

# UTJECAJ ANALIZE TEKSTA NA UČINKOVITOST KAMPANJE „NAPLATA PO KLIKU“ (PPC)

---

Petrak, Ljiljana

Master's thesis / Specijalistički diplomski stručni

2019

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Algebra  
University College / Visoko učilište Algebra**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:225:694006>

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

Download date / Datum preuzimanja: **2024-11-09**



Repository / Repozitorij:

[Algebra University - Repository of Algebra University](#)



**VISOKO UČILIŠTE ALGEBRA**

DIPLOMSKI RAD

**UTJECAJ ANALIZE TEKSTA NA  
UČINKOVITOST KAMPANJE „NAPLATA  
PO KLIKU“ (PPC)**

Ljiljana Petrak

Zagreb, rujan 2019.



# **Predgovor**

Zahvaljujem se mentoru Zlatanu Moriću na doprinosu i pomoći pri izradi diplomskog rada, svojoj obitelji na podršci, kolegici Tatjani Habulin na pomoći i stručnim savjetima te Krešimiru Richu, ravnatelju Poliklinike Identalia koji je omogućio realizaciju ovog rada.

**Prilikom uvezivanja rada, Umjesto ove stranice ne zaboravite umetnuti original potvrde o prihvaćanju teme diplomskog rada kojeg ste preuzeli u studentskoj referadi**

## Sažetak

Svrha ovog rada je utvrditi mogu li se saznanje i rezultati analize teksta nad prikupljenim tekstualnim upitima već akviziranih, visoko profitabilnih, klijenata iskoristiti za poboljšanje kampanja po modelu naplate po kliku i sukladno tome privući veći broj profitabilnih klijenata. U prvom dijelu rada se, pomoću analize teksta provedene na tekstualnim upitima profitabilnih klijenata, dolazi do saznanja o načinu komunikacije profitabilne ciljne skupine kroz najčešće korištene termine i n-grame prilikom potražnje usluge. U drugom dijelu rada ta saznanja se primjenjuju na komunikaciju unutar oglasa u kampanjama po modelu naplate po kliku. Testiranjem u trajanju od mjesec dana, promatra se potom mogu li oglasi kampanje kreirane na temelju rezultata prethodne analize teksta, preciznije nacijsati veći broj profitabilnih i perspektivnih klijenata spram oglasa kreiranih na temelju ekspretize. Testiranje novih oglasa napravljeno je u Google Adwords računu tvrtke Identalia. Rezultati su potom uspoređeni kroz marketinške metrike (broj prikaza, stopa prikaza, konverzije, kategorija profitabilnosti i perspektivnosti potencijalnog klijenta) koje će dati uvid u uspješnost rezultata. Za analizu teksta korišten je program otvorenog koda Orange a za testiranje kampanja i rezultata korištena je Google Adwords platforma i CRM sustav tvrtke Identalia.

**Ključne riječi:** analiza teksta, kampanje po modelu naplate po kliku, Orange, Google oglašavanje

## Summary

The purpose of this paper is to determine whether the findings and the results of the text mining of the collected textual inquiries of the already acquired profitable clients can be used for the improvement of the pay-per-click campaigns and, accordingly, attract a greater number of profitable clients. In the first part of the paper, using the text mining conducted on the textual inquiries of profitable clients, certain conclusions have been made about the communication manner of the profitable target audience through the most commonly used terms and n-grams when requesting a certain service. In the second part of the paper, these findings and conclusions are applied to the communication in ads in the pay-per-click campaigns. The testing period, which lasts for a month, is then used to examine if the ads of the campaign created on the basis of the previously mentioned text mining can target a larger number of profitable and perspective clients more accurately than the ads created on the basis of expertise. The testing of the new ads is performed in the Google AdWords account of the company Identalia. The results were then compared using the marketing metrics (impression, click-through rate, conversion rate, profitability and perspectiveness of the potential client) which will provide an insight into the results. The open source software Orange was used for text mining, and Google AdWords platform and the CRM software of Identalia were used for the analysis of the results.

**Keywords:** text mining, PPC campaigns, Orange, Google Ads

# Sadržaj

1. Uvod .....	1
2. Kontekst i okruženje poslovanja .....	<b>Error! Bookmark not defined.</b>
3. Poslovni slučaj, hipoteza i metodologija istraživanja.....	5
3.1 Poslovni slučaj i cilj projekta .....	5
3.2 Hipoteza.....	7
3.3 Metodologija istraživanja .....	9
4. Rudarenje teksta iz tekstualnih pupita profitabilnih klijenata .....	11
4.1 Podaci i alati korišteni za analizu .....	13
4.2 Priprema podataka za analizu i proces pred-proceisranja teksta .....	13
4.3 Proces analize teksta i interpretacija dobivenih rezultata.....	19
5. Metrike učinkovitosti oglasa po modelu naplata po kliku (PPC).....	26
5.1 Broj prikaza (eng.impression) .....	27
5.2 Stopa klikova (eng. click trough rate, CTR).....	27
5.3 Stopa konverzije (eng. conversion rate) .....	28
5.4 Kategorija profitabilnosti i perspektivnosti potencijalnog klijenta .....	29
6. Kreiranje oglasa na Google Adwords platformi.....	30
6.1 Metoda i način izrade teksta za oglase na temelju ekspertize .....	32
6.2 Metoda i način izrade teksta za oglase na temelju rezultata analize teksta .....	34
6.3 Testiranje oglasa na Google Adwords platformi.....	35
7. Mjerenje uspješnosti i analiza rezultata.....	37
7.1 Rezultati kampanja i oglasa temeljenih na ekspertizi.....	37
7.2 Rezultati kampanja i oglasa temeljenih na analizi teksta .....	38
7.3 Usporedba i interpretacija rezultata .....	40



8. Zaključak .....	42
Popis kratica .....	43
Popis slika.....	44
Popis tablica.....	45
Literatura .....	46

# 1. Uvod

Analiza velikih skupova podataka i različiti alati namijenjeni u te svrhe s ciljem poboljšanja učinkovitosti marketinških kampanja sve su dostupniji i manjim poduzećima. Na internetu se mogu pronaći mnogi besplatni alati otvorenog koda za dubinsku analizu podataka, što manjim poduzećima otvara mnoge mogućnosti. Ipak, i dalje postoje situacije koje predstavljaju izazove kao što je pronalazak svrhe i načina na koji se veliki skupovi podataka mogu iskoristiti za poboljšanje performansi poslovanja ili nedostatak resursa i znanja potrebnih za ostvarivanje tih mogućnosti.

U ovom diplomskom radu opisan je konkretan poslovni slučaj i primjer poboljšanja marketinške kampanje za tražilice po modelu naplate po kliku ostvaren pomoću analize velikih skupova tekstualnih podataka. Za analizu teksta korišten je besplatan program otvorenog koda Orange a znanja izvučena iz podataka korištena su za optimizaciju marketinških kampanja za tražilicu Google.

## 2. Kontekst i okruženje poslovanja

Poliklinika Identalia je dio grupacije koju sačinjava nekoliko tvrtki koje posluju unutar sektora dentalnog turizma. Poliklinika je specijalizirana za velike rekonstrukcije čeljusti zubnim implantatima i zubnim krunicama te u sklopu poslovanja posjeduje vlastiti digitalni zubotehnički laboratorij. Dentalni turizam uključuje brigu i liječenje pacijenata iz inozemstva pa se stoga na raspolaganje svakom klijentu, kako bi mu usluga bila maksimalno olakšana, stavlja besplatni smještaj, besplatan pregled i organizirani prijevoz. Sve navedeno za tvrtku posljedično uzrokuje visoke lateralne troškove za isporuku usluge. Klinika posluje na hrvatskom i šest stranih tržišta. U svojim marketinškim kampanjama tvrtka oglase uglavnom usmjerava na strana tržišta. Potencijalnim klijentima dana je mogućnost kontakta putem različitih komunikacijskih kanala: *online messaging* sistemom (whatsapp, viber, facebook messenger)<sup>1</sup>, putem obrasca na web stranici te putem besplatnog telefonskog poziva. Klinika najveći iznos financijskih sredstava unutar dostupnih budžeta za marketinško oglašavanje troši na oglašavanje putem internetskih pretraživača (eng. search engines)<sup>2</sup> s obzirom da kampanje namijenjene podizanju svjesnosti (eng. awareness) i brendiranju klinike zahtijevaju velika novčana ulaganja koja na širokom geografskom teritoriju trenutno rezultiraju niskom isplativošću<sup>3</sup>. Stoga nastojanje da se na tražilicama pokuša što preciznije nacičiti potencijalnog klijenta postaje nužno. Kada se zaprimi tekstualni upit korisnika, on se automatiziranim putem slijeva u CRM<sup>4</sup> sustav (eng. customer relationship management) gdje u organiziranom profilu odlazi u prodajni tunel<sup>5</sup> na daljnju obradu od strane prodajnog tima. Upit je unaprijed kategoriziran prema profitabilnosti perspektivnosti potencijalnog klijenta koje su automatski izračunate na temelju postojećih

---

<sup>1</sup> Whatsapp, viber, facebook messenger - sustavi za slanje tekstualnih poruka

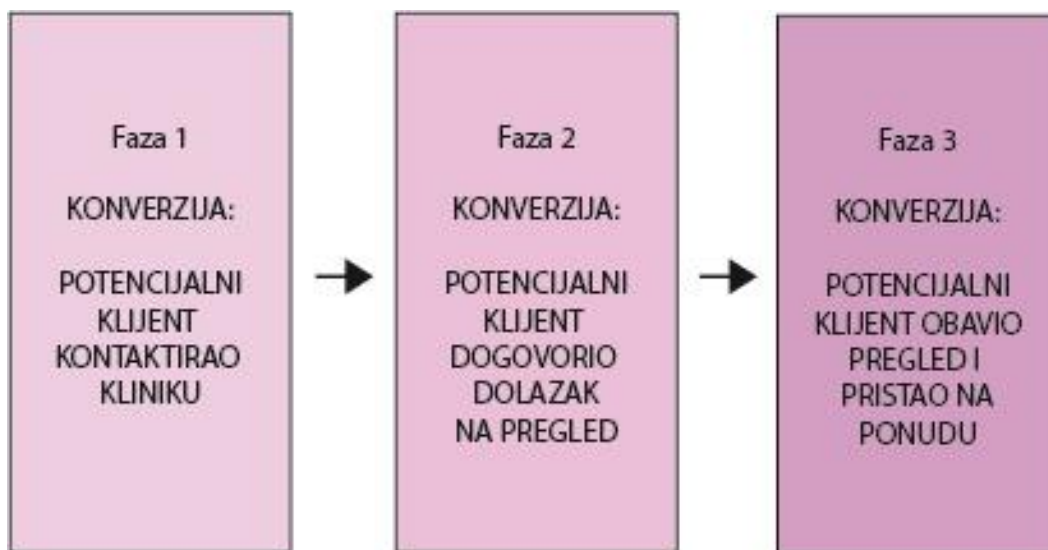
<sup>2</sup> Internet pretraživač (eng. Search engine)- specijalno web mjesto za pretraživanje sadržaja na internetu i pronalazak informacija pohranjenih na drugim web mjestima. Informacije koje pretraživač pronalazi mogu biti u obliku web stranice, slike, videa ili dokumenta a pronalazi ih na način da upisuje riječ, pojam ili rečenicu u tražilicu te u kratkom vremenu povratno dobiva brojne poveznice na traženi pojam.

<sup>3</sup> Prema dostupnim internim podacima Poliklinike Identalia

<sup>4</sup> CRM sustav (eng. Customer relationship management) je sustav za upravljanje odnosa sa klijentima unutar kojeg se mogu prikupljati razni podaci o korisnicima iz više kanala komunikacije sa kojima se na taj način omogućava upravljanje odnosima sa sadašnjim i budućim kupcima. Sustav služi za organizaciju informacija o klijentima, koja potom omogućava automatizaciju prodaje, marketinga i službe za korisnike.

<sup>5</sup> Prodajni tunel – marketinški pojam koji opisuje idealni proces i organizirane korake koje prolazi korisnik kako bi od potencijalnog klijenta postao klijent. Najčešće se sastoji od 4 faze. U prvoj prodajnoj fazi se nalaze svi oni koji jesu ili nisu potencijalni klijenti, u drugu prodajnu fazu prelaze samo potencijalni klijenti koji su pokazali interes za proizvod/uslugu. U trećoj fazi se nalaze potencijalni kupci koji su poslali upit za proizvod/uslugu a u zadnjoj fazi se nalaze svi oni koji su postali kupci proizvoda/usluge.

meta podataka pristiglih sa upitom a verificiraju se nakon dodatno unesenih atributa u CRM sustav prilikom provjere informacija o klijentu putem direktnog kontakta s njim. Svaki potencijalni korisnik prolazi sveukupno 3 faze unutar marketinškog tunela<sup>6</sup> kako bi konačno postao klijentom klinike. (Slika 1)



Slika 1. Koraci i konverzije unutar marketinškog tunela tvrtke

Prva faza uključuje konverziju (eng. conversion)<sup>7</sup> na razini ostvarivanja prvog kontakta koji se odvija putem raznih marketinških komunikacijskih kanala. Potom slijedi konverzija na razini dogovora dolaska na prvi liječnički pregled u kliniku. Nakon toga zadnja razina i konačna konverzija u pacijenta ostvaruje se samim prihvaćanjem predložene ponude nakon izvršenog prvog pregleda. Fokus marketinških aktivnosti stoga nije usmjeren samo na prvu razinu marketinškog tunela, nego uzima u obzir sve informacije i podatke sa preostale dvije razine koje zapravo postaju mjerilo kvalitete prikupljenih konverzija na prvoj razini. To znači da, ako ista konverzija na prvoj razini ne rezultira i konverzijom na trećoj razini ona se smatra zanemarivom. Primarni marketinški napor usmjeren je u većoj mjeri na

---

<sup>6</sup> Marketinški tunel je marketinški pojam za proces i korake koje korisnik prolazi od prvog kontakta sa tvrtkom do krajnjeg korisnika. Najčešće se sastoji od 4 faze a to su faza svjesnosti o postojanju brenda, faza interesa u kojoj potencijalni kupac pokazuje interes za proizvod/uslugu tvrtke, faza želje u kojoj potencijalni kupac evaluira proizvod/uslugu/brand te faza akcije u kojoj potencijalni kupac kupuje proizvod/uslugu.

<sup>7</sup> Konverzija (eng. Conversion) je radnja u kojoj korisnik stupa u interakciju sa tvrtkom i koju je tvrtka obilježila kao radnju vrijednu za svoje poslovanje. To može biti npr. klik na oglas, poslani upit, prihvaćanje ponude itd.

prikupljanje kvalitetnih upita od strane perspektivnih potencijalnih klijenata, a ne samo na kvantitativnu količinu upita.

Fokus analize ovog poslovnog slučaja je talijansko tržište koje je posljednjih nekoliko godina sve zasićenije, što sve većim brojem konkurentskih klinika i franšiznih lanaca u Italiji koji mogu ponuditi kompetitivnu cijenu, što porastom konkurencije na hrvatskom tržištu koja se bavi dentalnim turizmom sa talijanskim ciljnim tržištem [1].

Svi navedeni čimbenici znatno utječu na pozitivne odluke potencijalnih klijenata o obavljanju dentalnih zahvata u inozemstvu. Trenutno je primarni motivator kod donošenja odluke klijenta za odlazak izvan vlastite države na stomatološke zahvate – cijena zahvata, odnosno niža ponuda spram one u vlastitoj državi, na čemu se do sada i temelji generalni koncept dentalnog turizma u Hrvatskoj.

Cilj projekta je pokušati se što efikasnije istaknuti među oglasima konkurencije na internetskim pretraživačima a jedan od mogućih načina koji je opisan u ovom radu jest taj da se komunikacija unutar oglasa maksimalno prilagodi očekivanjima i potražnji potencijalnog klijenta.

# 3. Poslovni slučaj, hipoteza i metodologija istraživanja

## 3.1 Poslovni slučaj i cilj projekta

Predmet istraživanja ovog rada su modeli internet kampanje za tražilice (eng. search engines) po modelu naplate po kliku (eng. pay per click, skraćeno PPC)<sup>8</sup> te optimizacija i poboljšanje učinkovitosti teksta oglasa u svrhu preciznijeg ciljanja i akviziranja većeg broja profitabilnih klijenta.

Model kampanje naplate po kliku je način oglašavanja u kojem oglašivač plaća prikazivanje svojih oglasa isključivo kao rezultat pretraživanja korisnika interneta po određenim ključnim riječima.[2] Oglasi se dakle ciljano prikazuju korisnicima samo u trenutku unosa određenih ključnih riječi (eng. keywords)<sup>9</sup> u internet tražilicu prilikom pretraživanja određenog proizvoda, usluge ili informacije. Prema novijim istraživanjima: 77% korisnika vjeruje plaćenim oglasima na tražilicama dok će njih 33% kliknuti na oglas iz razloga što direktno odgovara na njihovo postavljeno pitanje. [3] Za tvrtku je stoga od iznimne važnosti efikasno povezati ključnu riječ koju korisnik upisuje u tražilicu prilikom pretraživanja usluge sa što učinkovitijim tekstom oglasa koji će potencijalnom klijentu što preciznije odgovoriti na zatraženu informaciju. Tekst oglasa bi dakle trebao zainteresirati korisnika da detaljnije procijeni usluge tvrtke na web sjedištu koje će moći posjetiti klikom na oglas koji mu je prikazan u trenutku pretraživanja. Osim atraktivnog teksta koji za svrhu ima prikupljanje klikova, potrebno je uzeti u obzir i karakteristike samog poslovanja tvrtke koje funkcionira po kompleksnoj shemi koja je prikazana u prethodnom poglavlju. Pošto tvrtka posluje unutar sektora dentalnog turizma, nudi široku paletu usluga i ima visoke troškove u lateralnim uslugama (visoko kvalificirano stručno osoblje, smještajne jedinice, organizacija prijevoza i druge turističke usluge koje su klijentu stavljene na raspolaganje besplatno prilikom boravka u klinici) može se dogoditi da ROI<sup>10</sup> po klijentu bude nizak ili čak negativan ukoliko

---

<sup>8</sup> Model kampanje naplata po kliku (eng. pay-per-click, PPC) – model internetskih kampanja za internet pretraživače u kojima je oglašivaču naplaćeno samo onda kada pretraživač klikne na prikazani oglas.

<sup>9</sup> Ključne riječi (eng. keywords) – riječi koje korisnik upisuje u tražilicu prilikom pretraživanja proizvoda ili usluge. Oglašivač može zakupiti ključne riječi kako bi se na upisanu riječ mogao prikazati povezan oglas.

<sup>10</sup> ROI (Return On Investment) – je povrat na uloženo. Računa se na način da se od prihoda koji su ostvarili određeni oglasi oduzmu ukupni troškovi koji su ostvarili ti isti oglasi te se zatim podijele sa ukupnim troškovima koji su ostvarili oglasi. (ROI=Prihod-trošak/trošak)

je klijent kupio usluge koje iznosom ne mogu pokriti cjelokupni trošak kompletno izdane usluge. Takva organizacija poslovnih procesa upućuje na potrebu tvrtke da identificira i teži isključivo ka profitabilnom klijentu ukoliko želi zadržati i ukupnu profitabilnost samog poslovanja. Postavlja se pitanje može li tvrtka već na samim tražilicama identificirati i ciljati isključivo profitabilnog klijenta. Ostvarivanjem te mogućnosti smanjili bi se troškovi poslovanja ako se povećanjem priljeva upita od strane profitabilnog klijenta posljedično smanji broj upita neprofitabilnih klijenata. Veći broj upita profitabilnih klijenata i manji broj upita neprofitabilnih klijenata značio bi za tvrtku i posljedično smanjenje troškova resursa potrebnih za inicijalnu obradu klijenta u prvoj i drugoj fazi prodajnog tunela koji je opisan u prethodnom poglavlju.

Definicije profitabilnosti i perspektivnosti klijenta već su uvedene u poslovanje i kao takve nalaze se unutar CRM (eng. customer relationship management) sustava tvrtke. Prema definiranoj metrici profitabilnosti koju sustav izračunava pomoću kompjutoriziranih i automatiziranih procesa i bilježi u profil klijenta, moguće je precizno identificirati već postojeće mase akviziranih profitabilnih klijenata. Pozicije na skali profitabilnosti i perspektivnosti klijenta unutar CRM sustava tvrtke u krajnjem obliku bilježe se u profilu svakog klijenta i nosi određene brojeve oznake. Unutar sustava omogućena je detaljna segmentacija klijenata koja dozvoljava i uvid u informacije o akviziciji istih. Npr. moguće je dobiti uvid u informaciju o ključnim riječima koje koristi isključivo profitabilni klijent prilikom pretraživanja usluge u tražilicama te informacije o tome iz koje je točno kampanje klijent ušao u CRM sustav za obradu klijenata.

Jednom kada smo izdvojili ključne riječi korištene samo od strane profitabilnog klijenta prilikom potražnje usluge na tražilicama i sukladno tome izračunali njihovu frekvenciju ponavljanja, u Google Adwords<sup>11</sup> sučelju tvrtke, provjerili smo u kojim kampanjama se te iste ključne riječi nalaze i koja je njihova učinkovitost u odnosu na broj klikova i broj konverzija.

Na temelju gore navedenih informacija identificirane su 2 kampanje unutar kojih se nalaze najčešće korištene ključne riječi od strane profitabilnog klijenta prilikom pretraživanja usluge.

---

<sup>11</sup> Google Adwords je platforma za plaćeno oglašavanje unutar rezultata pretraživanja na najpoznatijoj tražilici Google. Više o Google adwords platformi pisano je u poglavlju 6

Tvrtka želi veći priljev upita od strane profitabilnog klijenta pa se postavlja pitanje možemo li i kako, osim podizanjem novčanog ulaganja u navedene kampanje, postići veći broj konverzija od strane profitabilnog klijenta optimizacijom oglasa. Alati za plaćeno oglašavanje poput Google adwordsa dozvoljavaju optimiziranje oglasa na više razina<sup>12</sup> no u ovom diplomskom radu fokus je na optimizaciji plaćenih oglasa na razini teksta. Tekst od kojeg se sastoji naslov i opis oglasa zapravo je odgovor na pitanje koje je pretraživač postavio te stoga ima i najveći utjecaj na zainteresiranost klijenta za uslugu [1]. Cilj ovog poslovnog projekta je optimizacijom plaćenih oglasa za tražilicu Google pokušati privući veći broj profitabilnih i perspektivnih klijenata. Detaljnije informacije o kreiranju i optimizaciji oglasa unutar Google Adwords alata opisane su u poglavlju 6.

## 3.2 Hipoteza

Uzevši u obzir ponašanje korisnika na tražilicama, odnosno njihovo visoko povjerenje i sklonost klikova na plaćeni oglas [1], postavlja se pitanje može li potencijalni profitabilni klijent u našim oglasima pronaći idealan odgovor na upit koji potražuje na temelju ključne riječi koje je unio u tražilicu. Način na koji korisnici upotrebljavaju internetsku tražilicu je zapravo transformacija vlastitih misli u riječi. Takav proces te transformacije može se razlikovati od osobe do osobe, pa tako ista informacija može biti pretraživana na više različitih načina, odnosno pomoću različitih ključnih riječi. Ono što slijedi nakon što korisnik unese u tražilicu ključne riječi kojima će pretraživati informaciju je puno važnije, a radi se o ispunjenju vlastitih očekivanja. Korisnik želi pronaći prikaz točno onog odgovora kojeg je zamislio dobiti prilikom procesa transformacije misli u riječi. U prirodi čovjeka je predviđanje i očekivanje zamišljenog potencijalnog ishoda svaki puta kada napravi neku radnju. Po istom principu će dakle funkcionirati i prilikom pretraživanja informacija putem internetske tražilice. Kada se korisnik služi tražilicom kako bi pretražio određenu informaciju i svoje misli transformira u riječi, on pri tome automatski stvara određena očekivanja o tome što želi pronaći unutar rezultata pretrage. Od tražilice zahtijeva relevantne prijedloge koji odgovaraju na njegov zahtjev te u kojima može pronaći traženu informaciju. Ukoliko korisnik nije pronašao očekivani rezultat i prijedlog, pokušat će presložiti svoje misli i transformirati ih u novi set riječi koje će ponovno unijeti u tražilicu. U ovom trenutku

---

<sup>12</sup> Razine na kojima se može optimizirati Google kampanja su npr. postavke cilja, ciljne publike, odabir oblika, ekstenzija i kreativni tekst oglasa itd..



dolazi do izražaja važnost efikasno napisanog oglasa jer korisnik unutar samo nekoliko trenutaka donosi odluku o tome hoće li kliknuti na prvi ponuđeni rezultat pretrage, birati između više ponuđenih prijedloga ili se vratiti na početak i ponovno drugim setom ključnih riječi krenuti u potragu za traženim rezultatom.[4] Efikasan tekst je stoga jedan od najbitnijih elemenata samog oglasa i igra važnu ulogu za njegovu učinkovitost.

Temelj za pisanje učinkovitog teksta za oglas je dobro poznavanje vlastitog klijenta [5] kojeg se iz marketinške perspektive može upoznati putem ankete ili analize postojećih podataka o klijentima. Za stvaranje učinkovitih tekstova za poboljšanje oglasa, a potom i povezivanje tih istih oglasa sa najčešće korištenijim ključnim riječima od strane profitabilnog klijenta, potrebne su sljedeće informacije o profitabilnom klijentu: koje usluge takav klijent najčešće traži, koje izraze najčešće koristi, kako komunicira svoju situaciju (o oralnom zdravlju) te na koji način komunicira zahtjev za uslugom. Te informacije pokušat će se pronaći analizom podataka koje tvrtka već posjeduje, npr. unutar tekstualnih upita koje klijent šalje prilikom traženja usluge od tvrtke. Iz baze podataka moguće je precizno izdvojiti samo one upite koji pripadaju profitabilnom klijentu i pretpostavlja se da je iz teksta tih upita moguće izvući relevantna znanja. Tekst se nalazi u nestrukturiranom obliku i u velikim količinama pa je stoga u ovom slučaju najprikladnija metoda izvlačenja znanja iz tekstualnih podataka rudarenje teksta<sup>13</sup>. Analizom teksta primijenjenoj na izdvojenim upitima pokušati će se doći do informacija o tome na koji način profitabilni klijenti komuniciraju jednom kada dođu na web sjedište tvrtke i tamo postavite tekstualni upit za uslugu putem web obrasca. U tekstu će se promatrati koje izraze najčešće koriste i koje ih usluge najčešće zanimaju. Metoda rudarenja teksta je opširnije opisana u sljedećem poglavlju.

Hipoteza ovog diplomskog rada je sljedeća: pomoću znanja i informacija prikupljenima metodom rudarenja teksta primijenjenoj tekstualnim upitima profitabilnih klijenata mogu se kreirati učinkovitiji tekstovi za oglase na tražilicama koji će pozitivno utjecati na metrike uspješnosti te sukladno tome povećati broj konverzija od strane profitabilnih i perspektivnih klijenata

---

<sup>13</sup> Rudarenje teksta (eng. text mining) Rudarenje teksta odnosi se na analiziranje velikih skupova nestrukturiranih tekstualnih podataka te na prikupljanje skrivenog znanja unutar istih. Više informacija o rudarenju teksta mogu se pronaći u poglavlju 4.

Cilj ovog diplomskog rada je pokazati na koji način se saznanje i rezultati dobiveni analizom teksta nad prikupljenim tekstualnim upitima već akviziranih profitabilnih klijenata mogu iskoristiti za poboljšanje učinkovitosti oglasa na web tražilicama.

Krajnji cilj poslovnog projekta je optimiziranim oglasima na razini teksta preciznije ciljati i privući veći broj profitabilnih i perspektivnih klijenata na web sjedište tvrtke i pretvoriti ih u zadovoljne klijente.

### 3.3 Metodologija istraživanja

Tvrtka teži ka profitabilnom klijentu te je stoga poslovni zadatak optimizirati kampanje modela naplate po kliku kako bi tvrtka privukla veći broj profitabilnih i perspektivnih klijenata. U ovom diplomskom radu testirati će se teza mogu li oglasi kreirani na temelju rezultata iz analize teksta provedene nad upitima profitabilnih klijenata efikasnije privući broj klikova i povećati broj konverzija profitabilnih klijenata. Rezultati analize teksta bit će primijenjeni za kreiranje teksta za oglase unutar kampanja koje funkcioniraju po modelu naplate po kliku (PPC).

Metodološki rad je podijeljen u dvije cjeline. Prva cjelina odnosi se na rudarenje teksta u tekstualnim upitima profitabilnih klijenata. Analizom će se pokušati dobiti kvalitetan uvid u karakteristike i način komunikacije profitabilne ciljne skupine kroz najčešće korištene termine i n-grame<sup>14</sup> prilikom potražnje usluge. Ishod analize je tablica sa najčešće korištenim terminima i n-gramima unutar tekstualnih upita. Za analizu i rudarenje teksta koristit će se program otvorenog koda Orange<sup>15</sup> a sam proces je detaljnije opisan u sljedećem poglavlju.

U drugoj cjelini primijenit će se rezultati analize teksta za kreiranje novih tekstova za oglase unutar kampanja koje funkcioniraju po modelu naplate po kliku (PPC). Testiranjem u trajanju od 30 dana ustanoviti će se jesu li oglasi kreirani na temelju rezultata analize teksta u mogućnosti preciznije naciljati veći broj profitabilnih klijenata i ostvariti više perspektivnih upita spram oglasa kreiranih na temelju ekspertize te na taj način smanjiti nepotrebno raspršivanje troškova oglašavanja. Testiranje novih oglasa temeljenih na analizi teksta spram starih oglasa temeljenih na ekspertizi sproved će se unutar postojećih kampanja

---

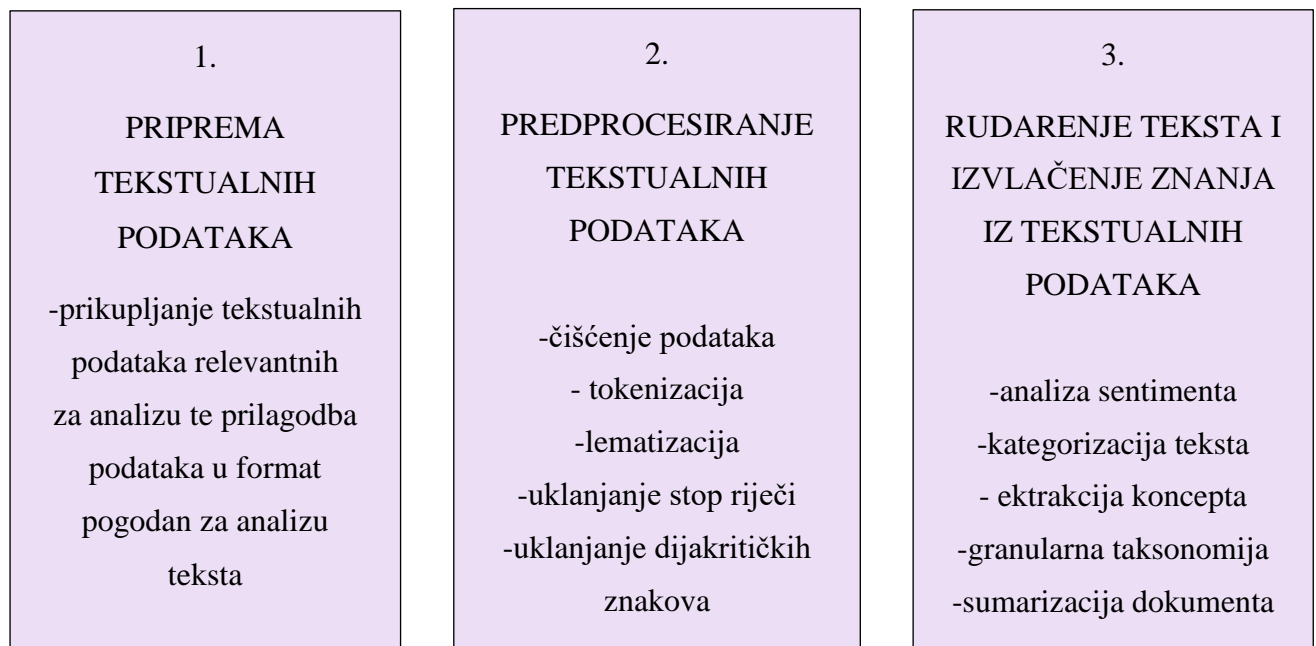
<sup>14</sup> N-grami su dvije ili više riječi koje se u tekstu najčešće pojavljuju zajedno

<sup>15</sup> Open source software Orange je programski paket otvorenog koda (eng. Open source) koji služi za dubinsku analizu i vizualizaciju podataka te za kreiranje prediktivnih modela. U ovom radu bit će korišten za analizu i vizualizaciju tekstualnih podataka.

u Google Adwords računu tvrtke Identialia za koje se prethodnom analizom ključnih riječi u CRM sustavu ustanovilo da privlače najviše profitabilnih klijenata. Rezultati će se potom usporediti kroz metrike (broj prikaza, stopa prikaza, stopa konverzije, kategorija profitabilnosti i perspektivnosti potencijalnog klijenta) koje će dati uvid u uspješnost rezultata. Za testiranje kampanja i rezultata učinkovitosti oglasa koristit će se platforma Google Adwords i CRM sustav tvrtke.

## 4. Rudarenje teksta iz tekstualnih upita profitabilnih klijenata

Prvi korak u postupku analize teksta je dobro razumijevanje postavljenog zadatka i poslovnog zahtjeva. Poslovni zadatak je povećati priljev upita od strane potencijalnih profitabilnih perspektivnih klijenata. Hipoteza glasi da je analizom teksta provedenom nad tekstualnim upitima profitabilnih klijenata moguće doći do saznanja koja potom mogu poboljšati performanse oglasa i preciznije naciljati veći broj potencijalno profitabilnih i perspektivnih klijenata. Rudarenje teksta odnosi se na prikupljanje skrivenog znanja unutar nestrukturiranih tekstualnih podataka. To je set procesa koji pomoću kompjuterskih algoritama pomaže identificirati uzorke i relevantne veze unutar tekstualnih podataka kako bismo dobili vrijedne i strukturirane informacije o samom tekstu.[6] Rudarenje tekstualnih podataka najčešće se odvija u 3 faze (slika2):



Slika 2. Faze unutar procesa rudarenja tekstualnih podataka te popis koraka koji se odnose na svaku fazu [6]

Prva faza procesa rudarenja teksta odnosi se na prikupljanje relevantnih tekstualnih podataka te na njihovu prilagodbu i smještanje u format pogodan za čitanje unutar alata u kojima će provesti analiza. Slijedi faza pred-procesiranja tekstualnih podataka koja je nužna za korektan ishod analize i koja podrazumijeva: čišćenje podataka, uklanjanje dijakritičkih

znakova, uklanjanje stop riječi, tokenizaciju i lematizaciju. Uklanjanje stop riječi odnosi se na postupak kojim se iz teksta eliminiraju sve one riječi koje ne pridonose značenju analize i rješenju poslovnog slučaja, već samo služe rečeničnoj strukturi (npr. prilozi, prijedlozi, veznici...). Ukoliko postoje situacije u kojima iz poslovne perspektive ovakve riječi unutar analize igraju značajnu ulogu za smisao rečenice, kontekst priče i pronalazak poslovnog rješenja, one se ne smiju zanemariti. Tokenizacijom se znakovni niz dijeli na niz riječi, odnosno tokena. Radnja je izvršena na temelju prepoznavanja određenih znakova unutar teksta koji npr. mogu biti interpunkcije ili razmaci. Tokenizacija se može vršiti na razini riječi ili na razini rečenice pa stoga jednim tokenom u procesu tokenizacije može biti obilježena jedna riječ ili cijela rečenica. Lematizacija je radnja koja svodi riječ/token na osnovnu formu. Izvršena je pomoću prepoznavanja deklinacija, ostalih promjena korijena riječi ili odbacivanjem sufiksa na kraju riječi. [7]

Zadnja faza u procesu rudarenja teksta je faza ekstrakcije informacija i znanja iz predprocesiranog teksta pomoću algoritama za analizu teksta. Tehnike za ekstrakciju znanja iz tekstualnih podataka ukratko su objašnjene u sljedećoj tablici:

Kategorizacija teksta	Svrstavanje teksta u predodređene kategorije npr. tekst o sportu, tekst o tuzirnu itd.
Klasteriranje teksta	Smještanje podataka u grupe na temelju određenih zajedničkih sličnosti i vrijednosti kojima su obilježeni. Npr. pronalazak skupva riječi koje se najčešće u tekstu pojavljuju zajedno.
Ekstrakcija koncepta/entiteta	Pronalazak i klasifikacija elemenata unutar teksta u predodređene kategorije kao što su npr. osobe, kompanije, geografkes logakcije i sl.
Granikular taksonomi	Izrada riječnika prilikom koje algoritam izvlači riječi koje se pojavljuju i grupira ih u skupine.
Analiza sentimenta	Identificiranje i vađenje subjektivnih informacija iz teksta. To mogu biti npr. emocije ili vjerovanja. Vrš se tako da se tokeni tagiraju određenim obilježjima.
Entitetni relacijski model	Automatizirano učenje o vezama unutar i između vrsta podataka.

Tablica 1. tehnike za ekstrakciju znanja iz tekstualnih podataka u fazi rudarenja teksta [8]

## 4.1 Podaci i alati korišteni za analizu

Kako bi podaci bili pogodni za korištenje i kako bi ih alati mogli učitati potrebno ih je pripremiti na specifičan i funkcionalan način. Već strukturirana baza podataka s informacijama o korisnicima preuzeta je od Poliklinike Identalia i anonimizirana tehnikom šifriranja riječi pomoću alata Microsoft Excel. Prilikom anonimizacije su funkcijom „SUBSTITUTE“ unutar teksta izmijenjene samo riječi smatrane relevantnima za poslovanje. U analizu su uključeni svi nestrukturirani tekstualni podaci iz upita pristiglih u periodu od 1.1.2017 do 1.1.2019 od strane osoba koje su već postale klijenti klinike, koji su označeni visokim kategorijama profitabilnosti te dolaze preko kanala iz web obrasca ili putem internet *messaginga*. Klijenti koji nisu označeni kao profitabilni nisu uzeti u obzir prilikom analize jer je cilj projekta otkriti način komunikacije profitabilnih klijenata i pošto će novi oglasi biti namijenjeni testiranju u kampanjama za koje smo prethodno ustanovili da generiraju najveći broj profitabilnih klijenata. U tom slučaju rudarenje teksta provedeno nad upitima nisko-profitabilnih klijenata predstavljalo bi za tvrtku gubitak vremena i resursa. Podaci o profitabilnim klijentima su vezani samo za klijente sa talijanskog tržišta te je izvorni tekst na talijanskom jeziku.

Za potrebe ovog rada, analizu teksta, testiranje oglasa i interpretaciju rezultata koristit će se sljedeći alati: za analizu teksta koristit će se Microsoft Excel i nekomercijalni alat otvorenog koda za analizu podataka Orange. Za kreiranje i testiranje efikasnosti oglasa koristit će se Google Adwords platforma za plaćeno oglašavanje te CRM sustav u vlasništvu Poliklinike Identalia za analizu dobivenih rezultata.

## 4.2 Priprema podataka za analizu i proces pred-procesiranja teksta

Prvi dio procesa pripreme podataka za analizu je segmentacija podataka koja je napravljena već pri samom SQL upitu<sup>16</sup> na bazu podataka tvrtke. Prilikom segmentacije su prema kategoriji profitabilnosti iz baze izvučeni tekstualni upiti svih klijenata koji su označeni visokom profitabilnošću. Podaci su u krajnjem obliku dobiveni u Excel tablici u kojoj se

---

<sup>16</sup> SQL (eng. Structured Query Language) upit -SQL je programski jezik koji se koristi za upravljanje podacima i bazama podataka, a najčešća operacija koja se provodi u SQL programskom jeziku je upit. SQL upit na bazu izvodi se putem programskog jezika a zatražene podatke vraća u obliku strukturirane tablice.

nalazi 1 stupac koji sadrži 6079 redaka teksta iz upita profitabilnih klijenata. Tablica u prvotnom obliku nije bila prikladna za direktno učitavanje u software Orange iz tehničkih razloga. Tekst iz upita bio je razdijeljen na više stupaca, postojala je prisutnost velikog broja izraza putem brojke te unutar podataka nisu bili obrisani dupli upiti.

Podaci su najprije obrađeni u alatu Excel na sljedeći način: spojeni su stupci putem funkcije CONCATENATE kako bi svaki cjeloviti tekstualni upit tvorio jedan redak unutar istog stupca.

Skraćenice i numerički izrazi pronađeni su ručnom analizom teksta i pretvoreni u cjelovitu riječ funkcijom FIND AND REPLACE (npr. X u talijanskom jeziku je slang za riječ „per“ (za) dok je npr. brojka napisana numeričkom oznakom 1 pretvorena u tekst: jedan).

Strukturirana tablica sa tekstualnim upitima spremljena je potom u format CSV (hrv. vrijednosti odvojene zarezom, eng. Comma separated values) koji je pogodan za učitavanje u software Orange

„Orange alat ima mogućnost čitanja predodređenih formata dokumenata kao što su .tab koji je također i nativni format datoteka u Orangeu, .xls/.xlsx, te tekstualne .txt datoteke odvojene zarezom (CSV). Prije obrade bilo kojeg nestrukturiranog skupa podataka potrebno ga je pravilno pripremiti i prilagoditi za obradu. Uobičajeno strukturiranje podataka unutar dokumenta je u obliku tabličnog prikaza“[9].

Za analizu teksta potrebni su tekstualni upiti koji su prethodno strukturirani u tablični prikaz unutar alata Excel i spremljeni u format CSV. Nakon što su podaci provjereni, tako spremljeni dokument se učitava u program Orange kroz alatnu jedinicu „Corpus“ odabirom iz lokalno spremljene datoteke. Pregled svih alatnih jedinica korištenih za analizu teksta za ovaj poslovni slučaj u alatu Orange: (slika 3)



Slika 3 Alatne jedinice unutar programa Orange korištene u ovom radu

Opis alatnih jedinica korištenih za analizu teksta unutar programa Orange:

- CORPUS – alatna jedinica za učitavanje podataka u program Orange u obliku vrijednosti odvojenima zarezom ( CSV)
- CORPUS VIEWER – alatna jedinica koja omogućava provjeru ispravnosti atributa i količine podataka unutar učitane tablice.
- PREPROCESS TEXT – alatna jedinica unutar koje se vrši pred-procesiranje teksta i koja omogućava automatsku eliminaciju diakritičkih znakova, lematizaciju, eliminaciju stop riječi i razdvajanje teksta na određene n-grame.
- WORD CLOUD – alatna jedinica za vizualizaciju teksta unutar koje se nalazi oblak sastavljen od najčešće ponavljanih riječi. Frekvencija ponavljanja riječi prikazana je dimenzijom slova. Tako će riječi koje se najčešće ponavljaju biti najveće, a one koje imaju najmanju frekvenciju biti će najmanje na grafičkom prikazu.
- CONCORDANCE – alatna jedinica koja služi za ručnu obradu analize sentimenta. Označavanjem određene riječi, možemo promatrati kontekst oko iste te tako dobiti jasniji uvid u informacije.
- DATA TABLE – alatna jedinica unutar koje možemo dobiti tablični prikaz rezultata.

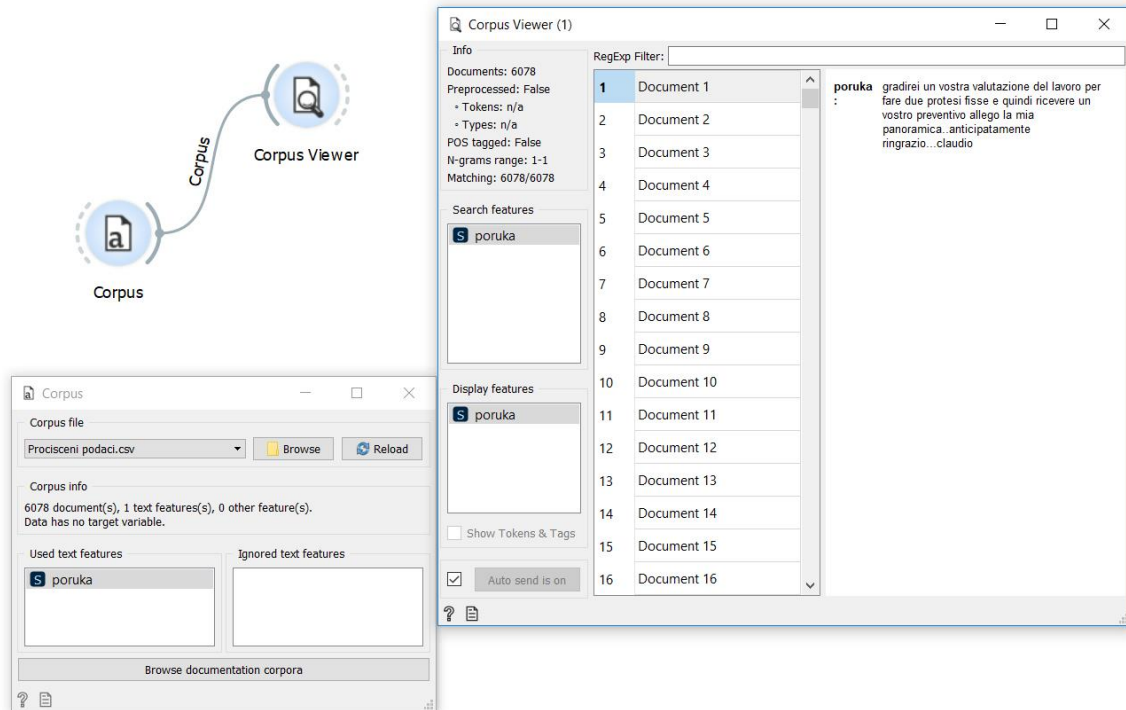
Prvi korak analize teksta u programu Orange je pred-procesiranje teksta. Pred-procesiranje teksta je nužan korak prije same analize pošto se izvorni tekst nalazi u sirovom obliku i nije prilagođen u format kako bi ga programski jezik i algoritam programa mogao razumijeti. Tekstualni podaci su stoga transformirani u format koristan za analizu kroz prethodno opisane procese tokenizacije, uklanjanja diakritičkih znakova te procesom svođenja svih riječi na njihov korijen. Odabir tehnike ili skupa tehnika za pred-procesiranje teksta ovisi o zahtjevima poslovnog slučaja i rezultatu koji se želi postići.

Proces pred-procesiranja teksta:

Na radnu ploču unutar programa Orange najprije je položena alatna jedinica „Corpus“ putem koje je zatim učitana tablica sa prethodno pripremljenim tekstualnim podacima. Tablicu i ispravnost učitavanja njezine strukture u program moguće je provjeriti pomoću alatne jedinice „Corpus Viewer“ koja daje uvid u osnovne informacije o podacima, broju dokumenata i pred-procesiranju podataka. Na taj način provjerava se je li program uspješno učitao tablicu, jesu li podaci u tekstualnom obliku te jesu li ispravno organizirani i označeni.



Na slici 4 može se uočiti da je dokument koji se koristi za ovaj poslovni slučaj, nakon što smo povezali alatnu jedinicu „Corpus“ sa alatnom jedinicom „Corpus Viewer“, sadržan od jednog stupca i 6078 redaka te da su svi podaci tekstualnog oblika. (slika 4).



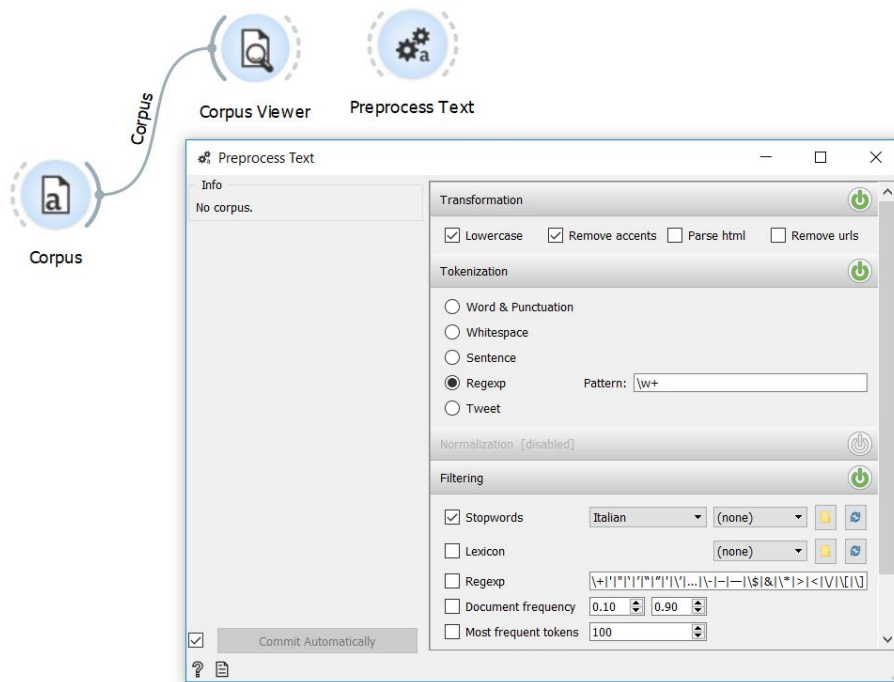
Slika 4. Unošenje i provjera ispravnosti podataka u programu Orange unutar alatnih jedinica „Corpus“ i Corpus Viewer“

Slijedi pred-procesiranje i čišćenje teksta pomoću alatne jedinice „Text Preprocessing“ putem koje na jednostavan način redom vršimo uklanjanje dijakritičkih znakova, svođenje riječi na mala slova, uklanjanje stop riječi i tokenizaciju.

Kada položimo na radnu ploču alatnu jedinicu „Text Preprocessing“ i otvorimo dane mogućnosti primjećujemo da se u prvom koraku vrši transformacija teksta (eng. transformation), odnosno svodimo čitav tekst na mala slova i uklanjamo dijakritičke znakove. Unutar tog koraka potrebno je označiti funkciju „lowercase“ za svođenje teksta na mala slova te „remove accents“ za eliminaciju dijakritičkih znakova. (slika 5)

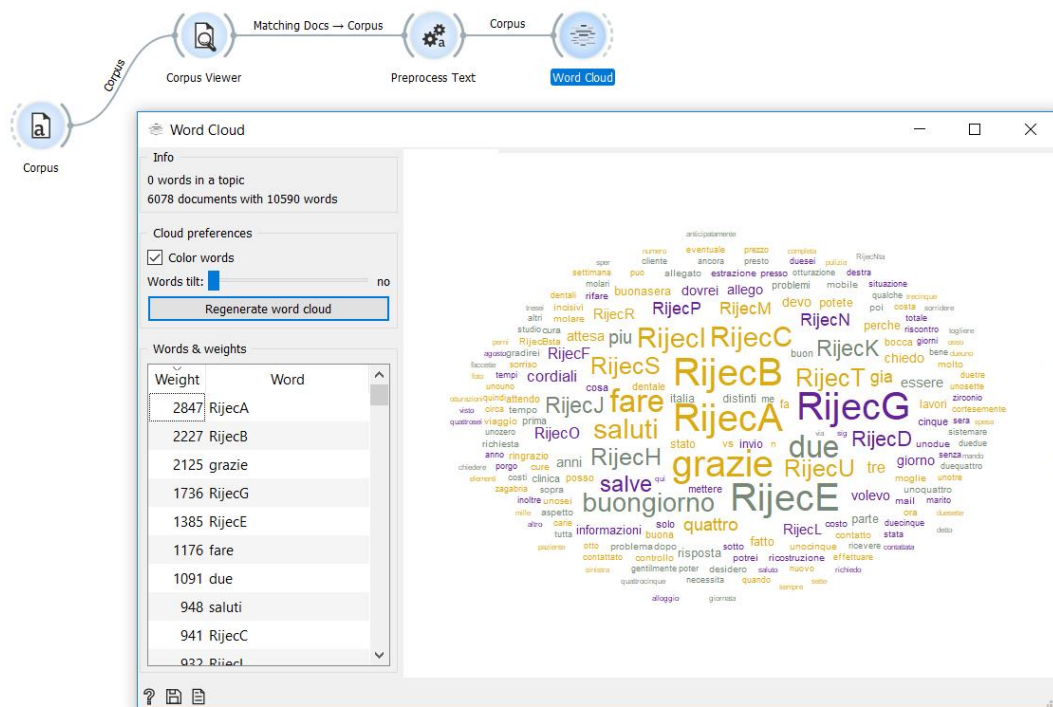
U drugom koraku odvija se tokenizacija koja se može vršiti na razini rečenice, razmaka, interpunkcije ili riječi. U ovom poslovnom slučaju tokenizacija je izvršena na razini riječi pa je stoga označena funkcija „Regexp“ (slika 5) .

U trećem koraku, pod nazivom „Filtering“ (hrv. filtriranje), vršimo uklanjanje stop riječi pa je iz tog razloga kvačicom označena funkcija „stop words“. Na ovom koraku potrebno je označiti i jezik teksta na koji se odnosi analiza, pošto alat ima unesene podatke o najčešćim stop riječima za različite jezike. Odmah pored funkcije „stop words“ naznačeno je da se radi o talijanskom jeziku. (slika 5).



Slika 5 Pred-procesiranje teksta u alatnoj jedinici „Preprocess Text“

Po izvršenom pred-procesiranju teksta možemo na radnoj ploči povezati alatnu jedinicu „Preprocess Text“ sa alatnom jedinicom „Word Cloud“ kako bi se dobio primarni rezultat i vizualizacija najčešće ponavljanih riječi u obliku grafičkog prikaza tekstualnog oblaka. Vizualizacija tekstualnih podataka putem alatne jedinice „Word Cloud“ prikazuje frekvenciju ponavljanja riječi koja je opisana dimenzijom znakova, od najveće do najmanje. Tako će više puta ponavljana riječ unutar svih tekstualnih dokumenata biti svojom dimenzijom veća, a ona koja je manje ponavljana će biti manja. Na dobivenom rezultatu potrebno je provjeriti postoje li riječi koje se dimenzijom prikazuju kao relevantne a ne odnose se na rješenje poslovnog slučaja. Takve riječi potrebno je naknadno ukloniti. (slika 6)



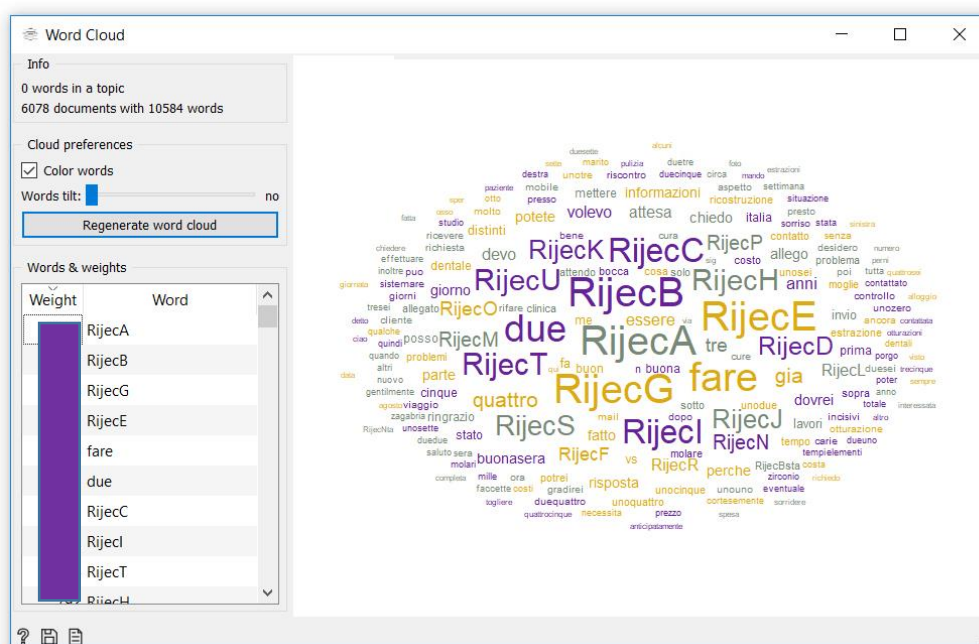
Slika 6. Prva provjera rezultata pred procesiranog teksta unutar alatne jedinice „Word Cloud“

Preciznom analizom nad grafičkim prikazom podataka i pripadajućom tablicom uočene su riječi koje je potrebno naknadno ukloniti pošto nisu relevantne za rješavanje ovog poslovnog slučaja te mogu utjecati na točnost rezultata. Radi se o riječima poput pozdrava ili zahvale (tal. *buongiorno, saluti, salve, grazie*). U programu Excel je stoga u sljedećem koraku kreirana nova tablica sa popisom svih riječi za koje je uočeno da je potrebno naknadno ukloniti. Tablica je zatim spremljena u format text.tab (tab delimited) koji je pogodan za učitavanje u program Orange. Spomenut tablicu je potrebno unijeti unutar alatne jedinice za pred-procesiranje teksta „Text Preprocessing“ na koraku „Filtering“ pored funkcije „Stop Words“.

Ovaj proces se ponavlja sve dok se iz teksta ne uklone sve riječi koje nisu relevantne za poslovni slučaj. Tim korakom završili smo pred-procesiranje teksta te možemo napraviti evaluaciju rezultata iz poslovne perspektive.

## 4.3 Proces analize teksta i interpretacija dobivenih rezultata

Nakon pripreme teksta i pred-procesiranja slijedi analiza teksta te interpretacija i vizualizacija dobivenih rezultata u programu Orange. Prva informacija koja je od važnosti za poslovni slučaj je otkrivanje trendova ponavljanja određenih riječi unutar tekstualnih upita kako bismo provjerili koje riječi najčešće koristi profitabilni klijenti prilikom slanja upita za uslugu. Alatna jedinica “Word Cloud “ daje nam grafički i tablični prikaz riječi koje se najviše ponavljaju unutar teksta na temelju dimenzije riječi, dok boja nema značenje i služi za lakšu čitljivost rezultata. (slika 7)



Slika 7 Vizualizacija frekvencije ponavljanja unigrama unutar teksta iz upita profitabilnih klijenata. Unutar prozora alatne jedinice „Word Cloud“, s lijeve strane moguće je provjeriti informacije o tekstualnim podacima, broju tokena i broju analiziranih dokumenata (Slika 7). Ispod nabrojanih informacija nalazi se tablica s popisom i točnom frekvencijom ponavljanja svakog tokena dok se s desne strane nalazi grafički prikaz u obliku oblaka. Iz dobivenih rezultata kreirat će se tablica u koju će se unijeti prvih 21 tokena s najvećom frekvencijom ponavljanja i koji će se razvrstati u 3 kategorije. U jedan stupac će se izdvojiti svi tokeni koji nose nazive usluga kako bismo dobili uvid u one koje se najčešće spominju unutar

tekstualnih upita. U drugi stupac će se izdvojiti svi tokeni koji nose oblik glagola i koji zapravo opisuju potražnju usluge, dok će se u treći stupac izdvojiti svi tokeni koji nose oblik pridjeva ili imenica koje opisuju traženu uslugu.

<b>Usluga</b>	<b>Potražnja usluge</b>	<b>Opis usluge</b>
Riječ A	Riječ G	Riječ C
Riječ E	Riječ H	Riječ D
Riječ K	Riječ J	Riječ I
Riječ U	Riječ R	Riječ B
Riječ S		Riječ T
Riječ M		Riječ N
Riječ P		Riječ O
		Riječ F
		Riječ L

Tablica 2 Najviše ponavljane riječi unutar teksta oglasa razvrstane su u tri kategorije: usluga, način potražnje usluge i opis tražene usluge

Tokeni su strukturirani u tablični format koji se sastoji od tri navedene kategorije iz potreba daljnje analize profitabilnih klijenata unutar tvrtke, a za potrebe ovog rada napravljen je još jedan korak kako bismo došli do relevantnih zaključaka i pripremili rezultate za daljnje korištenje pri kreiranju novih tekstova za oglase. Radi se, naime, o analizi već postojećih tekstova oglasa unutar Google Adwords kampanja za koje smo utvrdili da nose najveći broj profitabilnih klijenata. Pregledom teksta postojećih oglasa potrebno je zapravo utvrditi postoje li tokeni koji se trenutno ne nalaze unutar teksta oglasa i nikada nisu korišteni za oglašavanje. Svi ne-korišteni tokeni sačinjavaju finalnu tablicu koja će se potom koristiti za kreiranje novih oglasa. Kreiranje teksta oglasa na temelju ishoda ove analize detaljnije je opisano u poglavlju 6.

Usluga	Potražnja usluge	Opis usluge
Riječ A	Riječ J	Riječ C
		Riječ D

Tablica 3 Tablica sa ispisanim relevantnim tokenima po kreiranim kategorijama koji do sada nisu korišteni za tekstove oglasa

Usporedbom ishoda analize postojećih oglasa sa ishodom analize dobivene rudarenjem teksta utvrđeno je sljedeće:

- 1) Unutar kategorije “Usluga” otkriven je jedan token (Riječ A) koji nije korišten niti u jednom naslovu postojećih oglasa.
- 2) Unutar kategorije “Potražnja usluge” otkriven je 1 token (Riječ J) koji nije korišten niti u jednom postojećem oglasu.
- 3) Unutar kategorije “Opis usluge” otkrivena 2 tokena (Riječ C i Riječ D) koja nisu korištena unutar teksta već postojećih oglasa.

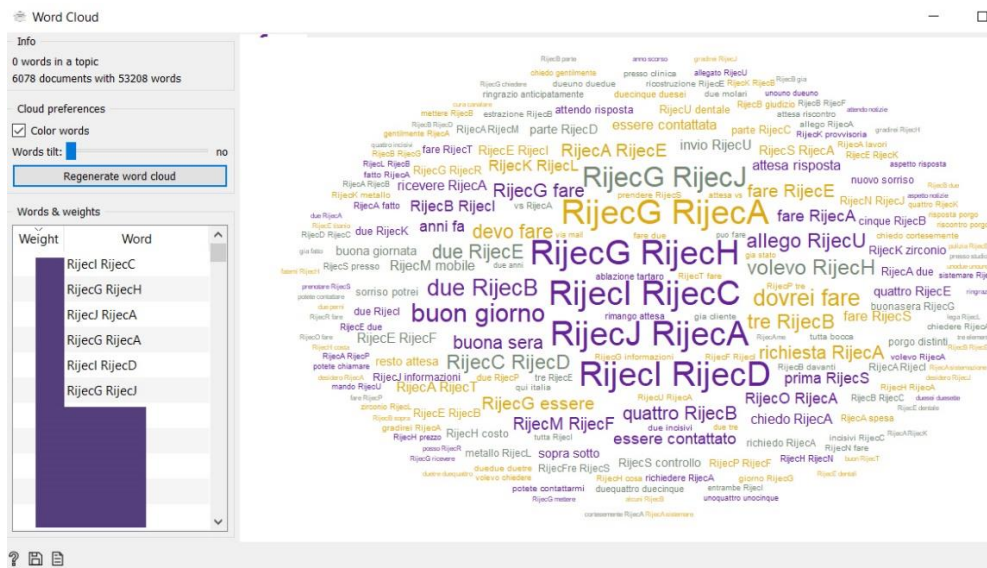
Sljedeći korak je analiza i vizualizacija N-grama. N-gram je skupna sekvenca n izraza iz dane sekvence teksta[10] a u ovom radu koristit će se kako bi se provjerilo jesu li kombinacije od 2, 3 ili 4 tokena koji se u tekstu najčešće pojavljuju zajedno, iskoristivi i kao kombinacije riječi za kreiranje tekstova za nove oglase.

Kada se tekst analizira na temelju jedne riječi, radi se o unigramu. Kada se analiziraju dvije riječi koje se najčešće nalaze zajedno unutar teksta radi se o bigramu, a ukoliko se analiziraju tri susjedne riječi radi se o trigramu, itd.

U ovom poslovnom zadatku pokušavamo putem rezultata analiziranih na temelju N-grama i vizualiziranih u alatnoj jedinici „Word Cloud“ otkriti relevantne kombinacije tokena koje otkrivaju više informacija o načinu komunikacije profitabilnog klijenta te su također iskoristive kao kombinacija izraza za kreiranje tekstova za nove oglase. Sve kombinacije tokena će se kasnije unijeti u tablicu i koristiti za kreiranje tekstova za nove oglase.

Bigrame i trigrame unutar programa Orange dobivamo na način da unesemo zahtijev unutar alatne jedinice „Preprocess Text“ na koraku „N-gram Range“. Za dobivanje bigrama potrebno je na tom koraku u prvo i u drugo polje upisati broj 2, a za trigrame broj 3.

Rezultati su sljedeći:



Slika 8 Vizualizacija bigrama u programu Orange unutar teksta upita profitabilnih klijenata pomoću alatne jedinice „Word Cloud“

Kada se radi o trigramima, ponovljen je jednaki postupak kao i za analizu bigrama. Potom je kreirana tablica sa najčešće ponavljanim tokenima u slučaju bigrama i trigrama (Tablica 4). Analiza nije nastavljena prema četirigramu jer se pokazalo da frekvencija ponavljanja skupa od 4 riječi nije relevantna za uzimanje u obzir za daljnju obradu i analizu.

Relevantne kombinacije tokena u bigramu	Relevantne kombinacije tokena u trigramu
Riječ I + Riječ C	Riječ G + Riječ J + Riječ A
Riječ G + Riječ H	Riječ I + Riječ C + Riječ D
Riječ J + Riječ A	Riječ B + Riječ I + Riječ C
Riječ I + Riječ A	Riječ E + Riječ I + Riječ C
Riječ I + Riječ D	Riječ A + Riječ I + Riječ C

Tablica 4 Relevantne kombinacije tokena u bigramu i trigramu po frekvenciji ponavljanja

Najbitniji tokeni za izradu novih tekstualnih oglasa su tokeni izdvojeni u tablici 3 i odnose se na izraze koji do sada još nisu korišteni u tekstovima postojećih oglasa i koji će biti testirani u ovom radu. Rezultatom analize teksta došlo se do saznanja o usluzi kod koje postoji prostor za napredak. Odnosno taj izraz nije bio prethodno prisutan unutar teksta oglasa.

<b>Relevantne kombinacije za token „Riječ A“ u bigramu:</b>	<b>Relevantne kombinacije za token „Riječ A“ u trigramu:</b>
Riječ A + Riječ R	Riječ A + Riječ G + Riječ J
Riječ A + Riječ G	Riječ A + Riječ I + Riječ C
Riječ A + Riječ E	Riječ A + Riječ N + Riječ J
Riječ A + Riječ O	

Tablica 5 relevantne kombinacije tokena u bigramu i trigramu za token „Riječ A“

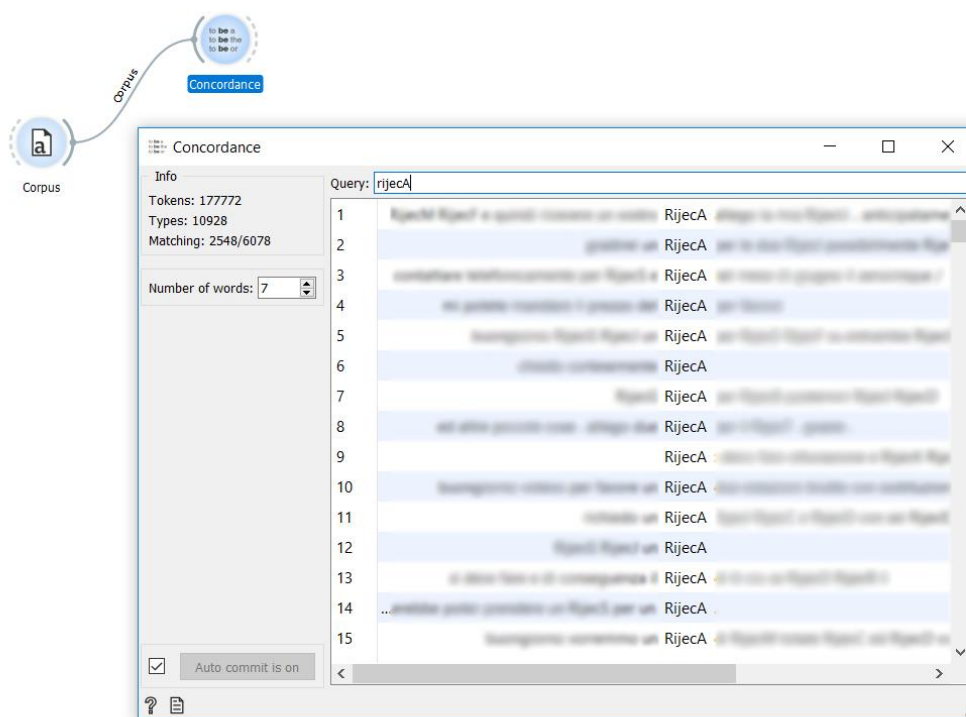
Za token iz kategorije „usluga“ (Riječ A) izrađena je kombinacija bigrama i trigrama dok ostali tokeni služe za opis i potražnju usluge. Sve relevantne kombinacije za token „Riječ A“ unesene su u sljedećoj tablici. (tablica 5)

S obzirom da su Google Adwords kampanje i Ad-setovi najčešće izgrađene oko pojedinih usluga i popratnih ključnih riječi, javlja se potreba za analizom konteksta te iste usluge unutar tekstualnih upita. Kontekst u kojem se ta usluga nalazi analizirat će se pomoću alatne jedinice „Concordance“. Tim korakom dobiva se uvid u informacije o načinu korištenja te riječi unutar teksta upita i kao rezultat se lako mogu vizualizirati susjedne riječi koje čine kontekst vezan uz taj specifičan token.

Alatnu jedinicu „Concordance“ potrebno je na radnoj ploči unutar programa Orange povezati direktno sa alatnom jedinicom „Corpus“ kako bismo dobili uvid u cjeloviti upit i ispravan kontekst upita koji nije pred-procesiran i u kojem nisu uklonjene stop riječi kako bi se zadržalo značenje rečenice u izvornom obliku. Analiza odabranog tokena se vrši unutar



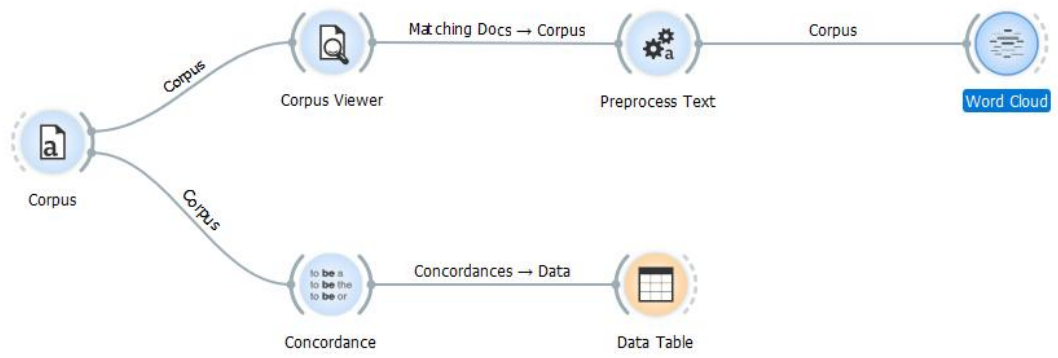
alatne jedinice „Concordance“ (slika 9) koja će u jednom stupcu prikazati analizirani token i sve riječi koje se nalaze oko istog.



Slika 9 Analiza konteksta u kojem se nalazi token Riječ A u alatnoj jedinici „Concordance“

Tablica sa svim dokumentima vezanima za analizirani token i njegov cjeloviti upit dobiva se povezivanjem alatne jedinice „Concordance“ sa alatnom jedinicom „Data Table“. Analiza konteksta pomoću alatne jedinice „Concordance“ provedena je ručno pošto se ne radi o velikom skupu podataka, a zaključci koji su doneseni na temelju ove analize su subjektivni te neće utjecati na daljnji tijek ovog rada.

Ova jednostavna analiza teksta provedena nad zaprimljenim tekstualnim upitima profitabilnih klijenata daje uvid o tome kako komunicira profitabilan klijent, koje termine najčešće koristi i s druge strane koristi li tvrtka te iste izraze u svojoj komunikaciji i oglašavanju. Rezultati su upisani u tablice koje će se primijeniti u poglavlju 6 za kreiranje novih tekstualnih oglasa. Krajnji prikaz povezivanja i korištenja alatnih jedinica za izvršenu analizu teksta u programu Orange prikazan je na slici 10.



Slika 10 Alatne jedinice korištene za analizu teksta unutar programa Orange i cijelokupan proces analize teksta

## 5. Metrike učinkovitosti oglasa po modelu naplata po kliku (PPC)

*Pay per click* (PPC) je model oglašavanja na internetskim tražilicama koji se koristi kako bi privukao korisnike na web sjedište određene tvrtke. Tvrtki se u tom slučaju naplaćuje samo onda kada korisnik klikne na prikazani oglas dok se prikazivanje samog oglasa se ne naplaćuje. Oblik kampanje po modelu naplate po kliku je zapravo jedan od najmjerljivijih oblika online oglašavanja za oglašivača na internetskim tražilicama. Primarni fokus ovih marketinških kampanja temelji se na rezultatima koji se mogu izmjeriti kroz razne marketinške metrike. Sve metrike za promatranje i mjerenje uspješnosti kampanje odabiru se prema prethodno postavljenom cilju projekta ili same kampanje. Većina tražilica dozvoljava više postavki na razini cilja, a one najčešće korištene su konverzije (eng. *conversion*) ili promet na web sjedištu (eng. *traffic*). Promet na web sjedištu je kampanja usmjerena prema korisniku u kojoj je cilj da što više korisnika posjeti određeno web sjedište.

U ovom radu cilj tvrtke je upit profitabilnog klijenta, odnosno konverzija na web stranici, pa je sukladno tome postavljeni cilj kampanje - konverzija. Konverzija u ovom radu za tvrtku znači broj klijenata koji su na web sjedištu ispunili obrazac i postavili upit za neku od usluga u ponudi tvrtke.

Učinkovitost je s druge strane pojam koji je često korišten unutar poslovnih procesa tvrtke i koji treba biti mjerljiv unutar postavljenih okvira ili mjerila. Mjerenje učinkovitosti marketinških kampanja za tvrtku znači evaluaciju rezultata spram očekivanih i postavljenih ciljeva. Mjerenje učinka i učinkovitosti marketinških kampanja se sastoji od metrika i postavljanja ključnih pokazatelja uspješnosti (eng. *key performance indicator*, skraćeno KPI) kako bi se procijenio učinak online kampanje. Struktura organizacije kampanja i sukladno praćenje rezultata na kraju zapravo prikazuju sve što je relevantno za poslovni cilj i služi kao osnova pri donošenju budućih pravilnih poslovnih odluka.

Cilj ovog istraživanja je povećati broj konverzija od strane profitabilnih i perspektivnih klijenata te će se uspješnost i učinkovitost novih oglasa mjeriti pomoću sljedećih metrika:

- Metrika o broju prikaza novih oglasa (eng. *impression*) pokazuje u kojoj mjeri je oglas relevantan za potencijalnog klijenta na način da se prati u kojoj količini se oglas prikazuje nakon što je korisnik upisao ključne riječi vezane za taj oglas.

- Metrika stopa klikova (eng. *click trough rate*, skraćeno CTR) pokazuje koliko je tekst oglasa efikasan, odnosno pratiti će se koliko je puta potencijalni klijent kliknuo na oglas spram broja prikaza istog oglasa.
- Potom će se mjeriti stopa konverzija (eng. *conversion rate*) pristiglih iz testiranih oglasa, te će se na kraju projekta u njima promatrati postotak potencijalno profitabilnih i perspektivnih klijenata.
- Metrike koje govore o profitabilnosti i perspektivnosti klijenta nalaze se u CRM sustavu tvrtke.

Metrike za mjerenje uspješnosti detaljnije su objašnjene u nastavku poglavlja.

## 5.1 Broj prikaza (eng. *impression*)

Broj prikaza je najosnovnija interakcija koju korisnik ostvaruje sa nekim oglasom, web sjedištem ili ostalim sadržajem. Ukratko, to je broj pregleda oglasa ili web sjedišta od strane korisnika u određenom vremenskom periodu. Svrha korištenja ove metrike i mjerenja prikaza oglasa u ovom radu je provjeriti relevantnost samih oglasa koje će Google ocijeniti pomoću algoritma. Google će češće prikazivati jedan oglas spram drugih oglasa u istoj kampanji, ukoliko na temelju broja klikova i konverzija ocijeni relevantnost istog za publiku. Na taj način mogu se izdvojiti oni oglasi koji su najrelevantniji za određeni *Ad set*. Pomoću broja prikaza ocijeniti će se najrelevantniji oglasi u setu novih oglasa. Broj prikaza oglasa je također značajna metrika ukoliko prikaz oglasa dovodi do klikova, koji će potom dovesti do mogućnosti konačne konverzije putem web obrasca. Pomoću korelacije broja prikaza sa drugim mjerenjima mogu se otkriti relevantna saznanja. Npr. broj prikaza u korelaciji sa brojem klikova na oglas govori nam o efikasnosti oglasa. To je metrika stope klikova (CTR) koja je objašnjena u sljedećem pod poglavlju.

## 5.2 Stopa klikova (eng. *click trough rate, CTR*)

Važna metrika za procijene uspješnosti oglasa je stopa klikova (CTR). Radi se o broju prikaza oglasa u odnosu na broj klikova te je u krajnjem obliku izražena u postotku. Pomoću ove metrike, u ovom radu, ocijenit će se kako ciljana publika reagira na nove oglase te će se donijeti procijene o uspješnosti novog teksta oglasa spram teksta postojećih oglasa. Stopa klikova je najčešće pokazatelj o efikasnosti dizajna ili teksta oglasa te ukoliko se npr. uoči velik broj broja prikaza oglasa spram malog broja klikova na isti, utoliko je potrebno

promijeniti neki segment dizajna ili komunikacije. Kada je riječ o suprotnoj situaciji, odnosno omjer broja prikaza oglasa je manji spram broja klikova na isti, to je pokazatelj da kampanja ide u dobrom smjeru. Korisnici tada često kliknu na oglas kada im je isti prikazan.

Stopa klikova računa se na sljedeći način:

$$\frac{\text{Broj klikova na oglas}}{\text{Broj prikaza oglasa}} * 100 = \text{Stopa klicanja na oglas}$$

Slika 11 Formula za računanje metrike stopa klikova

Broj klikova na oglas podijelimo sa brojem prikaza istog oglasa i dobiveni broj potom pomnožimo sa sto. Ishod je prikazan u postotku.

Stopa klikova je jedna od najbitnijih metrika kada se radi o kampanjama po modelu naplate po kliku, jer je upravo klik na oglas taj koji uzrokuje trošak za tvrtku. Svaka kompanija za sebe određuje vlastite KPI kada se govori o stopi klikova jer ona može ovisiti o raznim aspektima i karakteristikama samog poslovanja te postavljenim ciljevima. U ovom radu za kompaniju značajna stopa klikova iznosi između 10% i 15%.

Efikasana sadržaj i tekst oglasa najčešće povećavaju stopu klikova jer će korisnik kliknuti na onaj oglas koji optimalno odgovara na pitanje koje je postavio u tražilicu. U ovom radu pomoću metrike stopa klikova mjerit ćemo uspješnost teksta oglasa. Ukoliko će stopa klikova biti veća spram prijašnje, tekst oglasa možemo smatrati uspješnim.

Često je stopa klikova u korelaciji sa stopom konverzije (eng. *Conversion rate*) te bi u pravilu povećana stopa klikova trebala rezultirati i većim brojem konverzija. Iz tog razloga bi veća stopa klikova postignuta sa jednakim budžetom značila i da je tekst oglasa učinkovitiji i donosi više konverzija.

### 5.3 Stopa konverzije (eng. *conversion rate*)

Stopa konverzije je postotak korisnika koji dolaskom na web sjedište ispune neku radnju koju je tvrtka označila kao krajnji cilj. To može biti npr. prodaja, upis u program vjernosti,

instalacija aplikacije i sl. Stopa konverzije računa se tako da se broj ukupnih konverzija podjeli sa ukupnim brojem interakcija na oglas. Interakcije na oglas ovise o razini interesa korisnika, atraktivnosti ponude i jednostavnosti procesa koji zapravo vodi do ciljne konverzije. U ovom radu, za kompaniju, značajna stopa konverzije nije rangirana na osnovi postotka nego ovisi o tome koliko profitabilnih klijenata u krajnjoj točki kampanja generira. Stopom konverzije mjerit će se usklađenost između efikasnog teksta oglasa i broja upita. Naime, prema pretpostavci, broj upita, a samim time stopa konverzije bi trebala porasti u skladu sa porastom stope klikova. Ukoliko će stopa konverzije ostati jednaka a stopa klikova porasti, utoliko se može reći da cilj ovog projekta nije postignut.

## **5.4 Kategorija profitabilnosti i perspektivnosti potencijalnog klijenta**

Već je spomenuto da se tekstualni upit korisnika automatskim načinom slijeva u CRM sustav tvrtke u organiziranom profilu te takav odlazi u prodajni tunel na daljnju obradu od strane prodajnog tima. Upit je unaprijed kategoriziran prema procjeni profitabilnosti i perspektivnosti potencijalnog klijenta koje su automatski izračunate na temelju postojećih podataka pristiglih sa upitom a potvrđuju se nakon dodatno unesenih atributa u CRM sustavu putem provjere informacija o klijentu prilikom direktnog razgovora. Svaka nova osoba, koja je poslala upit i koja još nije postala klijent klinike dobiva u profil oznaku kao potencijalno profitabilan ili potencijalno neprofitan klijent. Za svakog klijenta, kako je prethodno navedeno, može se točno odrediti iz koje kampanje pristiže i koje je ključne riječi koristio pri pretraživanju usluge.

Za svakog klijenta automatski je izračunata i perspektivnost moguće konverzije na drugu fazu unutar marketinškog tunela tvrtke. U ovom diplomskom radu vrijednosti o perspektivnosti i profitabilnosti klijenta biti će izražene u postotku kao zajednički prosijek svih potencijalnih klijenata iz pojedine kampanje.

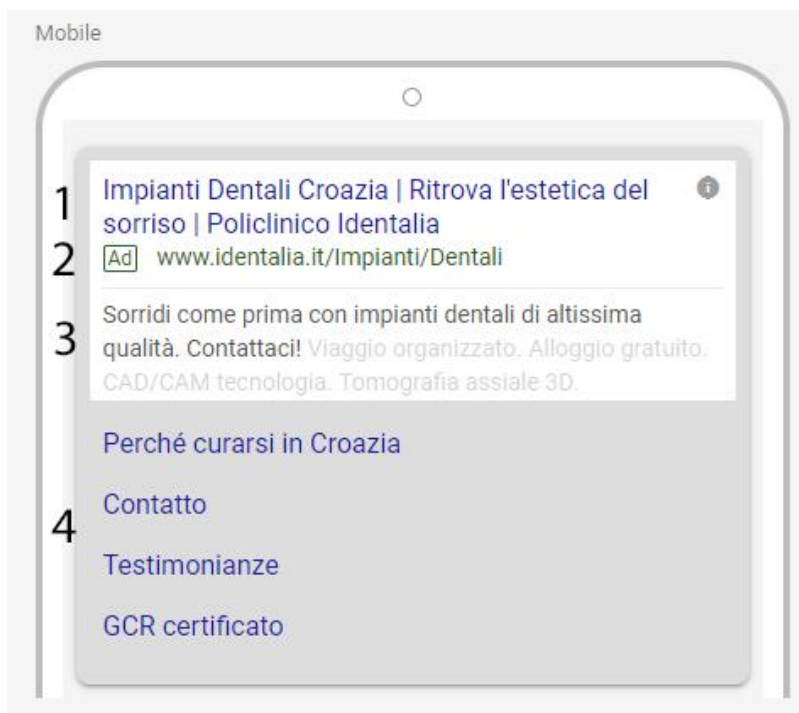
## 6. Kreiranje oglasa na Google Adwords platformi

Na internetu i u raznovrsnoj stručnoj literaturi mogu se pronaći različiti načini kreiranja kampanja za tražilice i kampanje po modelu naplata po kliku. Kreiranje kampanje koje se odnosi konkretno na Google Adwords platformu odvija se na tri razine a kreće od definiranja cilja i oblika isporuke oglasa na razini kampanje. Nakon tog koraka se prema ključnim riječima definiraju tematske grupacije, odnosno Ad grupe. Svaka Ad grupa okarakterizirana je skupinom ključnih riječi za koje se kreiraju oglasi. Za svaku skupinu ključnih riječi stoga se pišu i kreiraju oglasi koji bi se, kada korisnik upiše u tražilicu ključne riječi koje pripadaju toj specifičnoj ad grupi, trebali prikazivati kao rezultat pretrage. (slika 13)



Slika 12 Hijerarhija unutar Google Adwords kampanje za tražilicu Google [11]

Oblik plaćenog Google Adwords oglasa za Google pretraživač koji se testira u ovom radu sastoji se od naslova, opisa i ekstenzija. Premijer oglasa nalazi se na slici 13.



Slika 13 Primjer Google Adwords oglasa unuar računa tvrtke Identialia

Dijelovi Google Adwords oglasa: (slika 13)

- 1) Naslov oglasa: Google Adwords dozvoljava korištenje maksimalno tri naslova po oglasu i 30 karakternih znakova ne uzimajući u obzir razmake između riječi. Na ovom primjeru iskorišten je maksimalan broj naslova
- 2) Tematika stranice i oznaka za oglas: Ispod naslova uvijek je prikazan URL<sup>17</sup> web sjedišta na koje preusmjerava oglas i oznaka “Ad” koja korisniku daje do znanja da se radi o plaćenom oglasu
- 3) Opis oglasa: Google Adwords dozvoljava maksimalno dva opisa i sveukupno 90 karakternih znakova u tu svrhu, ne uzimajući u obzir razmake između riječi. Na ovom primjeru iskorišten je jedan od dva moguća opisa.
- 4) Ekstenzije: Google Adwrods dozvoljava ekstenzije na oglasima koje korisniku olakšavaju pristup užim informacijama o traženoj usluzi ili proizvodu. To mogu biti npr. ekstenzije sa brojem telefona, linkovima na određena web sjedišta, ekstenzija koja prikazuje geografsku lokaciju, ekstenzija za cijene, aplikacije ili razne promocije. U ovom slučaju korištene su ekstenzije koje služe kao prečac prema

---

<sup>17</sup> URL – usklađeni lokator sadržaja (eng. uniform resource locator), unikatna oznaka u obliku web linka koju se dodjeljuje sadržaju na internetu i koja služi kao poveznica između korisnika i traženog sadržaja. URL link može se dodijeliti bilo kojoj datoteci koja se nalazi na internetu.



specifičnim pod stranicama sa web sjedišta tvrtke. To su tzv. ne-klikabilne “call out” ekstenzije koje nisu prikazane svakom korisniku nego se aktiviraju samo onda kada je upisana specifična kombinacija ključnih riječi. *Call out* ekstenzije su prikazane na slici 13 sivom bojom u dijelu opisa oglasa.

U ovom diplomskom radu, fokus je na tekstu. Ako promatramo sliku br. 13 fokus analize odnosi se na dijelove oglasa koji se nalaze pred broja 1 i 3. Tekst je zapravo jedan od najbitnijih elemenata zaslužnih za privlačenje klikova na oglas te bi samim time trebao biti u skladu sa očekivanjima korisnika koji pomoću tražilice pretražuje određenu informaciju, uslugu ili proizvod. Oglas je u svojem krajnjem obliku zapravo jedini dio cjelokupne kampanje koji krajnji korisnik zapravo vidi i koristi te na temelju kojeg odlučuje hoće li ili neće posjetiti povezano web sjedište.

Nakon što se završi postupak grupiranja ključnih riječi i prema tome odgovarajući Ad setovi, moguće je jednostavnije kreirati sadržaj koji će reflektirati te ključne riječi za koje je namijenjen oglas. To znači da se ključne riječi koje opisuju Ad set trebaju spominjati i u tekstu oglasa unutar tog istog Ad seta, kako bi se oglas prikazao svaki puta kada pretraživač u tražilicu upiše pripadajuću ključnu riječ. Time oglas postaje efikasniji. Postoji više načina grupiranja ključnih riječi što uvelike ovisi o samom poslovanju, a najčešće jer riječ o Ad grupama određenima prema ključnim riječima koje reflektiraju pojedini proizvod ili uslugu koju tvrtka nudi i koji bi potencijalni klijent mogao pretraživati.

## **6.1 Metoda i način izrade teksta oglase na temelju ekspertize**

Tekst za oglase namijenjene za kampanje po modelu naplate po kliku do sada su u tvrtki kreirani slijedeći najprije neke od tehničkih uputstava od strane Google platforme.

- 1) Ključne riječi upisane od strane korisnika trebaju se pojavljivati u tekstu oglasa
- 2) Ključne riječi upisane od strane korisnika trebaju se ponavljati u URL-u stranice i u tekstu na web sjedištu
- 3) Tekst oglasa mora biti usklađen sa tekstom na web sjedištu na tematskoj razini i relevantan za korisnika, odnosno pružati mu odgovor na njegovo postavljeno pitanje.

Također su uzimane u obzir i neke od temeljnih preporuka za kreativno stvaranje teksta za oglase kako bi se što bolje istakli među oglasima konkurencije i kako bi oglasi mogli generirati rezultate iz Google-ovog priručnika su sljedeći:

1. Pisati privlačne i originalne tekstove kojima je u fokusu benefit za korisnika pošto isti reagira na oglas koji odgovara na njegove potrebe.
2. Kreativno povezati ključne riječi pošto korisnik reagira na oglas koji najefikasnije odgovara na njegovu potrebu.
3. Izbjegavati generičke izraze i koristiti specifične pozive na akciju kako bi korisniku ukazali najlakši put do odgovora.
4. Pisati tekstove u obliku odgovora te izbjegavati oglas u formi pitanja pošto korisnici koriste pretraživač kako bi našli odgovore na svoja pitanja.
5. Velika pažnja pridaje se naslovu kako bi bio prilagodljiv svim digitalnim uređajima
6. Koristiti najmanje 3 ekstenzije koje su bitne za poslovanje. Naime, Google tvrdi da oglasi sa više ekstenzija imaju bolju učinkovitost od oglasa sa npr. samo jednom ekstenzijom. Ekstenzije dodaju oglasu korisne informacije koje korisniku mogu biti od važnosti prilikom odabira na koji oglas će kliknuti. [12]

Osim uputstava za efikasno kreiranje teksta oglasa, od velike važnosti je i detaljno poznavanje poslovanja i samog klijenta. Ta dva segmenta su zapravo vodilja za kreiranje efikasnog teksta za oglase.

Upoznavanje sa klijentom u tvrtki se vrši na temelju različitih kvantitativnih i kvalitativnih analiza, no ono što predstavlja najveći izazov jest povezati dobivene informacije i izraziti ih pismenim putem na stranom jeziku. Kako bi se poštivalo pravilo koje nalaže da se komunikacija oglasa treba prilagoditi načinu na koji se izražava klijent, potrebno je poznavati način izražavanja ciljne skupine i detaljno poznavanje talijanskog jezika, pošto se u ovom slučaju radi o klijentima iz talijanskog govornog područja.

Tekstualni oglasi izrađeni na temelju ekspertize zanemarivali su do sada poznavanje jezika i način komunikacije korisnika ukoliko nije bilo dostupnih metoda i resursa. Izrađivani su na temelju generičkog pristupa usluzi koji je potom preveden na talijanski jezik i naknadno mijenjan, prilagođavan i poboljšavan kontinuiranim testiranjem oglasa kroz vrijeme. Riječi korištene za kreiranje oglasa preuzimane su osim iz internih analiza tvrtke također iz analize tzv. pojmova pretraživanja (eng. *Search terms*) dostupnih unutar Google Adwords sučelja. Radi se o popisu pojmova i kombinacija riječi koje su korisnici koristili u slučaju kada je

na njihov upit prikazan oglas tvrtke. U tim situacijama su se pojmovi upisani u tražilicu od strane korisnika poklopili sa jednom ili više ključnih riječi korištenih za oglašavanje usluge. Ova informacija koju na raspolaganje stavlja Google Adwords je praktična i primjenjiva za promatranje i analizu ponašanja korisnika u procesu transformacije misli u riječi prilikom potražnje usluge. S druge strane ta ista informacija ne govori nam dovoljno o samim očekivanjima korisnika koji su također bitni. To znači da nemamo informacije o tome što pretraživač očekuje, da će pretvorbom svojih misli u riječi i njihovim unosom u tražilicu, pronaći kao željeni odgovor.

## **6.2 Metoda i način izrade teksta za oglase na temelju rezultata analize teksta**

U ovom dijelu rada, u tekstualnim podacima, pokušalo se doći do odgovora o tome na koji način komunicira ciljna publika profitabilnih klijenata i što zapravo takav klijent želi pronaći kao odgovor na svoj upit unutar tražilice. Kada se radi o oglasima temeljenima na ekspertizi, ne možemo sa sigurnošću reći da su zahvaćene i isprobane sve moguće kombinacije riječi i izraza koje koristi potencijalni visoko profitabilni klijent.

U petom poglavlju ovog rada opisan je proces analize teksta proveden na tekstualnim upitima profitabilnih klijenata. Ishod analize je tablica u koju su uneseni tokeni koji su se ponavljali najvećom frekvencijom unutar tekstualnih upita a nikada nisu korišteni za tekstove oglasa. Tokeni su razdvojeni u 3 kategorije određene prema usluzi, opisu usluge i potražnji usluge. (Tablica 3)

Izrada teksta za oglas je subjektivni i kreativni proces koji ovisi o mnogo faktora prilikom kojeg je potrebno uzeti u obzir ciljeve kompanije, ciljnu skupinu, kupovni proces koji potencijalni klijent prolazi određenom trenutku te ostale faktore koji mogu utjecati na uspješnost same komunikacije unutar oglasa. Kampanje koje se nalaze u Google Ads računa tvrtke već su organizirane uzimajući u obzir sve dostupne informacije, a u ovom diplomskom radu, u procesu izrade oglasa, fokus je samo na optimizaciji komunikacije prema profitabilnoj i perspektivnoj ciljnoj skupini. U poglavlju 4 se u tablici br. 2 prikazuju najčešće ponavljani tokeni unutar teksta iz upita postojećih profitabilnih klijenata i koji su razvrstani u tri prethodno spomenute kategorije. U sljedećem koraku provjereno je jesu li ti tokeni prisutni u već postojećim oglasima unutar Google Adwords sučelja. Analizom su otkriveni i izdvojeni oni tokeni koji još nisu korišteni u tekstovima oglasa. Ti tokeni će se

integrirati u komunikaciju i testirat će se njihova učinkovitost. Uspješnost optimizacije kampanje ocijenit će se kroz metrike navedene u poglavlju 5.

Usporedbom analize postojećih oglasa i analize ishoda rudarenja teksta utvrđeno je sljedeće:

- 1) Unutar kategorije “usluga” otkriven je jedan token (Riječ A) koji nije korišten niti u jednom naslovu postojećih oglasa.
- 2) Unutar kategorije “potražnja usluge” otkriven je 1 token (Riječ J) koji nije korišten niti u jednom postojećem oglasu.
- 3) Unutar kategorije “opis usluge” otkrivena su 2 tokena (Riječ C i Riječ D) koja nisu korištena unutar već postojećih tekstova oglasa.

Tekst za oglase namijenjen za testiranje kreiran uzimajući u obzir tehnička i kreativna uputstva kao i pri izradi teksta oglasa na temelju ekspertize. Prilikom kreiranja teksta za nove oglase stavljena su u fokus četiri tokena iz tablice 3, za koje je uočeno da se do sada nisu nalazili unutar postojećih tekstova oglasa.

Za obje kampanje pripremljena su na temelju procijene odjela marketinga tvrtke, sljedeći oglasi sa pripadajućim kombinacijama tokena. Tokeni koji nikada nisu korišteni napisani su crvenom bojom unutar tablice 6.

Oglas:	Kombinacija korištenih tokena iz analize teksta:
Oglas A	Riječ P + Riječ A + Riječ B
Oglas B	Riječ P + Riječ B + Riječ J + Riječ C + Riječ D + Riječ A
Oglas C	Riječ P + Riječ B + Riječ J + Riječ C + Riječ A

Tablica 6 - Set oglasa sa pripadajućim kombinacijama tokena

### 6.3 Testiranje oglasa na Google Adwords platformi

U poglavlju 3 govorili smo o kampanjama za koje se analizom utvrdilo da, na temelju određenih ključnih riječi koje korisnik unosi u tražilicu Google, za tvrtku nose profitabilne klijente. Radi se o dvije kampanje koje će se u ovom radu koristiti za testiranje novih tekstova oglasa i vodit će se pod nazivima “Kampanja 1” i “Kampanja 2”.

Unutar obje kampanje nalaze se po tri Ad seta. U sva tri Adseta unutar obje kampanje unesena su četiri ogalsa kreirana kombinacijom tokena dobivenih iz analize teksta. Tekstovi su dodatno prilagođeni za uslugu i ključen riječi oko kojih je izgrađen pojedini Ad. Rezultate ćemo promatrati na razini kampanje, a oglase na temelju pozicije unutar Ad seta prema metrici stope klikova.

Usporedba rezultata vrši se pomoću podataka iz edvije vremenske serije. Prvi vremenski period je onaj koji se odnosi na 30 dana prije početka testiranja novih oglasa unutar izdvojenih kampanja. To je period između 22. lipnja 2019. i 21. srpnja 2019. Za obje kampanje očitane su vrijednosti za metrike korištene za procijene učinkovitosti oglasa koje su unesene u tablicu. Tablica sa vrijednostima nalazi se u poglavlju 6. Drugi vremenski period u kojem se očitavaju vrijednosti za metrike korištene za procijene učinkovitosti je od 22. Srpnja 2019. do 20. Kolovoza 2019. Promatrat će se promijene u vrijednostima metrika na kraju 2 vremenska perioda u trajanju od 30 dana.

Drugačija usporedba podataka učinkovitosti oglasa nije bila moguća iz vremenske perspektive pošto kampanja pod nazivom “Kampanja 1” nije postojala prošle godine u isto vrijeme i kreirana je u travnju 2019. S druge strane ako bismo u istom vremenskom periodu, na iste ključne riječi i unutar iste kampanje podizali novi Ad set sa novim oglsima, postali bismo konkurencija sami sebi na Google tražilici. Odnosno novi Ad seti koji se aktiviraju na jednake ključne riječi kao i postojeći *Ad seti* međusobno bi se natjecali, što nije u cilju kompanije. Postavke unutar obje kampanje i pripadajućih Ad seta ostale su jednake u cijelom periodu testiranja. Nisu mijenjani parametri *targetiranja* ciljne skupine, uloženi sredstava i krajnjih web sjedišta na koje vode oglasi iz Ad seta. Niti jedan oglas, prethodno kreiran na temelju ekspertize, nije dodatno optimiziran ili izmijenjen kako bi se dobio što precizniji rezultat testiranja.

Uspješnost oglasa iz perspektive profitabilnosti i perspektivnosti prikupljenih upita bit će izražena u postotku.

## 7. Mjerenje uspješnosti i analiza rezultata

Uspješnost teksta oglasa kreiranih na temelju ishoda analize teksta mjeri se usporedbom rezultata dvije jednake kampanje ali iz dva različita vremenska perioda. Period bez oglasa kreiranih na temelju analize teksta i period kada su ti oglasi uneseni unutar Ad setova navedenih kampanja. Dva osnovna kriterija i pravila uspješnosti izraženi korelacijom između metrika koje ćemo promatrati na sljedeći način:

- 1) Ukoliko na kraju testiranja stopa klikova povećana a broj impresija smanjen ili približno jednak prijašnjem, utoliko je tekst oglasa je smatran efikasnijim. Ako se uz to povećao i broj konverzija od strane perspektivnih klijenata, možemo rezultat testiranja smatrati uspješnim.
- 2) Ukoliko se na kraju testiranja stopa klikova povećala i sukladno tome povećao se broj prikaza oglasa, a stopa upita perspektivnih klijenata porasla u jednakom omjeru, ne možemo tvrditi da se radi isključivo o tekstu oglasa koji je utjecao na poboljšanje efikasnosti.

### 7.1 Rezultati kampanja i oglasa temeljenih na ekspertizi

Najprije su očitani dosadašnji rezultati kampanja koji su postignuti do 30 dana unatrag u periodu od 22. lipnja 2019. do 21. srpnja 2019. kroz metrike stopa prikaza, stopa klikova i stopa konverzije za obje kampanje. Radi se također o kampanjama koje nisu svrstane među najuspješnije kampanje tvrtke no specifične su po tome što nose rezultat sa najviše profitabilnih klijenata uz specifične ključne riječi. Stoga je u ovom radu stavljen naglasak na njihovu optimizaciju tih dvaju kampanja. Rezultati za prvi vremenski period navedeni su u sljedećoj tablici:

<b>Metrike:</b>	<b>Kampanja 1</b>	<b>Kampanja 2</b>
Broj prikaza	38.141	31.481
Stopa klikova	13.27 %	10.18 %
Stopa konverzije	0.26 %	0.16 %

Tablica 7 Metrike: broj prikaza, stopa klikova i stopa konverzije za kampanje „Kampanja 1“ i „Kampanja 2“ prije testiranja učinkovitosti oglasa temeljenih na ishodu analize teksta

Za isti period, prosječan broj potencijalnih klijenata koji su unutar CRM sustava označeni kao profitabilni klijenti i prosjek perspektivnosti po potencijalnom klijentu izraženi su u postotku:

Kampanja 1: potencijalna prosječna profitabilnost klijenta iznosi 60%. Perspektivnost po klijentu iznosi u prosijeku 6.64 %.

Kampanja 2: potencijalna prosječna profitabilnost klijenta iznosi 100 %. Perspektivnost po klijentu iznosi u prosijeku 12.33%.

Vrijednosti rezultata kampanja iz ovog vremenskog perioda služit će kao parametar za usporedbu i procijene uspješnosti teksta oglasa kreiranih na temelju ishoda analize teksta. Metrika stope klikova u ovom radu je ključna metrika za ocijene uspješnosti teksta, dok je metrika stope konverzije i postotak potencijalnih profitabilnih klijenata metrika za uspješnost poslovnog cilja.

## 7.2 Rezultati kampanja i oglasa temeljenih na analizi teksta

U periodu od 30 dana testiranja oglasi unutar obje nisu izmijenjivani.

Rezultati iz kampanje „Kampanja 1“ nisu pokazali niti pozitivne a niti negativne ishode. Google Ads naime automatski izmjenjuje oglase kako bi češće prikazivao samo one s najboljom izvedbom [13]. Oglasi temeljeni na analizi teksta unutar kampanje „Kampanja 1“ nisu prikazivani i smatrani su stoga nerelevantnima za pretraživača od strane Google tražilice. Za oglase temeljene na analizi teksta i pozicionirane unutar kampanje „Kampanja 1“ može se reći da nisu imali nikakav utjecaj na poboljšanje performansi same kampanje. Kampanja „Kampanja 2“ je s druge strane postigla iznimne rezultate koji su navedeni u tablici 8:

<b>Metrike:</b>	<b>Kampanja 1</b>	<b>Kampanja 2</b>
Broj prikaza	/	29.346
Stopa klikova	/	15.63%
Stopa konverzije	/	0.35

Tablica 8 Metrike: broj prikaza, stopa klikova i stopa konverzije za kampanje „Kampanja 1“ i „Kampanja 2“ nakon testiranja učinkovitosti oglasa temeljenih na ishodu analize teksta

Broj prikaza oglasa unutar kampanje „Kampanja 2“ je podjednak broju prikaza oglasa iz prethodnog testnog perioda dok je stopa klikova porasla sa 10.18% na 15.63% pa stoga možemo zaključiti da je tekst oglasa kreiran na temelju analize teksta iz upita profitabilnih klijenata uspješan te je podigao efikasnost i performanse kampanje. Najveći pozitivni porast ima metrika stopa konverzije koja je sa 0.16% porasla na 0.35% što je porast za više od 100% unutra 30 dana.

Ako se unutar kampanje „Kampanja 2“ spustimo na razinu Ad setova, najbolji performans pokazao je Ad set 1 i pripadajući set ključnih riječi. Taj Ad set davao je pozitivne rezultate već u prva 3 dana testnog perioda, a oglasi kreirani na temelju analize teksta su pozicionirani na prvo mjesto s najvećim brojem prikazivanja, najvećom stopom klikova i stopom konverzija spram ostalih oglasa unutar Ad seta. (slika 15)

Ad	Clicks	Impr.	CTR	Avg. CPC	Cost	Conversions	Cost / conv.	Conv. rate	Label
OGLAS B	1,322		9.83%					2.31%	Test text mi...
OGLAS C	14,972		20.64%					0.32%	Test text mi...

Slika 14 Sučelje iz Google Adwords računa tvrtke koje prikazuje oglase temeljene na analizi teksta unutar Ad seta 1 u kampanji „Kampanja 2“

Za testni period, prosječan broj potencijalnih klijenata koji su unutar CRM sustava označeni kao profitabilni klijenti i prosijek perspektivnosti po potencijalnom klijentu izraženi su u postotku:

Kampanja 1: /

Kampanja 2: potencijalna prosječna profitabilnost klijenta i dalje iznosi 100%. perspektivnost po klijentu u prosijeku iznosi 14.46%.

Iz dobivenih rezultata zaključeno je kako postoji specifična korelacija između Ad seta 1 iz kampanje „Kampanja 2“ koji nosi određen skup ključnih riječi i specifičnih tokena



dobivenih iz analize teksta i korištenih za kreiranje oglasa. Ovaj Ad set je ostaje daljnji predmet promatranja, testiranja i analize unutar tvrtke.

### 7.3 Usporedba i interpretacija rezultata

Oglasi čiji je tekst kreiran na ishodu analize teksta nisu imali utjecaj na performanse kampanje „Kampanja 1“. Oglasi se nisu niti prikazivali ukoliko, u korelaciji sa ključnim riječima, nisu smatrani relevantnima za pretraživača. Ti oglasi nisu uzimani u obzir za usporedbu rezultata. Uzroci ne prikazivanja navedenih oglasa mogu biti višestruki te u ovom radu ta tematika nije dalje obrađena.

Kampanja „Kampanja 2“ i Ad set 1 unutar iste kampanje pokazali su pozitivne rezultate. Ustanovljeno je da postoji pozitivna korelacija između ključnih riječi iz Ad seta 1 i teksta novih oglasa unutar istog Ad seta. Usporedba vrijednosti metrika i rezultata na razini kampanje „Kampanja 2“ iz 2 vremenska perioda nalazi se na tablici 9.

<b>Kampanja 2 / metrike:</b>	<b>Prvi period:</b>	<b>Drugi period:</b>
Broj prikaza	31.481	29.346
Stopa klikova	10.18 %	15.63%
Stopa konverzije	0.16 %	0.35 %

Tablica 9. Usporedba vrijednosti metrika broj impresija, stopa klikova i stopa konverzije na razini kampanje za oba perioda unutar testiranja učinkovitosti teksta oglasa za kampanju „Kampanja 2“

Uzevši u obzir pravilo da ako na kraju testiranja stopa klikova bude povećana a broj prikaza smanjen ili približno jednak prijašnjoj vrijednosti, tekst oglasa koji je kreiran pomoću analize teksta, unutar ovog Ad seta možemo smatrati efikasnijim. Stopa konverzija porasla je za više od 100% te su naknadno unutar CRM sustava provjerene vrijednosti o postotku potencijalno profitabilnih klijenata i prosijeku perspektivnosti po klijentu unutar kampanje.

Postotak potencijalno profitabilnih klijenata ostao je jednak nakon testiranja novih oglasa, odnosno i dalje iznosi 100%, no prosječna perspektivnost prema konverziji u drugu fazu unutar marketinškog tunela tvrtke porasla je sa 12.33% na 14.46%. Samim time možemo

zaključiti kako su tekstovi oglasa temeljeni na ishodu analize teksta imali pozitivan utjecaj na performanse kampanje.

Kombinacija tokena iz oglasa „Oglas B“ koji su navedeni unutar tablice 6 imala je najbolje performanse spram svih oglasa unutar kampanje „Kampanja 2“ prema stopi konverzije, dok je „Oglas C“ imao najbolje performanse prema stopi klikova. Oglas „Oglas A“ nije pokazao performanse niti u jednom Ad setu.

## 8. Zaključak

Unutar dva vremenska perioda, svaki u trajanju od 30 dana testirane su 2 kampanje za koje se ustanovilo da u svojim *Ad setima* sadrže najviše ključnih riječi korištenih od strane profitabilnog klijenta. Kako bismo maksimalno ispunili očekivanjima profitabilnog klijenta prilikom pretraživanja stomatološke usluge, rudarenjem teksta iz upita postojećih profitabilnih klijenata trtke došlo se do saznanja o načinu komunikacije ciljne skupine. Analizom su najprije izdvojeni svi tokeni koji su često korišteni u tekstu upita profitabilnih klijenata. Potom su usporedbom sa postojećim oglasima u posebnu tablicu izdvojeni samo oni tokeni koji do sada još nisu bili korišteni za kreiranje teksta i naslova za oglase. Na temelju tih tokena kreiran je set oglasa koji je testiran u obje kampanje na razini svih Ad seta. Testni period 1 u ovom radu uzima u obzir vremenski period od 30 dana prije testiranja oglasa temeljenih na analizi teksta. Testni period 2 odnosi se na vremenski period od 30 dana nakon pokretanja oglasa temeljenih na analizi teksta. Na temelju metrika broj prikaza, stopa klikova, stopa konverzije, profitabilnost i perspektivnost klijenta procijenila se uspješnost i efikasnost teksta oglasa temeljenih na analizi teksta.

Prethodno je definirano da se uspješnost oglasa mjeri većim postotkom broja konverzija od strane profitabilnog i perspektivnog klijenta iz testnog perioda 2 spram testnog perioda 1. Uspješnost teksta oglasa mjeri se povećanjem stope klikova i broja impresija u testnom periodu 2 spram testnog perioda 1.

Mjerenje je pokazalo da su oglasi temeljeni na analizi teksta imali pozitivan utjecaj na performanse kampanje. Jedna od dvije testirane kampanje pokazala je pozitivan rast u svim ključnim metrikama za projekt, osim za stopu prikaza. Ustanovljeno je da se analizom teksta provedenom nad upitima postojećih profitabilnih klijenata može doći do relevantnih saznanja o načinu komunikacije profitabilne ciljne skupine. Ta saznanja u ovom slučaju donijela su pozitivan rast za stopu klikova te porast u perspektivnosti potencijalnih klijenata koji su upit poslali putem te kampanje.

Iz ovog slučaja može se potvrditi da su saznanja dobivena analizom teksta iz upita akviziranih klijenata primjenjiva za kreiranje boljih i učinkovitijih tekstova oglasa za internetske tražilice te da će ti isti oglasi biti u mogućnosti privući više profitabilnih i perspektivnih klijenata za tvrtku.

## Popis kratica

CRM	<i>Customer relationship management</i>	sust. za upravljanje odnosa s klijentima
SQL	<i>Structured query language</i>	programski jezik za baze podataka
URL	<i>Uniform resource locator</i>	usklađeni lokator sadržaja
CTR	<i>Click trough rate</i>	marketinška metrika stopa klikova
PPC	<i>Pay per click</i>	model naplate po kliku
ROI	<i>Return on investment</i>	povrat fin. sredstava na uloženo
CSV	<i>Coma separated value</i>	oblik tekstualne datoteke

## Popis slika

Slika 1 . Koraci i konverzije unutar marketinškog tunela tvrtke .....	3
Slika 2 Faze unutar procesa rudarenja tekstualnih podataka te popis koraka koji se odnose na svaku fazu .....	11
Slika 3 Alatne jedinice unutar programa Orange korištene u ovom radu .....	14
Slika 4 Unošenje i provjera ispravnosti podataka u programu Orange unutar alatnih jedinica „Corpus“ i Corpus Viewer“ .....	<b>Error! Bookmark not defined.</b>
Slika 5 Pred-procesiranje teksta u alatnoj jedinici „Preprocess Text“ .....	<b>Error! Bookmark not defined.</b>
Slika 6 Prva provjera rezultata pred procesiranog teksta unutar alatne jedinice „Word Cloud“ .....	<b>Error! Bookmark not defined.</b>
Slika 7 Vizualizacija frekvencije ponavljanja unigrama unutar teksta iz upita profitabilnih klijenata .....	<b>Error! Bookmark not defined.</b>
Slika 8 Vizualizacija bigrama u programu Orange unutar teksta upita profitabilnih klijenata pomoću alatne jedinice „Word Cloud“ .....	22
Slika 9 Analiza konteksta u kojem se nalazi token Riječ A u alatnoj jedinici „Concordance“ .....	24
Slika 10 Alatne jedinice korištene za analizu teksta unutar programa Orange i cijelokupan proces analize teksta .....	25
Slika 11 Formula za računanje metrike stopa klikova .....	28
Slika 12 Hijerarhija unutar Google Adwords kampanje za tražilicu Google .....	30
Slika 13 Primjer Google Adwords oglasa unuar računa tvrtke Identialia .....	31
Slika 14 Sučelje iz Google Adwords računa tvrtke koje prikazuje oglase temeljene na analizi teksta unutar Ad seta 1 u kampanji „Kampanja 2“ .....	39

## Popis tablica

Tablica 1 Tehnike za ekstrakciju znanja iz tekstualnih podataka u fazi rudarenja teksta ...	12
Tablica 2 Najviše ponavljane riječi unutar teksta oglasa razvrstane su u tri kategorije: usluga, način potražnje usluge i opis tražene usluge .....	20
Tablica 3 Tablica sa ispisanim relevantnim tokenima po kreiranim kategorijama koji do sada nisu korišteni za tekstove oglasa .....	21
Tablica 4 Relevantne komb. tokena u bigramu i trigramu po frekvenciji ponavljanja .....	23
Tablica 5 Relevantne kombinacije tokena u bigramu i trigramu za token „Riječ A“ .....	23
Tablica 6 Set oglasa sa pripadajućim kombinacijama tokena	<b>Error! Bookmark not defined.</b>
Tablica 7 Metrike: broj prikaza, stopa klikova i stopa konverzije za kampanje „Kampanja 1“ i „Kampanja 2“ prije testiranja učinkovitosti oglasa temeljenih na ishodu analize teksta ..	37
Tablica 8 Metrike: broj prikaza, stopa klikova i stopa konverzije za kampanje „Kampanja 1“ i „Kampanja 2“ nakon testiranja učinkovitosti oglasa temeljenih na ishodu analize teksta	38
Tablica 9 Usporedba vrijednosti metrika broj impresija, stopa klikova i stopa konverzije na razini kampanje za oba perioda unutar testiranja učinkovitosti teksta oglasa za kampanju „Kampanja 2“ .....	<b>Error! Bookmark not defined.</b>

## Literatura

- [1] IDENTALLIA, Informacije iz internih istraživanja tvrtke Identallia
- [2] PROMORESPEKT, <https://www.promorespekt.hr/ppc-pay-per-click-oglašavanje/> (srpanj 2019)
- [3] CLUTCH, <https://clutch.co/seo-firms/resources/search-engine-marketing-why-people-click-paid-search-ads> (kolovoz 2019).
- [4] Advanced Google AdWords, John Wiley & Sons, Inc; Goddes B.; Indianapolis; str. 7-11
- [5] GOOGLE ADS SUPPORT, [www.support.google.com](http://www.support.google.com) (kolovoz 2019)
- [6] EDUCBA, <https://www.educba.com/text-mining> (kolovoz, 2019)
- [7] Poboljšanje dijagnostike prekida pretplatničkog odnosa pomoću dubinske analize internetskog sadržaja, Diplomski rad, Zdravko Kunić, Visoko učilište Algebra, srpanj 2015
- [8] AMAZON BLOG, <https://aws.amazon.com/blogs/big-data/applying-machine-learning-to-text-mining-with-amazon-s3-and-rapidminer/> (rujan 2019)
- [9] Primjerna tehnika rudarenja podacima u telekomunikacijama, Diplomski rad, Zoran Kundek, Sveučilište Jurja Strossmayera u Osijeku, Fakultet elektrotehnike, računarstva i informacijskih tehnologija, 2017
- [10] Rudarenje Teksta: Analiza sentimenta, završni rad, Igor Dubrek, Visoko učilište Algebra, ožujak 2016
- [11] WORDSTREAM, <https://www.wordstream.com/blog/ws/2017/05/02/adwords-account-structure-guide> (kolovoz 2019)
- [12] GOOGLE ADS SUPPORT, <https://support.google.com/google-ads/answer/6167122?hl=en> (kolovoz 2019)
- [13] GOOGLE ADS SUPPORT, <https://support.google.com/google-ads/answer/1704392?hl=hr> (kolovoz 2019)

Student vlastoručno potpisuje diplomski rad iza zaključka s datumom i oznakom mjesta završetka rada te naznakom:

*„Pod punom odgovornošću pismeno potvrđujem da je ovo moj autorski rad čiji niti jedan dio nije nastao kopiranjem ili plagiranjem tuđeg sadržaja. Prilikom izrade rada koristio sam tuđe materijale navedene u popisu literature ali nisam kopirao niti jedan njihov dio, osim citata za koje sam naveo autora i izvor te ih jasno označio znakovima navodnika. U slučaju da se u bilo kojem trenutku dokaže suprotno, spreman sam snositi sve posljedice uključivo i poništenje javne isprave stečene dijelom i na temelju ovoga rada“.*

*U Zagrebu, datum.*

*Ljiljana Petrak*