

# Analiza značaja brandova sportske obuće obradom velikih tekstualnih podataka

---

Slamić Tarade, Sara

Master's thesis / Diplomski rad

2023

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University North / Sveučilište Sjever**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:122:872745>

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

Download date / Datum preuzimanja: **2024-11-07**



Repository / Repozitorij:

[University North Digital Repository](#)



**SVEUČILIŠTE SJEVER**

**SVEUČILIŠNI CENTAR VARAŽDIN**

**Diplomski sveučilišni studij Odnosi s javnostima**



DIPLOMSKI RAD br. 266/OJ/2023

**ANALIZA ZNAČAJA BRANDOVA SPORTSKE  
OBUĆE OBRADOM VELIKIH TEKSTUALNIH  
PODATAKA**

Sara Slamić Tarade

Varaždin, srpanj 2023.

**SVEUČILIŠTE SJEVER**

**SVEUČILIŠNI CENTAR VARAŽDIN**

**Diplomski sveučilišni studij Odnosi s javnostima**



DIPLOMSKI RAD br. 266/OJ/2023

**Analiza značaja brandova sportske obuće obradom  
velikih tekstualnih podataka**

Studentica:

Sara Slamić Tarade, 1584/336D

Mentorica:

doc. dr. sc. Dijana Vuković

Varaždin, srpanj 2023.

# Prijava diplomskog rada

## Definiranje teme diplomskog rada i povjerenstva

ODJEL Odjel za odnose s javnostima

STUDIJ diplomski sveučilišni studij Odnosi s javnostima

PRISTUPNIK Sara Slamić Tarade

MATIČNI BROJ 0301004631

DATUM 14. 6. 2023.

KOLEGIJ Integrirana tržišna komunikacija

NASLOV RADA Analiza značaja brandova sportske obuće obradom velikih tekstualnih podataka

NASLOV RADA NA ENGL. JEZIKU Analysis of the significance of sports footwear brands by processing large textual data

MENTOR dr.sc. Dijana Vuković

ZVANJE Docentica

ČLANOVI POVJERENSTVA

1. doc. dr. sc. Ivana Stanić- predsjednica
2. doc.dr.sc. Gordana Lesinger - članica
3. prof. dr. sc. Dijana Vuković - mentorica
4. izv.prof.dr.sc. Darijo Čerepinko - zamjenski član
- 5.

## Zadatak diplomskog rada

BROJ 266/OJ/2023

OPIS

Analize teksta predstavlja razumijevanje kulturnih, ekonomskih, socijalnih ili ideoloških aspekata sadržanih u suštini nekog teksta i kako su oni povezani s određenim kontekstom. Danas su tekstualni podaci sveprisutni i često im se slobodno pristupa s više izvora kao što su dobro poznate platforme društvenih medija Facebook i Twitter, blogovi, tematski forumi poput TripAdvisora i tradicionalni mediji kao što su novinske vijesti. Potrošači izražavaju svoje osjećaje i mišljenja u vezi s proizvodom ili uslugom na više načina, a njihov odnos prema markama često se može izdvojiti iz komunikacije na društvenim mrežama. Interakcije potrošača kako međusobne tako i s tvrtkama mogu utjecati na pridobivanje potencijalnih potrošača te učinak na buduće proizvode koje će tvrtke razvijati. U radu je potrebno:

- \* definirati elemente marke;
- \* pojasniti analizu značaja upravljana markom na tržištu;
- \* primjenu nove paradigme ocjene značaja marke koja se temelji na tekstualnim sadržajima s interneta vezanim uz robne marke sportske obuće kao komplementaran način klasičnim metodama istraživanja tržišta kao što su upitnici, intervjui, fokus grupe i slično.
- \* definirati zaključak rada.

ZADATAK URUČEN 23. 06. 2023.



POTPIS MENTORA  
*[Signature]*

## Sadržaj

1. UVOD .....	1
1.1. Predmet i ciljevi rada .....	2
1.2. Izvori i metode prikupljanja podataka .....	2
1.3. Struktura i sadržaj rada .....	4
2. MARKA U SUVREMENOM OKRUŽENJU .....	5
2.1. Pojam i definicija marke .....	5
2.2. Osnovni elementi marke .....	5
2.3. Identitet marke .....	7
2.3.1. Vrijednost i mjere vrijednosti marke .....	8
2.3.2. Imidž i prepoznatljivost marke .....	12
2.4. Izgradnja marke .....	14
2.5. Upravljanje markom .....	15
3. ANALIZA ZNAČAJA UPRAVLJANJA MARKOM.....	16
3.1. Teorijski aspekti upravljanja markom .....	17
3.2. Potreba za mjerenjem vrijednosti marke .....	17
3.3. Prikaz metoda analize značaja marke .....	18
3.3.1. <i>Analiza teksta i konteksta</i> .....	19
3.3.2. <i>NLP metode određivanja frekvencije pojavnosti riječi</i> .....	22
3.3.3. <i>NLP metode za tematsko modeliranje</i> .....	24
3.3.4. SBS – Semantic Brand Score ili semantička ocjena marke .....	25
3.3.5. NLP metoda za analizu sentimenta .....	30
3.3.6. Programski alati za NLP analize teksta prethodno opisanim metodama .....	34
4. REZULTATI ISTRAŽIVANJA.....	36
4.1. Istraživački model.....	36
4.2. Predmet i problem istraživanja .....	36
4.3. Ciljevi istraživanja .....	37
4.4. Hipoteze istraživanja .....	37
4.5. Analiza rezultata određivanja ključnih robnih marki .....	38
4.5.1. TF-IDF analiza .....	39
4.5.2. Tematsko modeliranje .....	41
4.5.3. SBS analiza .....	44
4.5.4. Analiza sentimenta .....	47
4.6. Potvrda postavljenih hipoteza istraživanja .....	51

4.7. Ograničenja istraživanja .....	55
5. ZAKLJUČAK .....	57
Popis literature.....	60
Popis slika .....	63
Popis tablica .....	63
Prilog 1 .....	64

## SAŽETAK

Robna marka u suvremenom okruženju razlikuje proizvode ili usluge tvrtke od onih njezinih konkurenata. Snažna robna marka može pomoći tvrtki u izgradnji povjerenja kod kupaca, budući da oni smatraju da je marka pouzdana i vrijedna povjerenja. U ovom diplomskom radu obrađuje se analiza značaja brandova sportske obuće obradom velikih tekstualnih podataka. U radu se koristi NLP metoda za analizu velikih tekstualnih sadržaja koji su dostupni na internetu. Napravljena je analiza najznačajnije robne marke sportske obuće kao što su: Nike, Adidas, Puma, Under Armour, Reebok i Asics. Predmet samog istraživanja vezan je uz analizu značaja marki i ocjenu mišljenja potrošača sportske obuće koje se temelji na obradi velikih tekstualnih podataka preuzetih s Interneta. Ciljevi istraživanja vezani su uz novi način metrike i vrednovanja značaja robnih marki sportske obuće primjenom NLP metoda za analizu velikih tekstualnih sadržaja dostupnih na internetu. Dobiveni rezultati dokazuju da novi način metrike i vrednovanja značaja robnih marki može u značajnoj mjeri poboljšati postojeće metode vrednovanja robnih marki.

Ključne riječi: brand, marka, NLP metoda, analiza teksta, sportska obuća

## **ABSTRACT**

The brand in a modern environment distinguishes a company's products or services from those of its competitors. A strong brand can help build trust with customers, as they perceive the brand to be reliable and trustworthy. This graduate thesis focuses on the analysis of the significance of sports footwear brands through the processing of large textual data. The study utilizes NLP (Natural Language Processing) methods to analyze extensive textual content available on the internet. The analysis is conducted on the most prominent sports footwear brands such as Nike, Adidas, Puma, Under Armour, Reebok, and Asics. The research subject is related to the analysis of brand significance and the evaluation of consumer opinions on sports footwear, based on the processing of large textual data retrieved from the internet. The research objectives are tied to a new approach to metrics and evaluation of brand significance in sports footwear, employing NLP methods to analyze extensive textual content available on the internet. The obtained results demonstrate that this new approach to metrics and evaluation can significantly improve existing methods of brand evaluation.

Keywords: brand, NLP method, text analysis, sports footwear



# 1. UVOD

Tema ovog diplomskog rada bavi se analizom značaja odnosno vrijednosti robnih marki sportske obuće. Istaknuti su glavni elementi robne marke kao što su identitet, imidž i kapital odnosno vrijednost robne marke. Opisani su tradicionalni modeli vezana uz kapital marke i korištene marketinške metode u razvoju i vrednovanju robne marke. Naglašen je proces učinkovite izgradnje robne marke kao jedan od važnih uvjeta za pozicioniranje poduzeća na tržištu kroz izgradnju lojalnosti kupaca i povećanje vrijednosti proizvoda ili usluga. Posebno je istaknut značaj upravljanja robnom markom s ciljem stvaranje snažne robne marke koja može potaknuti lojalnost kupaca, generirati veći profit i pružiti održivu konkurentsku prednost. Učinkovito upravljanje robnom markom uključuje niz aktivnosti poput stvaranja identiteta robne marke, pozicioniranja robne marke na tržištu, osiguravanja dosljednosti u porukama robne marke i iskorištavanja vrijednosti robne marke za poticanje rasta i profitabilnosti.

Ključno istraživanje temelji se na primjeni inovativnih metoda određivanja značaja i vrijednosti robne marke koje koriste tehnike obrade prirodnog jezika (Natural Language Processing ili ukratko NLP) analizom sadržaja velikog tekstualnog korpusa skinutog s web stranica koje se bave robnim markama sportske obuće. NLP je široko razvijena grana umjetne inteligencije koja koristi različiti spektara tehnika za obradu i analize tekstualnih sadržaja iz bilo kojih izvora. Primjenom tih tehnika izvršena su istraživanja značaja i vrijednosti robnih marki sportske obuće obradom velikih tekstualnih podataka prikupljenih s web stranica, mrežnih foruma i blogova te prikupljenih ocjene potrošača istraživanih marki. Cijeli proces određivanja značaja robnih marki sportske obuće primjenom NLP uključuje nekoliko koraka. Prvo su analizom tekstualnog korpusa pomoću metode frekvencijske analize i metode modeliranja tema iz neoznačenih tekstualnih sadržaja identificirane ključne robne marke sportske obuće za koje je zatim izvršeno rangiranje značaja tih robnih marki primjenom inovativne metode semantičke ocjene značaja (Semantic Brand Score). U zadnjem koraku ovog istraživanja izvršila se analiza sentimenta potrošača spram najbolje rangirane robne marke. Postignuti rezultati istraživanja o značaju robnih marki sportske obuće potvrdili su postavljene ciljeve da primjena NLP metoda za analizu velikih tekstualnih sadržaja dostupnih na internetu predstavljaju novi način metrike u procesu vrednovanja značaja robnih marki što u značajnoj mjeri proširuje mogućnosti postojećih metode vrednovanja robnih marki.

S obzirom na velike količine tekstualnih sadržaja na web stranicama i društvenim medijima u značajnoj mjeri proširuju se mogućnosti prikupljanja informacija o stavovima potrošača i javnosti vezanim uz robne marke sportske obuće.

## 1.1. Predmet i ciljevi rada

Predmet ovog diplomskog rada vezan je uz obradu značaja brandova sportske obuće koji se temelji na obradi velikih tekstualnih podataka prikupljenih s web stranica, mrežnih foruma i blogova te prikupljenih ocjene potrošača istraživanih marki. Ciljevi rada vezani su uz primjenu nove paradigme ocjene značaja marke koja se temelji na tekstualnim sadržajima s interneta vezanim uz robne marke sportske obuće kao komplementaran način klasičnim metodama istraživanja tržišta kao što su upitnici, intervjui, fokus grupe i slično. Jedan od ciljeva je pokazati da ovaj način analize značaja marki na tržištu značajno proširuje mogućnost marketinških i stručnjaka za upravljanje markom i u određenoj mjeri je manje pristran i ograničen nego što su upitnici i intervjui. Također se temeljem te analize pokazuje kako se mogu napraviti poboljšanja u strategiji nastupa u odnosima s javnošću vezano uz robne marke sportske obuće.

## 1.2. Izvori i metode prikupljanja podataka

Za potrebe istraživanja u ovom diplomskom radu prikupljeni su tekstualni sadržaji 57 različitih web stranica (javno i besplatno dostupnih) kao što su Forbes, Finance Yahoo, LinkedIn, Money CNN te različitih blogova. Većina sadržaja je iz 2021. godine no ima ih i iz ranijih godina. U nastavku je prikazan fragment samo jednog manjeg broja tih web stranica (Tablici 1), a kompletna lista je dana u Prilogu 1. Za svaku web stranicu u tablici je prikazan naziv te stranice i poveznica na nju.

Tablica 1. Fragment web stranica s kojih su prikupljeni sadržaji

ID	Naslov Web stranice	Poveznica
	<u>2</u>	<u>3</u>
ID02	<a href="#">20 Best Running Shoes Brands to Know - The Trend Spotter</a>	<a href="https://www.thetrendspotter.net/best-running-shoes-brands/">https://www.thetrendspotter.net/best-running-shoes-brands/</a>
ID07	<a href="#">The Most Influential Athletes with Athletic Apparel Endorsement Deals   Opendorse</a>	<a href="https://opendorse.com/blog/the-most-influential-athletes-with-athletic-apparel-endorsement-deals/">https://opendorse.com/blog/the-most-influential-athletes-with-athletic-apparel-endorsement-deals/</a>
ID13	<a href="#">Sports Apparel Online Retailing Market Drive Big Growth   Nike, Adidas, PUMA, Under</a>	<a href="https://reportedtimes.com/sports-apparel-online-retailing-market-drive-big-growth-nike-adidas-puma-under-armour/">https://reportedtimes.com/sports-apparel-online-retailing-market-drive-big-growth-nike-adidas-puma-under-armour/</a>

	<a href="#">Armour · Wall Street Call</a>	
ID20	<a href="#">Sports brands, don't forget that women love sneakers too!</a>	<a href="https://www.heuritech.com/blog/articles/sneakers-are-for-women-too/">https://www.heuritech.com/blog/articles/sneakers-are-for-women-too/</a>
ID44	<a href="#">The Forbes Fab 40: Puma Debuts On 2019 List Of The World's Most Valuable Sports Brands</a>	<a href="https://www.forbes.com/sites/mikeozanian/2019/10/16/the-forbes-fab-40-puma-debuts-on-2019-list-of-the-worlds-most-valuable-sports-brands/?sh=39101ca5d356">https://www.forbes.com/sites/mikeozanian/2019/10/16/the-forbes-fab-40-puma-debuts-on-2019-list-of-the-worlds-most-valuable-sports-brands/?sh=39101ca5d356</a>
ID53	<a href="#">21 Best Running Shoes Brands to Know - The Trend Spotter</a>	<a href="https://www.thetrendspotter.net/best-running-shoes-brands/">https://www.thetrendspotter.net/best-running-shoes-brands/</a>

(izvor: rad autora temeljem provedenog istraživanja)

Za prikupljanje tekstualnog sadržaja korištena je tehnika tzv. web „grebanja“ sadržaja (**eng.** *web scraping*) pomoću Python biblioteka. Tako skinuti tekstualni sadržaji pohranjeni su u Excel i CSV datoteke kao paragrafi. Svaki paragraf predstavlja jedan dokument. Ukupan broj paragrafa odnosno dokumenata je 1660. Temeljem toga stvoren je istraživački korpus koji se sastoji od velikog broja različitih tekstualnih dokumenata. Prikupljeni korpus dokumenata koristite se za NLP analize koje su obrađene u točki 3. Za potrebe analize sentimenta prikupljena su mišljenja potrošača s web stranica tvrtki koje prikupljaju podatke za razne proizvode i robne marke kao što su ConsumerAffairs i SiteJabber. Tekstualna mišljenja proizvođača i ocjene (broj zvjezdica između 1 i 5) također su prikupljena primjenom web „grebanja“ i pohranjena u CSV i Excel datoteke. Mišljenja potrošača prikupljena su samo za robnu marku za koju su NLP analize pokazale da je rangirana kao najznačajnija.

### 1.3. Struktura i sadržaj rada

Ovaj diplomski rada podijeljen je u pet poglavlja i u nastavku je ukratko opisan sadržaj svake točke.

U prvom poglavlju naziva *Uvod* dana su uvodna razmatranja u kojima se ističe predmet istraživanja i ciljevi koji se žele postići. Ukratko su opisani izvori i metode prikupljanja podataka. Opisana je struktura i sadržaj rada.

Drugo poglavlje naziva *Marka u suvremenom okruženju* obrađuje problematiku teorijske analize značaja robne marke u suvremenom okruženju. Dana je definicija marke te njenih najznačajnijih karakteristika kao što identitet, imidž i vrijednost odnosno kapital marke. Posebno su istaknuti najznačajniji modeli za određivanje vrijednosti marke koji se koriste u klasičnim pristupima marketinških i stručnjaka za razvoj marki. Naglašeni su glavni principi izgradnje i upravljanja robnom markom.

*Analiza značaja upravljanja markom* naziv je trećeg poglavlja u kojem se dodatno diskutira problematika upravljanja markom i određivanja vrijednosti marke utemeljenoj na inovativnom pristupu određivanja značaja marke primjenom NLP metoda na analizi korpusa tekstualnih sadržaja s interneta vezanih uz sportsku obuču. Opisane su NLP metode za frekvencijsku analizu, tematsko modeliranje, semantičku ocjenu značaja marki primjenom inovativne metode SBS ocjene te metoda analize sentimenta potrošača spram robnih marki.

U četvrtom poglavlju prikazani su i diskutirani rezultati istraživanja utemeljeni na NLP tehnikama i modelima primijenjenih na tekstualni korpus sadržaja s interneta vezanih uz sportsku obuču temeljem čega su izdvojene i ocijenjene najznačajnije marke. U toj točki su također potvrđene postavljene hipoteze istraživanja usporedbom značaja robnih marki dobivenih NLP analizom s financijskim pokazateljima robnih marki koje su istraživane. Istaknuta su i neka ograničenja istraživanja.

U petom poglavlju dan je zaključak koji naglašava doprinose rezultata istraživanja važne za razvoj i upravljanje robnom markom i dane su smjernice kako se kombiniranjem određivanja vrijednosti i značaja marke pomoću NLP analiza i klasičnih modela brendiranja mogu unaprijediti marketinške operacije važne za razvoj marke.

## 2. MARKA U SUVREMENOM OKRUŽENJU

Robna marka u suvremenom okruženju razlikuje proizvode ili usluge tvrtke od onih njezinih konkurenata. Snažna robna marka može pomoći tvrtki u izgradnji povjerenja kod kupaca, budući da oni smatraju da je marka pouzdana i vrijedna povjerenja. Snažna robna marka može biti važan čimbenik i u financijskoj vrijednosti tvrtke, jer može utjecati na sposobnost tvrtke da privuče i zadrži kupce, kao i na sposobnost da naplaćuje više cijene za svoje proizvode ili usluge.

### 2.1. Pojam i definicija marke

Keller (2003) u prvi plan ističe definiciju marke (**eng. brand**) prema American Marketing Association (AMA) kao „naziv, izraz, znak, simbol ili dizajn, ili kao njihovu kombinaciju radi identifikacije robe i usluge jednog ili grupe prodavača pri čemu ih to razlikuje od konkurencije“.

Keller dalje naglašava da svaki puta kad marketing služba proizvođača kreira novo ime ili logotip za neki novi proizvod da marketinška služba zapravo ima na umu stvaranje robne marke, no ujedno i proširuje samu osnovnu definiciju prema AMI u smislu stvaranja nove količine svijesti, ugleda i istaknutosti na tržištu. Treba naglasiti dva gledišta značenja robne marke, potrošačevo i proizvođačevo. Za potrošača robna marka znači dobru identifikaciju izvora proizvoda, povezivanje s proizvođačem i prepoznavanje signala kvalitete.

Proizvođaču robna marka stvara mogućnost zaštite jedinstvenih značajki, mogućnost povratnih informacija koje prima od zadovoljnih korisnika i konkurenata, a posljedično i financijsku prednost. Kotler i Pfoertsch (2006, str.12-13) dodatno naglašavaju da je marka jamstvo ukupne percepcije svega što se vidi ili zna o proizvodu. Razvoj te percepcije je proces razvoja robne marke (**eng. branding**) koji se temelji na raznim marketinškim operacijama, ali s osnovnim ciljem stvaranja percepcije kod potrošača. Ako usvojimo činjenicu da je marka entitet onda treba prihvatiti da proces razvoja marke zapravo treba naučiti potrošača o svemu onome što definira marku, tj. daje joj značenje i posebnost koju će on prepoznati (Keller, 2003).

### 2.2. Osnovni elementi marke

Robnu marku čine različiti dijelovi koje možemo nazvati elementi marke. Kotler i Keller (2009) te elemente još imenuju i kao identitete koji čine neku marku.

Prema njima elemente marke čine „naziv marke, logotip, simboli, boje, likovi, slogani, prateći zvučni efekti, dizajn i značajke proizvoda, pakiranje i web stranice“. Za većina tih elemenata možemo reći da su opipljivi prikaz marke i u pravilu su zaštićeni. Proizvođači također puno pažnje pridaju sigurnosnoj zaštiti navedenih elemenata od različitih oblika krivotvorenja. Neke jake robne marke mogu koristiti sve navedene elemente, a neke samo najvažnije za njihov proizvod ili uslugu. U nastavku će se ukratko pojasniti glavni elementi robne marke.

1. *Ime marke* predstavlja element koji već u startu treba jasno razlikovati proizvod od proizvoda konkurencije (Keller 2013). Ključno pitanje kod stvaranja robne marke je originalnost imena. Nadalje Keller (2013, str. 152-154) naglašava da ime marke mora biti nedvosmisleno, razumljivo i jasno kako u značenju tako i u izgovoru. Naravno uvijek ostaje otvoreno pitanje da li ime marke jednako dobro zvuči npr. i na engleskom i na francuskom govornom području.
2. *Logotipi i simboli marke*, poslije jedinstvenog imena, moraju biti sveobuhvatni i jasni. Svojim vizualnim izgledom oni trebaju ne samo identificirati nego i razlikovati marku od konkurencije i razvijati svijesti o robnoj marki. Imaju dugu povijesnu tradiciju, npr. habsburški orao Austro-Ugarskog carstva. Mogu biti samo slovni znakovi vezani uz ime tvrtke, ali prikazani na specifičan način u prepoznatljivom obliku npr. Dunhill ili Kit Ket, no mogu biti i potpuno apstraktni npr. poznata zvijezda Mercedesa. Dodatna prednost logotipa u razvoju marke ogleda se u njihovoj svestranosti jer s obzirom na to da su često neverbalni lako ih je prenositi kroz više kategorija proizvoda. Specifičnosti su im fontovi, veličina, boja i druge značajke koje ih čine prepoznatljivim (Keller, 2013, str. 156-157).
3. *Različite boje* daju različito značenje marki i stoga se trebaju pažljivo birati, a često mogu imati prikriveno značenje kao na primjer crvena boja označava ljubav (Coca-Cola).
4. *Slogani* trebaju iskazati filozofiju kompanije i u pravilu se radi o kratkim frazama koje pružaju uvjerljivu informaciju o marki (Keller, 2009). Često su sastavni dio reklamnih poruka npr. poznata Snickersova poruka „Gladan si? Uzmi Snickers!“, ali mogu se vidjeti i na pakiranjima. Kao i u slučaju imena marke i slogani su vrlo moćno skraćeno sredstvo kod izgradnje vrijednosti marke i pomoći potrošačima da shvate što marku čini posebnom (Keller, 2013, str. 159).
5. *Pakiranje* ne služi samo za fizičku zaštitu proizvoda već određuje cjelokupni identitet i na njemu u pravilu se ističu prethodno navedeni elementi.

6. S obzirom na današnji veliki značaja Interneta, kao važan element treba istaknuti *web stranice i platforme društvenih mreža* bez kojih je danas ne zamislivo razvijanje robne marke. S obzirom na široku paletu društvenih mreža (*Facebook, Instagram, LinkedIn, Twitter, Pinterest* i druge) svaka od njih zahtjeva različite elemente (video, slike, dizajn, sadržaj) za privlačenje korisnika. Web stranice su vezane uz odgovarajuću internetsku domenu koristeći URL (uniform resource locator) koji točno određuje lokaciju u Web prostoru, a vrlo često i štiti robnu marku jer se tvrtke jako brinu o zaštiti svoje domene (Keller, 2013, str. 156).

Može se zaključiti prema Kelleru da su navedeni elementi: ime marke, logotip, slogan ključni elementi koji u kombinaciji sa ostalima ističu samu prepoznatljivost robne marke.

### **2.3. Identitet marke**

Identitet marke skup je karakteristika koje definiraju marku i razlikuju je od drugih marki. Uključuje elemente kao što su naziv marke, logotip, shema boja, slogane i druge vizualne elemente koji komuniciraju osobnost, vrijednosti i pozicioniranje brenda na tržištu. Snažan identitet robne marke pomaže u stvaranju dosljedne slike marke i čini marku lako prepoznatljivom kupcima (Aaker, 1996). Identitet robne marke pomaže u stvaranju dosljedne i nezaboravne slike marke, što kupcima olakšava prepoznavanje i pamćenje marke. Također pomaže u stvaranju osjećaja povjerenja i lojalnosti među kupcima, jer je vjerojatnije da će vjerovati i poslovati s markom koja ima snažan i dosljedan identitet (Ries & Ries, 2002). Razvijanje identiteta marke uključuje pažljivo razmatranje ciljne publike, vrijednosti i jedinstvene prodajne ponude marke, kao i duboko razumijevanje konkurentskog okruženja. Zahtijeva promišljen pristup dizajnu i slanju poruka koji je usklađen s cjelokupnom strategijom i ciljevima marke (Keller, 2013, str. 142-170).

Pojasnimo nekoliko primjera elemenata identiteta marke:

1. *Ime robne marke* - važan je dio njezina identiteta. Pamtljivo i jedinstveno ime marke može pomoći da se marka razlikuje od konkurencije i olakšati kupcima da ga upamte.
2. *Logo* - predstavlja vizualni prikaz marke i često se koristi za identifikaciju marke. Snažan logotip može pomoći u stvaranju vizualnog identiteta robne marke i učiniti je lako prepoznatljivom kupcima.
3. *Shema boja* - boje koje se koriste u vizualnom identitetu marke mogu pomoći u stvaranju osjećaja razlikovanja i prepoznavanja. Različite boje mogu izazvati različite

emocije i asocijacije, stoga je važno odabrati boje koje su u skladu s vrijednostima i osobnošću brenda.

4. *Tipografija* - tipovi i stilovi fontova koji se koriste u vizualnom identitetu marke mogu pomoći u stvaranju jedinstvenog i dosljednog izgleda.

5. *Slike i grafika* - korištene u vizualnom identitetu marke mogu pomoći u prenošenju poruke i vrijednosti marke.

6. *Ton komunikacije* - način na koji brend komunicira, uključujući njegov jezik, stil i ton, može pomoći u utvrđivanju njegove osobnosti i vrijednosti.

Jedna od prepoznatljivih robnih marki je Coca - Cola čiji identitet se gradi još od početka 20 stoljeća i neprekidno se i dalje razvija što se može vidjeti iz nekoliko primjera elemenata identiteta:

1. Coca - Cola je poznata po svojoj prepoznatljivoj crveno-bijeloj shemi boja, koja se koristi na ambalaži, reklamnim materijalima i materijalima za brendiranje.
2. Logotip Coca - Cole ispisan je prepoznatljivim kurzivnim pismom koje prepoznaju ljudi diljem svijeta.
3. Coca - Cola ima jedinstven i prepoznatljiv oblik boce, s oblikovanim dizajnom koji je odmah prepoznatljiv.
4. Slogan "Prava stvar" poznati je slogan koji Coca - Cola koristi dugi niz godina, a pomaže u jačanju identiteta brenda kao visokokvalitetnog, autentičnog proizvoda.
5. Isticanje slika životnog stila – Coca - Colin marketing često prikazuje ljude koji uživaju u piću u društvenim okruženjima, prikazujući zabavan i bezbrižan način života.

Navedeni primjeri pokazuju koliku važnost identitet i prepoznatljivost robne marke ima na uspješnost poslovanja Coca - Cole i vjernost njezinih potrošača kroz sve ove godine

### *2.3.1. Vrijednost i mjere vrijednosti marke*

Temelje kapitala i vrijednosti marke postavili su Aaker (1991) i Keller (2013) i kada se govori o razvoju marke u suvremenim uvjetima njihova razmatranja se koriste u svim marketinškim operacijama.

U današnje vrijeme vrijednost renomirane marke je puno veća budući da je puno teže danas graditi marku nego što je to bilo prije nekoliko desetljeća. Danas je trošak oglašavanja i distribucije skuplji pa većina marki trebaju biti praktični, a i svakodnevno raste broj novih marki, a i samim time konkurencija postaje „žešća“ budući da se otprilike 3000 robnih marki uvede svake godine u supermarketete. (Aaker,1991).



Kapital robne marke nije ništa drugo nego vrijednost marke. Ali koncept pridavanja vrijednosti brendu je vrlo zanimljiv i ima dublje značenje. Općenito gledano, vjeruje se da dobar kapital robne marke osigurava bolju cijenu, a time i bolji prihod tvrtki. Postoje brojni načini na koje se izračunava vrijednost marke. Provedeni su brojni načini i opsežna istraživanja kako bi se razumio sam koncept vrijednosti marke.

Na temelju brojnih istraživanja utvrđeno je da je marka jedna od najvažnijih nematerijalnih dobara poduzeća koja pomaže u poboljšanju financijskog poslovanja poduzeća. Na temelju različitih modela jednakosti robne marke, vidljivo je da kapital marke određuju različite komponente kao što su kvaliteta, performanse, svijest o marki i odanost.

Modeli kapitala robne marke osmišljeni su kako bi ustanovili način na koji se stvara vrijednost marke. Svaki od modela kapitala marke nudi dubok uvid u koncept vrijednosti marke i načine njegove procjene. Modeli vrijednosnog kapitala koriste se za osmišljavanje marketinških strategija u različitim fazama. Neke od značajnih radnji koje se mogu poduzeti korištenjem ovih udjela u model robne marke su poboljšanje percepcije proizvoda, stjecanje lojalnijih kupaca, postizanje konkurentne prednosti itd.

U daljnjem tekstu ukratko su opisane osnovne razlika između dva najčešća modela vrijednosnog kapitala marke, a to su Aakerov model i Kellerov model.

#### *2.3.1.1. Aakerov model*

Aakerov model kapitala marke (Aaker, 1996) okvir je koji pomaže razumjeti vrijednost marke. Aakerov model vrijednosti robne marke proširuje izbor kupca. Kupac će biti voljan povezati se s markom koja nudi višu kvalitetu, a samim time i veće zadovoljstvo. Prema Aakeru, vrijednost marke određuje pet komponenti:

1. *Odanost robnoj marki* - odnosi se na razinu lojalnosti kupaca i predanosti marki. Vjerojatnije je da će kupci koji su lojalni robnoj marki nastaviti s kupnjom te marke i preporučiti je drugima.
2. *Svijest o robnoj marki* - odnosi se na mjeru u kojoj su kupci upoznati s markom i mogu je prepoznati. Visoka svijest o marki može kupcima olakšati pamćenje i odabir marke.
3. *Percipirana kvaliteta* - odnosi na percepciju kupca o kvaliteti proizvoda ili usluga vezanih uz marku. Kupci mogu marku smatrati visokokvalitetnom ako dosljedno isporučuje proizvode ili usluge koji ispunjavaju ili premašuju njihova očekivanja.
4. *Ostale asocijacije* - odnosi se na percepciju kupaca o brendu na temelju njegovih atributa, kao što su njegove vrijednosti, osobnost i druge karakteristike.

5. *Ostala vlasnička imovina marke* - uključuje stvari poput zaštitnih znakova, patenata i drugog intelektualnog vlasništva koje marki daje konkurentsku prednost.

Značaj Aakerovog modela u najvećoj mjeri je utemeljena na potrošaču budući da će kupac biti voljan povezati se s markom koja nudi kvalitetu. Ukoliko se dosljedno slijedi Aakerov model kupci će iskazati svoju lojalnost kroz nastavak kupnje ali i kroz daljnju preporuku drugima.

#### 2.3.1.2. *Kellerov model*

Kellerov model vrijednosti robne marke posebno se fokusira na perspektivu potrošača. Prema Kelleru (Keller, 2013), postoje tri ključne komponente kapitala marke:

1. *Istaknutost marke* - odnosi se na mjeru do koje marka pada na pamet kada kupci donose odluku o kupnji. Kupci će vjerojatnije uzeti u obzir marke koje su istaknutije.
2. *Izvedba marke* - uključuje kupčevu procjenu marke na temelju njezinih funkcionalnih atributa i prednosti koje pruža.
3. *Imidž marke* - predstavlja percepciju kupca o marki na temelju njegovih psiholoških i simboličkih atributa, kao što su njegove vrijednosti, osobnost i druge karakteristike.

Kellerov (Keller, 2013, str. 108-112) model vrijednosti marke može se predstaviti i kroz dobivanje odgovora na četiri pitanja čiji odgovori su prikazani kao smjernice za izgradnju i mjerenje kapitala marke:

1. *Tko si ti?* - odnosi se na izgradnju snažnog identiteta brenda te stvaranje svijesti o samom brendu. Ako ljudi nisu čuli ili vidjeli proizvod, taj proizvod je teško onda prodati. Važno je uvijek poznavati kupce kojima se obraćamo i znati njihova očekivanja od marke budući da tako gradimo identitet robne marke. Brend bi se trebao isticati od konkurencije kako bi kupci mogli prepoznati marku.
2. *Što si ti?* - sljedeći korak odnosi se na komunikaciju s krajnjim korisnicima a to su kupci o tome što brend znači i čemu služi. U ovom koraku bi trebali objasniti karakteristike proizvoda. Što znači da brend treba biti pouzdan, nuditi dobru uslugu, imati dobar dizajn te nuditi razumnu cijenu. Važno je objasniti kako brend može zadovoljiti potrebe kupaca te ih potaknuti da se povežu s brendom na psihološkoj razini. To se može postići korištenjem raznih marketinških strategija kao što su: izravna promocija, dijeljenje korisničkih iskustava.
3. *Što ja mislim o tebi?* - u ovoj fazi dobiva se odgovor marke. Reakcija marke može biti ili osjećaj ili prosudba o samom proizvodu. Potrošači uvijek imaju osjećaj ili prosudbu o proizvodu kojeg koriste. Kada proizvod ispuni očekivanja korisnika, on izaziva pozitivan osjećaj o našem brendu. Proizvod mora biti privlačan, zadovoljiti potrebe potrošača i biti jedinstven u usporedbi s proizvodima konkurencije.

4. Koje je vaše mišljenje? – odnosi se na interakciju gdje se jača odnos između samog brenda i kupca. Odaziv marke koji je došao iz ranije faze sada se pretvara u intenzivnu i emocionalnu vezu između marke i kupca. Ovo je ujedno i završna faza koju je najteže postići budući da treba postići da kupac bude u dobrom odnosu s robnom markom što onda rezultira ponovljenom kupnjom tog proizvoda i stvaranjem osjećaja lojalnosti kod kupca.

### 2.3.1.3. Mjerenje vrijednosti marke

Postoji niz različitih metoda koje tvrtke mogu koristiti za mjerenje vrijednosti svojih marki (Heding, Knudtzen & Bjerre, 2009). Neke od najčešćih metoda uključuju:

1. *Tržišni pristup* - ova metoda uključuju procjenu vrijednosti marke na temelju njezinih tržišnih performansi, kao što je udio na tržištu ili cijena koju su kupci spremni platiti za proizvode te marke. Detaljnije, ova metoda uključuje procjenu financijske vrijednosti marke. Razmatra različite čimbenike kao što su svijest o robnoj marki, odanost robnoj marki, tržišni udio i financijski učinak kako bi se izračunala vrijednost robne marke. Interbrand, BrandZ i Forbes neke su od popularnih tvrtki za procjenu robnih marki. Kao primjer istaknimo vrijednost marke Adidas od \$15,9 milijardi prema Interbrand ocjeni 100 najznačajnijih marki za 2022. Prema toj listi marka Adidas je na 42 mjestu što je poboljšanje od 9 mjesta s obzirom na 2021. godinu (Interbrand 2022).
2. *Troškovno utemeljen pristup* – metoda koje uključuju procjenu vrijednosti marke na temelju troškova povezanih s njezinim stvaranjem i održavanjem, kao što su troškovi marketinga i oglašavanja.
3. *Prihodovno utemeljeni pristup* - metoda koje uključuju procjenu vrijednosti marke na temelju prihoda ili dobiti koji generira.
4. *Pristup utemeljen na potrošačima* – metoda koja uključuje procjenu vrijednosti marke na temelju stavova i ponašanja potrošača, kao što je njihova lojalnost marki i spremnost da preporuče marku drugima. Ova metrika pokazuje u kojoj su mjeri kupci voljni nastaviti poslovati s markom. Može se mjeriti kroz ponovljene kupnje, stope zadržavanja kupaca ili dugotrajnu vrijednost kupca. Primjer često korištene metrike je Net Promoter Score (NPS) koja određuje vjerojatnost da kupci preporuče marku svojim prijateljima ili obitelji. Omogućuje pokazatelj lojalnosti marki i zadovoljstva kupaca.
5. *Metrika angažmana na društvenim medijima* - mjeri prisutnost marke na platformama društvenih medija i opseg u kojem korisnici stupaju u interakciju sa sadržajem brenda. Mjerni podaci kao što su lajkovi, dijeljenja, komentari i pratitelji koriste se za mjerenje angažmana na društvenim mrežama. Na primjer što se tiče angažmana marke Adidas na

društvenim mrežama ističe se snažna prisutnost s preko 27,2 milijuna pratitelja na Instagramu ( Instagram Stats & Analytics for adidas, 2023).

Konkretna metrika i metode koje se koriste mogu varirati ovisno o industriji i specifičnim ciljevima marke. Također je važno napomenuti da vrijednost marke nije statičan koncept, a vrijednost marke može se mijenjati tijekom vremena zbog niza čimbenika, kao što su promjene u preferencijama potrošača, tržišnim uvjetima i izvedbi marke.

Ovo je samo nekoliko primjera metoda za mjerenje vrijednosti i kapitala marke. Konkretna metrika i metode koje se koriste mogu varirati ovisno o industriji i specifičnim ciljevima marke.

Ovo je samo nekoliko primjera metoda za mjerenje vrijednosti i kapitala marke. Konkretna metrika i metode koje se koriste mogu varirati ovisno o industriji i specifičnim ciljevima marke.

### 2.3.2. *Imidž i prepoznatljivost marke*

Imidž robne marke sastavni je dio vrijednosti robne marke jer potrošačima prenosi vrijednost marke. Kotler (2001. str. 273) definira imidž kao "skup uvjerenja, ideja i dojmova koje osoba ima o objektu". Dakle, kada govorimo o imidžu marke, radi se o mentalnom predstavljanju brenda na temelju potrošačeve percepcije, ideje i dojma.

Imidž marke određuje očekivanja koja stvaramo od određenog brenda. Jedan od primjera za to može biti Porche, on ima imidž proizvođača luksuznih automobila, što znači da se ne može od te marke očekivati da proizvedu obiteljski karavan, jer to nije u skladu s percepcijom i očekivanjima koje kupci imaju o Porcheu.

Imidž marke odnosi se na cjelokupnu percepciju koju kupci imaju o marki, na temelju njihovih iskustava i interakcija s njom (Keller, 2013, str.76-78) obuhvaća emocionalne i psihološke asocijacije koje kupci stvaraju s brendom, uključujući njegovu osobnost, vrijednosti i ugled. Oblikuju ga različiti čimbenici, uključujući proizvode ili usluge marke, marketinške komunikacije i osobna iskustva kupaca s markom. Snažan imidž marke može pomoći u izgradnji lojalnosti kupaca i povećati percipiranu vrijednost proizvoda ili usluga marke. Prepoznatljivost robne marke, s druge strane, odnosi se na sposobnost kupaca da identificiraju i povežu marku s njenim imenom, logotipom ili drugim vizualnim elementima. Prepoznatljivost marke je ono što razlikuje neki proizvod ili uslugu od konkurencije. To je kontinuirani proces u kojem se kupac redovito izlaže pozitivnim karakteristikama marke, tako da se kod kupca stvara određena percepcija ili slika o brendu. To je mjera koliko lako i brzo kupci mogu prepoznati marku kada je vide, čak i ako nisu imali prethodnog iskustva s njom.

I imidž marke i prepoznatljivost marke važni su aspekti upravljanja markom. Snažan imidž marke može pomoći u stvaranju pozitivnog i nezaboravnog dojma o marki, što može dovesti

do povećane lojalnosti kupaca i zagovaranja marke. Učinkovito prepoznavanje robne marke, s druge strane, može pomoći u povećanju svijesti o robnoj marki i olakšati kupcima prepoznavanje i odabir određene marke u odnosu na konkurente. Marketinški stručnjaci rade na razvoju i održavanju imidža i prepoznatljivosti robne marke kroz različite marketinške napore i napore u markiranju, kao što su oglašavanje, odnosi s javnošću i društveni mediji. Nastoje stvoriti dosljedan i kohezivan identitet marke koji odjekuje kod kupaca i osnažuje vrijednosti marke i jedinstvenu prodajnu ponudu (Keller, 2003). Imidž marke mogu oblikovati različiti čimbenici koji će biti argumentirani s primjerima robne marke Coca - Cola:

1. *Proizvodi ili usluge* - kvaliteta i karakteristike proizvoda ili usluga marke mogu imati značajan utjecaj na imidž marke. Kupci mogu marku doživjeti kao visokokvalitetnu ako dosljedno isporučuje proizvode ili usluge koji ispunjavaju ili čak premašuju njihova očekivanja. Na primjer Coca-Cola je poznata po svom prepoznatljivom gaziranom bezalkoholnom piću, koje ima prepoznatljiv okus i široko je dostupno diljem svijeta. Tvrtka je također proširila svoju liniju proizvoda kako bi uključila niz drugih pića, poput sokova, vode i sportskih napitaka.

2. *Marketinške komunikacije* - poruke i slike koje se koriste u marketinškim komunikacijama robne marke mogu pomoći u oblikovanju njezine slike u svijesti kupaca. Svima znamo da je Coca-Cola dobro poznata po svojim marketinškim kampanjama koje često sadrže privlačne slogane i nezaboravne slike. Neki primjeri marketinških kampanja Coca-Cole uključuju "*The Coca-Cola Side of Life*", "*Open Happiness*" i "*Taste the Feeling*".

3. *Iskustva kupaca* - na imidž robne marke mogu utjecati osobna iskustva kupca s markom, uključujući interakcije s korisničkom službom, fizičko okruženje trgovina ili lokacija marke i cjelokupno korisničko iskustvo. Primjerice Coca-Cola je dostupna na velikom broju lokacija, uključujući automate za prodaju, trgovine i supermarkete, što je čini lako dostupnom kupcima. Tvrtka također sponzorira događaje i sponzorstva, poput glazbenih festivala i sportskih događaja, koji mogu poboljšati korisničko iskustvo.

4. *Javna percepcija* - na imidž marke također može utjecati njegov ugled i javna percepcija, koja se može oblikovati medijskom pokrivenošću, usmenom predajom i drugim oblicima komunikacije. Tako nam je poznato da Coca-Cola ima jaku reputaciju i široko je prepoznata diljem svijeta. Tvrtka ima dugu povijest filantropije i često se povezuje s pozitivnim vrijednostima kao što su sreća i zajedništvo.

5. *Kulturni kontekst* - na imidž marke može utjecati kulturni kontekst u kojem ona djeluje, budući da različite kulture mogu imati različite asocijacije i percepcije marke. Zbog svoje globalne prisutnosti Coca-Cola je prepoznata u mnogim različitim kulturama diljem svijeta. U

nekim kulturama Coca-Cola se može povezivati sa slavljinama i posebnim prilikama, dok se u drugima na nju može gledati kao na svakodnevno piće.

6. *Konkurenti* - imidž robne marke može se oblikovati u odnosu na njezinu konkurenciju, jer kupci mogu usporediti marku s drugim opcijama na tržištu. Znamo da Coca-Cola ima niz konkurenata na tržištu gaziranih bezalkoholnih pića, uključujući Pepsi i neke druge manje brendove. Na imidž marke Coca-Cola može utjecati način na koji se ona doživljava u odnosu na ove konkurente.

Iz primjera koji su istaknuti u prethodnim nabrajanjima vidi se kako Coca-Cola gradi svoju prepoznatljivost s ciljem isticanja razlika u odnosu na njenu konkurenciju.

## **2.4. Izgradnja marke**

Izgradnja marke je proces stvaranja jedinstvenog imena, identiteta i imidža za proizvod ili uslugu u svijesti kupaca. Uključuje stvaranje snažne, pozitivne percepcije marke i njezino razlikovanje od konkurenata. Učinkovita izgradnja robne marke može pomoći poduzeću da se pozicionira na tržištu, izgradi lojalnost kupaca i poveća vrijednost svojih proizvoda ili usluga (Aaker 1996, Keller 2009). U nastavku se navode neki ključni elementi izgradnje robne marke:

1. *Razvijanje strategije marke* uključuje definiranje svrhe marke, ciljne publike, vrijednosti i pozicioniranja na tržištu.
2. *Stvaranje identiteta marke* uključuje vizualne elemente marke, kao što su logo, shema boja i tipografija, kao i ton glasa i poruke marke.
3. *Izgradnja svijesti o robnoj marki* podrazumijeva predstavljanje robne marke što je moguće većem broju ljudi putem marketinga i oglašavanja, odnosa s javnošću i društvenih mreža.
4. *Stvaranje dosljednog iskustva marke* osigurava da je svaka dodirna točka s markom, od samog proizvoda ili usluge do iskustva korisničke službe, u skladu s identitetom i vrijednostima marke.
5. *Mjerenje i procjena izvedbe marke* uključuje praćenje metrika kao što su svijest o marki, lojalnost kupaca i prodaja kako bi se ocijenila učinkovitost napora za izgradnju marke.

Želja svakog poduzeća, odnosno vlasnika marke da kroz izgradnju marke osigurati čvrsti temelj za ostvarivanje poslovne strategije kompanije. Prije svega marka se mora usmjeriti na izgradnju svijesti kao temeljni cilj kojemu je potrebno dati relevantnu prednost. Proces izgradnje marke

je proces koji opravdava ovaj atribut putem širenja svijesti o konstantnim prednostima korištenja određenih proizvoda ili usluga široj populaciji.

## 2.5. Upravljanje markom

Upravljanje markom je proces izgradnje, mjerenja i upravljanja vrijednošću marke. Uključuje stvaranje snažnog identiteta marke, pozicioniranje marke na tržištu i prenošenje poruke marke kupcima (Keller 2013, Kotler, Pfoertsch, 2009). Jaka robna marka može pomoći tvrtki da se razlikuje od konkurencije, izgradi lojalnost kupaca i poveća životnu vrijednost kupaca (Kotler, Keller, 2009). Neke od ključnih aktivnosti uključenih u upravljanje markom uključuju:

- *Razvoj strategije marke* - podrazumijeva identificiranje ciljnog tržišta, određivanje pozicioniranja marke na tržištu i definiranje poruka i osobnosti marke.
- *Stvaranje identiteta marke* - uključuje razvoj vizualnih elemenata kao što su logotip i smjernice marke, kao i definiranje tona glasa i poruka marke.
- *Komuniciranje poruke robne marke* – realizira se kroz korištenje različitih kanala, kao što su oglašavanje, društveni mediji i odnosi s javnošću, za promicanje marke i dopiranje do kupaca.
- *Upravljanje reputacijom marke* – ostvaruje se praćenjem i upravljanjem povratnim informacijama kupaca, odgovaranje na pritužbe kupaca i poduzimanje koraka za rješavanje bilo kakvih problema koji mogu utjecati na reputaciju marke.
- *Mjerenje učinka marke* – realizira se kroz praćenje ključnih metrika kao što su svijest o marki, zadovoljstvo kupaca i prodaja kako bi se razumjelo kako marka funkcionira i identificirala područja za poboljšanje.

Može se istaknuti primjer upravljanja robnom markom Coca - Cola: Coca-Cola ima ikonski imidž robne marke koji je prepoznat u cijelom svijetu. Tvrtka ulaže velika sredstva u upravljanje markom kako bi održala ovaj imidž i održala svoju marku relevantnom. Coca-Cola koristi različite marketinške strategije, kao što su oglašavanje, društveni mediji i sponzorstva, kako bi izgradila svijest o robnoj marki i uključila se u svoju ciljanu publiku.

### 3. ANALIZA ZNAČAJA UPRAVLJANJA MARKOM

Upravljanje markom odnosi se na proces stvaranja, razvoja, održavanja i povećanja ugleda marke na tržištu. Detaljnije o tome napisano je u točki 2. ovog rada, pa se podsjetimo nekih važnijih činjenica. Učinkovito upravljanje robnom markom uključuje niz aktivnosti poput stvaranja identiteta robne marke, pozicioniranja robne marke na tržištu, osiguravanja dosljednosti u porukama robne marke i iskorištavanja vrijednosti robne marke za poticanje rasta i profitabilnosti. Općenito, upravljanje robnom markom ključno je za stvaranje snažne marke koja može potaknuti lojalnost kupaca, generirati veći profit i pružiti održivu konkurentsku prednost. Važnost upravljanja markom može se sagledati kroz sljedeće elemente:

- odanost kupaca (pozitivna percepcija robne marke povećava vjerojatnost da će joj kupci ostati lojalni i preporučiti je drugima),
- vrijednost marke (jaka robna marka može zahtijevati vrhunsku cijenu, povećati lojalnost kupaca i generirati veći profit),
- konkurentsku prednost (snažna marka može pomoći poduzeću da se istakne na pretrpanom tržištu i privuče nove kupce),
- upravljanje reputacijom ( ključno je u današnjem digitalnom dobu. Jaka robna marka može pomoći u izgradnji pozitivne reputacije, dok negativna reputacija može naštetiti imidžu robne marke i utjecati na njen konačni rezultat ) i
- inovacije (praćenjem trendova i razumijevanjem potreba kupaca, tvrtke mogu osigurati da njihov brend ostane relevantan i konkurentan).

Temelje iznesenih činjenica vezanih uz upravljanje markom postavili su Aaker i Keller devedesetih godina što je diskutirano u točki 2. i u današnje doba ih marketinški stručnjaci koriste za razvoj brendova. No kako je digitalna tranzicija zahvatila sve segmente društva zbog naglog razvoja interneta i društvenih medija došlo je do značajnih pomaka u marketinškim komunikacijama između tvrtki i potrošača. Iz tih razloga u svakodnevicu ušle metode koje se temelje na analizi teksta kao povratna informacija od strane potrošača (mišljenja i ocjene brendova na društvenim mrežama i blogovima). S obzirom na veliku količinu informacija u tekstualnim sadržajima na internetu, analizu teksta nije više moguće provoditi „pješke“ od strane ljudi već se za te namjene moraju koristiti računala. Zato se zadnjih desetak godina intenzivno razvijaju metode utemeljene na obradi prirodnog jezika ( **eng.** *Natural Language Processing – NLP*) kao grane umjetne inteligencije (Beysolow II, 2018).



### **3.1. Teorijski aspekti upravljanja markom**

Upravljanje robnom markom kritični je aspekt marketinga koji se fokusira na stvaranje i održavanje pozitivnog imidža robne marke u svijesti potrošača. Dva popularna modela upravljanja markom su Aakerov model identiteta marke i Kellerov model vrijednosti marke. Ovi modeli pružaju teoretski okvir za razumijevanje različitih komponenti marke i kako se one mogu iskoristiti za stvaranje snažnog identiteta marke i lojalnosti kupaca.

Nedavno su metode obrade prirodnog jezika (NLP) stekle popularnost u upravljanju robnom markom zbog svoje sposobnosti da analiziraju velike količine podataka i izvuku značajne uvide o ponašanju i percepcijama potrošača. NLP tehnike kao što su analiza sentimenta i tematsko modeliranje mogu se koristiti za dublje razumijevanje stavova potrošača prema marki, njenim proizvodima ili uslugama i njihovim marketinškim naporima.

Dok se Aakerovi i Kellerovi modeli usredotočuju na unutarnje i vanjske komponente marke, NLP metode pružaju pristup upravljanju markom koji se više temelji na podacima koji mogu nadopuniti i poboljšati te modele. Analizirajući recenzije potrošača, objave na društvenim mrežama i druge izvore povratnih informacija kupaca, NLP može pružiti uvid u učinkovitost marketinških napora brenda, kvalitetu njegovih proizvoda ili usluga i cjelokupni osjećaj prema brendu. NLP metode mogu se koristiti za analizu komunikacije potrošača na društvenim mrežama i izvođenje metrika za mjerenje ključnih vrijednosti marke, kao što su svijest o marki, sentiment i angažman.

Zaključno, dok Aakerovi i Kellerovi modeli upravljanja robnom markom pružaju teorijsku osnovu za razumijevanje komponenti marke i kako se one mogu iskoristiti za stvaranje lojalnosti kupaca, NLP metode nude pristup upravljanju robnom markom koji se više temelji na podacima koji mogu pružiti dublje uvide u ponašanje i percepciju potrošača i javnost. Kombinirajući ove pristupe, marketinški stručnjaci mogu razviti sveobuhvatnije razumijevanje svoje robne marke i njenog utjecaja na potrošače.

### **3.2. Potreba za mjerenjem vrijednosti marke**

Mjerenje vrijednosti marke važno je jer omogućuje tvrtkama da procijene učinak svojih napora u brendiranju i donesu odluke o budućim ulaganjima u upravljanje robnom markom na temelju podataka. Općenito, mjerenje vrijednosti brenda ključno je za razumijevanje utjecaja napora brendiranja na razvoj tvrtke i donošenje afirmativnih odluka o budućim ulaganjima u upravljanje brendom (Aaker 1991, Keller 2013).

Korištenjem niza metrika za procjenu vrijednosti brenda, tvrtke mogu prepoznati područja za poboljšanje i s vremenom izgraditi jače i vrijednije marke. Vrijednost marke sveobuhvatna je mjera vrijednosti koju marka donosi tvrtki. Uključuje mjere kao što su svijest o robnoj marki, lojalnost robnoj marki, povezanost s robnom markom i financijski učinak. Izračun vrijednosti marke uključuje analizu niza izvora podataka i korištenje specijaliziranih modela za procjenu vrijednosti marke.

Tvrtke koriste različite metrike za mjerenje vrijednosti marke, pa tako svijest o marki radi spoznaje u kojoj su mjeri potrošači upoznati s markom mjere provođenjem anketa ili korištenjem alata kao što je analitika društvenih mreža za praćenje spominjanja robne marke i angažmana. Visoka razina svijesti o robnoj marki ukazuje na to da je marka najvažnija za potrošače, što može dovesti do povećanja prodaje i lojalnosti kupaca.

Za utvrđivanje odanosti marke tvrtke mjere stupanj do kojeg su potrošači privrženi određenom brendu. To se može mjeriti provođenjem anketa ili praćenjem ponašanja kupaca kao što su ponovljene kupnje ili preporuke. Visoke razine lojalnosti brendu pokazuju da je brend uspješno izgradio emocionalnu vezu sa svojim kupcima, što može dovesti do dugoročne profitabilnosti. Također je potrebno utvrditi asocijacije na robnu marku koje se odnose se na mentalne veze koje potrošači uspostavljaju između marke i određenih atributa, kao što su kvaliteta, pouzdanost ili inovativnost. Te se povezanosti mogu mjeriti provođenjem anketa ili korištenjem alata kao što je analiza teksta za analizu povratnih informacija kupaca. Pozitivne asocijacije na marku mogu pomoći u razlikovanju robne marke od konkurencije i potaknuti lojalnost kupaca.

Ne manje važan je financijska učinak koji mjeri utjecaj marke na konačni rezultat tvrtke. To se može mjeriti analizom podataka o prodaji, tržišnog udjela ili profitabilnosti. Jaka robna marka može pomoći u održavanju vrhunskih cijena, povećanju prodaje i poboljšanju profitabilnosti.

Pored prethodno navedenih tradicionalnih metoda mjerenja vrijednosti robne marke, koje se redovito koriste u marketinškim operacijama vezanim uz marku, koristit će se i suvremene metode koje na osnovu analize tekstualnih sadržaja dostupnih na internetu primjenom NLP definiraju novu metriku za određivanje vrijednosti marke. U pravilu se danas ta dva pristupa kombiniraju s ciljem što bolje procjene učinka brendiranja.

### **3.3. Prikaz metoda analize značaja marke**

U nastavku će se prikazati mogućnosti metoda analize značaja marke utemeljene na analizi teksta i konteksta primjenom obrade prirodnog jezika (NLP) tekstualnih sadržaja dostupnih na internetu. Danas su tekstualni podaci sveprisutni i često im se slobodno pristupa s

više izvora kao što su dobro poznate platforme društvenih medija Facebook i Twitter, blogovi, tematski forumi poput TripAdvisora i tradicionalni mediji kao što su novinske vijesti. Potrošači izražavaju svoje osjećaje i mišljenja u vezi s proizvodom ili uslugom na više načina, a njihov odnos prema markama često se može izdvojiti iz komunikacije na društvenim mrežama. Interakcije potrošača kako međusobne tako i s tvrtkama mogu utjecati na pridobivanje potencijalnih potrošača te učinak na buduće proizvode koje će tvrtke razvijati.

Nagli bum tekstualnih podataka (Big Data) potaknuo je razvoj novih automatiziranih pristupa kako bi se mogli analizirati veliki tekstualni korpusi i iz njih izvlačiti značenje (Colladon, 2018). Istodobno, postoji značajan interes za analizu vrijednosti i važnosti marki, s obzirom na tvrtku i potrošače. Kapital marke utemeljena na potrošačima definirao je Keller (1993) kao „diferencijalni učinak poznavanje robne marke i reakcije potrošača na marketing marke“. Autor je također postavio imidž marke i svijest kao dvije glavne dimenzije znanja o marki. Ove dimenzije su u mnogim slučajevima procjenjuju pomoću anketa, studija slučaja, intervjua i/ili fokusnih grupa. U ovom radu primijenit će se metode NLP-a na analizu i promišljanje teksta. Naravno svaka analiza teksta podrazumijeva i razumijevanje konteksta odnosno jednom riječi analiza teksta koja će se koristiti odnosi se na semantiku, a manje na sintaksu.

### 3.3.1. Analiza teksta i konteksta

Analize teksta predstavlja razumijevanje kulturnih, ekonomskih, socijalnih ili ideoloških aspekata sadržanih u suštini nekog teksta i kako su oni povezani s određenim kontekstom. Analiza teksta podrazumijeva opisivanje karakteristika teksta i interpretaciju u cilju dobivanja odgovora na neka konkretna pitanja. Potrebno je naglasiti da tekstualni sadržaji spadaju skupinu takozvanih nestrukturiranih podataka, za razliku od naprimjer podataka koje možemo predstaviti tablicama (baze podataka, Excel datoteke), zahtijevaju sasvim drugi pristup.

U današnje doba ekstremne količine tekstualnih podataka, potrebno je objediniti razna znanja i vještine koje će omogućiti transformaciju složenih i velikih nestrukturiranih podataka (big data) u korisne i upravljive informacije. Zbog toga je proces analize sadržaja teksta nezamisliv bez pomoći računalnih sustava i alata koji omogućuju automatizaciju procesa, pojednostavljaju promatranje obrazaca i kategorizaciju sadržaja prema prethodno definiranim kategorijama. Ta metodologija omogućuje izdvajanje različitih značajki sadržaja kao što su imena, zemlje, e-mail adrese, autori, robne marke i drugo. S druge strane analizu teksta može se gledati kroz prizmu promišljanja teksta (**eng. text mining**) utemeljenoj na interdisciplinarnoj metodologiji

pronalaženja informacija koristeći strojno učenje, statističke metode i računalnu lingvistiku (Sarkar, 2016).

Komunikacija tekstem između komunikatora vrlo često se ostvaruje kontekstom koji se može razvrstati na ko-tekst ili suradnički tekst, situacijski kontekst i kulturološki kontekst (Lihong, 2012) i na tu temu lingvisti su napravili mnoštvo različitih istraživanja o odnosu teksta i konteksta.

Tekst se može opisati modelom kojeg čine osnovne leksičke jedinice riječi i interpunkcijski znakovi. Za označavanje tih jedinica koristi se izraz token tekstualnih jedinica. Tokeni se u nekom tekstu nalaze u određenom redosljedom pa se tekst može modelirati pomoću konačnog niza svih riječi i interpunkcijskih znakova. U tekstovima ima više različitih informacija strukturnog karaktera kao što su naslovi, odlomci i drugo, ali i značajke prikaza po kojima se razlikuju jedni dijelovi teksta od drugih.

Cjelinu tekstualnog modela uvijek se može postaviti u nekakav kontekst koji pobliže određuje smisao teksta. Kontekstualna analiza se može definirati određenim pravilima koja omogućuju identificiranje i po potrebi označavanje nekih dijelova koji prethode nekom pojmu (**eng. left context**) ili slijede nakon njega (**eng. right context**). Ako se stvar pojednostavi često neki pojam označen kao imenica uz sebe veže neki pridjev koju mu daje odgovarajući kontekst. Elementi kontekstualnih pravila mogu biti pozitivni i negativni pa to omogućuje analizu sentimenta vezanog za neki pojam, događaj ili brand (Wonsever, Minel, 2001).

Glavni problem koji se susreće kod primjene NLP metoda i strojnog učenja za analizu teksta je što računala za razliku od čovjeka ne mogu računati s riječima već samo brojevima. Iz tog razloga prije bilo kakve analize teksta nakon pred obrade sadržaja, riječi teksta moraju se pretvoriti u brojeve koji se u pravilu predstavljaju višedimenzionalnim vektorima. Taj proces pretvaranja riječi u brojeve NLP eksperti razvili su tehnike koje se zovu ugradnja riječi (**eng. word embedding**) koje riječi pretvaraju u numeričku reprezentaciju. Ugradnjom riječi mapiraju se riječi kao numerički vektori stvarnih vrijednosti. To se čini označavanjem svake riječi u nizu (ili rečenici) i pretvaranjem u vektorski prostor. Cilj ugrađivanja riječi je odrediti semantičko značenje riječi u nizu teksta. To se ostvaruje dodjeljivanjem sličnog numeričkog prikaza riječima koje imaju slična značenja. Navedimo neke algoritme za ugradnju riječi koje se danas koriste (Beysolow II, 2018). Neke od njih će biti detaljnije opisane u nastavku jer se koriste u ovom radu.

- TF-IDF - je algoritam strojnog učenja (ML) koji se temelji na statističkoj mjeri pronalaženja relevantnosti riječi u tekstu. Tekst može biti u obliku dokumenta (paragrafi, rečenice i sl.) ili skupa raznih dokumenata (korpus). Ovaj algoritam je

kombinacija dviju metrika: Frekvencija termina (TF) i Inverzna frekvencija dokumenta (IDF).

- Word2Vec algoritam - namijenjen je za određivanje semantičkih informacija. Word2Vec je tehnika ugrađivanja riječi za rješavanje naprednih NLP problema. Može iterirati preko velikog korpusa teksta kako bi naučio asocijacije ili ovisnosti među riječima. Word2Vec pronalazi sličnosti među riječima pomoću metrike kosinusne sličnosti. Ako je kut kosinusa 1, to znači da se riječi preklapaju. Ako je kut kosinusa 90, to znači da su riječi neovisne ili nemaju kontekstualne sličnosti. Sličnim riječima dodjeljuje slične vektorske reprezentacije. Word2Vec nudi dvije varijante temeljene na neuralnim mrežama: Continuous Bag of Words (CBOW) i Skip-gram. U CBOW, model neuronske mreže uzima različite riječi kao ulaz i predviđa ciljnu riječ koja je usko povezana s kontekstom ulaznih riječi. S druge strane, Skip-gram arhitektura uzima jednu riječ kao ulaz i predviđa njezine blisko povezane kontekstualne riječi.
- GloVe proširuje rad tehnike Word2Vec za određivanje globalnih kontekstualnih informacija u tekstualnom korpusu izračunavanjem globalne matrice zajedničkog pojavljivanja riječi i riječi. Word2Vec određuje samo lokalni kontekst riječi tijekom strojnog učenja i uzima u obzir samo susjedne riječi kako bi uhvatio kontekst. GloVe razmatra cijeli korpus i stvara veliku matricu koja može odrediti istovremeno pojavljivanje riječi unutar korpusa.
- BERT - prikazi dvosmjernog kodera iz transformatora. Google ga je predstavio 2019. BERT pripada klasi jezičnih algoritama temeljenih na NLP-u poznatih kao transformeri. BERT je masivan model transformera temeljen na dubinskom dvosmjernom enkoderu koji je unaprijed obučen i dolazi u dvije varijante. BERT-Base ima 110 milijuna parametara, a BERT-Large 340 milijuna parametara. Na njegovim temeljima danas je razvije puno novih jezičnih modela pod nazivom Large Language Models (LLM).

U nastavku će biti opisane 4 NLP tehnike analize teksta i konteksta koje su korištene nad istraživačkim tekstualnim korpusom kojeg čine tekstovi skinuti s različitih web stranica vezanih uz sportsku obuču. Također će biti ukratko opisane programske aplikacije za provođenje eksperimenata navedenih NLP tehnika.

- NLP tehnika za frekvencijsko modeliranje
- NLP tehnika za tematsko modeliranje

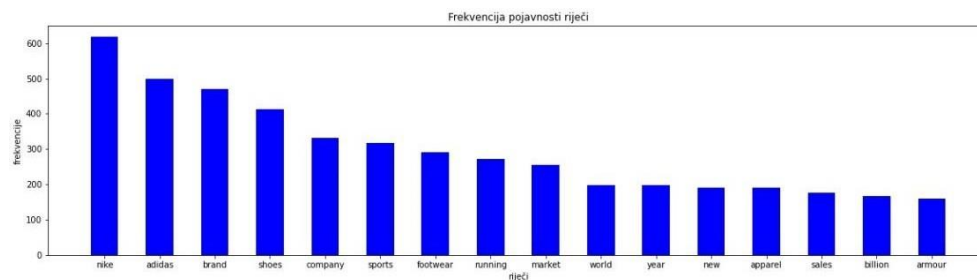
- Tehnika semantičkog modeliranja i ocjena značaja robnih marki (*Semantic Brand Scores*)
- Analiza sentimenta potrošača određene robne marke

Prije nego što se započne analiza teksta navedenim tehnikama potrebno je tekstualni korpus koji se analizira pripremiti primjenom metoda pred obrade teksta (**eng. preprocessing**). Sirove tekstualne sadržaje prije bilo kakve analize nužno je pripremiti i obraditi na način koji će biti razumljiv računalnim algoritmima koji moraju izvlačiti ciljane značajke koje se mogu dobiti iz teksta (Nitin, 2015). Najznačajniji koraci pred obrade teksta su: filtriranje sadržaja, izbacivanje interpunkcija i numeričkih podataka, pretvaranje svih riječi (tokena) u riječi s malim slovima, normalizacija koja uključuje micanje sufiksa i prefiksa iz korijena riječi (**eng. stemming**) i preslikavanje riječi na korijenski oblik (**eng. lemmatization**) i na kraju izbacivanje zaustavnih riječi koje nemaju neko bitno značenje u tekstu (**eng. stopwords**). Na primjer kada izbacimo zaustavne riječi iz teksta nestaju riječi kao što su 'for' i 'this', 'I', 'am', 'the', 'a' i sl.. Nakon izbacivanja sufiksa neke riječi se svode na osnovni oblik npr. 'personal' se pretvara u 'person', a ako bi u tekstu imali još riječi sa sličnim korijenom sve bi one bile pretvorene u 'person' i u određivanju frekvencije pojavnosti sve se broje kao jedna teista riječ. Nešto detaljnije o pred obradi teksta prikazat će se u opisu rezultata analize teksta.

### 3.3.2. NLP metode određivanja frekvencije pojavnosti riječi

Nakon što je izvršena pred obrada tekstualnog korpusa određuje se frekvencija pojavnosti riječi u korpusu. Ova analiza može nam pokazati koje riječi su najzastupljenije u analiziranom korpusu kako to pokazuje primjer na slici 1 što može biti određena informacija o čemu se u analiziranim dokumentima radi ali ne i dovoljna ako je istraživanje usmjereno na pojavnost naziva robnih marki za koje iz analize teksta želimo utvrditi koliko često se te riječi pojavljuju u velikom tekstualnom korpusu. Iako se u tekstu pojavljuju dvije marke, među prvih 16 riječi nema drugih marki već se pojavljuju riječi vezane uz sportsku obuću općenito. Zbog toga se umjesto određivanja frekvencije riječi jednostavnim brojanjem koriste druge tehnike koje daju puno bolju informaciji koliko su neke riječi zastupljene u analiziranom tekstualnom korpusu. Za tu tehniku razvijen je klasifikacijski algoritam strojnog učenja (ML) koji se temelji na statističkoj mjeri pronalaženja relevantnosti riječi u tekstu pod nazivom TF-IDF (Sarkar, 2016, str.181-193). Tekst može biti u obliku dokumenta (paragrafi, rečenice i sl.) ili skupa raznih

dokumenata (korpus). Ovaj algoritam je kombinacija dviju metrika: Frekvencija termina (TF) i Inverzna frekvencija dokumenta (IDF).



Slika 1 Frekvencija pojavnosti riječi

(izvor: rad autora temeljem provedenog istraživanja)

TF rezultat temelji se na učestalosti riječi u dokumentu. Riječi se broje prema broju pojavljivanja u dokumentima. TF se izračunava dijeljenjem broja pojavljivanja riječi ( $i$ ) s ukupnim brojem ( $N$ ) riječi u dokumentu ( $j$ ) prema formuli.

$$TF(i) = \text{frequency}(i,j) / \log(N(j))$$

No u dokumentima (korpusu) može biti puno uobičajenih riječi koje se često pojavljuju, a zapravo ne nose onu informaciju koja nas zanima, na primjer 'brand' ili 'shoes'. Njihova visoka pojavnost daje nam samo informaciju da u korpusu tekstova se radi o obući ali to nije dovoljna informacija ako su predmet istraživanja marke sportske obuće.

Iz tog razloga IDF rezultat izračunava rijetkost riječi. To je važno jer TF daje veću težinu riječima koje se pojavljuju češće. Međutim, riječi koje se rijetko koriste u korpusu mogu sadržavati značajnije informacije, na primjer 'Adidas' ili 'Puma'. IDF označava upravo ovakve podatke. Može se izračunati dijeljenjem ukupnog broja ( $N$ ) dokumenata ( $d$ ) s brojem dokumenata koji sadrže riječ ( $i$ ) prema formuli.

$$IDF(i) = \log(N(d) / \text{frequency}(d,i))$$

Logaritam se koristi za ublažavanje učinka vrlo uobičajenih riječi (riječi koje se pojavljuju u gotovo svim dokumentima), koje bi inače dominirale rezultatom.

Konačna mjera pojavnosti riječi u korpusu koja se označava s TF-IDF dobije se pomoću formule

$$TF-IDF(i) = TF(i) * IDF(i)$$

i ona omogućuje klasifikaciju riječi u tekstu koje su od interesa.

### 3.3.3. NLP metode za tematsko modeliranje

Modeliranje tema (**eng.** *Topic modelling*) je analiza pronalaženja ključnih tema kroz skup tekstualnih sadržaja zvanih korpus. Proučavanje tema pomaže istraživačima razumjeti skrivene (**eng.** *Latent*) semantičke strukture u korpusu teksta. Za modeliranje tema koristi se LDA model (Latent Dirichlet Allocation). LDA model spada u nenadzirane (unsupervised) algoritme strojnog učenja (Blei et al., 2003). LDA model se temelji na pretpostavci da unutar analiziranog korpusa sadržaja postoje dokumenti koji imaju slične riječi vezane za istu temu (**eng.** *topic*) ili da imaju grupe riječi koje se najčešće pojavljuju u nekim temama. Poznavanje tema korisno je i za klasifikaciju postojećih tekstualnih podataka i za generiranje novih podataka. Kod modeliranja tema postoji konvencionalna terminologija koje se koriste u analizi teksta kako to navode Blei et al. (2003) terminologija je prikazana na sljedeći način:

- riječ je osnovna jedinica tekstualnih podataka i indeksira se kao rječnik  $\{1, \dots, V\}$  gdje svaka riječ ima svoju poziciju definiranom s indeksom  $id$ ;
- dokument je niz od  $N$  riječi definiranih kao  $w = (w_1, w_2, \dots, w_N)$ , gdje je  $w_n$   $n$ -ta riječ u nekom dokumentu;
- korpus je zbirka  $M$  dokumenata definiranih kao  $D = (w_1, w_2, \dots, w_M)$ ;

LDA model nenadzirani je generativna probabilistička algoritam strojnog učenja koja se koristi za izdvajanje:

- broja tema u podacima (korpus)
- raspodjela tema u svakom tekstu (dokumentu)
- raspodjela riječi (pojmova) u svakoj temi

Tema se definira kao ponavljajući obrazac pojmova koji se istovremeno pojavljuju u korpusu. LDA model svaki dokument u korpusu tretira kao mješavinu tema. LDA je hijerarhijski Bayesov model na tri razine, matematički dosta kompleksan pa umjesto dubokog ulaska u matematičke detalje, lakše je demonstrirati model na primjeru.

LDA model generira distribuciju tema za svaki dokument u korpusu. Na primjer, dokument se može raspodijeliti na dvije teme: 40% u temi "Voće", 60% u temi "Povrće":

Svaka tema je distribucija svih riječi iz skupa rječnika  $\{ "slatko", "mrkva", "jabuka", "zeleno" \}$ . Ovisno o temi, vidimo neke riječi s većom vjerojatnošću. Na primjer, intuitivno bi se riječ "zeleno" pojavila s većom vjerojatnošću u temi "povrće" nego u temi "voće".



Definiraju se dvije latentne varijable koje pomažu u procesu stvaranja teksta od korpusa do dokumenta pa do riječi. Latentna varijabla je varijabla koju ne promatramo izravno iz podataka. Međutim, otkriva skrivene strukture u podacima i može biti korisno u izgradnji vjerojatnog modela. Prva latentna varijabla koja se označava s  $\theta$ , raspodjela je tema po svakom dokumentu (40% „Voće“, 60% „Povrće“). Druga latentna varijabla je  $Z$  ( $Z \in \{1,2..T\}$ ), koja predstavlja temu svake riječi. Za svaki dokument  $d$  postoji distribucija teme  $\theta_d$ . Za svaku riječ 'i' u dokumentu  $d$  ( $w_{d_i}$ ) generira se na temelju raspodjele teme  $\theta_d$ , tema  $Z$  za ovu riječ  $d_i$  i raspodjele riječi po temi  $Z_{d_i}$ . Ako pretpostavimo da je prva riječ u dokumentu  $d$  "zelena". Generira se ( $\theta_d$ ) tako da se prvo navede raspodjela tema: 60% "Voće" i 40% "Povrće" a potom za prvu riječ uzorkujemo temu „Povrće“ ( $Z_{d_1}$ ). U temi "Povrće" uzorkujemo riječ "zeleno" ( $w_{d_1}$ ) iz distribucije teme (Beysolow II, 2018, str. 77-99).

Ono što treba još jednom naglasiti da je LDA model nenadzirana tehnika što znači da ne postoji nikakva informacija o nazivima tema u korpusu već se samo pretpostavlja mogući broj tema (Slamić Tarade, 2020). Implementacija LDA tehnike realizirana je u raznim programskim okruženjima npr. kao biblioteke u Pythonu ili R. Za određivanje tema potrebno je zadati sljedeće parametre: broj tema ( $K$ ), broj riječi u temi ( $N$ ) i broj dokumenata za analizu. Točnost izdvajanja relevantnih tematskih pojmova regulira se s parametrima alfa i beta. Parametar alfa predstavlja distribuciju tema po dokumentima. Visoka vrijednost za alfa parametar znači da će svaki dokument vjerojatno sadržavati mješavinu više tema, a ne bilo koju pojedinačnu temu. Nizak alfa znači da je vjerojatnije da će dokument biti predstavljen sa samo nekoliko tema. Parametar beta predstavlja distribucija riječi po temi. Visoka vrijednost za beta  $\beta$  znači da će svaka tema vjerojatno sadržavati mješavinu većine riječi, a ne samo bilo koju riječ, dok niska vrijednost za beta znači da tema može sadržavati mješavinu samo nekoliko riječi (Blei et al., 2003). U slučaju da pojmovi koji su izdvojeni nisu relevantni navedene parametre treba varirati dok se ne dobije željeni rezultat. Nakon dobivanja rezultata raspodjele  $N$  riječi po  $K$  tema, korisnik mora sam imenovati teme sukladno pojavnosti riječi u temi.

#### 3.3.4. SBS – Semantic Brand Score ili semantička ocjena marke

Semantic Brand Score (SBS) ili semantička ocjena marke novi je način procjene značaja brendova u različitim kontekstima ispitivanjem velike količine tekstualnih podataka razvijen od strane autora Fronzetti Colladon (2018). Za razliku od tradicionalnih metoda koje se oslanjaju na ankete malih skupina potrošača, SBS se može izračunati iz bilo kojeg izvora tekstualnih dokumenata, uključujući društvene medije, blogove, internetske forume, mišljenja

potrošača, e-poštu i novinske članke. Cilj je dobiti prave uvide i signale analizom spontanih izraza potrošača i drugih dionika marke. Ovaj pristup minimizira pristranosti svojstvene upitnicima, gdje sudionici znaju da ih se pomno ispituje. Osim toga, SBS se može prilagoditi za proučavanje relevantnosti određenih riječi ili grupa riječi, ne nužno ograničenih na „robne marke“, a također se može prilagoditi različitim jezicima. Izraz "brand" može se odnositi na različite stvari, kao što je ime političara ili skupina riječi koje predstavljaju određenu ideju ili korporativnu vrijednost, a SBS može se koristiti za procjenu promjene u dinamici koja se događa kada nova marka zamjenjuje staru. SBS je također koristan u usporedbi važnosti marke s konkurentima ili analizi vremenskih trendova jedne marke. U nekim slučajevima rezultat se koristio u prediktivne svrhe, kao što je predviđanje izbornih ishoda na temelju važnosti brendova političkih kandidata u online medijima (Fronzetti Colladon, 2020).

Semantic Brand Score (SBS) metrika je koja koristi tehnike obrade prirodnog jezika (NLP) za analizu mrežnog sadržaja povezanog s markama.

Autor (Fronzetti Colladon, 2018) demonstrira učinkovitost SBS-a uspoređujući ga s tradicionalnim pokazateljima vrijednosti marke, kao što su tržišna kapitalizacija i kapital marke. Rezultati pokazuju da SBS pruža točniju i sveobuhvatniju mjeru vrijednosti marke, budući da uzima u obzir čimbenike kao što su raspoloženje kupaca, poslovna izvješća i relevantnost marke koji nisu obuhvaćeni tradicionalnim mjerenjima. Autor također navodi da SBS može biti koristan alat za tvrtke i investitore za mjerenje i praćenje vrijednosti njihovih marki tijekom vremena te za prepoznavanje područja za poboljšanje u strategijama njihovih marki. No potrebno je naglasiti da SBS nije savršena mjera te da su potrebna daljnja istraživanja kako bi se doradila i poboljšala metodologija. Na primjer SBS ništa ne govori o sentimentu potrošača spram marke ili marki i stoga se nakon određivanja SBS mjere mora napraviti i analiza sentimenta.

Semantička analiza primjenom metode SBS temelji se na matematički dobro poznatoj strukturi podataka poznatoj pod nazivom grafovi ili mreže. Ne ulazeći duboko u matematičko opisivanje grafova, graf se može definirati kao skup vrhova (**eng.** *vertex*) i bridova (**eng.** *edges*) i može se zapisati sljedećim izrazom  $G = \{V, E\}$  (Balakrishna, 1997). Grafovi mogu biti usmjereni (brid je usmjeren od vrha A prema B, ali nije obrnuto) ili neusmjereni (brid povezuje vrhove A i B u oba smjera). Primjena grafova je vrlo široka, od određivanja ruta obilaska trgovačkog putnika, optimizacije ruta leta putničkih zrakoplova do danas vrlo rasprostranjenih socijalnih grafova. Socijalni grafovi pokazuju međusobnu povezanost sudionika društvenih mreža ili različitih druženja npr. različiti turniri u sportu. Analizom socijalnih grafova mogu se utvrditi najutjecajnije osobe neke društvene mreže, dubina povezanosti preko zajedničkih prijatelja kao

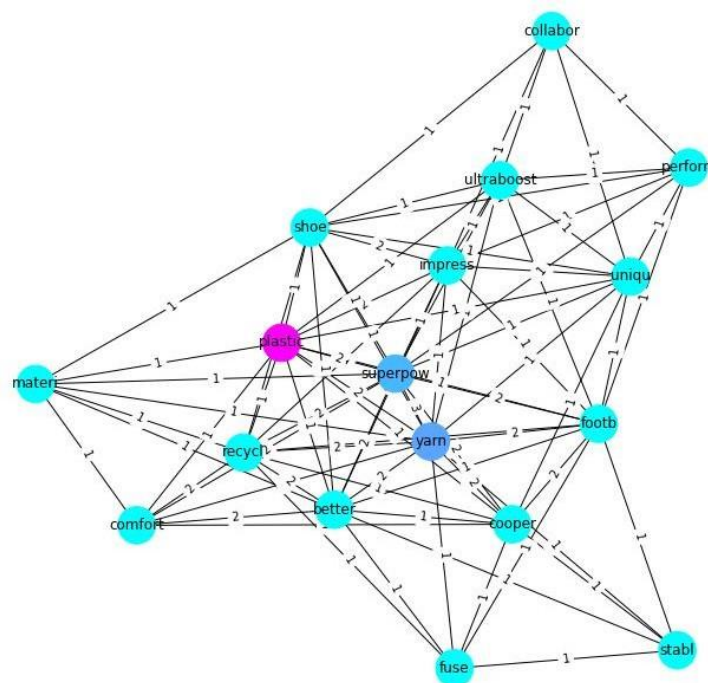
i drugi pokazatelji. Grafovima se također mogu prikazati tekstualni sadržaji gdje su vrhovi riječi, a bridovi su veze sa supojavnim riječima (**eng.** *co-occurrence words*) i oni će biti primijenjeni u SBS semantičkoj analizi u ovom radu.

Kao što je prethodno rečeno Semantic Brand Score (SBS) nova je mjera važnosti marke koja se izračunava na temelju tekstualnih sadržaj, kombinirajući metode društvenih ili socijalnih grafova (mreža) i semantičke analize. Ova metrika je fleksibilna jer se može koristiti u različitim kontekstima i među proizvodima, tržištima i jezicima. Može se koristiti za podršku procesima donošenja odluka unutar poduzeća; na primjer, može se primijeniti za predviđanje cijene dionice poduzeća ili za procjenu važnosti marke u odnosu na konkurente. S jedne strane, SBS se odnosi na poznate konstrukte vrijednosti marke, s druge strane, nudi nove perspektive za učinkovito strateško upravljanje robnim markama u eri velikih podataka (Big Data). Danas menadžeri robnih marki i marketinški stručnjaci mogu mnogo iskoristiti velike količina podataka za otkrivanje obrazaca i trendova u percepcijama potrošača, poslovne i opće javnosti, praćenje pozitivnih ili negativnih asocijacija na robne marke s obzirom na željene teme. U nastavku će se ukratko opisati postupak stvaranja grafa riječi, a potom će biti definirani glavni elementi SBS analize koja se temelji na tri dimenzije.

Metoda Semantic Brand Score (SBS) uključuje pretvaranje prethodno obrađenih izvornih dokumenata u graf istovremenih pojavljivanja riječi (supojavne riječi). Nakon što su izvorni dokumenti koji se analiziraju prethodno obrađeni do razine razbijanja teksta na tokene tj. pojedinačne riječi u rečenici, određivanje SBS-a prolazi kroz njihovu transformaciju u graf supojavljanja riječi. Graf čini (n) vrhova  $G = \{g_1, g_2, g_3, \dots, g_n\}$  – gdje svaki predstavlja riječ i (m) bridova. Brid  $a_{ij}$  koji počinje u vrhu (i) i prethodi vrhu (j) u barem jednom tekstualnom dokumentu. Graf se formira povezivanjem vrhova (riječi) ako jedna riječ prethodi drugoj unutar raspona od pet ili sedam riječi (korisnik može mijenjati taj broj radi postizanja boljeg rezultata) u najmanje jednom tekstualnom dokumentu. Riječi koje se pojavljuju više puta proizvest će više bridova u grafu. Prije izračunavanja rezultata za svaku SBS dimenziju, graf se simetrizira, uklanjaju se petlje (slučaj kada se vrh povezuje s bridom sam sa sobom), a višestruki bridovi između dva čvora zamjenjuju su s jednim bridom ponderiranim s brojem višestrukih bridova. Tako dobiveni graf se zove otežani graf je bridovi dobivaju težine koje pokazuju koliko puta je jedna riječ povezana s drugom riječi u više dokumenata. Analitičari mogu ukloniti bridove s malim težinama koji ukazuju na rijetke istodobne pojave. Primjer za opisani postupak stvaranja grafa riječi na tekstu iz tri rečenice je prikazan na slici 2. Tekst je formiran kao lista rečenica, a svaka rečenica kao lista tokeniziranih riječi u uglatim zagradama uz napomenu da je tekst prethodno obrađen i sve riječi su prikazane malim slovima.

```
text = [['collabor', 'shoe', 'perform', 'impress', 'ultraboost', 'uniqu', 'superpow', 'footb', 'plastic',
'yarn', 'cooper'], ['comfort', 'superpow', 'materi', 'recycl', 'plastic', 'better', 'yarn', 'shoe', 'impress'],
['comfort', 'superpow', 'recycl', 'better', 'yarn', 'cooper', 'footb', 'fuse', 'stabl']]
```

Graf za navedeni tekst pokazuje kako su riječi povezane s drugim riječima i s kojim težinama bridova kada je neka riječ s drugom riječi povezana više puta. Na primjer riječ „better“ je s riječi „comfort“ povezaa u cjelokupnom tekstu dva puta, a s riječi „cooper“ jednom. U radu će se pokazati važnost nekih riječi koje proglasimo brandom s obzirom na riječi koje ju okružuju direktnim vezama preko bridova.



Slika 2 Primjer grafa

(izvor: rad autora temeljem provedenog istraživanja)

Uvjet za određivanje SBS ocjene je poznavanje riječi koje se označavaju kao marke odnosno brendovi pa se zadaju kao ulazni podatak.

SBS ocjena procjenjuje značaj brenda, koji čini temelj vrijednosti brenda (Fronzetti Colladon, 2018). Ta mjera je zapravo djelomično pod utjecajem utvrđenih koncepcija vrijednosti marke, te koncepata imidža marke i svijesti o marki kao što je to prikazano, na primjer, u Kellerovom istraživanju (Keller, 1993). Važnost brenda mjeri se u 3 dimenzije: prevalencija (**eng. prevalence**), raznolikost (**eng. diversity**) i povezanost (**eng. connectivity**). Prevalencija mjeri frekvenciju korištenja imena robne marke, odnosno koliko je puta marka izravno spomenuta.

Raznolikost mjeri raznolikost riječi povezanih s markom. Povezivost predstavlja sposobnost marke da premosti veze između drugih riječi ili grupa riječi (ponekad se vide kao teme izričaja). Prevalencija se odnosi na prisutnost određene riječi (marke) u tekstu i izračunava se kao njezina ukupna frekvencija riječi  $f(g_i)$ , za dati skup dokumenata. Da bi se moglo uspoređivati ovu vrijednost za različite tekstualne korpuse različitih veličina, potrebno je podijeliti apsolutnu frekvenciju termina s ukupnim brojem riječi u tekstu (označeno kao  $totW$ ). Računanje se radi prema formuli:

$$PREV(g_i) = f(g_i) \text{ odnosno } PREV'(g_i) = f(g_i)/totW$$

Za izračunavanje vrijednosti ove dimenzije nije potrebna transformacija dokumenata u graf supojavnih riječi, kako je to prethodno opisano, jer je samo bitno u tekstu prebrojiti koliko puta se neki termin pojavio.

Vrijednost druge dimenzije raznolikost izražava heterogenost riječi koje okružuju riječ koja je definirana kao marka i daje odgovor na pitanje „koliko različitih susjeda u grafu ima marka?”. Veća raznolikost u semantičkom kontekstu u kojem se koristi robna marka označava veću heterogenost. U teoriji grafova raznolikost se izračunava kao centralnost stupnjeva (**eng. degree centrality**) (Balakrishna, 1997) koja pokazuje u grafu supojavnih riječi s koliko vrhova (riječi) je marka direktno povezana preko bridova. Ovaj broj je veći kada se marka pojavljuje zajedno s mnogo različitih riječi. Nasuprot tome, ako je diskurs (izričaj) oko marke konvergira prema malom skupu riječi, njegova će raznolikost biti manja. U sljedeća formula  $d(g_i)$  označava stupanj vrha marke  $g_i$ . Raznolikost se računa prema formuli:

$$DIV(g_i) = d(g_i)$$

$$DIV'(g_i) = d(g_i)/(n-1)$$

odnosno nakon normalizacije gdje je  $(n)$  ukupan broj vrhova grafa. Za razliku od prevalencije za određivanje raznolikost potrebno je formirati graf supojavnih riječi cijelog tekstualnog korpusa.

Treća dimenzija je povezivost i ona ovisi o poziciji marke u grafu. Vrijednost povezivosti je viša kada kroz vrh koji je označen s riječi robne marke češće prolaze najkraći putevi koji međusobno povezuju druge riječi unutar teksta i stoga je dublje ukorijenjen u izričaj. U ovom slučaju povezivost marke možemo smatrati kao most preko kojeg prolaze najkraći putevi

između dviju drugih riječi cijelog tekstualnog korpusa. U teoriji grafova povezivost možemo definirati kao centralnost između (**eng.** betweenness centrality). Ta centralnost se široko koristi u društvenim mrežama kao mjera posredovanja utjecaja ili kontrole informacija koja nadilazi direktne veze između osoba u društvenoj mreži. Povezivost marke  $g_i$  određuje se pomoću formule:

$$CON(g_i) = \sum_{j < k} (d_{jk}(g_i)/d_{jk})$$

Gdje je  $d_{jk}$  broj najkraćih puteva koji povezuju generički par vrhova  $g_j$  i  $g_k$ , a  $d_{jk}(g_i)$  je broj tih puteva koji u sebi sadrže vrh  $g_i$ . Ova formula se može također normalizirati pa dobijemo formulu.

$$CON(g_i) = \sum_{j < k} (d_{jk}(g_i)/d_{jk}) / [(n-1)(n-2)/2]$$

Konačna vrijednost značaja marke SBS dobiva se zbrajanjem standardiziranih vrijednosti prevalencije, raznolikosti i povezanosti, kako bi se dala ista važnost za svaku dimenziju, čak i u slučaju različitih varijanci izračunava se prema formuli:

$$SBS(g_i) = \frac{PREV(g_i) - \overline{PREV}}{std(PREV)} + \frac{DIV(g_i) - \overline{DIV}}{std(DIV)} + \frac{CON(g_i) - \overline{CON}}{std(CON)}$$

pri čemu se sve tri vrijednosti standardiziraju prije zbrajanja tako da se od svake vrijednosti za prevalenciju, raznolikost i povezivost za vrh ( $g_i$ ) oduzme njihova srednja vrijednost i podjeli se sa standardnom devijacijom. U gornjoj formuli svaka od tri komponente ima istu težinu jer se smatra se da sve tri dimenzije imaju istu važnost.

### 3.3.5. NLP metoda za analizu sentimenta

Kao što je prethodno opisano SBS ocjena značajnosti robnih marki ne govori nam ništa o raspoloženju potrošača spram robne marke te je iz tog razloga potrebno napraviti dodatna istraživanja. Jedan od načina utvrđivanja osjećaja potrošača spram robne marke je analiza sentimenta. Proces analize sentimenta uključuje ispitivanje konteksta pisanog ili izgovorenog materijala kako bi se identificirale subjektivne informacije, koje mogu pomoći tvrtkama da razumiju kako se njihova marka, proizvod ili usluga percipiraju na društvenim medijima (Kauffmann et al., 2019). Dok se osnovna analiza sentimenta i metrika brojanja obično koriste za analizu društvenih medija, ovaj pristup možda neće otkriti vrijedne uvide. Kako bi uhvatili

ove vrijedne informacije, stručnjaci za robne marke bi trebali razmotriti korištenje naprednih tehnika umjetne inteligencije, kao što su algoritmi dubinskog učenja, kako bi proveli temeljitija istraživanja. Kako bi se to postiglo, preporučuje se da se mišljenja potrošača o robnoj marki klasificiraju na temelju sljedećih kriterija:

- Ključni aspekti proizvoda i usluga marke do kojih je kupcima stalo.
- Temeljne namjere i reakcije korisnika u vezi s tim aspektima.

što je u skladu sa svim stavovima izrečenim u točki 2 ovog rada na temelju istraživanja i modela autora Aaker i Keller.

Sentiment se uobičajeno kategorizira kao pozitivan (**eng. positive**) i negativan (**eng. negative**), ali može se uvesti i treća kategorija neutralan (**eng. neutral**). Postoji više različitih NLP pristupa za ocjenu sentimenta koji se temelje na strojnom učenju odnosno na klasifikaciji dokumenata. Dokumenti su u pravilu mišljenja potrošača prikupljena sa stranica društvenih mreža (*Facebook, Twitter*), web stranica tvrtki koje se bave prodajom robnih marki (Amazon, eBay) ili specijaliziranih web stranica koje prikupljaju mišljenja potrošača (Brand24, ConsumerAffairs, Sitejaber). Klasifikacija dokumenata (svaki dokument je jedno mišljenje potrošača) može biti nenadzirana (**eng. unsupervised**) i nadzirana (**eng. supervised**) kada su mišljenja označena (**eng. labeled**) s klasama pozitivan ili negativan (Sarkar, 2016., str.342-352). Kada imamo samo dvije klase za klasifikaciju onda se koriste uobičajeno binarni klasifikatori. Najčešće primjenjivani algoritmi za binarnu klasifikaciju su Naive Bayes, LogisticRegression i. Za klasifikaciju dokumenata s više klasa koriste se algoritmi dubokog strojnog učenja. U ovom radu će za analizu sentimenta biti primijenjen algoritam koji koristi stablo odlučivanja (Decision Tree). Ovaj algoritam se sastoji od Decision Tree Learner-a i Decision Tree Predictor-a. Algoritam prvo uči na označenim (pozitivan ili negativan) mišljenjima potrošača (za učenje se uzima veći skup mišljenja) nakon čega se dobije model koji je ulaz u prediktor. Prediktor na manjem skupu (testni model) neoznačenih mišljenja na osnovu naučenog modela predviđa koja mišljenja su pozitivna, a koja negativna. Kvaliteta predikcije se ocjenjuje statistikama kao što su točnost, preciznost, opoziv, broju korektno i ne korektno klasificiranih dokumenata. Kvaliteta klasifikacije se također u pravilu prikazuje i grafički pomoću matrice konfuzije i ROC krivulje. U nastavku su ukratko objašnjeni svi navedeni kvalitativni pokazatelji (Géron, 2019., str. 87-102).

Uobičajeno za procjenu nadziranih klasifikatora strojnog učenja koristi se matrica zabune (**eng. Confusion Matrix**) koja vizualno pokazuje koliko je model sposoban u predviđanju klase testnog promatranja. Sastoji se od tablice događaja stvarnih oznaka klasa u usporedbi s

predviđanjima modela. Binarni klasifikator ima samo dva moguća ishoda, pozitivan (+) ili negativan (-).

Kod binarnih klasifikatora predviđanje se može dogoditi na dva načina:

- TP (True Positive) - kada pozitivna varijabla modela predviđa da je stvarno pozitivna,
- TN (True Negative) – model predviđa da je negativna varijabla stvarno negativna.

U bilo kojem od ovih slučajeva rezultat zadovoljava. Uostalom, krajnji cilj bilo kojeg klasifikatora je ispravno razlikovati pozitivne i negativne slučajeve i u skladu s tim označiti ih.

Na sličan način, pogrešna klasifikacija može se dogoditi na jedan u ova dva slučaja:

- FN (False Negative) – za pozitivnu varijabla modela lažno predviđa da je negativna,
- FP (False Positive) – za negativnu varijablu model predviđa da je lažno pozitivna.

Pogrešna klasifikacija nikada nije poželjna u algoritmima strojnog učenja. Ovisno o skupu podataka i fenomenu koji se proučava, lažno negativno ili lažno pozitivno može imati teže implikacije nad drugima. Vizualno se matrica konfuzije prikazuje na slijedeći način (Slika 3).

		Predicted (Predviđeno)	
		- (Negative)	+ (Positive)
True (Stvarno)	- (Negative)	TN (True Negative)	FP (False Positive)
	+ (Positive)	FN (False Negative)	TP (True Positive)

Slika 3 Matrica konfuzije

(izvor: rad autora temeljem provedenog istraživanja)

Iz matrice konfuzije može se odrediti sljedeće metrike: točnost (Accuracy), preciznost (Precision), opoziv ili osjetljivost (Recall ili True Positive Rate), specifičnost (Specificity), lažno pozitivnu brzinu (False Positive Rate) i F1 Score. Točnost se određuje slijedećom formulom:

$$\text{Točnost (Accuracy)} = \frac{TP + TN}{TP + TN + FP + FN}$$



Preciznost je mjera "korektnosti" jer se promatra omjer između ispravno predviđenih pozitivnih i svih predviđenih pozitivnih rezultata (stvarnih i lažnih).

$$\text{Preciznost (Precision)} = \text{TP} / (\text{TP} + \text{FP})$$

Opoziv (ili osjetljivost, istinska pozitivna stopa) je mjera "cjelovitosti" jer promatra omjer između pravilno predviđenih pozitivnih i svih stvarnih pozitivnih rezultata:

$$\text{Opoziv (Recall ili TPR)} = \text{TP} / (\text{TP} + \text{FN})$$

gdje je TPR True Positive Rate.

Specifičnost (**eng.** *Specificity*) je mjera koliko dobro klasifikacijski model predviđa klasu koju smo deklarirali kao negativnu klasu. Ovo je, na neki način, isto što i Recall samo za za negativnu klasu.

$$(\text{Specifičnost}) \text{ Specificity} = \text{TN} / (\text{TN} + \text{FP})$$

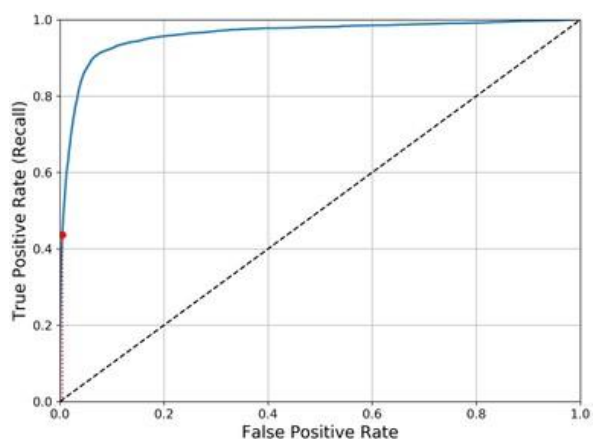
Na kraju F1 Score kombinira Recall i Precision u jednu metriku učinkovitosti klasifikacije i stoga ova ocjena uzima u obzir i lažno pozitivne i lažno negativne.

$$\text{F1 Score} = 2 * (\text{Recall} * \text{Precision}) / (\text{Recall} + \text{Precision}).$$

Druga grafička metrika performansi klasifikacije je ROC krivulja (Receiver Operating Characteristics). Ukupna performansa klasifikatora, daje se na svim mogućim pragovima, koji formiraju ROC krivulju. Za određivanje ROC krivulje potrebno je pored False Positive Rate (Recall) poznavati i False Positive Rate (FPR) koja se određuje prema formuli.

$$\text{FPR} = \text{FP} / (\text{FP} + \text{TN})$$

ROC krivulja koriste se kako bi se vidjelo koliko naš klasifikator može razdvojiti pozitivne i negativne primjere i identificirati najbolji prag za njihovo odvajanje. Jedan primjer ROC krivulje je dan na slici 4.



Slika 4 Primjer ROC krivulje

(izvor: Géron Aurélien ,2019, *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow*.)

Prema ROC krivulji performanse klasifikatora su to bolje što se plava linija na slici brže približava jedinici. Crtkana linija je prag ispod kojeg se klasifikator karakterizira kao netočan.

### 3.3.6. Programski alati za NLP analize teksta prethodno opisanim metodama

Za provedbu istraživanja korištena je Open Source aplikacija KNIME Analytics Platform (2023) pomoću koje je napravljeni svi modeli istraživanja. Temelj ove aplikacije je koncept „low code no code“ tj. programiranje s modulima koji ne zahtijevaju pisanje programskog koda ili se koristi minimalno kodiranje. Ovaj koncept je pogodan za istraživače koji nisu jako vični u programiranju i pisanju koda u nekom od programskih jezika (Python, C++, R). Snaga ove aplikacije je u brojnim proširenjima za provedbu statističkih analiza, strojnog učenja i obradu teksta. Pored toga aplikacija po potrebi omogućuje pisanje manjih dijelova programskog koda korištenjem skripti pisanih u programskim jezicima Python i R radi proširenja mogućnosti modela. Ova platforma ima jako razvijen Community Hub sa velikim brojem primjera iz bilo kojeg područja analitike, strojnog učenja, NLP-a i drugih primjena i zahvaljujući tome krivulja učenja je relativno jednostavna. Platforma omogućuje razvoj modela bilo koje složenosti i testiranje na osobnom računalu, a u slučaju potrebe za velikim računalnim resursima zbog velike količine sadržaja koji se obrađuju moguće je isti model izvoditi na serverima tvrtke KNIME uz plaćanje.

Osim primjene KNIME analitičke platforme tijekom istraživanja i dobivanja rezultata korištene su različite biblioteke programskog jezika Python kao što su: NetworkX za rad s grafovima kod SBS analize, Beautiful Soup i Request za potrebe web scraping-a tekstualnih sadržaja s interneta, NLTK za pred obradu tekstualnih sadržaja, Matplotlib za crtanje grafova.

Za potrebe pohrane tekstualnih sadržaja s interneta kao i za obradu nekih rezultata korišten je Excel i formati datoteka CSV, XSLX.

## **4. REZULTATI ISTRAŽIVANJA**

Na temelju teorijskih razmatranja upravljanja i vrednovanja robnih marki te opisa NLP metoda korištenih u ovom radu izvršen je najznačajniji dio ovog rada tj. dobivanje rezultata analize značaja robnih marki primjenom NLP metoda nad velikim internetskim tekstualnim sadržajima.

### **4.1. Istraživački model**

Istraživački model može se sažeti u sljedećim koracima:

- a) Prikupljeni tekstualni sadržaji iz navedenih internetskih izvora moraju se pripremiti za konačnu obradu i analize. Priprema predstavlja čišćenje i normalizaciju tekstualnih sadržaja (isti oblik prikaza) te druge pred obrade nužne za dobivanje relevantnih podatak.
- b) Primjenom analitičkog alata KNIME programirani su modeli NLP analize opisani u točki 3. Programiranje s ovim alatom podrazumijeva izradu modela bez kodiranja ili s minimalnim kodiranjem.
- c) Nakon što je napravljena analiza NLP metodama TF-IDF i tematsko modeliranje izdvojeni su najznačajnije robne marke sportske obuće koje se pojavljuju u cjelokupnom tekstualnom korpusu provedena je semantička analiza značaja marki primjenom SBS koncepta. SBS analiza je napravljena pomoću osnovnog programa napisanog u Python programskom jeziku koji tekstualne sadržaje pretvara u graf riječi.
- d) Za robnu marku sportske obuće koja je SBS metodom rangirana kao najbolja učinjena je analiza sentimenta i pomoću programa KNIME napravljena je evaluacija performansi klasifikacije sentimenta s obzirom na dvije klase, pozitivna i negativna mišljena potrošača.

### **4.2. Predmet i problem istraživanja**

Predmet samog istraživanja vezan je uz analizu značaja marki i ocjenu mišljenja potrošača sportske obuće koje se temelji na obradi velikih tekstualnih podataka preuzetih s interneta.

Problemi istraživanja su:

- a) izdvajanje ključnih robnih marki sportske obuće koje se temelje na metodama TF-IDF frekvencijske analize i modeliranja tema (**eng. Topics modeling**).
- b) određivanje ocjene značaja marke korištenjem inovativne metrike pod nazivom Semantička ocjena branda (**eng. Semantic Brand Score**) kroz tri dimenzije - pojavnost (**eng. prevalence**), raznolikost (**eng. diversity**) i povezanost (**eng. connectivity**).
- c) analiza sentimenta potrošača za marku sportske obuće koja je u prethodno navedenim analizama rangirana kao najbolja kroz dvije klase mišljenja potrošača, pozitivne i negativne.

### 4.3. Ciljevi istraživanja

Ciljevi istraživanja vezani su uz novi način metrike i vrednovanja značaja robnih marki sportske obuće primjenom NLP metoda za analizu velikih tekstualnih sadržaja dostupnih na internetu. S obzirom na velike količine tekstualnih sadržaja na web stranicama i društvenim medijima mogućnost prikupljanja informacija i stavovima potrošača vezanim uz robne marke sportske obuće proširuju se u neograničenim razmjerima. Temeljem analize značaja marke utvrditi će se kako se mogu napraviti poboljšanja u marketinškim i strategijama odnosa s javnošću vezanim uz robne marke sportske obuće. Jedan od ciljeva je pokazati da istraživanja tržišta i robnih marki mogu biti komplementarna klasičnim metodama kao što su financijska izvješća, upitnici i intervjui i kombiniranjem obaju pristupa dobije se sinergijski učinak.

### 4.4. Hipoteze istraživanja

S obzirom da se istraživanje temelji na novoj paradigmi određivanja značaja robnih marki koja se razlikuje od ciljanih upitnika ili intervjua dokazivanje postavljenih hipoteza biti će nešto drugačijeg karaktera. Postavljene su dvije hipoteze:

**Hipoteza H1:** Ocjena važnosti branda NLP metodama analize je važna za definiranje strategije marke i mogućih poboljšanja. Veći značaj marke povlači veću vrijednost prihoda.

**Hipoteza H2:** Važnost branda na pozitivan je način povezana s ocjenom značaja marke te povjerenjem i lojalnošću.

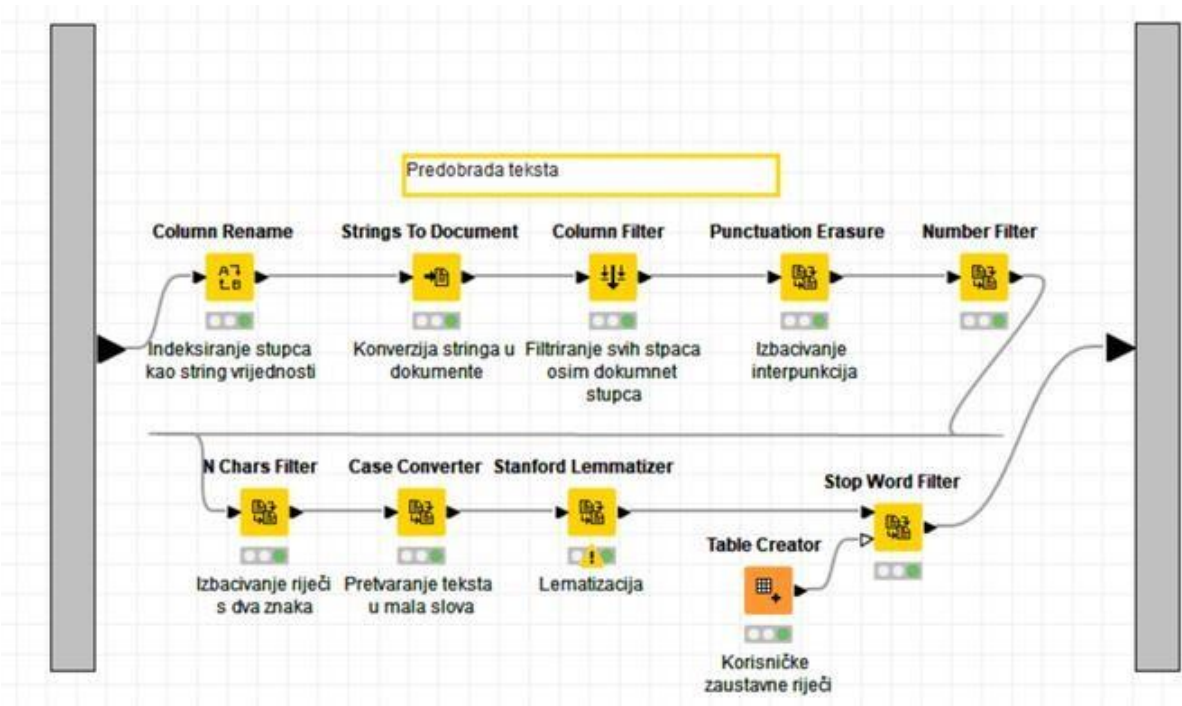
Za dokazivanje ili odbacivanje hipoteza prikupiti će se godišnji statistički podaci o prihodima (**eng. revenue**) marki. Usporediti će se rezultati SBS metrike s podacima o prihodima

istraživanih robnih marki sportske obuće. Utvrditi će se mišljenja potrošača analizom sentimenta za najbolje rangiranu robnu marku sportske obuće.

#### **4.5. Analiza rezultata određivanja ključnih robnih marki**

Za modeliranje teme (**eng** *topics modeling*) radi označavanja najvažnijih marki koji se spominju u tekstovima primijenjene su metode NLP-a za određivanje frekvencija pojavnosti riječi vezanih uz brandove sportske obuće i metoda LDA (**eng**. *Latent Dirichlet Allocation*) nenadziranog (**eng**. *unsupervised*) algoritma strojnog učenja koja kroz analizu korpusa sadržaja određuje teme koji imaju slične povezane riječi ili određuje grupe riječi koje se najčešće pojavljuju u nekim temama. Ove analize su napravljene primjenom analitičke platforme KNIME.

Tekstualni korpus naveden u točki 1 i pohranjen u Excel i CSV datoteke prije bilo kakve analize potrebno je očistiti od različitih znakova koji ne nose nikakvu informaciju i nakon toga dodatno pred obradom (**eng**. *preprocessing*) svesti na formu koja je pogodna za analizu. U pred obradi sve riječi se pretvaraju riječi pisane malim slovom, miču se interpunkcijski znakovi, radi se tokenizacija tj. razbijanje rečenica na riječi, radi se filtriranje zaustavnih riječi (**eng**. *stopwords*) tj. Riječi koje bitno ne doprinose značenju teksta (I, me, am, my, myself, we, our, ours, ourselves, you, you're, the, a i druge), po potrebi se radi lematizacija tj. svođenje riječi na korijenski oblik ili svođenje na osnovni oblik (**eng**. *stemming*). Proces pred obrade primjenom čvorova analitičkog alata KNIME prikazan je na slici 5. Pored prethodno navedenih operacija sa slike 5 vidimo da se dodatno filtriraju svi brojevi te riječi s manje od 3 slova i dodatno su preko tabele unesene korisničke zaustavne riječi.



Slika 5 Predobrada teksta

(izvor: rad autora temeljem provedenog istraživanja)

Nakon ovih radnji izvršene su analize i dobiveni su željeni rezultati ovog istraživanja opisani narednim pod točkama.

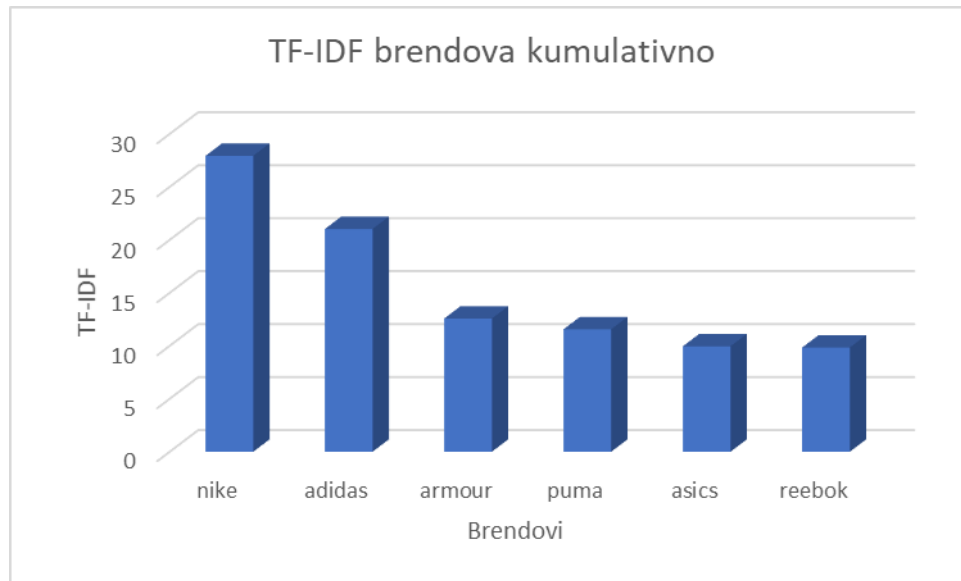
#### 4.5.1. TF-IDF analiza

KNIME model za TF-IDF frekvencijsku analizu prikazan je na slici 6 gdje vidimo da nakon učitavanja teksta i pred obrade kreira se takozvana „vreća“ riječi (**eng. Bag of Words ili BoW**), Vreća riječi je prikaz teksta koji opisuje pojavljivanje riječi unutar dokumenta i to uključuje dvije stvari: stvaranje rječnika poznatih riječi i mjeru prisutnosti poznatih riječi. Naziva se "vreća" riječi jer se odbacuju sve informacije o redoslijedu ili strukturi riječi u dokumentu. Model se bavi samo time pojavljuju li se poznate riječi u dokumentu, a ne gdje u dokumentu. To je bitno za slijedeći korak određivanje frekvencije riječi TF i inverzne frekvencije dokumenata IDF. Konačni rezultat se dobije na izlazu čvora Java Snippet koji računa produkt TF\*IDF.





prikazuje poredak kumulativnih vrijednost (suma kroz sve dokumente) TF-IDF za svaku robnu marku.



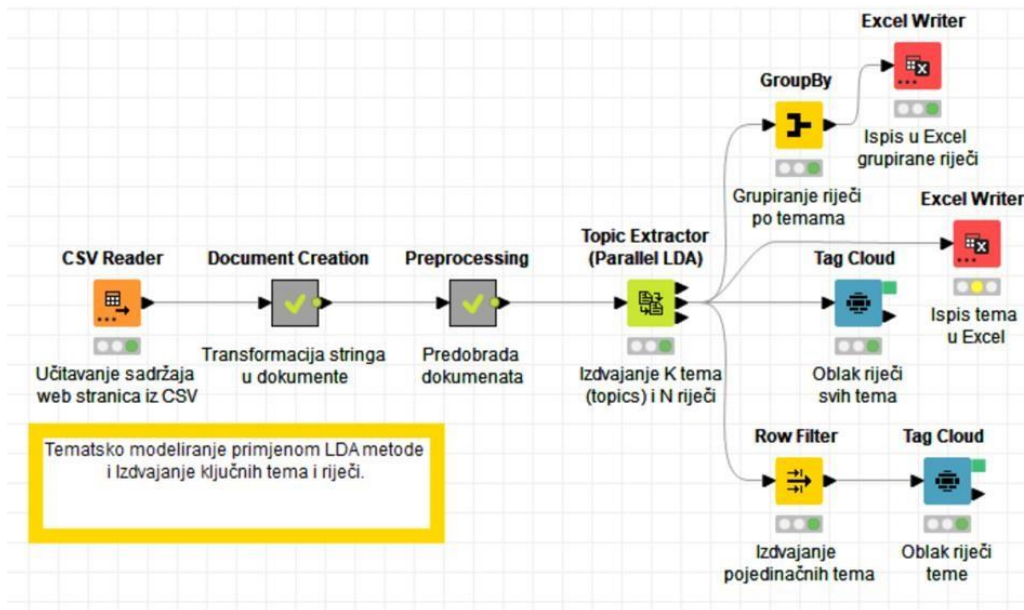
Slika 8 TF-IDF kumulativno

(izvor: rad autora temeljem provedenog istraživanja)

Zaključno, TF-IDF frekvencijska analiza, za razliku od običnih frekvencija pojavnosti, daje nam informaciju koji su to najznačajnije robne marke u tekstovima koje smo skinuli s 57 web stranica pri čemu tekstove nismo trebali posebno označiti (labelirati). Ovaj rezultata se potvrđuje u slijedećoj NLP analizi tj. tematskom modeliranju.

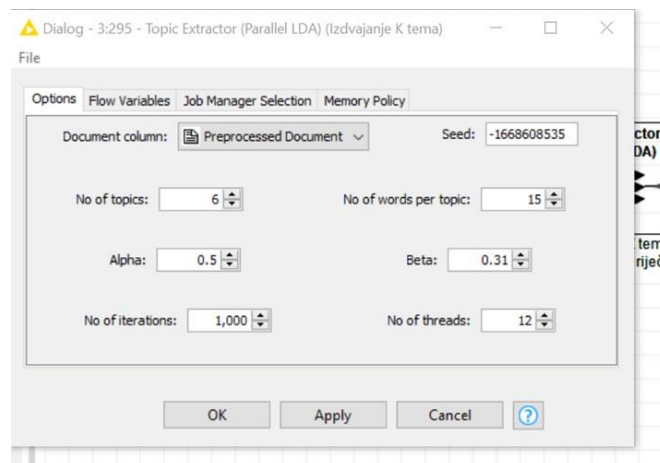
#### 4.5.2. Tematsko modeliranje

Za tematsko modeliranje je rečeno da je to metoda nenadziranog učenja i određivanja tema i riječi koje su najznačajnije za određene teme. KNIME model je prikazan na slici 9. Sa slike vidimo da proces prolazi već opisane korake do modula koji se zove Topic Exrtactor i koji implementira algoritam nenadziranog učenja Latent Dirichlet Allocation. Prije samog određivanja tema potrebno je postaviti ključne parametre prema slici 10. U ovom istraživanju je broj tema  $K = 6$ , broj riječi po temi  $N = 15$  i zadani su parametri  $\alpha = 0.5$  i  $\beta = 0.31$ . Broj tema je eksperimentalno određen jer su testiranja za  $K = 10$  i  $K = 8$  pokazala da se teme i ključne riječi vezane uz teme ponavljaju i da odabrani  $K = 6$  daje prihvatljivo rješenje. Parametri  $\alpha$  i  $\beta$  su također u više eksperimenata varirani i postavljene vrijednost su dale najbolji rezultat. Rezultati se mogu prikazati na više načina i biti će ukratko obrazloženi.



Slika 9 Model za tematsko modeliranje pomoću LDA metode

(izvor: rad autora temeljem provedenog istraživanja)



Slika 10 Postavljanje parametara za LDA analizu

(izvor: rad autora temeljem provedenog istraživanja)

Oblak riječi (slika 11) vizualno kumulativno pokazuje najznačajnije riječi koje se javljaju u ovih šest tema među kojima se opet ističu robne marke Nike, Adidas, Puma, Under Armour, Reebok i Asics. Naravno tu se pojavljuju i druge riječi koje daju kontekst temama odnosno robnim markama (performance, market, growth, billion).



Slika 11 Oblak riječi u temama

(izvor: rad autora temeljem provedenog istraživanja)

U tablicama 2 i 3 prikazane su teme i raspodjela najznačajnijih 15 riječi vezanih uz teme. Ovdje treba naglasiti da LDA metoda ne može imenovati teme već to mora napraviti korisnik na osnovu riječi koje su pridružene temi. Frekvencije riječi u temama su normirane na vrijednosti između 0 i 1.

Tablica 2. LDA teme od 1 do 3

Tema 1: Mišljenja korisnika		Tema 2: Sportska obuća i sportovi		Tema 3: Dizajn i tehnologije	
Ključne riječi	Frekvencija (normirana)	Ključne riječi	Frekvencija (normirana)	Ključne riječi	Frekvencija (normirana)
nike	0.19	adidas	0.15	cushioning	0.10
reviews	0.12	sportswear	0.11	upper	0.08
market	0.09	puma	0.08	sport	0.08
adidas	0.09	founder	0.08	technology	0.08
analysis	0.08	reebok	0.08	design	0.07
women	0.06	athletic	0.07	performance	0.07
categories	0.06	accessories	0.05	comfort	0.07
customers	0.05	performance	0.05	trail	0.06
report	0.04	range	0.05	lightweight	0.06
price	0.04	athletes	0.05	foam	0.06
example	0.04	popular	0.05	nike	0.06
data	0.04	sport	0.05	asics	0.06
impact	0.04	nike	0.04	foot	0.05
sneakers	0.03	name	0.04	runners	0.05
cloud	0.03	fashion	0.04	midsole	0.05

(izvor: rad autora temeljem provedenog istraživanja)

U tablici 2 tema 1 imenovana je kao „Mišljenje korisnika“ na osnovu najistaknutijih riječi kao što su: „nike“, „reviews“, „woman“, „customers“. Tema 2 imenovana je kao „Sportska obuća i sportovi“ što se vidi iz raspodjele ključnih riječi: „adidas“, „puma“, „reebok“. Treća tema odnosi se na „Dizajn i tehnologije“ a nju obilježavaju riječi: „cushioning“, „upper“, „technology“, „design“, „comfort“.

Tablica 3.LDA teme od 4 do 6

Tema4: Brendovi i igrači		Tema 5: Brendovi, tržište i korisnici		Tema 6: Tržište i financije	
Ključne riječi	Frekvencija (normirana)	Ključne riječi	Frekvencija (normirana)	Ključne riječi	Frekvencija (normirana)
nike	0.20	nike	0.16	nike	0.15
adidas	0.10	adidas	0.14	sales	0.13
sneaker	0.08	armour	0.10	billion	0.12
players	0.07	consumer	0.06	adidas	0.11
jordan	0.07	fashion	0.06	market	0.10
football	0.06	performance	0.06	growth	0.08
million	0.06	consumers	0.06	revenue	0.06
deal	0.05	innovation	0.06	million	0.05
star	0.05	market	0.05	armour	0.04
athletes	0.05	business	0.05	puma	0.04
signature	0.05	industry	0.05	share	0.03
kanye	0.04	digital	0.04	north	0.03
athlete	0.04	companies	0.04	america	0.03
sport	0.04	trend	0.04	based	0.03
female	0.04	women	0.04	business	0.02

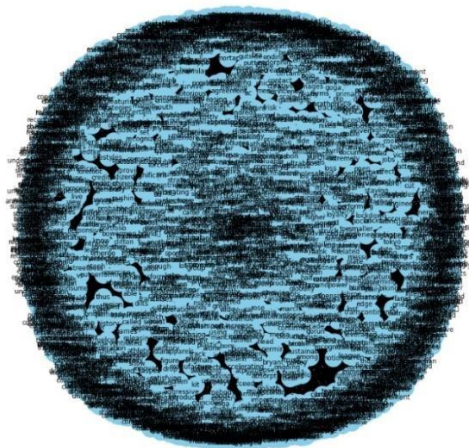
(izvor: rad autora temeljem provedenog istraživanja)

U tablici 3 treba istaknuti sljedeće tri teme: „Brendovi i igrači“, „Brendovi, tržište i korisnici“ i „Tržište i financije“ Tako temu 4 obilježavaju riječi: „nike“, „adidas“, „sneaker“. Tema 5 vezana je za glavne brendove kao što su: „nike“, „adidas“, „armour“ i njihovu povezanost s potrošačima i tržištem. U temi 6 prevladavaju riječi: „nike“, „sales“, „billion“ „growth“.

#### 4.5.3. SBS analiza

Određivanje ocjene važnosti brandova dobiveno je korištenjem inovativne metrike pod nazivom semantička ocjena branda (SBS – Semantic Brand Score) koja ocjenjuje riječi koje predstavljaju brandove kroz tri svojstva – pojavnost (**eng. prevalence**), raznolikost (**eng. diversity**) i povezanost (**eng. connectivity**). Ova metoda koristi teoriju grafova (mreža) na način da tekstove razbije na rečenice, rečenice na riječi i zatim radi graf (mrežu) povezanosti tih riječi. Riječi predstavljaju čvorove u grafu, a veze između riječi imaju težine koliko puta je neka riječ

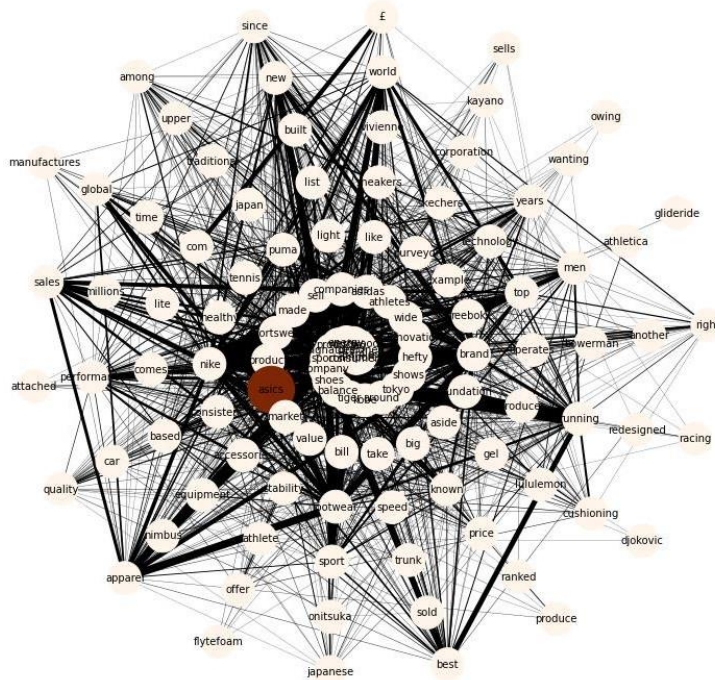
povezana s nekom drugom riječi. Nakon toga radi se analiza grafa (mreže) koristeći algoritme određivanja tzv. centralnosti na grafu. Za SBS analizu korišten je osnovni program u Pythonu preuzet je od autora Fronzetti Colladon (2019). Program je dodatno doraden i dopunjen s učitavanjem tekstualnih sadržaja iz CSV datoteke, izdvajanjem rečenica iz dokumenata (tokenizacija rečenica) i na kraju izdvajanjem riječi (tokenizacija riječi). Potom je napravljena standardna pred obrada i čišćenje teksta od nepotrebnih elemenata kao što su URL adrese, e-mail adrese i slično. Kao ulazni parametar zadana je lista imena robnih marki za koje se radi SBS ocjena brands = ['puma', 'adidas', 'nike', "armour", "asics", "reebok"], a koje su u prethodne dvije NLP izdvojene kao najznačajnije u cjelokupnom tekstualnom korpusu. Potom je izvršeno formiranje grafa riječi primjenom Python NetworkX biblioteke i pristupilo se izračunavanju SBS ocjene. Na slici 12 je prikazan izgled grafa koji nam zapravo ne može dati bitnu informaciju jer je sastavljen od velikog broja riječi (više od 40.000 koje se ponavljaju). Iz tog razloga su u grafu izbačene sve riječi koje povezuje brid s težinom manjom od tri tako da je graf sa 7180 vrhova i 168.529 bridova sveden na graf koji ima 4373 vrha i 35.433 brida.



*Slika 12 Graf riječi prije filtriranja*

(izvor: rad autora temeljem provedenog istraživanja)

Sada se može prikazati nešto čitljiviji graf ali samo za marku Asics s obzirom da ona ima relativno najmanje povezanih riječi s njom. Taj graf je prikazan na slici 13. Težine bridova na ovoj slici nisu prikazane brojačano već deblje linije označavaju bridove s većom težinom.



Slika 13 Izgled grafa nakon filtriranja za marku Asics

(izvor: rad autora temeljem provedenog istraživanja)

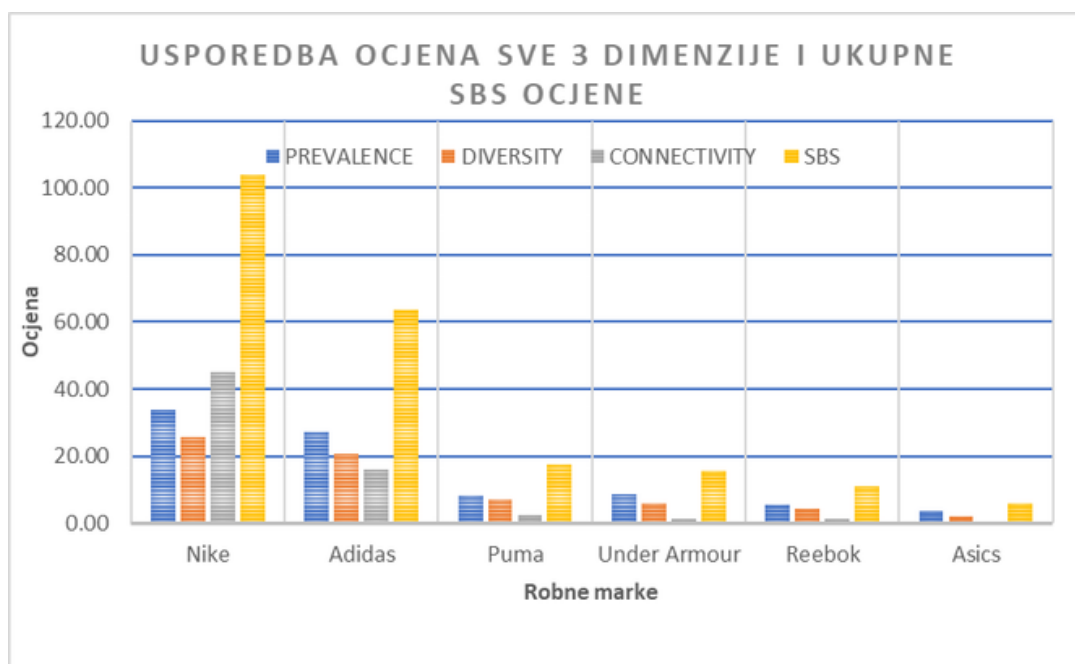
Prethodne slike su samo vizualizacija grafa koji je korišten za određivanje značaja robnih marki kroz tri dimenzije i one ne daju nikakvu informaciju o SBS ocjeni. Nakon izvršenih izračuna dobiven je rezultat SBS ocjene prikazan u tablici 4.

Tablica 4 Ocjene sve tri dimenzije i ukupna SBS ocjena

	PREVALENCE	DIVERSITY	CONNECTIVITY	SBS
Nike	33.59	25.47	44.92	103.98
Adidas	27.01	20.62	16.06	63.69
Puma	7.98	6.82	2.52	17.32
Under Armour	8.41	5.68	1.31	15.41
Reebok	5.51	4.23	1.21	10.94
Asics	3.53	1.94	0.55	6.02

(izvor: rad autora temeljem provedenog istraživanja)

Odnosi između dimenzija prevalencije, diversity connectiviti i ukupne ocjene SBS značaja robnih marki vizualno su prikazani na grafikonu na slici 14.



Slika 14 SBS ocjene semantičke analize

(izvor: rad autora temeljem provedenog istraživanja)

Prethodni rezultati za metriku koju pruža semantička analiza primjenom inovativnog koncepta određivanja SBS ocjene ne samo da pokazuje značaj marki nego i međusobni odnos između marki koje su bile predmet ocjene. Slika prikazuje da marke Nike i Adidas imaju najveći značaj u analiziranom korpusu web sadržaja vezanih uz sportsku obuću s time da marka Nike značajno odskoče od marke Adidas. Ovakva analiza je značajna za marketinške stručnjake iz dva razloga, prvo što mogu vidjeti u kakvom odnosu je njiva marka spram konkurencije i drugo da je za vidljivost marke izuzetno bitno biti u što više različitih objava na internetu. SBS analiza bi bila još bolji pokazatelj značaja robne marke ako bi se radila i kroz vremensku dimenziju na primjer da se radi analiza tekstova objavljenih u godišnjem (za više godina) i polugodišnjem razdoblju jer bi se tada vidio i trend promjena.

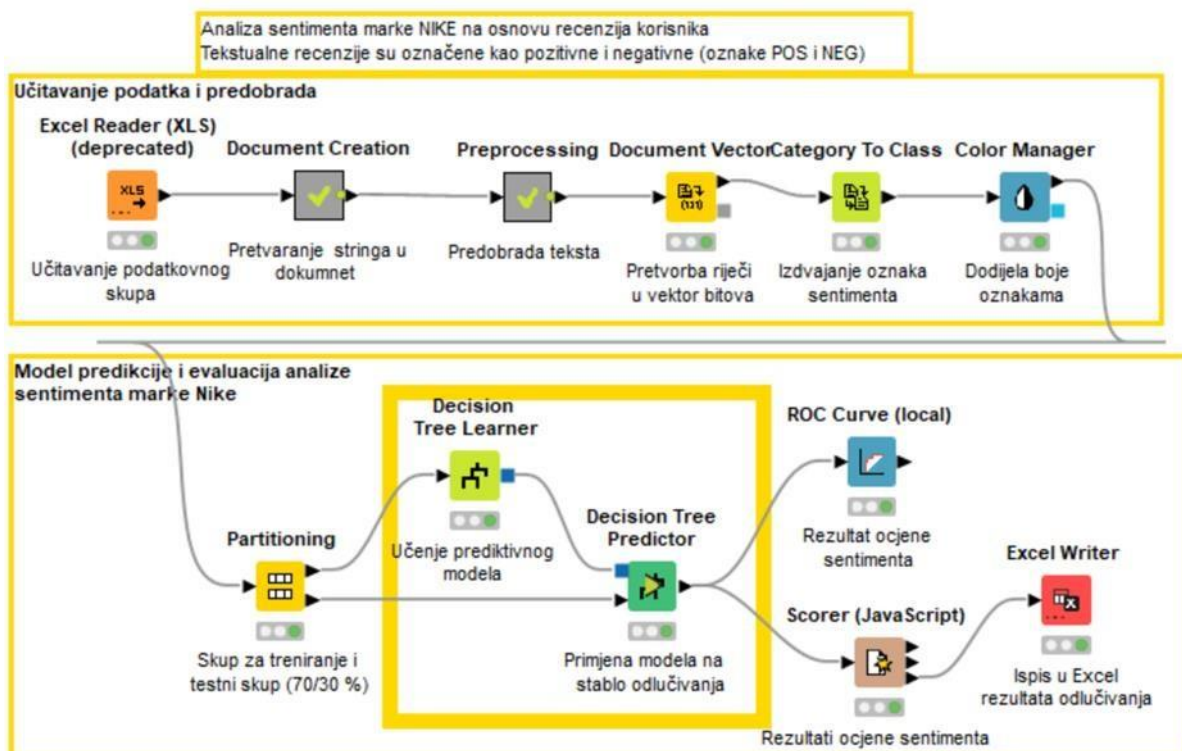
#### 4.5.4. Analiza sentimenta

S obzirom na to da je SBS analiza robnu marku Nike rangirala kao najznačajniju za tu marku je napravljena analiza sentimenta. Recenzije i ocjene iskustva korisnika Nike sportske obuće odnose se kako na kvalitetu proizvoda tako i na usluge u najvećoj mjeri online prodaje i korisničke podrške tvrtke Nike. Recenzije korisnika s ocjenama skinute su dvije s web stranica dviju organizacija koje se bave prikupljanjem mišljenja korisnika o različitim markama. To su ConsumerAffairs (2021) i SiteJabber (2021):

Za analizu sentimenta prikupljeno je 424 ocjena korisnika koje su kategorizirane na pozitivne (POS) i negativne (NEG). Klasifikacija je napravljena na osnovu ocjena korisnika tako da su ocjene 1, 2 i 3 kategorizirane kao negativne (NEG), a ocjene 4 i 5 kao pozitivne (POS).

Na osnovu toga kreirana je Excel odnosno CSV datoteka u kojoj je u jednom stupcu tekst komentara (recenzija) korisnika, a u drugom stupcu je klasa ocijene POS (pozitivna) ili NEG (negativna). Na osnovu tog skupa podataka pristupa se analizi sentimenta koji se temelji na tzv. binarnoj klasifikaciji primjenom stabla odlučivanja (**eng. decision tree**). Nakon učitavanja podatkovnog skupa radi se pred obrada i čišćenje teksta komentara. Nakon toga se podatkovni skup dijeli na dva skupa slučajnim odabirom na skup podatak za učenje i skup podataka za testiranje u omjeru 70/30 %. Ovaj postupak se temelji na nadziranom (**eng. supervised**) učenju jer su tekstualni komentari označeni (**eng. labeled**) i za to se koristi modul Decision Tree Learner. Po završetku procesa učenja, s testnim skupom podataka radi se predikcija na osnovu teksta recenzija ali bez oznaka (**eng. labels**) na osnovu naučenog modela i za to se koristi modul Decision Tree Predictor. Zadnji korak analize sentimenta je kvalitativna i kvantitativna statistička ocjena modela predviđanja na osnovu teksta recenzija korisnika koje nisu označene s kategorijama POS ili NEG.

KNIME model u obliku radnog procesa (**eng. workflow**) za analizu sentimenta prikazan je na slici 15. Programiranje je obavljeno grafički pomoću raspoloživih čvorova.



Slika 15 KNIME radni tijek

(izvor: rad autora temeljem provedenog istraživanja)



Nakon učitavanja Excel tablice s mišljenjima potrošača i oznakama POS ili NEG kreiraju se dokumenti koji idu na predobradu teksta. Potom se oznake POS i NEG pretvaraju u klase i tako obrađeni dokumenti idu čvor koji dokumente dijeli na one za treniranje (učenje) i one za testiranje kvalitete predikcije u odnosu 70/30 %.

Ulazne ocjene potrošača prikazane su na slici 16. Premda je u cjelini robna marka ocjenjena pozitivno razlika u odnosu na negativna mišljenja nije velika. Razlog tome je što je većina mišljenja usmjerena na usluge tvrtke Nike, a manje na proizvode i u mišljenjima potrošači nisu zadovoljni s razinom usluga vezanih uz online podršku korisnicima ili reklamacije.



Slika 16 Mišljenja potrošača

(izvor: rad autora temeljem provedenog istraživanja)

No kako su oznake POS i NEG mišljenjima dodijeljene po principu da su ocjene 1, 2 i 3 označene negativnim, a ocjene 4 i 5 pozitivnim napravljena je analiza prema modelu na slici X koji treba napraviti klasifikaciju i odrediti njezinu učinkovitost. Izlazni rezultati dobiju se na dva čvora, ROC curve i Scorer (statističke ocjene). Na slici 17 prikazana je matrica konfuzije za koju je rečeno da je vizualni pokazatelj koliko je model sposoban u predviđanju klasa na osnovu testnih promatranja.

## Scorer View

Confusion Matrix



	NEG (Predicted)		POS (Predicted)		
NEG (Actual)	TN	46	FP	16	74.19%
POS (Actual)	FN	12	TP	53	81.54%
	79.31%		76.81%		

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa ( $\kappa$ )	Correctly Classified	Incorrectly Classified
77.95%	22.05%	0.558	99	28

Slika 17 Matrica konfuzije

(izvor: rad autora temeljem provedenog istraživanja)

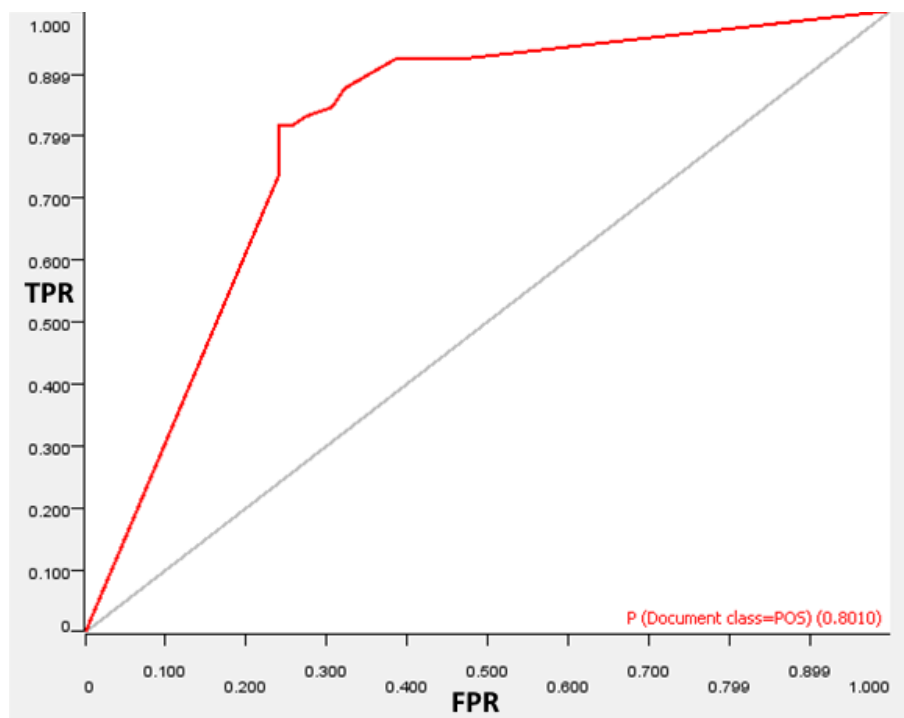
Statistički pokazatelji kvalitete predikcije na osnovu rezultata matrice konfuzije objašnjeni u točki 3.3.5 prikazani su u tablici 5.

Tablica 5 Statistički pokazatelji kvalitete predikcije

Label	Recall	Precision	Sensitivity	Specificity	F1-measure
NEG	0.741935	0.793103	0.741935	0.815385	0.766666667
POS	0.815385	0.768116	0.815385	0.741935	0.791044776

(izvor: rad autora temeljem provedenog istraživanja)

Slijedeća vizualna ocjena kvalitete je ROC krivulja prikazana na slici 18.



Slika 18 ROC krivulja na izlazu prediktora

(izvor: rad autora temeljem provedenog istraživanja)

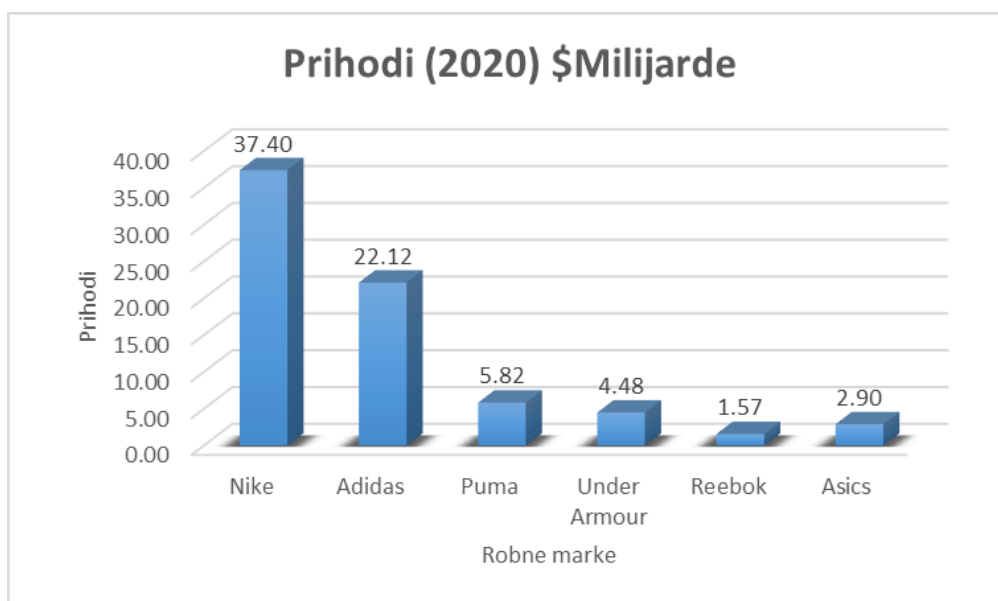
Zaključno, prethodni rezultati pokazuju sljedeće:

- Od ukupnog broja od 125 testnih (neoznačenih) mišljenja potrošača klasifikator je točno predvidio 99 slučajeva i 26 neispravnih slučajeva klasa POS i NEG.
- Ukupna točnost (korektnost) iznosi 77.95 % što je s obzirom na relativno mali broj mišljenja potrošača vrlo prihvatljiva vrijednost.
- Parametri Recall i Precision su također značajno visoki pa je i F1-mjera koja objedinjuje prethodne dvije mjere prihvatljiva (veća od 75 % za POS i NEG klase).
- ROC krivulja koja prikazuje promjenu True Positive Rate u odnosu na False Pozitive rate za klasu POS relativno brzo teži k jedinici i vjerojatnost dobre klasifikacije iznosi  $P = 0.80$ .

S obzirom na postignute rezultate može se reći da je model dobiven na izlazu Decision Tree Learner-a prihvatljivo dobar za predviđanje klasa POS ili NEG, te bi mogao poslužiti za automatsko označavanje (labeliranje) tekstualnih mišljenja potrošača iz drugih izvora (Twitter, Reddit).

#### 4.6. Potvrda postavljenih hipoteza istraživanja

Kao što je prethodno navedeno hipoteza H1 će se dokazivati prihodima robnih marki. Na slici 19 prikazani su prihodi istraživanih robnih marki za 2020. godinu.



Slika 19 Grafikon prihoda robnih marki za 2020. godinu

(izvor: <https://www.feedough.com/10-biggest-nike-competitors/>, 28.03.2022.)

Tablica 6 Tablica usporednih vrijednosti SBS ocjene i prihoda robnih marki

Marka	Originalne vrijednosti		Normalizirane vrijednosti	
	Prihodi (2020) \$Milijarde	SBS	Prihodi (2020) \$Milijarde	SBS
Nike	37.4	103.98	1.00	1.00
Adidas	22.12	63.69	0.57	0.59
Puma	5.82	17.32	0.12	0.12
Under Armour	4.48	15.41	0.08	0.10
Reebok	1.57	10.94	0.00	0.05
Asics	2.9	6.02	0.04	0.00

(izvor: rad autora temeljem provedenog istraživanja)

U tablici 6 prikazane su usporedne vrijednosti prihoda robnih marki i SBS ocjene za istraživane i rangirane robne marke. Već iz tablice vidi se korelacija između SBS ocjene i prihoda. No dodatno je napravljen i t-Test dva uzorka za srednju vrijednost. Kako su prihodi i SBS ocjena dobiveni iz dva različita izvora zbog razlika u skali nije moguće napraviti t-Test na originalnim podacima. Iz tog razloga prije određivanja t-Testa napravljena je normalizacija vrijednosti

pomoću formule  $X_{norm} = (X - X_{min}) / (X_{max} - X_{min})$ , gdje su  $X_{min}$  i  $X_{max}$  minimalna i maksimalna vrijednost iz tablice. Normalizirane vrijednosti su također prikazane u tablici 6.

t-Test je napravljen u Excelu (Data analysis) i rezultirao je vrijednostima prikazanim u tablici 7.

Tablica 7 t-Test dva uzorka za srednju vrijednost na osnovu normaliziranih vrijednosti

<i>t-Test: Paired Two Sample for Means</i>	<i>Variable Prihodi</i>	<i>Variable SBS</i>
Mean	0.302	0.308
Variance	0.161	0.160
Observations	6.000	6.000
<b>Pearson Correlation</b>	<b>0.997</b>	
Hypothesized Mean Difference	0.000	
df	5.000	
t Stat	-0.566	
<b>P(T&lt;=t) one-tail</b>	<b>0.298</b>	
t Critical one-tail	2.015	
<b>P(T&lt;=t) two-tail</b>	<b>0.596</b>	
t Critical two-tail	2.571	

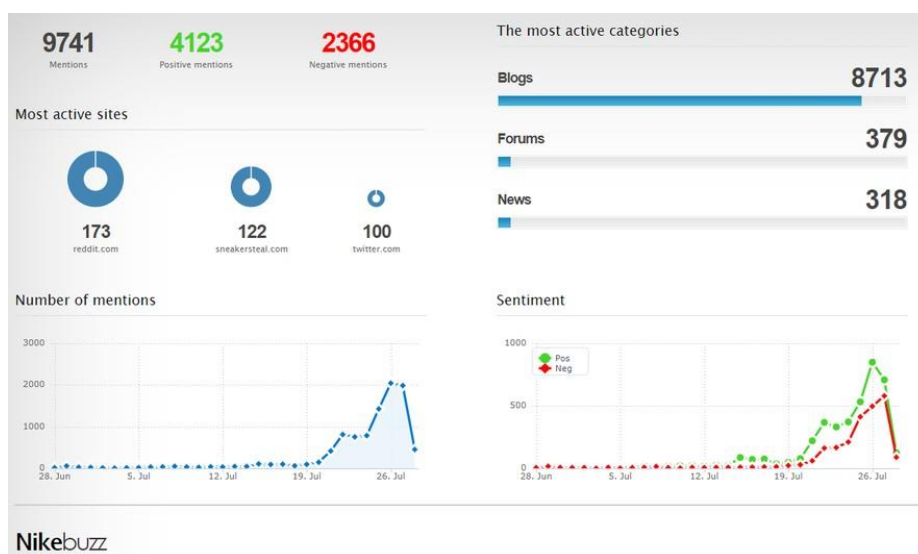
(izvor: rad autora temeljem provedenog istraživanja)

Vidi se da je p-vrijednost s jednostranim testom je 0.298, a p-vrijednost s dvostranim testom je 0.596. U oba slučaja, p-vrijednost je značajno veća od alfa vrijednosti 0.05. U ovom slučaju, p-vrijednost je veća od alfa vrijednosti, tako da je nulta hipoteza ( $H_0$ ) ISTINITA, tj. slabi su dokazi protiv nulte hipoteze. To znači da je to vrlo bliska podatkovna veza između dvije podatkovne skupine. Također i parametar Pearson Correlation koji iznosi 0.997 pokazuje visoku korelaciju između SBS ocjena značaja marke i finansijskih vrijednosti tj. prihoda za 2020. godinu tvrtki istraživanih robnih marki. Temeljem ovoga možemo potvrditi postavljen hipotezu  $H_1$ .

Za dokazivanje hipoteze  $H_2$  koristi se analiza sentimenta na osnovu mišljenja potrošača za robnu marku Nike. Analiza je pokazala da najbolje rangirana marka SBS metrikom ima pozitivne osjećaje u 54% posto slučajeva naspram 49% negativnih za proizvode i usluge robne marke. Provedena analiza klasifikacije pozitivnih i negativnih osjećaja spram marke Nike je

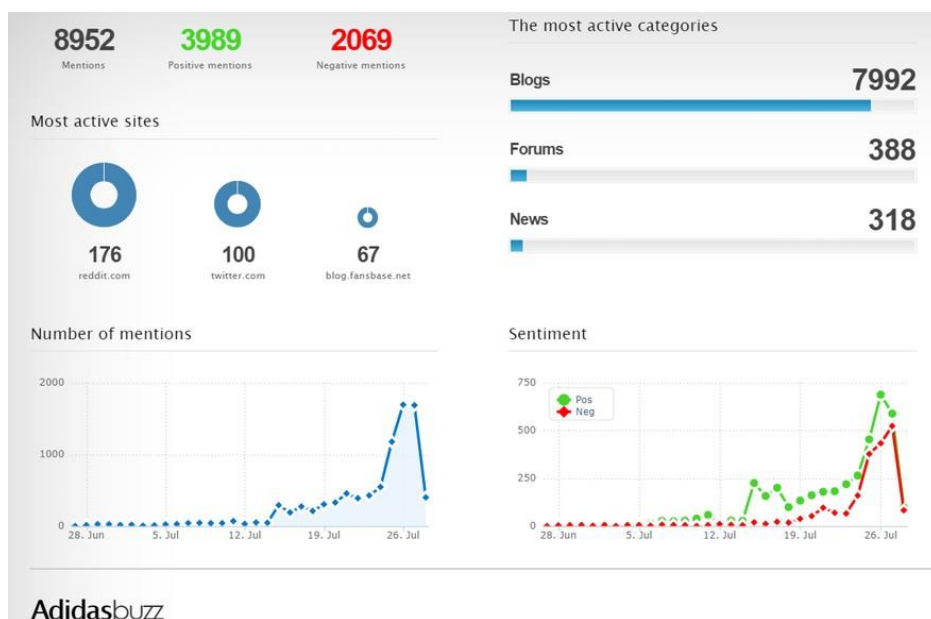
pokazala da se s visokom preciznošću (oko 80%) mogu ocijeniti mišljenja potrošača na osnovu tekstualnih mišljenja dostupnih na web stranicama koje se bave prikupljanjem mišljenja za različite robne marke.

S obzirom na to da nije rađena analiza sentimenta za ostale robne marke za usporedbu dvije najbolje rangirane marke Nike i Adidas korištena je analiza sentimenta tvrtke Brand24 za razdoblje 28.6. – 26.7. 2021. godine. Tvrtka Brand24 profesionalno se bavi analizama ocjena potrošača različitih robnih marki. Detaljnija izvješća nije moguće dobiti bez pretplate, tako da je na slikama 20 i 21 prikazan samo jedan fragment koji je bio besplatno dostupan.



Slika 20 Brand24 analiza sentimenta marke Nike

(izvor: Brand24, 2021)



Slika 21 Brand24 analiza sentimenta marke Adidas

(izvor: Brand24, 2021)

Za obje robne marke za istraživano razdoblje vidi se da je pozitivan sentiment veći od negativnog. Za robnu marku Nike on iznosi 63.5% pozitivnih naspram 36.5% negativnih, a za robnu marku Adidas 65.8% pozitivnih naspram 34.2% negativnih.

Zaključno možemo reći da se ovom analizom može potvrditi postavljena hipoteza H2, da robne marke koje imaju veći značaj dobiven SBS ocjenom imaju i pozitivna mišljenja potrošača o robnoj marki.

#### 4.7. Ograničenja istraživanja

Kao i druga istraživanja i ovo ima svoja ograničenja. Ona prvenstveno proizlaze iz činjenice da je istraživani korpus tekstualnih sadržaja relativno mali. Za postizanje još boljih rezultata NLP analiza poželjno je obuhvati nekoliko stotina pa i tisuća web stranica, posebno onih koje objavljuju različite poslovne sadržaje kao što su Bloomberg, Financial Times, Business Insider, MarketWatch, Forbes i drugi. No kako neki od ovi sadržaji nisu dostupni besplatno u ovom istraživanju nisu korišteni. Isto tako je korišten i ograničeni broj blogova s obzirom da većina od njih ne dozvoljava web scraping.

Slična situacija je i sa izvorima vezanim uz recenzije potrošača jer se relevantnim stranicama može pristupiti samo korištenjem API-a uz registraciju. Naravno ova ograničenja za

profesionalne organizacije koje se bave marketingom i brendiranjem nisu prepreka jer oni mogu uložiti značajna sredstva za pretplate relevantnih izvora na internetu.

Osim navedenih ograničenja, treba spomenuti i osobna ograničenja s obzirom na raspoloživu računalnu snagu jer su sve NLP analize koje su rađene računalno intenzivne kako po pitanju vremena izvođenja tako i raspoložive radne memorije.



## 5. ZAKLJUČAK

Rezultati provedenih istraživanja o značaju robnih marki sportske obuće potvrdili su postavljene ciljeve da primjena NLP metoda za analizu velikih tekstualnih sadržaja dostupnih na internetu dokazuje da novi način metrike i vrednovanja značaja robnih marki može u značajnoj mjeri poboljšati postojeće metode vrednovanja robnih marki. S obzirom na velike količine tekstualnih sadržaja na web stranicama i društvenim medijima mogućnosti prikupljanja informacija o stavovima potrošača i javnosti vezanim uz robne marke sportske obuće proširuju se u neograničenim razmjerima. U prvom koraku NLP metode analize TF-IDF i modeliranje tema iznjedrile su iz korpusa tekstualnih sadržaja o sportskoj obući najznačajnije robne marke: Nike, Adidas, Puma, Under Armour, Reebok i Asics. Nakon toga je za navedene robne marke napravljena semantička analiza značaj primjenom inovativne metrike Semantic Brand Score (SBS) koja je rezultirala rang listom navedenih šest marki. Na kraju je za najbolje rangiranu marku Nike napravljena analiza sentimenta na osnovu mišljenja potrošača o proizvodima i uslugama tvrtke Nike. Dobiveni rezultati su korišteni za potvrđivanje postavljenih hipoteza istraživanja. Provedene analize potvrđuju da NLP metode istraživanja tržišta i robnih marki mogu biti komplementarne klasičnim metodama kao što su financijska izvješća, upitnici i intervjui i kombiniranjem obaju pristupa dobije se sinergijski učinak za poboljšanja u području brendiranja, odnosa s javnosti i marketinških strategija vezanim uz robne marke sportske obuće. Temelje iznesenih činjenica vezanih uz upravljanje markom postavili su Aaker i Keller devedesetih godina što je diskutirano u točki 2. i u današnje doba ih marketinški stručnjaci koriste za razvoj brendova. No kako je digitalna tranzicija zahvatila sve segmente društva zbog naglog razvoja interneta i društvenih medija došlo je do značajnih pomaka u marketinškim komunikacijama između tvrtki i potrošača. Iz tih razloga u svakodnevicu ušle metode koje se temelje na analizi teksta kao povratna informacija od strane potrošača (mišljenja i ocjene brendova na društvenim mrežama i blogovima). S obzirom na veliku količinu informacija u tekstualnim sadržajima na internetu, analizu teksta nije više moguće provoditi „pješke“ od strane ljudi već se za te namjene moraju koristiti računala. Zato se zadnjih desetak godina intenzivno razvijaju metode utemeljene na obradi prirodnog jezika (Natural Language Processing – NLP) kao grane umjetne inteligencije. Primijenjene NLP tehnike u ovom istraživanju kao što su frekvencijska analiza, SBS analiza, analiza sentimenta i tematsko modeliranje mogu se koristiti za dublje razumijevanje stavova potrošača prema marki, njenim proizvodima ili uslugama i njihovim marketinškim naporima. Dok se Aakerovi i Kellerovi modeli usredotočuju na

unutarnje i vanjske komponente marke, NLP metode pružaju pristup upravljanju markom koji se više temelji na podacima koji mogu nadopuniti i poboljšati te modele. Analizirajući recenzije potrošača, objave na društvenim mrežama i druge izvore povratnih informacija kupaca, NLP može pružiti uvid u učinkovitost marketinških napora u razvoju marke, kvalitetu njegovih proizvoda ili usluga i cjelokupni osjećaj prema brendu. NLP metode mogu se koristiti za analizu komunikacije potrošača na društvenim mrežama i izvođenje metrika za mjerenje ključnih vrijednosti marke, kao što su svijest o marki, sentiment i angažman.

Kombiniranje istraživanja važnosti brendova pomoću NLP metoda s klasičnim modelima Aakera i Kellera može se sažeti u nekoliko sljedećih koraka:

- Prikupiti tekstualne sadržaje relevantne za ciljeve istraživanja iz različitih izvora kao što su platforme društvenih medija, blogovi i internetski forumi, koristeći alate za scraping weba i napraviti sva potrebne čišćenja i pred obrade tekstova za NLP analizu.
- Primijenite NLP tehnike kao što su prepoznavanje entiteta, modeliranje tema, analiza sentimenta, semantička SBS analiza. Analiza sentimenta pomaže u prepoznavanju pozitivnih i negativnih osjećaja povezanih s markom. Modeliranje tema pomaže u identificiranju tema o kojima se raspravlja u tekstualnim sadržajima koji se odnose na marku. Prepoznavanje entiteta pomaže u identificiranju entiteta spomenutih u tekstualnim podacima koji se odnose na marku, kao što su proizvodi i konkurenti. SBS analiza daje rang listu ocjena značaja istraživanih robnih marki u tekstualnim sadržajima na internetu.
- Koristiti klasične modele Aakera i Kellera radi analize rezultata dobivenih NLP tehnikama. Aakerov model fokusira se na identitet marke i uključuje osobnost marke, viziju marke i kulturu marke. Kellerov model usredotočen je na kapital marke i uključuje svijest o marki, lojalnost marki i povezanost s markom.
- Kombinirati rezultate NLP tehnika i klasičnih modela Aakera i Kellera radi dobivanja šireg uvida u važnost marke. Na osnovu toga identificirati područja snage i slabosti marke, koja se mogu koristiti za razvoj marketinških strategija i poboljšanje učinka brendiranja marke.
- Upotrijebiti uvide stečene istraživanjem za razvoj marketinških operacija koje su u skladu sa značajem, identitetom i kapitalom marke. Unaprijediti strategije za poboljšanje svijesti o marki, lojalnosti i asocijacijama.
- Upotrijebite digitalne marketinške kanale kao što su društveni mediji i online oglašavanje za promicanje robne marke i interakciju s ciljanom publikom.

Zaključno, kombiniranjem NLP tehnika s klasičnim marketinškim i modelima brendiranja, tvrtke mogu steći dublji uvid u važnost svoje robne marke i razviti učinkovite marketinške operacije koje značajno poboljšavaju performanse marke.

## Popis literature

### Knjige:

1. Balakrishna V.K. (1997), *Theory And Problems Of Graph Theory*, The McGraw-Hill Companies, Inc., ISBN 0-07-005489-4.
2. Beysolow Taweh II (2018), *Applied Natural Language Processing with Python*, publisher Apress, ISBN-13 (electronic): 978-1-4842-3733-5.
3. Géron Aurélien (2019), *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow*, Published by O'Reilly Media, ISBN: 9781492032649.
4. Heding, T., Knudtzen, C. F. & Bjerre, M. (2009). *Brand management: Research, theory and practice*. Routledge, ISBN 0-203-99617-8 Master e-book
5. Keller, Lane, K. (2013), *Strategic brand management* (4th ed.), Pearson Education Limited, ISBN: 13: 978-0-273-77941-4
6. Kotler, P., Pföertsch, W., & Michi, I. (2006). *B2B brand management*. In *B2B Brand Management*, Springer, ISBN-10 3-540-25360-0.
7. Kotler, P., Keller, K., (2009). *Marketing Management*, Pearson Education Limited, ISBN 13: 978-1-292-092 -71-3 (PDF).
8. Nitin Hardeniya (2015), *NLTK Essentials*, Packt Publishing Ltd., ISBN 978-1-78439-690-9
9. Ries, A., & Ries, L. (2002), *The 22 immutable laws of branding*, New York: HarperCollins, ISBN 0-06-008518-5.
10. Sarkar Dipanjan (2016), *Text Analytics with Python: A Practical Real-World Approach to Gaining Actionable insights from Your Data*, Apress, ISBN-13 (electronic): 978-1-4842-2388-8.

### Članci:

1. Aaker, David A. (1996). Building strong brands, New York: *The Free Press*, ISBN-0-02-900151-X, pp. 7 - 25.
2. Blei, David M.; Ng, Andrew Y.; Jordan, Michael I (2003). Lafferty, John (ed.). "Latent Dirichlet Allocation". *Journal of Machine Learning Research*. 3 (4–5): pp. 993–1022. doi:10.1162/jmlr.2003.3.4-5.993, 12.6.2020
3. Colladon Fronzetti, A. (2018). The Semantic Brand Score. *Journal of Business Research*, 88(September 2017), 150–160. <https://doi.org/10.1016/j.jbusres.2018.03.026>

4. Colladon Fronzetti, A. (2020). Forecasting election results by studying brand importance in online news. *International Journal of Forecasting*, 36(2), 414–427. <https://doi.org/10.1016/j.ijforecast.2019.05.013> 21.05.2023.
5. Kauffmann Erick, Peral Jesús, Gil David, Ferrández Antonio , Sellers Ricardo and Higinio Mora, (2019), Managing Marketing Decision-Making with Sentiment Analysis: An Evaluation of the Main Product Features Using Text Data Mining, *Sustainability 11*, 4235, doi:10.3390/su11154235.
6. Keller, K. L. (2003). Understanding brands, branding and brand equity. *Interactive Marketing Vol. 5, No. 1*, pp. 7–20, <https://doi.org/10.1057/palgrave.im.4340213> , 21.05.2023.
7. Keller, K. L. (2009). Building strong brands in a modern marketing communications environment. *Journal of Marketing Communications*, 15(2–3), 139–155. <https://doi.org/10.1080/13527260902757530> , 21.05.2023.
8. Keller, K. L. (1993). Conceptualizing, Measuring, and Managing Customer-Based Brand Equity. *Journal of Marketing*, 57(1), 1–22.
9. Lihong Shen, 2012, Context and Text, *Theory and Practice in Language Studies*, Vol. 2, No. 12, pp. 2663-2669, ISSN 1799-2591. , 21.05.2023.
10. Slamić Tarade Sara (2020), Istraživanje branda korištenjem analize teksta i konteksta, *Polytechnic & Design*, Vol. 8, No. 2, DOI: 10.19279/TVZ.PD.2020-8-2-02.
11. Wonsever Dina, Minel Jean-Luc (2001), Contextual Rules for Text Analysis, Conference Paper, Lecture Notes in Intelligence Artificial, DOI: 10.1007/3-540-44686-9\_52.

Web stranice:

1. Brand24, <https://brand24.com/> , pristupljeno 27.7.2021.
2. Colladon Fronzetti, A. (2019). Calculating the Semantic Brand Score with Python. Published in Towards Data Science, <https://towardsdatascience.com/calculating-the-semantic-brand-score-with-python-3f94fb8372a6> , 27.12.2022.
3. ConsumerAffairs (2021), [https://www.consumeraffairs.com/sporting\\_goods/nike.html](https://www.consumeraffairs.com/sporting_goods/nike.html), 29. 10.2021.
4. InstagramStats&Analyticsforadidas (2023), <https://hypeauditor.com/instagram/adidas/>, 30.3.2023.
5. Interbrand (2022), Best Global Brands, <https://www.rankingthebrands.com/The-Brand-Rankings.aspx?rankingID=37>, 28.01.2023.

6. KNIME Analytics Platform (2023), <https://www.knime.com/software-overview> ,  
10.01.2023.
7. SiteJabber (2021), <https://www.sitejabber.com/reviews/nike.com>, 29.10.2021.

## Popis slika

Slika 1 Frekvencija pojavnosti riječi .....	23
Slika 2 Primjer grafa.....	28
Slika 3 Matrica konfuzije (izvor: autor) .....	32
Slika 4 Primjer ROC krivulje (izvor: Géron, 2019.) .....	34
Slika 5 Predobrada teksta .....	39
Slika 6 TF-IDF model .....	40
Slika 7 TF-IDF oblak riječi .....	40
Slika 8 TF-IDF kumulativno .....	41
Slika 9 Model za tematsko modeliranje pomoću LDA metode .....	42
Slika 10 Postavljanje parametara za LDA analizu .....	42
Slika 11 Oblak riječi u temama .....	43
Slika 12 Graf riječi prije filtriranja .....	45
Slika 13 Izgled grafa nakon filtriranja za marku Asics .....	46
Slika 14 SBS ocjene semantičke analize .....	47
Slika 15 KNIME radni tijek .....	48
Slika 16 Mišljenja potrošača .....	49
Slika 17 Matrica konfuzije .....	50
Slika 18 ROC krivulja na izlazu prediktora .....	51
Slika 19 Grafikon prihoda robnih marki za 2020. godinu (izvor: <a href="https://www.feedough.com/10-biggest-nike-competitors/">https://www.feedough.com/10-biggest-nike-competitors/</a> ) .....	52
Slika 20 Brand24 analiza sentimenta marke Nike (izvor: Brand24, 2021) .....	54
Slika 21 Brand24 analiza sentimenta marke Adidas (izvor: Brand24, 2021) .....	55

## Popis tablica

Tablica 1. Fragment web stranica s kojih su prikupljeni sadržaji .....	2
Tablica 2. LDA teme od 1 do 3 .....	43
Tablica 3.LDA teme od 4 do 6 .....	44
Tablica 4 Ocjene sve tri dimenzije i ukupna SBS ocjena.....	46
Tablica 5 Statistički pokazatelji kvalitete predikcije.....	50
Tablica 6 Tablica usporednih vrijednosti SBS ocjene i prihoda robnih marki.....	52
Tablica 7 t-Test dva uzorka za srednju vrijednost na osnovu normaliziranih vrijednosti.....	53

## Prilog 1

ID	Naslov Web stranice	Poveznica
	<a href="#">2</a>	<a href="#">3</a>
ID01	<a href="#">Top 10 Shoe Companies in the World 2020   Top Shoe Brands   Footwear Market Report - Technavio</a>	<a href="https://blog.technavio.com/blog/top-10-shoe-companies">https://blog.technavio.com/blog/top-10-shoe-companies</a>
ID02	<a href="#">20 Best Running Shoes Brands to Know - The Trend Spotter</a>	<a href="https://www.thetrendspotter.net/best-running-shoes-brands/">https://www.thetrendspotter.net/best-running-shoes-brands/</a>
ID03	<a href="#">Top 15 Sports Shoes Brands in the US 2019   Athletic Footwear Market - Technavio</a>	<a href="https://blog.technavio.com/blog/top-15-vendors-athletic-footwear-market-us">https://blog.technavio.com/blog/top-15-vendors-athletic-footwear-market-us</a>
ID04	<a href="#">The Best Sportswear Brands In The World Today: 2021 Edition</a>	<a href="https://www.apetogentleman.com/best-sportswear-brands/">https://www.apetogentleman.com/best-sportswear-brands/</a>
ID05	<a href="#">Top 10 sportswear popular brands in the world of sports</a>	<a href="https://www.sportzbusiness.com/top-10-sportswear-popular-brands-in-the-world-of-sports/">https://www.sportzbusiness.com/top-10-sportswear-popular-brands-in-the-world-of-sports/</a>
ID06	<a href="#">Top 10 Biggest Sportswear Brands in the World - All Top Everything</a>	<a href="https://www.alltopeverything.com/top-10-sportswear-brands/?amp">https://www.alltopeverything.com/top-10-sportswear-brands/?amp</a>
ID07	<a href="#">The Most Influential Athletes with Athletic Apparel Endorsement Deals   Opendorse</a>	<a href="https://opendorse.com/blog/the-most-influential-athletes-with-athletic-apparel-endorsement-deals/">https://opendorse.com/blog/the-most-influential-athletes-with-athletic-apparel-endorsement-deals/</a>
ID08	<a href="#">Nike vs. Adidas vs. Under Armour: Which has the best product pages?   Econsultancy</a>	<a href="https://econsultancy.com/nike-vs-adidas-vs-under-armour-which-has-the-best-product-pages/">https://econsultancy.com/nike-vs-adidas-vs-under-armour-which-has-the-best-product-pages/</a>
ID09	<a href="#">Nike, Adidas, and Under Armour Make Big Changes. What Does It Mean?   Complex</a>	<a href="https://www.complex.com/sneakers/2019/10/nike-adidas-under-armour-executive-changes-analysis">https://www.complex.com/sneakers/2019/10/nike-adidas-under-armour-executive-changes-analysis</a>
ID10	<a href="#">Why Under Armour Stock Gained 10% in June   The Motley Fool</a>	<a href="https://www.fool.com/investing/2019/07/06/why-under-armour-stock-gained-10-in-june.aspx?source=awin&amp;awc=12195_1622759808_34e247d319c40c11dcd2f820867ce0c&amp;utm_source=aw&amp;utm_campaign=78888">https://www.fool.com/investing/2019/07/06/why-under-armour-stock-gained-10-in-june.aspx?source=awin&amp;awc=12195_1622759808_34e247d319c40c11dcd2f820867ce0c&amp;utm_source=aw&amp;utm_campaign=78888</a>
ID11	<a href="#">The best activewear brands and luxury gym clothes for men in 2021   Health   Wellbeing   Luxury London</a>	<a href="https://luxurylondon.co.uk/wellbeing/health/the-best-activewear-brands-most-stylish-gym-clothes-workout-gear-for-men">https://luxurylondon.co.uk/wellbeing/health/the-best-activewear-brands-most-stylish-gym-clothes-workout-gear-for-men</a>
ID12	<a href="#">How Nike became No. 1. And how it plans to stay there</a>	<a href="https://www.cnn.com/2019/06/21/how-nike-became-no-1-and-how-it-plans-to-stay-there.html">https://www.cnn.com/2019/06/21/how-nike-became-no-1-and-how-it-plans-to-stay-there.html</a>



ID13	<a href="#">Sports Apparel Online Retailing Market Drive Big Growth   Nike, Adidas, PUMA, Under Armour - Wall Street Call</a>	<a href="https://reportedtimes.com/sports-apparel-online-retailing-market-drive-big-growth-nike-adidas-puma-under-armour/">https://reportedtimes.com/sports-apparel-online-retailing-market-drive-big-growth-nike-adidas-puma-under-armour/</a>
ID14	<a href="#">FDRA   Under Armour surpasses Adidas to become No. 2 sports brand</a>	<a href="https://fdra.org/latest-news/under-armour-surpasses-adidas-to-become-no-2-sports-brand/">https://fdra.org/latest-news/under-armour-surpasses-adidas-to-become-no-2-sports-brand/</a>
ID15	<a href="#">Puma's Business Is Booming— But Everybody's Talking About Nike &amp; Adidas – Footwear News</a>	<a href="https://footwearnews.com/2019/business/features/puma-vs-nike-adidas-under-armour-sales-1202811310/">https://footwearnews.com/2019/business/features/puma-vs-nike-adidas-under-armour-sales-1202811310/</a>
ID16	<a href="#">Nike vs Adidas: Who's winning?   Data Science Blog</a>	<a href="https://nycdatascience.com/blog/student-works/web-scraping/nike-vs-adidas-whos-winning/">https://nycdatascience.com/blog/student-works/web-scraping/nike-vs-adidas-whos-winning/</a>
ID17	<a href="#">Exploration and Analysis of Nike Shoes   Data Science Blog</a>	<a href="https://nycdatascience.com/blog/student-works/exploration-and-analysis-of-nike-shoes/">https://nycdatascience.com/blog/student-works/exploration-and-analysis-of-nike-shoes/</a>
ID18	<a href="#">How Ethical and Sustainable Are Your Favourite Sportswear brands?   The Green Hub</a>	<a href="https://thegreenhubonline.com/2019/09/18/how-ethical-and-sustainable-are-your-favourite-sportswear-brands/">https://thegreenhubonline.com/2019/09/18/how-ethical-and-sustainable-are-your-favourite-sportswear-brands/</a>
ID19	<a href="#">Top 10 Sports Brands that Use Salesforce Commerce Cloud - ArganoUV (United Virtualities) – Tech Consultants &amp; Salesforce Commerce Cloud Experts</a>	<a href="https://weareuv.com/top-10-sports-brands-that-use-salesforce-commerce-cloud/">https://weareuv.com/top-10-sports-brands-that-use-salesforce-commerce-cloud/</a>
ID20	<a href="#">Sports brands, don't forget that women love sneakers too!</a>	<a href="https://www.heuritech.com/blog/articles/sneakers-are-for-women-too/">https://www.heuritech.com/blog/articles/sneakers-are-for-women-too/</a>
ID21	<a href="#">Top 25 Sport Shoe Brands In The World • The Good Look Book</a>	<a href="https://cashkaro.com/thegoodlookbook/top-sport-shoe-brands-in-the-world/">https://cashkaro.com/thegoodlookbook/top-sport-shoe-brands-in-the-world/</a>
ID22	<a href="#">Top Nike Competitors - Competitors Analysis of Nike Inc   Marketing Tutor</a>	<a href="https://www.marketingtutor.net/top-nike-competitors/">https://www.marketingtutor.net/top-nike-competitors/</a>
ID23	<a href="#">Top 10 Richest Sports Brands In The World Right Now - Sports Show</a>	<a href="https://sportsshow.net/richest-sports-brands/">https://sportsshow.net/richest-sports-brands/</a>
ID24	<a href="#">SPORTS DRIPS: Major brands trying to save the planet</a>	<a href="https://www.iol.co.za/sport/opinion/sports-drips-major-brands-trying-to-save-the-planet-463c223d-5d65-4e4b-ae5a-7708daec3760">https://www.iol.co.za/sport/opinion/sports-drips-major-brands-trying-to-save-the-planet-463c223d-5d65-4e4b-ae5a-7708daec3760</a>
ID25	<a href="#">Sportswear Brand Wars - Q3 2019 Results and Analysis for Nike, Adidas, Puma, Under Armour and Skechers</a>	<a href="https://www.linkedin.com/pulse/sportswear-brand-wars-q3-2019-results-analysis-nike-siddharth-patil/">https://www.linkedin.com/pulse/sportswear-brand-wars-q3-2019-results-analysis-nike-siddharth-patil/</a>
ID26	<a href="#">The Best Sportswear Brands In The World   FashionBeans</a>	<a href="https://www.fashionbeans.com/article/best-sportswear-brands/">https://www.fashionbeans.com/article/best-sportswear-brands/</a>
ID27	<a href="#">Nike, Adidas, and Puma Flagship Sportswear Stores Compared</a>	<a href="https://www.businessinsider.com/nike-adidas-and-puma-flagship-stores-compared-photos-details-2019-8">https://www.businessinsider.com/nike-adidas-and-puma-flagship-stores-compared-photos-details-2019-8</a>

ID28	<a href="#">Adidas To Unload Reebok And Focus On Own Brand</a>	<a href="https://www.forbes.com/sites/roberthart/2021/02/16/adidas-to-unload-reebok-and-focus-on-own-brand/">https://www.forbes.com/sites/roberthart/2021/02/16/adidas-to-unload-reebok-and-focus-on-own-brand/</a>
ID29	<a href="#">Nike Vs Adidas, The Battle Is Still On! But Which One Is Better?</a>	<a href="https://travel.earth/adidas-vs-nike/amp/">https://travel.earth/adidas-vs-nike/amp/</a>
ID30	<a href="#">Clash of the Titans: Adidas and Nike continue to battle it out</a>	<a href="https://money.cnn.com/2018/04/16/news/companies/adidas-nike-sportswear/index.html">https://money.cnn.com/2018/04/16/news/companies/adidas-nike-sportswear/index.html</a>
ID31	<a href="#">The Most Popular Shoes And Brands Worn By Players Around The NBA - 2020 Edition   Baller Shoes DB</a>	<a href="https://ballershoesdb.com/blog/the-most-popular-shoes-and-brands-worn-by-players-around-the-nba-2020-edition/">https://ballershoesdb.com/blog/the-most-popular-shoes-and-brands-worn-by-players-around-the-nba-2020-edition/</a>
ID32	<a href="#">The Top 3 Brands in Sportswear- Adidas, Nike, Puma</a>	<a href="https://www.sportsleisurewear.com/The-Top-3-Brands-in-Sportswear.html">https://www.sportsleisurewear.com/The-Top-3-Brands-in-Sportswear.html</a>
ID33	<a href="#">Top 10 Best Branded Sports Shoes for Men [2021 Updated] - ZestMoney</a>	<a href="https://www.zestmoney.in/shop-smartly/branded-sport-shoes-for-men-india">https://www.zestmoney.in/shop-smartly/branded-sport-shoes-for-men-india</a>
ID34	<a href="#">Under Armour Overtakes Adidas in U.S. Sportswear Market - WSJ</a>	<a href="https://www.wsj.com/articles/under-armour-overtakes-adidas-in-u-s-sportswear-market-1420753934">https://www.wsj.com/articles/under-armour-overtakes-adidas-in-u-s-sportswear-market-1420753934</a>
ID35	<a href="#">Nike Outranks Adidas, Under Armour And Lululemon Where It Counts Most: Consumer Perception</a>	<a href="https://www.forbes.com/sites/pamdanziger/2019/05/23/nike-outranks-adidas-under-armour-and-lululemon-where-it-counts-most-consumer-perception/?sh=75b6a7b42478">https://www.forbes.com/sites/pamdanziger/2019/05/23/nike-outranks-adidas-under-armour-and-lululemon-where-it-counts-most-consumer-perception/?sh=75b6a7b42478</a>
ID36	<a href="#">Winning The Sports Retail Race: Under Armour And Nike Hit The Wall</a>	<a href="https://www.forbes.com/sites/pamdanziger/2017/11/01/winning-the-sports-retail-race-under-armour-and-nike-hit-the-wall/?sh=78286e471a96">https://www.forbes.com/sites/pamdanziger/2017/11/01/winning-the-sports-retail-race-under-armour-and-nike-hit-the-wall/?sh=78286e471a96</a>
ID37	<a href="#">Adidas Is Giving Nike A Run For Its Money</a>	<a href="https://www.forbes.com/sites/greatspeculations/2019/08/06/adidas-is-giving-nike-a-run-for-its-money/?sh=39893f422e00">https://www.forbes.com/sites/greatspeculations/2019/08/06/adidas-is-giving-nike-a-run-for-its-money/?sh=39893f422e00</a>
ID38	<a href="#">Adidas Steals The Buzz From Nike And Under Armour</a>	<a href="https://www.forbes.com/sites/panosmourdukoutas/2019/05/05/adidas-steals-the-buzz-from-nike-and-under-armour/?sh=1a19a8377764">https://www.forbes.com/sites/panosmourdukoutas/2019/05/05/adidas-steals-the-buzz-from-nike-and-under-armour/?sh=1a19a8377764</a>
ID39	<a href="#">Adidas or Nike? Which Retail Giant is Winning The Sneakers War?</a>	<a href="https://www.forbes.com/sites/callyrussell/2019/08/22/adidas-or-nike-which-retail-giant-is-winning-the-sneakers-war/?sh=7bda1ed812b7">https://www.forbes.com/sites/callyrussell/2019/08/22/adidas-or-nike-which-retail-giant-is-winning-the-sneakers-war/?sh=7bda1ed812b7</a>
ID40	<a href="#">Why Nike Will Outpace The Sports Apparel Market's Growth</a>	<a href="https://www.forbes.com/sites/greatspeculations/2013/05/13/why-nikes-growth-will-outpace-the-sports-apparel-markets/?sh=7fda94367380">https://www.forbes.com/sites/greatspeculations/2013/05/13/why-nikes-growth-will-outpace-the-sports-apparel-markets/?sh=7fda94367380</a>

ID41	<a href="#">Nike, Adidas, Puma, Under Armour: Sizing up the sportswear giants in 2020 - SportsPro Media</a>	<a href="https://www.sportspromedia.com/analysis/nike-adidas-puma-under-armour-sportswear-profit-loss-annual-reports-2020">https://www.sportspromedia.com/analysis/nike-adidas-puma-under-armour-sportswear-profit-loss-annual-reports-2020</a>
ID42	<a href="#">Uniqlo, Fila, Adidas Lead The ATP Finals Sponsorship Race</a>	<a href="https://www.forbes.com/sites/timnewcomb/2018/11/09/uniqlo-fila-adidas-lead-the-atp-finals-sponsorship-race/?sh=7cbefb8c12e8">https://www.forbes.com/sites/timnewcomb/2018/11/09/uniqlo-fila-adidas-lead-the-atp-finals-sponsorship-race/?sh=7cbefb8c12e8</a>
ID43	<a href="#">Adidas or Nike? Which Retail Giant is Winning The Sneakers War?</a>	<a href="https://www.forbes.com/sites/callyrussell/2019/08/22/adidas-or-nike-which-retail-giant-is-winning-the-sneakers-war/?sh=4a9f8dae12b7">https://www.forbes.com/sites/callyrussell/2019/08/22/adidas-or-nike-which-retail-giant-is-winning-the-sneakers-war/?sh=4a9f8dae12b7</a>
ID44	<a href="#">The Forbes Fab 40: Puma Debuts On 2019 List Of The World's Most Valuable Sports Brands</a>	<a href="https://www.forbes.com/sites/mikeozanian/2019/10/16/the-forbes-fab-40-puma-debuts-on-2019-list-of-the-worlds-most-valuable-sports-brands/?sh=39101ca5d356">https://www.forbes.com/sites/mikeozanian/2019/10/16/the-forbes-fab-40-puma-debuts-on-2019-list-of-the-worlds-most-valuable-sports-brands/?sh=39101ca5d356</a>
ID45	<a href="#">Global Brands - adidas Annual Report 2020</a>	<a href="https://report.adidas-group.com/2020/en/group-management-report-our-company/global-brands.html">https://report.adidas-group.com/2020/en/group-management-report-our-company/global-brands.html</a>
ID46	<a href="#">The Best Sport-Specific Shoe Designs For Every Type Of Activity</a>	<a href="https://www.forbes.com/sites/timnewcomb/2020/08/03/the-best-sport-specific-shoe-designs-for-every-type-of-activity/?sh=2b96293e1935">https://www.forbes.com/sites/timnewcomb/2020/08/03/the-best-sport-specific-shoe-designs-for-every-type-of-activity/?sh=2b96293e1935</a>
ID47	<a href="#">Most Valuable Sports Business Brands Worldwide</a>	<a href="https://finance.yahoo.com/news/most-valuable-sports-business-brands-090717745.html?guccounter=1">https://finance.yahoo.com/news/most-valuable-sports-business-brands-090717745.html?guccounter=1</a>
ID48	<a href="#">Nike, Adidas, Puma, Under Armour: Sizing up the sportswear giants in 2020 - SportsPro Media</a>	<a href="https://www.sportspromedia.com/analysis/nike-adidas-puma-under-armour-sportswear-profit-loss-annual-reports-2020">https://www.sportspromedia.com/analysis/nike-adidas-puma-under-armour-sportswear-profit-loss-annual-reports-2020</a>
ID49	<a href="#">How Nike used Brand Equity to become the Best Sports Apparel Brand Ever</a>	<a href="https://medium.com/@mariahparker/how-nike-used-brand-equity-to-become-the-best-sports-apparel-brand-ever-42e8c0cd553">https://medium.com/@mariahparker/how-nike-used-brand-equity-to-become-the-best-sports-apparel-brand-ever-42e8c0cd553</a>
ID50	<a href="#">The Top 3 Brands in Sportswear</a>	<a href="https://www.sportsleisurewear.com/The-Top-3-Brands-in-Sportswear.html">https://www.sportsleisurewear.com/The-Top-3-Brands-in-Sportswear.html</a>
ID51	<a href="#">The Best Running Shoes of 2021</a>	<a href="https://www.esquire.com/uk/style/shoes/g24739613/best-mens-running-shoes/">https://www.esquire.com/uk/style/shoes/g24739613/best-mens-running-shoes/</a>
ID52	<a href="#">Top 10 Shoe Companies in the World 2020   Top Shoe Brands   Footwear Market Report - Technavio</a>	<a href="https://blog.technavio.com/blog/top-10-shoe-companies">https://blog.technavio.com/blog/top-10-shoe-companies</a>
ID53	<a href="#">21 Best Running Shoes Brands to Know - The Trend Spotter</a>	<a href="https://www.thetrendspotter.net/best-running-shoes-brands/">https://www.thetrendspotter.net/best-running-shoes-brands/</a>

ID54	<a href="https://blog.technavio.com/blog/top-15-vendors-athletic-footwear-market-us">Top 15 Sports Shoes Brands in the US 2019   Athletic Footwear Market - Technavio</a>	<a href="https://blog.technavio.com/blog/top-15-vendors-athletic-footwear-market-us">https://blog.technavio.com/blog/top-15-vendors-athletic-footwear-market-us</a>
ID55	<a href="https://www.apetogentleman.com/best-sportswear-brands/">The Best Sportswear Brands In The World Today: 2021 Edition</a>	<a href="https://www.apetogentleman.com/best-sportswear-brands/">https://www.apetogentleman.com/best-sportswear-brands/</a>
ID56	<a href="https://www.sportzbusiness.com/top-10-sportswear-popular-brands-in-the-world-of-sports/">Top 10 sportswear popular brands in the world of sports</a>	<a href="https://www.sportzbusiness.com/top-10-sportswear-popular-brands-in-the-world-of-sports/">https://www.sportzbusiness.com/top-10-sportswear-popular-brands-in-the-world-of-sports/</a>
ID57	<a href="https://www.alltopeverything.com/top-10-sportswear-brands/?amp">Top 10 Biggest Sportswear Brands in the World - All Top Everything</a>	<a href="https://www.alltopeverything.com/top-10-sportswear-brands/?amp">https://www.alltopeverything.com/top-10-sportswear-brands/?amp</a>



### IZJAVA O AUTORSTVU

Završni/diplomski rad isključivo je autorsko djelo studenta koji je isti izradio te student odgovara za istinitost, izvornost i ispravnost teksta rada. U radu se ne smiju koristiti dijelovi tuđih radova (knjiga, članaka, doktorskih disertacija, magistarskih radova, izvora s interneta, i drugih izvora) bez navođenja izvora i autora navedenih radova. Svi dijelovi tuđih radova moraju biti pravilno navedeni i citirani. Dijelovi tuđih radova koji nisu pravilno citirani, smatraju se plagijatom, odnosno nezakonitim prisvajanjem tuđeg znanstvenog ili stručnoga rada. Sukladno navedenom studenti su dužni potpisati izjavu o autorstvu rada.

Ja, Sara Slamić Tarade (*ime i prezime*) pod punom moralnom, materijalnom i kaznenom odgovornošću, izjavljujem da sam isključivi autor/ica završnog/diplomskog (*obrisati nepotrebno*) rada pod naslovom Analiza značaja brandova sportske obuće obradom velikih tekstualnih podataka (*upisati naslov*) te da u navedenom radu nisu na nedozvoljeni način (bez pravilnog citiranja) korišteni dijelovi tuđih radova.

Student/ica:

(*upisati ime i prezime*)

Slamić Tarade

(*vlastoručni potpis*)

Sukladno čl. 83. Zakonu o znanstvenoj djelatnosti i visokom obrazovanju završne/diplomske radove sveučilišta su dužna trajno objaviti na javnoj internetskoj bazi sveučilišne knjižnice u sastavu sveučilišta te kopirati u javnu internetsku bazu završnih/diplomskih radova Nacionalne i sveučilišne knjižnice. Završni radovi istovrsnih umjetničkih studija koji se realiziraju kroz umjetnička ostvarenja objavljuju se na odgovarajući način.

Sukladno čl. 111. Zakona o autorskom pravu i srodnim pravima student se ne može protiviti da se njegov završni rad stvoren na bilo kojem studiju na visokom učilištu učini dostupnim javnosti na odgovarajućoj javnoj mrežnoj bazi sveučilišne knjižnice, knjižnice sastavnice sveučilišta, knjižnice veleučilišta ili visoke škole i/ili na javnoj mrežnoj bazi završnih radova Nacionalne i sveučilišne knjižnice, sukladno zakonu kojim se uređuje znanstvena i umjetnička djelatnost i visoko obrazovanje.