umjetne inteligencije u provođenju političkih marketinških promotivnih aktivnosti itekako može otvoriti i etičke dileme oko pitanja privatnosti birača, te eventualnih manipulacija prema njima.

Kišić, A.

Kako profilirati birače u političkim kampanjama primjenom umjetne inteligencije?

Kišić, A.

Kako profilirati birače u političkim kampanjama primjenom umjetne inteligencije?

Literatura

Alkhiary,D., Saleh, D., & Marie, M. (2024). Proposed Machine learning model for predicting Egyptian Parliament Election Results. *International Journal of Advanced Science and Computer Applications*. doi: 10.47679/ijasca.v4i1.61

Elo, G., Ghansah, B., Ephrem, K., & Kwaa-Aidoo, K. (2024). Critical Review of Stack Ensemble Classifier for the Prediction of Young Adults' Voting Patterns Based on Parents' Political Affiliations. *International Journal of an Emerging Transdiscipline*. doi: 10.28945/5254

Emeto, I., C., Elenwo, D., C., Anthony, U. W., Shawulu, J. C., Ajayi, C. O., & Galadima, A. (2024). Election Voting Trend Prediction System. *Asian Journal of Research in Computer Science*. doi: 10.9734/ajrcos/2024/v17i110508

Feng, G., Chen, K., Cai, H., & Li, Z. (2023). A Hybrid Method of Sentiment Analysis and Machine Learning Algorithm for the US Presidential Election Forecasting. In *2023 IEEE International Conference on Big Data (BigData) (1495-1500)*. IEEE.

Gowtham, T. J. (2024). Election result prediction. *Indian Scientific Journal of Research in Engineering and Management*. doi: 10.55041/ijsrem36223

Gu, F., Chen, K., Cai, H., & Li, Z. (2023). A Hybrid Method of Sentiment Analysis and Machine Learning Algorithm for the U.S. *Presidential Election Forecasting*. doi: 10.1109/bigdata59044.2023.10386950

Guan, L., & Mani, G. S. (2023). Election Forecasting Using Macroeconomic and Social Indicators via Machine Learning. *Journal of Student Research*. doi: 10.47611/jsrshs.v11i13.350

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359–366. [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8)

Kišić, A. (2023). *Modeli za predikciju ishoda političkih izbora korištenjem društvene mreže Facebook i algoritama strojnog učenja* (Doktorska disertacija, Sveučilište u Zagrebu, Fakultet organizacije i informatike).

Kišić, A. (2018). The Use of Social Media in Political Campaigns: The Case of Croatian Local Elections 2017. In *Central European Conference on Information and Intelligent Systems* (pp. 133-139). Faculty of Organization and Informatics Varazdin.

Kišić, A., & Kliček, B. (2021). Machine Learning Based Prediction of Croatian 2017 Local Elections. In *2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO) (1382-1386)*. IEEE.

Klepac, G., Mrsić, L., & Kopal, R. (2016). Efficient Risk Profiling Using Bayesian Networks and Particle Swarm Optimization Algorithm. In D. Jakóbczak (Ed.), *Analyzing Risk through Probabilistic Modeling in Operations Research* (pp. 91–124). Hershey, PA, USA: IGI Global. doi:10.4018/978-1-4666-9458-3.ch004

Klepac, G. (2017). Customer Profiling in Complex Analytical Environments Using Swarm Intelligence Algorithms // *Nature-Inspired Computing: Concepts, Methodologies, Tools, and Applications (3 Volumes) / Information Resources Management Association (USA) (ur.)*. USA: IGI-Global, 1391-1422.

Klepac, G. (2015). Particle Swarm Optimization Algorithm as a Tool for Profiling from Predictive Data Mining Models. In S. Bhattacharyya & P. Dutta (Eds.), *Handbook of Research on Swarm Intelligence in Engineering* (406–434). Hershey, PA, USA: IGI Global. doi:10.4018/978-1-4666-8291-7.ch013

Larose, D. T., & Larose, C. D. (2014). *Discovering knowledge in data: an introduction to data mining* (Vol. 4). John Wiley & Sons.

Mandhyan, G. R., & Bose, S. (2024). Decoding the ballot: predicting indian general elections with machine learning. *International journal of advanced research*. doi: 10.21474/ijar01/19477

Mishra, S., & Lee, J. (2023). Using Machine Learning Techniques to Predict United States House of Representatives Elections. *Journal of Student Research*. doi: 10.47611/jsrshs.v12i3.4697

Myilvahanan, K., Yashas, P., Pasha, S., Ismail, M., & Tharun, V. (2023, February). A Study on Election Prediction using Machine Learning Techniques. In *2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS)* (1518-1520).

Kišić, A.

Kako profilirati birače u političkim kampanjama primjenom umjetne inteligencije?

