DIPLOMSKO DELO

UPORABA PODATKOVNEGA RUDARJENJA PRI ANALIZI NAKUPNIH NAVAD KUPCEV BENCINSKIH SERISOV PODJETJA PETROL D.D. LJUBLJANA

Ljubljana, avgust 2008

MAJA SANCIN
IZJAVA

Študentka MAJA SANCIN izjavljam, da sem avtorica tega diplomskega dela, ki sem ga napisala pod mentorstvom MAG. ALEŠA POPOVIČA, in dovolim njegovo objavo na fakultetnih spletnih straneh.


Podpis: _____________________
KAZALO

1 OPREDELITEV IN POMEN PODATKOVNEGA RUDARJENJA ..............................................2
2 POMEMNIBI POJMI ........................................................................................................3
3 ŠEST NALOG PODATKOVNEGA RUDARJENJA .................................................................4
   3.1 Klasifikacije ..............................................................................................................4
   3.2 Ocenjevanje ..............................................................................................................4
   3.3 Napovedovanje .........................................................................................................4
   3.4 Opisovanje ................................................................................................................5
   3.5 Asociacije .................................................................................................................5
   3.6 Razvrščanje v skupine ..............................................................................................5
4 PODROČJA UPORABE PODATKOVNEGA RUDARJENJA ..................................................6
   4.1 Podatkovno rudarjenje v poslovnih vedah ................................................................7
      4.1.1 Izdelava profilov kupcev ..................................................................................7
      4.1.2 Segmentacija ....................................................................................................8
      4.1.3 Odzivnost strank ...............................................................................................8
      4.1.4 Upravljanje s tveganji .......................................................................................8
      4.1.5 Aktiviranje kupca .............................................................................................8
      4.1.6 Navzkrižna prodaja in modeli »več enakih izdelkov« .....................................8
      4.1.7 Analiza nakupne košarice .................................................................................9
      4.1.8 Obraba ali »mešalnik« ......................................................................................9
5 TEHNIKE PODATKOVNEGA RUDARJENJA ...................................................................9
   5.1 Najблиžji sosed ........................................................................................................9
   5.2 Asociativna pravila .................................................................................................10
   5.3 Odločitveno drevo ..................................................................................................11
   5.4 Razvrščanje v skupine ............................................................................................12
   5.5 Umetne nevronske mreže .......................................................................................12
6 PROCES PODATKOVNEGA RUDARJENJA ...................................................................13
7 ANALIZA PODATKOV IZ ANKETE PODJETJA PETROL d.d ........................................15
   7.1 Opis podjetja Petrol d.d. .........................................................................................15
   7.2 Opis ankete ..............................................................................................................16
   7.3 Orodje Visual Studio 2005 .......................................................................................17
   7.4 Analiza podatkov ....................................................................................................18
SKLEP ....................................................................................................................................22
LITERATURA IN VIRI ............................................................................................................24
PRILOGE
KAZALO SLIK

Slika 1: Odločitveno drevo za ugotavljanje kreditne sposobnosti........................................... 11
Slika 2: Preprosta nevronska mreža .......................................................................................... 13
Slika 3: Proces podatkovnega rudarjenja (CRISP-DM model)............................................... 14
Slika 4: Profil najboljših strank ................................................................................................ 20

KAZALO PRILOG

PRILOGA 1: Vprašalnik ............................................................................................................. 1
PRILOGA 2: Seznam podatkov iz baze podatkov .............................................................. 4
PRILOGA 3: Demografski podatki anketirancev .............................................................. 7
PRILOGA 4: Nakupno obnašanje anketirancev ................................................................. 9
PRILOGA 5: Analiza vpliva demografskih in socioekonomskih značilnosti kupcev na potrošnjo .................................................................................................................................. 10
PRILOGA 6: Analiza profila najboljših kupcev ....................................................................... 12
PRILOGA 7: Analiza vpliva mnenj strank z različnimi nakupnimi obnašanji o kakovosti za Petrol ........................................................................................................................................ 13
PRILOGA 8: Asociativna pravila: vpliv mnenja o zaupanju prodajalcem na potrošnjo samo goriva ........................................................................................................................................ 14
PRILOGA 9: Slovarček slovenskih prevodov tujih izrazov .................................................... 15
UVOD

Zmožnosti za ustvarjanje in shranjevanje podatkov se iz dneva v dan povečujejo. Eksplozivna rast shranjenih podatkov in čedalje ostrejša konkurenca v globaliziranem svetu ustvarjata potrebo po novih tehnikah in avtomatiziranih orodjih, ki nam lahko pomagajo pri pretvarjanju obsežne količine podatkov v informacije in znanje za učinkovito odločanje. Pri tem nam je v veliko pomoč podatkovno rudarjenje, s katerim odkrivamo informacije, vzorce in povezave v velikih količinah podatkov. Pridobljeno znanje lahko nato uporabimo za boljše (poslovno) odločanje in za zagotavljanje konkurenčne prednosti.


Namen diplomskega dela je v analizi literature s področja podatkovnega rudarjenja, uporabi pridobljenega znanja na primeru iz prakse in s tem v pridobivanju uporabnih informacij iz baze podatkov podjetja Petrol. Cic diplomskega dela je na podlagi teoretičnega znanja analizirati bazo podatkov Petrola in pridobiti koristne informacije, ki so lahko podlagu za nadaljnje odločanje (predvsem na področju trženja in prodaje) v podjetju Petrol. Pri tem je moj cilj analizirati ciljne kupce v vplive na večjo potrošnjo. Na podlagi tega lahko podjetje sestavi trženski splet in se usmeri v bolj donosne kupce.

Raziskovalni pristop gre pri mojem delu iz teoretičnega na praktično in iz splošnega na specifično. Najprej sem analizirala splošno literaturo o podatkovnem rudarjenju, nakar sem za vsako področje pregledala specifične članke (iskala sem jih v znanstvenih bazah člankov). Teoretično znanje sem nato skušala tudi praktično uporabiti.


Diplomska naloga je razdeljena na sedem poglavij, pri čemer jih je prvih šest teoretičnih, sedmo poglavje pa predstavlja praktičen del diplomske naloge. V prvem poglavju opredeljujem pojem podatkovnega rudarjenja in razložim, zakaj ima dandanes tako velik

1 OPREDELITEV IN POMEN PODATKOVNEGA RUDARJENJA


Velič pomen podatkovnega rudarjenja danes in v prihodnosti ponazarja že dejstvo, da količina zbranih podatkov eksponentno narašča. Pri tem si lahko postavimo vprašanje, kaj se posameznik lahko sploh nauči iz teh podatkov. Naraščajoče in ogromne količine podatkov so neobvladljive le z uporabo tradicionalnih načinov analiziranja podatkov. Zato so orodja za podatkovno rudarjenje pri odkrivanju znanja iz podatkov v veliko pomoč (Dürr, 2003, str. 3).

Larose (2005, str. 4) našteva naslednje razloge za znatno razvijanje področja podatkovnega rudarjenja:

- eksponentno večanje zbiranja podatkov;
- podjetja uporabljajo podatkovna skladišča, ta pa predstavljajo zanesljiv vir podatkov;
- naraščajoča dostopnost in uporaba interneta;

2
naraščajoči pritisk tekovalnosti v globaliziranem svetu;
apredki v tehnologiji in povečanje prostora na trdem disku.

2 POMEMBNI POJMI

Medtem ko ima model splošen značaj, so vzorci specifični. Vzorec je lokalna struktura, ki daje specifično izjavo o nekaj spremenljivkah ali nekaj podatkih (Laxman & Sastry, 2006, str. 175). Vzorec nam na primer pokaže kombinacijo dogodkov z visoko korelacijo.


Spremenljivke lahko ločimo po tem, koliko informacij vključujejo v sebi (Pyle, 2003, str. 314):

- **nominalne spremenljivke** so le označbe ali imena, uporabljenega za identificiranje posameznih entitet. Na primer so lahko to imena, serijske številke, kode itd. S takimi spremenljivkami ne moremo računati.
- **kategorične (diskretne) spremenljivke** identificirajo skupino entitet. Entitete, ki spadajo v določeno kategorijo, imajo določene skupne lastnosti, ki so značilne za to skupino. Na primer spadajo vse študente v kategorijo ljudi.
- **ordinalne spremenljivke** so kategorije, ki jih lahko racionalno naštejemo v določenem vrstnem redu. Primer za tako kategorijo vključuje vrstni red vroče > milo > hladno. Pri tem poznamo vrstni red, ne poznamo pa točne razdalje med vrednostmi. Teh vrednosti zato ne moremo seštevati ali odštevati.
- **intervalne spremenljivke** so ordinalne spremenljivke, pri katerih lahko tudi merimo razdaljo med vrednostmi. Torej niso le razvrščene, a so tudi merjene v fiksnih in enakih enotah. Primer intervalne spremenljivke je temperatura v stopinjah Celzija. Računanje razlike takih vrednosti je smiselno, seštevanje pa ni.
- **razmernostne spremenljivke** so spremenljivke, pri katerih obstaja veljavno razmerje in neka spodnja meja. S takimi števili ravnamo kot z realnimi števili in torej lahko z njimi računamo. Razdalja je dober primer razmernostne spremenljivke.

3 ŠEST NALOG PODATKOVNEGA RUDARJENJA

S šestimi nalogami podatkovnega rudarjenja, ki jih opisujem v tem poglavju, lahko rešujemo veliko ekonomskih, intelektualnih, poslovnih in drugih problemov.


3.1 Klasifikacije


3.2 Ocenjevanje


3.3 Napovedovanje

Napovedovanje (ang. *Prediction*) je zelo podobno klasificiranju in ocenjevanju. Pri napovedovanju klasificiramo na podlagi nekega napovedanega prihodnjega obnašanja ali na


3.4 Opisovanje


3.5 Asociacije


Analizo asociacij uporabljamo predvsem pri analizi nakupne košarice ali pri analizi podatkov iz transakcij (Han & Kamber, 2000, str. 13-14).

3.6 Razvrščanje v skupine

Razvrščanje v skupine (ang. Clustering) je segmentiranje heterogene populacije v več homogenih skupin. Razlika med razvrščanjem v skupine in klasifikacijami je ta, da pri razvrščanju v skupine želimo analizirati objekte, za katere ne poznamo razredov. Pri klasifikacijah spada vsak objekt v določen razred, ki ga dobimo s pomočjo modela, ki smo

Naloga razvrščanja v skupine je ponavadi prvi korak v raziskovanjih s pomočjo podatkovnega rudarjenja. Rezultate te naloge uporabljajo potem analitiki kot vhode za nadaljnja raziskovanja z različnimi tehnikami podatkovnega rudarjenja (Larose, 2005, str. 17).

4 PODROČJA UPORABE PODATKOVNEGA RUDARJENJA

Tehnike podatkovnega rudarjenja uporabljajo analitiki iz več področij znanosti: iz poslovnih ved, astronomije, medicine, molekularne biologije, geologije in še bi lahko naštevala.


Podatkovno rudarjenje je koristno tudi pri ugotavljanju prevar. Na primer lahko, s pomočjo podatkovnega rudarjenja, razkrijezmo zlorabe kreditnih kartic. Nekatere transakcije s kreditnimi karticami so neveljavne in pri teh ni problemov. Obstajajo pa tudi transakcije, ki so formalno korektna, a gre v bistvu za sleparije. S povečanjem števila izdanih kreditnih kartic in z napredki v elektronskem poslovanju postaja ugotavljanje prevar čedalje bolj pomembno. Pri tem znanje strokovnjakov ni dovolj, saj se metode prevar spreminjajo in je zato potrebno stalno učenje. Pri detekciji prevar s kreditnimi karticami so umetne nevronske mreže najbolj uspešna tehnika podatkovnega rudarjenja, ki zasači zlorabe na podlagi učenja po zgodovinskih podatkih. Uporabna so tudi asociativna pravila, s katerimi lahko izvemo, v katerih primerih je verjetnost prevare največja (Brause, Langsdorf & Hepp, 1999, str. 103-106).

Tudi na področju ugotavljanja in napovedovanja kriminalnih dejanj so orodja za podatkovno rudarjenje v veliko pomoč. Uporabljajo jih zvezne agencije (kot na primer ameriška CIA in FBI) v borbi zoper prometnim prekrškom, spolnim zlorabam, kraji, prevaram, požigom,
Na znanstvenem področju ima podatkovno rudarjenje velik pomen. V biologiji pomaga na primer pri identifikaciji na tisoče genov v vsakem genomu. V biomedicini uporabljajo podatkovno rudarjenje za analizo kemičnih lastnosti in tem pospešujejo odkritja na tem področju ter znižujejo stroške raziskovanj. V astronomiji so ustvarili popolnoma avtomatiziran sistem katalogiranja oddaljenih nebeških objektov, ki jih z vizualnim nadzorom ne moremo zaznati. Tehnike podatkovnega rudarjenja so se izkazale na vseh zgoraj omenjenih področjih in s svojo »spretnostjo« konkurirale strokovnjakom ali jih celo prekosile (Witten & Frank, 2005, str. 28).

Ker je diplomsko delo usmerjeno v uporabo podatkovnega rudarjenja na področju poslovnih ved, bom to podrobneje obravnavala v naslednjem podpoglavju. Znanje o primerih uporabe podatkovnega rudarjenja je nujno za določanje cilja analize in za izvedbo analize. V poslovnih podjetjih uporabljajo podatkovno rudarjenje za odkrivanje pomembnih informacij, s katerimi lahko bolje sprejemajo odločitve, zmanjšajo tveganja in ustvarjajo konkurenčno prednost.

4.1 Podatkovno rudarjenje v poslovnih vedah


4.1.1 Izdelava profilov kupcev

Poznavanje stranke (kupca) je dandanes zelo pomembno. Izdelava profilov kupcev (ang. Profile Analysis) je odličen način spoznavanja stranka in vključuje določitev skupnih značilnosti populacije, za katero se zanimamo. Pri tem vzamemo v poštev razne demografske značilnosti (kot na primer spol, starost, zakonski stan itd.) in druge karakteristike, ki so pomembne za posel (na primer povprečno višino prodaje). Profil stranke lahko ugotavljamo z analizo odvisnosti, z razvrščanjem v skupine in z opisovanjem konceptov. Na podlagi profila stranke se tržnik odloči o strategijah in taktikah zadovoljevanja strank. Izdelovanje profila stranke je posebej uporabno pri ciljnih tržnih segmentih (Shaw, Subramaniam, Tan & Welge, 2001, str. 133).
4.1.2 Segmentacija

4.1.3 Odzivnost strank

4.1.4 Upravljanje s tveganji

4.1.5 Aktiviranje kupca

4.1.6 Navzkrižna prodaja in modeli »več enakih izdelkov«
Navzkrižno prodajo je na primer učinkovito uporabilo podjetje Tesco iz Velike Britanije, ki je v glavnem trgovsko podjetje, a se je v koncu 20. stoletja diverzificiralo (vstopilo je tudi na področje finanč). S pomočjo kartice zvestobe so v podjetju zbirali podatke iz transakcij v prodajalnah. Podatke so potem analizirali s pomočjo podatkovnega rudarjenja in pridobljene informacije uporabili tudi pri prodaji finančnih storitev. Navzkrižna prodaja jim je omogočila pridobivanje strank za finančne storitve s 50% nižjimi stroški od drugih finančnih institucij (Mukund, 2003, str. 9).

4.1.7 Analiza nakupne košarice


4.1.8 Obraba ali »mešalnik«


5 TEHNIKE PODATKOVNEGA RUDARJENJA

Obstaja več tehnik (algoritmov oziroma metod) podatkovnega rudarjenja, s katerimi lahko opravljamo v tretjem poglavju opisanih šest nalog podatkovnega rudarjenja. Izbira tehnik je v praksi odvisna od poslovnega primera. V tem poglavju predstavljam pet najbolj znanih algoritmov podatkovnega rudarjenja: tehniko najbližjega sosed, asociativna pravila, odločitvena drevesa, razvrščanje v skupine in nevronske mreže.

5.1 Najблиžji sosed


Tehnico najbližjega soseda lahko na primer uporabi zdravnik, ko nov bolnik potrebuje zdravilo. Zdravnik mora bolniku predpisati zdravilo, s katerim so ozdravili bolniki s podobnim profilom. Primerja lahko starost, simptome, spol in druge spremenljivke in na podlagi tega dobi najbolj podobnega preteklega bolnika. Ta je v tem primeru »najблиžji sosed«. Zdravilo, ki ga je zdravnik njemu predpisal, bo sedaj predpisal tudi novemu pacientu.

5.2 Asociativna pravila

S tehnico asociativnih pravil iščemo atribute ali značilnosti, ki »gredo skupaj«. To pomeni, da skušamo razkriti asociacije med temi atributi tako, da odkrižemo pravila, s katerimi lahko izmerimo zvezo med dvema ali več atributi. Ta tehnika je predvsem uporabna pri analizi nakupne košarice (Larose, 2005, str. 180).


Primer asociativnega pravila je lahko naslednji: »Če kupec kupi brezalkoholno pijačo, potem kupi tudi čipse.« Pri tem pravilu je podpora 5%, zaupanje pa 70%.
Ob ugotavljanju pravil analitik določi minimalno podporo in minimalno zaupanje pravila. Omejitev nastavi zato, ker bi bilo v nasprotnem slučaju visoko število pravil, ki le redkodaj veljajo.

Problem analize nakupne košarice je ta, da pri tem pridemo večkrat do neuporabnih spoznanj. Primer je lahko pravilo: »če kupec kupi lončec, potem kupi tudi pokrovko«. Tako spoznanje je samo po sebi umevno in zato ni zanimivo. Analitik cilja vedno na prej nepoznana pravila, ki so skrita v podatkih in ki so lahko uporabna za nadaljnje odločanje in poslovanje, s katerimi lahko podjetje doseže konkurenčno prednost.

5.3 Odločitveno drevo


*Slika 1: Odločitveno drevo za ugotavljanje kreditne sposobnosti*

Drevesa odločitve so posebno dobra pri ugotavljanju segmentov populacije, ki se obnašajo na tak način, kot mi želimo (Parr Rud, 2001, str. 19). Na primer lahko analiziramo značilnosti naših kupcev, ki so unovčili poslan kupon za določen izdelek.

Na Sliki 1 (na strani 11) je prikazan primer odločitvenega drevesa za ugotavljanje kreditne sposobnosti kupcev. Razred 0 predstavlja stranke, ki niso kreditno sposobne, medtem ko so kupci iz Razreda 1 kreditno sposobni. Kot lahko opazimo, zaposleni z dohodki pod 100.000 denarnih enot in z dolgom nad 250.000 denarnih enot niso vredni kredit. Obratno velja za študente, za katere jamčijo starši.

5.4 Razvrščanje v skupine
Z razvrščanjem v skupine (ang. Clustering) dobimo skupine predmetov (ali ljudi), ki imajo podobne značilnosti. Naš cilj je, da se skupine med sabo čim bolj razlikujejo, znotraj skupine pa želimo doseči čim višjo homogenost.

Pri tehniki razvrščanja v skupine najprej dobimo diagram, ki kaže kako enote padejo v skupine. V najbolj preprostem slučaju lahko to preprosto naredimo tako, da upodobimo enote na podlagi dveh dimenzij in logično razdelimo sliko ter tako dobimo skupine (Witten & Frank, 2005, str. 81).


5.5 Umetne nevronske mreže
Inspiracija za umetne nevronske mreže (ang. Artificial Neural Networks) je bilo spoznanje kompleksnega delovanja sistema učenja živalskih možganov, ki jih sestavljajo povezani nizi nevronov. Čeprav ima neuron relativno preprosto strukturo, lahko goste mreže povezanih neuronov opravljajo zahtevne naloge učenja, kot so na primer klasifikacije in prepoznavanja vzorcev. Umetne nevronske mreže predstavljajo poskus na osnovni ravni posnemanja vrste nelinearnega učenja, ki je značilna za nevronske mreže, ki jih ustvarja narava (Larose, 2005, str. 128-129).

Na Sliki 2 (na strani 13) je prikazan enostaven primer nevronske mreže. Nevronska mreža lahko vodi le v eno smer in ne omogoča vrtenja podatkov. Sestavljena je iz dveh ali več

Da se lahko nevronska mreža uči, moramo preskrbeti skupok zgodovinskih vhodnih podatkov in ustreznih poznanih vrednosti izhodov. Nevronska mreža se uči tako, da skuša postopno razumeti povezavo med vhodnimi in izhodnimi podatki. Iterativno uravnava uteži za vsak testni primer, dokler povprečna napaka na kvadrat med ocenjeno in dejansko vrednostjo ne nahaja na predhodno določenem zadovoljivem nivoju (De Barcelos Tronto, Simoes da Silva & Sant'Anna, 2008, str. 359).

Slika 2: Preprosta nevronska mreža

Vhodni sloj sprejema vhode iz niza podatkov (na primer vrednosti atributov) in jih preprosto poda naprej. V naslednji fazi se vhodni podatki pretvorijo s pomočjo transformacijske funkcije (ki je običajno seštevek zmnožkov uteži in vrednosti spremenljivk) in gredo v naslednji sloj. Postopek se ponavlja, dokler podatki ne pridejo iz izhodnega sloja. Iz hod iz izhodnega sloja predstavlja rešitev problema.

6 PROCES PODATKOVNEGA RUDARJENJA


**Slika 3: Proces podatkovnega rudarjenja (CRISP-DM model)**

---

Na spletni strani CRISP-DM modela (http://www.crisp-dm.org/Process/index.htm) lahko najdemo opise posameznih faz:

1. **faza: razumevanje poslovanja ali raziskave**
   V prvi fazi smo usmerjeni v razumevanje ciljev projekta in zahtev iz poslovnega zornega kota. To potem pretvorimo v definicijo problema podatkovnega rudarjenja. Poleg tega moramo pripraviti preliminarno strategijo za doseganje zadanih ciljev.

2. **faza: razumevanje podatkov**
   Zbrati moramo podatke in jih raziskovati, da pridemo v stik z njimi in da jih lahko bolje razumemo. Poleg tega je potrebno oceniti kakovost podatkov.

3. **faza: priprava podatkov**
   Faza priprave podatkov je precej mučna. Podatki prihajajo iz različni virov podatkov in posledično obstajajo razlike v časovnih dobah, glavnih ključih, napakah, načinih shranjevanja, agregaciji podatkov itd. Zato je potrebno podatke zbrati, jih integrirati in prečistiti. Iz začetnih surovih podatkov moramo pripraviti končni skup podatkov, ki ga...
bomo uporabljali v vseh naslednjih fazah. Izbriati moramo vse tabele, enote in spremenljivke, ki jih želimo analizirati in ki so primerni za naš analiz. Če je potrebno, moramo pretvoriti nekatere spremenljivke v drugo obliko. Prečistiti moramo surove podatke, da so pripravljeni za orodja modeliranja.

4. faza: modeliranje
V četrti fazi moramo izbrati in aplicirati primerno tehniko modeliranja. Večkrat se zgodi, da je potrebno uporabiti več tehnik modeliranja za reševanje istega problema. Z namenom optimiziranja rezultatov moramo nato preveriti nastavitev parametrov. Če določena tehnika zahteva spremembe v podatkih, se lahko vnemo v fazo priprave podatkov.

5. faza: vrednotenje
Na začetku te faze imamo že izgrajen model (ali več modelov), ki z vidika analize podatkov zgleda visoko kvaliteten. Preden uporabimo model, moramo bolj podrobno ovrednotiti njegovo kvaliteto in učinkovitost. Zagotoviti moramo, da lahko z modelom dejansko dosežemo zadane poslovne ali raziskovalne cilje in da nismo pozabili česa pomembnega. Na koncu te faze se moramo odločiti o uporabi rezultatov podatkovnega rudarjenja.

6. faza: razvijanje
Projekt se ponavadi ne konča takrat, ko je model ustvarjen. Tudi ko je cilj projekta le povečanje znanja, je potrebno to znanje primerno organizirati in predstaviti, da se ga lahko potem uporabi. Zadnja faza projekta je lahko preprosta (ko na primer naredimo le poročilo) ali kompleksna (ko na primer implementiramo ponavljajoč proces podatkovnega rudarjenja). Kakor koli že, moramo na koncu razumeti, katera dejanja bomo morali izvesti, da bomo kreiran model tudi dejansko uporabili.

7 ANALIZA PODATKOV IZ ANKETE PODJETJA PETROL d.d.

7.1 Opis podjetja Petrol d.d.
Podjetje Petrol spada med največje slovenska podjetja: je največje slovensko energetsko podjetje, največji slovenski uvoznik in podjetje z najvišjimi prihodki v Sloveniji. Poleg tega spada med največje slovenske trgovske družbe in ponuja delo več kot 500 zaposlenim (Spletna stran podjetja Petrol, 2008).

Osrednja dejavnost podjetja je trgovanje z naftnimi derivati, zemeljskim in utekočinjenim plinom ter električno energijo. Pri tem podjetje ustvarja nekaj več kot 80% vseh svojih prihodkov. Preostalih nekaj manj kot 20% prihodkov pridobi s trgovanjem z blagom za
široko porabo in s storitvami. Na tem perspektivnem področju ima Petrol 10% realno letno stopnjo rasti (Spletna stran podjetja Petrol, 2008).

Podjetje Petrol stremi k poslovni odličnosti in spremlja ter spoštuje tržne zahteve s področja storitev, informacijske tehnologije in varovanja okolja (Spletna stran podjetja Petrol, 2008). V zadnjem desetletju skuša s pomočjo raziskav zagotoviti celovite informacije za vse faze nakupne aktivnosti potrošnikov. Raziskave dajejo informacije na primer o zadovoljstvu potrošnikov, o njihovem odzivu na oglaševanja in pospeševanja prodaje, o njihovem vedenju na področju oskrbe z energenti itd. Poleg tega v podjetju raziskujejo tudi potrebe ter zadovoljstvo drugih deležnikov podjetja (zaposlenih, delničarjev, medijev in javnosti) in značilnosti trgov. Vse te informacije so namenjene sprejemanju poslovnih odločitev (predvsem na področju trženja) in omogočajo vpeljavo strategije upravljanja odnosov s strankami (ang.: Customer Relationship Management, CRM) (Blažič, 2008, str. 5-19).


7.2 Opis ankete

V anketi je sodelovalo 201 imetnikov Magna kartice (63% od teh je bilo moških in 37% je bilo žensk). Povprečna starost je bila 43 let. V Prilogi 3 so nazorno prikazani demografski statistični podatki anketirancev. Iz grafov je razvidno, da je skupina anketirancev iz raznih zornih kotov precej homogena. Na primer je večina vprašanih odgovorila, da so samski (81%) in da so (so)lastniki stanovanja, v katerem živijo (80%).


7.3 **Orodje Visual Studio 2005**


7.4 Analiza podatkov

Pred analizo podatkov sem si postavila nekaj vprašanj in sovpadajočih hipotez. S pomočjo podatkovnega rudarjenja sem želela odkriti prave odgovore na zastavljena vprašanja.

Vprašanje 1: Ali demografski in socioekonomski podatki vplivajo na potrošnjo na prodajnih mestih Petrola? Če da, kateri?

- Podvprašanje 1.1: Kateri demografski in socioekonomski podatki vplivajo na potrošnjo goriva?
- Podvprašanje 1.2: Kateri demografski in socioekonomski podatki vplivajo na potrošnjo dodatne ponudbe Petrola?

Vprašanje je po mojem mnenju pomembno za boljše razumevanje kupcev in za nadaljnje trženjske akcije. Podjetje se s takimi informacijami lahko usmerja v bolj donosne kupce in izpušča manj donosne.

Prva hipoteza je v tem primeru ta, da demografski in socioekonomski podatki vplivajo na potrošnjo na prodajnih mestih Petrola. Druga hipoteza pa je, da v največji meri vpliva na potrošnjo dohodek gospodinjstva. Poleg potrditve hipoteze sem želela dobiti še druge značilnosti, ki vplivajo na celotno potrošnjo, na potrošnjo le goriva in na potrošnjo dodatne ponudbe. Te je le s pomočjo logičnega razmišljanja težko oceniti. Analizo sem izvedla s pomočjo tehnike odločitvenih dreves.

Pri izdelavi modela sem kot napovedne spremenljivke (Predict Only) določila potrošnjo v zadnjih šestih mesecih z Magna kartico na Petrolovih prodajnih mestih, potrošnjo goriva v zadnjih šestih mesecih z Magna kartico na Petrolovih prodajnih mestih in potrošnjo dodatne ponudbe v zadnjih šestih mesecih z Magna kartico na Petrolovih prodajnih mestih. Da bi izvedela, katere demografske in socioekonomskie spremenljivke vplivajo na potrošnjo, sem izbrala kot vhode (Input) vse možne vplivne spremenljivke: mesečni dohodek gospodinjstva, spol, starost, izobrazbo, zakonski stan, število vzdrževanjih članov in stanovanjske razmere.

V Prilogi 5 je na Sliki 10 odločitveno drevo, ki prikazuje na levi strani celotno populacijo in se nato proti desni deli glede na najbolj pomembne spremenljivke. Svetloba lista kaže na gostoto populacije v listu: najbolj temni listi imajo največji del populacije. Visual Studio 2005 prikaže le spremenljivke, ki v večji meri vplivajo na napovedovalno spremenljivko (zato v prikazu ni vseh spremenljivk). Najbolj pomembni vpliv na napovedovalno spremenljivko imajo vhodi, ki so bližji listu s celotno populacijo.

Iz rezultata analize lahko opazimo, da je prva hipoteza potrjena, druga pa ne. Demografski podatki res vplivajo na višino potrošnje, saj je orodje Visual Studio 2005 delilo populacijo in ustvarilo odločitveno drevo. Ko razčlenitev glede na vplive vhodov ni smiselna, celotna populacija ostane v prvem listu. Druga hipoteza je ovržena, saj orodje za podatkovno
rudarjenje ni delilo populacije glede na dohodek. To pomeni, da ta vhodna spremenljivka nima pomembnega vpliva na potrošnjo.


Na celotno potrošnjo v največji meri vpliva zakonski stan, nato pa še število vzdrževanih članov in stanovanjske razmere (glej Sliko 10 v Prilogi 5). Kupec, ki največ potroši, so samski in nimajo vzdrževanih članov.

Na potrošnjo za nakup goriva na Petrolovnih prodajnih mestih vpliva v največji meri število vzdrževanih članov in nato še spol (glej Sliko 11 v Prilogi 5). Za gorivo največ zapravijo moški, ki nimajo vzdrževanih članov.

Pri potrošnji za nakup dodatne Petrolove ponudbe spol ni pomembna spremenljivka (glej Sliko 12 v Prilogi 5). Največji je vpliv zakonskega stanu, pri čemer ločene ali ovdovele osebe porabijo manj denarja v te namene.

2. Vprašanje: Katere so demografske značilnosti najboljših Petrolovnih strank?


Da bi dobila najboljše kupce, sem se zanašala na Paretovo 20/80 pravilo, ki pravi, da 20% najboljših kupcev daje 80% prihodkov. Za to sem najprej izvozila podatke v Excel in razporedila kupce od tistega, ki na prodajnih mestih Petrola je v zadnjih šestih mesecih pred anketo največ potrošil vse do najslabšega kupca. Nato sem preverila Paretovo 20/80 pravilo in prišla do zaključka, da najboljših 20% kupcev prinašajo Petrolu 50% prihodkov. Čeprav ne dosegajo Paretovega pravila, so ti kupci najbolj pomembni. Podatke o 20% kupcev sem izvozila v Accessovo bazo podatkov. Nato sem kreirala povezavo na poizvedbo o najboljših kupcih (ki v tem primeru predstavljajo celotno populacijo). Za tehniko podatkovnega rudarjenja sem določila tehniko razvrščanja strank, saj ta na zelo nazoren način pokaže značilnosti populacije.
Da bi lahko videla vse značilnosti populacije, sem kot vhodne spremenljivke vključila vse spremenljivke, razen identifikacijske številke, saj je to glavni ključ. Določila sem še, katerega tipa je posamezna spremenljivka, in zagnala sem proces.

Kot lahko vidimo na Sliki 4, je hipoteza skoraj v celoti potrjena. Kar 97,4% najboljših kupcev Petrola je samskih in 74,4% jih je moškega spola. Število vzdrževanih članov pa ni značilna lastnost najboljših kupcev (variira med 0 in 3 in povprečje je 1,4). 89,7% najboljših kupcev je (so)lastnikov stanovanja, v katerem živijo. Večina (64,1%) jih je končala srednjo šolo, 30,8% pa višjo ali visoko šolo.

Med demografske in socioekonomiske podatke, pri katerih je vrednost spremenljivke razpršena, spadata še starost in dohodek gospodinjstva. Starost se giblje med 27 in 60 let, povprečna starost pa je 44 let.

Mnenja najboljših kupcev o Petrolovih prodajnih mestih so v splošnem zelo dobra (ta podatek velja tudi za druge kupce). Prav tako je večina zadovoljna z načinom obračuna transakcij z Magna kartico.

Slika 4: Profil najboljših strank

Na podlagi analize odgovorov strank, sem zasledila nekatere zanimive vzorce obnašanja strank, ki so po mojem mnenju uporabni za nadaljnje poslovanje in odločanje v podjetju Petrol. Ker je pri podatkovnemu rudarjenju bistvo prav odkrivanje prej neznanih, a uporabnih informacij, bom povzela svoje ugotovitve. Prva ugotovitev je naslednja: 76% najboljših kupcev (sploh) se ne strinja, da na Petrolovih prodajnih mestih kupuje le gorivo, in se (popolnoma) strinja, da ob nakupu goriva vedno kupi še kako malenkost. Istočasno se večina (sploh) ne strinja, da se pogosto na Petrolovih prodajnih mestih ustavi le za nakup izdelkov dodatne ponudbe. To lahko pomeni, da je pri nakupovanju na Petrolovih prodajnih mestih

Drugi pomembni vzorec obnašanja je po mojem mnenju ta, da večina populacije (89,7%) dobiva obvestila o ugodnostih, namenjenih imetnikom Magna kartice, a da le 17,9% najboljših kupcev izkorišča ugodnosti. Pri tem bi svetovala podjetju Petrol, naj raziskuje razloge za tako nizek odstotek, naj analizira želje in potrebe kupcev ter naj na podlagi tega ukrepa. Mogoče dane ugodnosti nimajo večje vrednosti v očeh kupcev in jih je zato potrebno prilagoditi.

3. vprašanje: Ali stranke, ki na Petrolovih prodajnih mestih kupujejo le gorivo, imajo podpovprečno mnenje o kakovosti za Petrol? In tiste stranke, ki kupujejo tudi druge izdelke ali hodijo večkrat le po izdelke dodatne ponudbe na Petrolova prodajna mesta, imajo nadpovprečno mnenje o kakovosti za Petrol?

Nekatere stranke se strinjajo ali popolnoma strinjajo, da na Petrolovih prodajnem mestu kupujejo samo gorivo (glej Prilogo 4). Pri tem me zanima, katera so mnenja teh strank o kakovosti Petrolovih prodajnih mestih. Rada bi izvedela, ali mnenje o kakovosti za Petrol vpliva na dejstvo, da kupci kupujejo le gorivo, kupujejo tudi malenkosti iz dodatne ponudbe ali gredo do Petrola tudi za dodatne nakupe. Odgovor na to vprašanje lahko pomaga Petrolu pri večanju prihodkov s poveščanjem zadovoljstva strank na določenih ključnih področjih. Moja hipoteza je pri tem ta, da boljše mnenje o kakovosti vpliva na nakup tudi drugih izdelkov.

Da bi pregledala, katera mnenja v največji meri vplivajo na zgoraj omenjene odločitve, sem izbrala tehniko odločitvenega drevesa. Kot vir podatkov sem uporabila celotno Petrolovo bazo podatkov. Med napovedne spremenljivke sem dala mnenja o naslednjih vprašanjih: ali kupci na Petrolovem prodajnem mestu kupujejo samo gorivo, ali ob nakupu goriva vedno kupijo še kako malenkosti in ali se pogosto ustavljajo na Petrolovih prodajnih mestih le za nakup izdelkov iz dodatne ponudbe. Med vhodne spremenljivke sem vključila vse odgovore na vprašanja o kakovosti za Petrol (glej vprašalnik v Prilogi 1).

Rezultati analize so pokazali, da na odločitev o nakupu izdelka iz dodatne ponudbe (z gorivom ali brez) ne vplivajo mnenja o kakovosti za Petrol.

Iz Slike 15 v Prilogi 7 lahko opazimo, da je mnenje o zaupanju prodajnim osebam Petrola (SQ_12) edina spremenljivka, ki ima določen vpliv na odločitev o nakupu goriva brez dodatnih izdelkov.

Da bi bolje raziskovala korelacijo med tema dvema spremenljivkama, sem jo preizkusila tudi s tehniko asociativnih pravil (glej Prilogo 8). Pri tem sem prišla do naslednjih pravil:

21
če se kupec popolnoma strinja, da lahko zaupa prodajalcem, potem se sploh ne strinja, da na Petrolovnih prodajnih mestih kupuje le gorivo (z zaupanjem 50,1% in podporo 16,3%);
če se kupec (popolnoma) ne strinja, da lahko zaupa prodajnim osebam, potem se delno strinja ali strinja, da na Petrolovnih prodajnih mestih kupuje samo gorivo (z zaupanjem 50,0% in podporo 37,9%).

SKLEP

Podatkovno rudarjenje je ob naraščanju količine shranjenih podatkov v veliko pomoč pri pridobivanju uporabnih informacij za poslovno odločanje. V teoretičnem delu sem razložila definicijo in pomen te vede ter opisala razne možnosti uporabe v praksi. Poleg tega sem predstavila šest nalog, glavnih pet tehnik in proces podatkovnega rudarjenja.


V diplomskem delu sem izdelala profil 20% najboljših strank. Ti prinašajo podjetju 50% prihodkov in so zato ciljni kupci. Tipični ciljni kupec podjetja Petrol je samski moški s srednjo ali visoko izobrazbo, ki je (so)lastnik stanovanja, v katerem živi. V take kuce mora podjetje vlagati največ energij in sredstev ter prilagoditi trženjski splet njihovim potrebam in željam. Ker velika večina ciljnih kucev ne izkorišča ugodnosti Magna kartice, podjetju priporočam, naj analizira želje ciljnih strank in temu skuša prilagoditi ugodnosti.

V zadnjem delu analize sem želela izvedeti, če obstaja vpliv mnenj o kakovosti za Petrol na to, ali kupci kupujejo na Petrolovnih prodajnih mestih le gorivo, tudi izdelke iz dodatne ponudbe in na to, ali gredo včasih na Petrol le za nakup izdelkov iz dodatne ponudbe. Taka informacija je pomembna, saj lahko pomaga Petrolu pri večanju prihodkov s povišanjem zadovoljstva strank na določenih ključnih področjih. Rezultati analize so pokazali, da na odločitev o nakupu izdelka iz dodatne ponudbe (z gorivom ali brez) ne vplivajo mnenja o
kakovosti za Petrol. Edina spremenljivka, ki ima določeno stopnjo korelacije z odločitvijo o nakupu goriva brez dodatnih izdelkov, je mnenje o zaupanju prodajnim osebam Petrola. Pri tem se kupec, ki se popolnoma strinja, da lahko zaupa prodajalcem, sploh ne strinja, da na Petrolovih prodajnih mestih kupuje le gorivo. Če pa se kupec (popolnoma) ne strinja, da lahko zaupa prodajnim osebam, potem se delno strinja ali strinja, da na Petrolovih prodajnih mestih kupuje samo gorivo.

V diplomskem delu sem dosegla zastavljene cilje. Na podlagi teoretičnega znanja sem raziskala bazo podatkov Petrola. Analizirala sem ciljne kupce in vplive na večjo potrošnjo. Pridobila sem koristne informacije, ki bi jih lahko podjetje Petrol uporabljalo za nadaljnje odločanje in na podlagi katerih bi lahko sestavilo trženjski splet ter se usmerilo v bolj donosne kupce (ob predpostavki, da bi bili podatki ažurni).
LITERATURA IN VIRI


PRILOGA 1: Vprašalnik

Merjenje zaznane kakovosti za Petrol

Navedli vam bomo nekaj trditev, vi pa nam povejte, v kolikšni meri se z njimi strinjate:
1. Oprema Petrolovih prodajnih mest je moderna.
2. Petrolova prodajna mesta delujejo privlačno.
3. Petrolova prodajna mesta so čista.
4. Petrolovi prodajalci so primerno oblečeni.
5. Petrolovi prodajalci so primerno urejeni.
7. Ob nakupu goriva vedno kupim še kako malenkost.
8. Na Petrolovih prodajnih mestih se pogosto ustavim samo zaradi nakupa izdelkov dodatne ponudbe (tj. vsega, kar ni gorivo).
10. Petrolovi prodajalci so vedno prijazni.
11. Petrolovi prodajalci mi znajo dobro svetovati.
13. Ker imajo Petrolovi prodajalci veliko dela, je popolnoma razumljivo, da mi ne morejo takoj pomagati.
15. Nerealno je pričakovati, da bodo Petrolovi prodajalci ustregli mojim posebnim željam.
16. Petrolovi prodajalci mi navadno niso pripravljeni pomagati.
17. Na blagajni Petrolovega prodajnega mesta sem hitro na vrsti.
18. Lokacija Petrolovega prodajnega mesta, ki ga najpogosteje obiščem, mi popolnoma ustreza.
19. Delovni čas Petrolovega prodajnega mesta, ki ga najpogosteje obiščem, mi popolnoma ustreza.

Merjenje pričakovanj

1. Izmed naštetoj lastnosti izberite tri, ki so za vas ob obisku prodajnega mesta na črpalki najpomembnejše:
   ➢ Oprema prodajnega mesta.
   ➢ Čistoča prodajnega mesta
   ➢ Urejenost prodajalcov.
   ➢ Prijaznost prodajalcov.
   ➢ Pestrost ponudbe.
   ➢ Hitrost postrežbe pri blagajni.
Lokacija črpalke.
Delovni čas črpalke.
2. Ali je za vas pomembna še kakšna lastnost, ki je nismo našteli?
   ➢ Da.
   ➢ Ne.

Ocena poslovanja z Magno:

1. Ali ste imetnik Magna kartice?
   ➢ Da.
   ➢ Ne.
2. Kako dolgo ste že imetnik Magna kartice (v letih)?
3. Ali ste v letošnjem letu dobili kakšno obvestilo o ugodnostih, namenjenih imetnikom Magna kartice?
   ➢ Da.
   ➢ Ne.
4. Ali ste v letošnjem letu izkoristili katero od ugodnosti, namenjenih imetnikom Magna kartice?
   ➢ Da.
   ➢ Ne.
5. V kolikšni meri ste zadovoljni z obstoječim pisnim načinom obveščanja o ugodnostih, namenjenih imetnikom Magna kartice?
6. V kolikšni meri ste zadovoljni z načinom obračuna opravljenih transakcij?
7. Ste že kdaj poklicali Magna center s pritožbo?
   ➢ Da.
   ➢ Ne.

Demografska in socioekonomska vprašanja:

1. Spol:
   ➢ Moški.
   ➢ Ženski.
2. Starost (v letih): ___________
3. Izobrazba:
   ➢ Nedokončana osnovna šola.
   ➢ Končana osnovna šola.
   ➢ Končana srednja šola.
   ➢ Diploma višje ali visoke stopnje.
   ➢ Magisterij ali doktorat.
4. Zakonski stan:
   ➢ Samski.
   ➢ Poročen oziroma živeč s partnerjem.
Ločen ali ovdovel.

5. Dohodki članov gospodinjstva v prejšnjem mesecu:
   - Do 100.000 SIT.
   - Od 100.000 do 200.000 SIT.
   - Od 200.000 do 400.000 SIT.
   - Od 400.000 do 600.000 SIT.
   - Od 600.000 do 800.000 SIT.
   - Od 800.000 do 1.000.000 SIT.
   - Brez dohodkov.

6. Število vzdrževanih družinskih članov: __________

7. Stanovanjske razmere:
   - Lastnik ali solastnik hiše ali stanovanja.
   - Najemnik hiše ali stanovanja.
   - Podnajemnik v hiši ali stanovanju.
### PRILOGA 2: Seznam podatkov iz baze podatkov

<table>
<thead>
<tr>
<th>Koda spremenljivke</th>
<th>Opis spremenljivke</th>
</tr>
</thead>
<tbody>
<tr>
<td>GENDER</td>
<td>Spol anketiranca</td>
</tr>
<tr>
<td>AGE</td>
<td>Starost anketiranca</td>
</tr>
<tr>
<td>RES_OWN</td>
<td>Stanovanjske razmere</td>
</tr>
<tr>
<td>ENROL</td>
<td>Leto včlanitve v Magna program</td>
</tr>
<tr>
<td>MEMBER</td>
<td>Dolžina članstva v Magna programu (v letih)</td>
</tr>
<tr>
<td>EDUCA</td>
<td>Izobrazba</td>
</tr>
<tr>
<td>MARRIAGE</td>
<td>Zakonski stan</td>
</tr>
<tr>
<td>DEPEND</td>
<td>Število vzdrževanih družinskih članov</td>
</tr>
<tr>
<td>INCOME</td>
<td>Dohodki članov gospodinjstva v prejšnjem mesecu (maj 2001)</td>
</tr>
<tr>
<td>EXP_1</td>
<td>Prva najbolj pomembna značilnost pri obisku prodajnega mesta.</td>
</tr>
<tr>
<td>EXP_2</td>
<td>Prva najbolj pomembna značilnost pri obisku prodajnega mesta.</td>
</tr>
<tr>
<td>EXP_3</td>
<td>Prva najbolj pomembna značilnost pri obisku prodajnega mesta.</td>
</tr>
<tr>
<td>EXP_ADD</td>
<td>Ali je za vas pomembna še kakšna lastnost, ki je nismo našeli?</td>
</tr>
<tr>
<td>EXP_DET</td>
<td>Če da, katera lastnost, ki je nismo našeli, je za vas pomembna?</td>
</tr>
<tr>
<td>B_INFO</td>
<td>Ali ste v letošnjem letu dobili kakšno obvestilo o ugodnostih, namenjenih imetnikom Magna kartice?</td>
</tr>
<tr>
<td>B_USE</td>
<td>Ali ste v letošnjem letu izkoristili katero od ugodnosti, namenjenih imetnikom Magna kartice?</td>
</tr>
<tr>
<td>S_INFO</td>
<td>V kolikšni meri ste zadovoljni z obstoječim pisnim načinom obveščanja o ugodnostih, namenjenih imetnikom Magna kartice?</td>
</tr>
<tr>
<td>S_TRAN</td>
<td>V kolikšni meri ste zadovoljni z načinom obračuna opravljenih transakcij?</td>
</tr>
<tr>
<td>COMPL</td>
<td>Ste že kdaj poklicali Magna center s pritožbo?</td>
</tr>
<tr>
<td>Koda spremenljivke</td>
<td>Opis spremenljivke</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>SQ_1</td>
<td>Oprema Petrolovih prodajnih mest je moderna.</td>
</tr>
<tr>
<td>SQ_2</td>
<td>Petrolova prodajna mesta delujejo privlačno.</td>
</tr>
<tr>
<td>SQ_3</td>
<td>Petrolova prodajna mesta so čista.</td>
</tr>
<tr>
<td>SQ_4</td>
<td>Petrolovi prodajalci so primerno oblečeni.</td>
</tr>
<tr>
<td>SQ_5</td>
<td>Petrolovi prodajalci so primerno urejeni.</td>
</tr>
<tr>
<td>SQ_6</td>
<td>Na Petrolovih prodajnih mestih kupujem samo gorivo.</td>
</tr>
<tr>
<td>SQ_7</td>
<td>Ob nakupu goriva vedno kupim še kako malenkost.</td>
</tr>
<tr>
<td>SQ_8</td>
<td>Na Petrolovih prodajnih mestih se pogosto ustavim samo zaradi nakupa izdelkov dodatne ponudbe (tj. vsega, kar ni gorivo).</td>
</tr>
<tr>
<td>SQ_9</td>
<td>Dodatna ponudba Petrola (tj. vse, kar ni gorivo) je zelo pesta.</td>
</tr>
<tr>
<td>SQ_10</td>
<td>Petrolovi prodajalci so vedno prijazni.</td>
</tr>
<tr>
<td>SQ_11</td>
<td>Petrolovi prodajalci mi znajo dobro svetovati.</td>
</tr>
<tr>
<td>SQ_12</td>
<td>Petrolovim prodajalcem lahko zaupam.</td>
</tr>
<tr>
<td>SQ_13</td>
<td>Ker imajo Petrolovi prodajalci veliko dela, je popolnoma razumljivo, da mi ne morejo takoj pomagati.</td>
</tr>
<tr>
<td>SQ_14</td>
<td>Ker imajo Petrolovi prodajalci veliko dela, je popolnoma razumljivo, da mi ne morejo posvetiti veliko časa.</td>
</tr>
<tr>
<td>SQ_15</td>
<td>Nerealno je pričakovati, da bodo Petrolovi prodajalci ustregli mojim posebnim željam.</td>
</tr>
<tr>
<td>SQ_16</td>
<td>Petrolovi prodajalci mi navadno niso pripravljeni pomagati.</td>
</tr>
<tr>
<td>SQ_17</td>
<td>Na blagajni Petrolovega prodajnega mesta sem hitro na vrsti.</td>
</tr>
<tr>
<td>SQ_18</td>
<td>Lokacija Petrolovega prodajnega mesta, ki ga najpogosteje obiščem, mi popolnoma ustreza.</td>
</tr>
<tr>
<td>SQ_19</td>
<td>Delovni čas Petrolovega prodajnega mesta, ki ga najpogosteje obiščem, mi popolnoma ustreza.</td>
</tr>
<tr>
<td>SQ_20</td>
<td>Nerealno je pričakovati, da bodo vsa Petrolova prodajna mesta odprta 24 ur na dan.</td>
</tr>
<tr>
<td>Koda spremenljivke</td>
<td>Opis spremenljivke</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>TOTAL</td>
<td>Celotni znesek, ki ga je anketiranec potrošil s kartico Magna v zadnjih šestih mesecih.</td>
</tr>
<tr>
<td>T_EXT</td>
<td>Celotni znesek, ki ga je anketiranec s kartico Magna v zadnjih šestih mesecih v prodajnih mestih, ki niso lastnina Petrola.</td>
</tr>
<tr>
<td>T_INT</td>
<td>Celotni znesek, ki ga je anketiranec potrošil na Petrolovih prodajnih mestih s kartico Magna v zadnjih šestih mesecih.</td>
</tr>
<tr>
<td>T_INT1</td>
<td>Celotni znesek, ki ga je anketiranec potrošil za gorivo na Petrolovih prodajnih mestih s kartico Magna v zadnjih šestih mesecih.</td>
</tr>
<tr>
<td>T_INT2</td>
<td>Celotni znesek, ki ga je anketiranec potrošil za dodatno blago na Petrolovih prodajnih mestih s kartico Magna v zadnjih šestih mesecih.</td>
</tr>
<tr>
<td>POS</td>
<td>Število Petrolovih prodajnih mest, ki jih je imetnik kartice obiskal v zadnjih šestih mesecih.</td>
</tr>
<tr>
<td>VISIT1</td>
<td>Število obiskov imetnika kartice na prvem najbolj obiskanem Petrolovenem prodajnem mestu.</td>
</tr>
<tr>
<td>VISIT2</td>
<td>Število obiskov imetnika kartice na drugem najbolj obiskanem Petrolovenem prodajnem mestu.</td>
</tr>
<tr>
<td>VISIT3</td>
<td>Število obiskov imetnika kartice na tretjem najbolj obiskanem Petrolovenem prodajnem mestu.</td>
</tr>
<tr>
<td>SUM2</td>
<td>Seštevek VISIT1 in VISIT2.</td>
</tr>
<tr>
<td>SUM3</td>
<td>Seštevek VISIT1, VISIT2 in VISIT3.</td>
</tr>
<tr>
<td>SUMT</td>
<td>Celotno število obiskov imetnika kartice na vseh prodajnih mestih Petrola v zadnjih šestih mesecih.</td>
</tr>
<tr>
<td>SHARE1</td>
<td>Odstotek VISIT1 v SUMT.</td>
</tr>
<tr>
<td>SHARE2</td>
<td>Odstotek SUM2 v SUMT.</td>
</tr>
<tr>
<td>SHARE3</td>
<td>Odstotek SUM3 v SUMT.</td>
</tr>
</tbody>
</table>
PRILOGA 3: Demografski podatki anketirancev

Slika 1: Razporeditev anketirancev po spolu

![Pie chart showing the distribution of respondents by gender. 63% are female and 37% are male.](image1)

Vir: Lastna izdelava

Slika 2: Razporeditev anketirancev po zakonskem stanu

![Pie chart showing the distribution of respondents by legal status. 81% have a household partner, 9% are single or separated, 2% are divorced or widowed, and 8% have no legal status.](image2)

Vir: Lastna izdelava

Slika 3: Razporeditev anketirancev po izobrazbi

![Pie chart showing the distribution of respondents by educational level. 63% have a lower secondary school diploma, 34% have an upper secondary school diploma, and 1% have a university degree or higher.](image3)

Vir: Lastna izdelava
**Slika 4: Razporeditev anketirancev po starosti**

![Pie chart showing age distribution](chart1.png)

*Vir: Lastna izdelava*

**Slika 5: Razporeditev anketirancev glede na stanovanjske razmere**

![Pie chart showing property ownership](chart2.png)

*Vir: Lastna izdelava*

**Slika 6: Število vzdrževanih družinskih članov**

![Bar chart showing number of members](chart3.png)

*Vir: Lastna izdelava*
PRILOGA 4: Nakupno obnašanje anketirancev

Slika 7: Delitev kupcev na podlagi tega, ali kupujejo samo gorivo

![Diagram](attachment:image7.png)

Vir: Lastna izdelava

Slika 8: Delitev kupcev na podlagi tega, ali pogosto kupujejo le izdelke dodatne ponudbe

![Diagram](attachment:image8.png)

Vir: Lastna izdelava

Slika 9: Delitev kupcev glede tega, ali kupujejo ob gorivu še kako malenkost

![Diagram](attachment:image9.png)

Vir: Lastna izdelava
PRILOGA 5: Analiza vpliva demografskih in socioekonomskih značilnosti kupcev na potrošnjo

Slika 10: Odločitveno drevo: vpliv demografskih in socioekonomskih značilnosti kupcev na celotno potrošnjo

Vir: Latsna izdelava

Slika 11: Odločitveno drevo: vpliv demografskih in socioekonomskih značilnosti kupcev na potrošnjo za nakup goriva

Vir: Latsna izdelava
Slika 12: Odločitveno drevo: vpliv demografskih in socioekonomskih značilnosti kupcev na potrošnjo za nakup izdelkov iz dodatne ponudbe

Vir: Lastna izdelava

Slika 13: Mreža odvisnosti

Vir: Lastna izdelava
PRILOGA 6: Analiza profila najboljših kupcev

Slika 14: Značilnosti populacije

Vir: Lastna izdelava
PRILOGA 7: Analiza vpliva mnenj strank z različnimi nakupnimi obnašanji o kakovosti za Petrol

Slika 15: Odločitveno drevo: vpliv zaupanja prodajalcem na stranke, ki kupujejo samo gorivo

Vir: Lastna izdelava

Slika 16: Mreža odvisnosti: Vpliv ocen kakovosti za Petrol na nakupno obnašanje potrošnikov

Vir: Lastna izdelava
PRILOGA 8: Asociativna pravila: vpliv mnenja o zaupanju prodajalcem na potrošnjo samo goriva

Slika 17: Mreža odvisnosti med zaupanjem prodajalcem in nakupom samo goriva

Vir: Lastna izdelava

Slika 18: Asociativna pravila

Vir: Lastna izdelava
<table>
<thead>
<tr>
<th>ANGLEŠKI IZRAZ</th>
<th>SLOVENSKI IZRAZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activation</td>
<td>aktiviranje strank</td>
</tr>
<tr>
<td>Artificial Intelligence</td>
<td>umetna inteligencia</td>
</tr>
<tr>
<td>Artificial Neural Network</td>
<td>umetna nevronska mreža</td>
</tr>
<tr>
<td>Association Analysis</td>
<td>analiza asociacij</td>
</tr>
<tr>
<td>Association Rules</td>
<td>asociativna pravila</td>
</tr>
<tr>
<td>Attrition</td>
<td>obraba</td>
</tr>
<tr>
<td>Binary Attributes</td>
<td>binarne spremenljivke</td>
</tr>
<tr>
<td>Categorical Variable</td>
<td>kategorična spremenljivka</td>
</tr>
<tr>
<td>Churn</td>
<td>mešalnik</td>
</tr>
<tr>
<td>Classification</td>
<td>klasificiranje</td>
</tr>
<tr>
<td>Clustering</td>
<td>razvrščanje v skupine</td>
</tr>
<tr>
<td>Confidence</td>
<td>zaupanje</td>
</tr>
<tr>
<td>Customer Relationship Management</td>
<td>upravljanje odnosov s strankami</td>
</tr>
<tr>
<td>Cross-Selling</td>
<td>navzkrižna prodaja</td>
</tr>
<tr>
<td>Data Mining</td>
<td>podatkovno rudarjenje</td>
</tr>
<tr>
<td>Data Visualization Techniques</td>
<td>vizualizacijska orodja</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>odločitveno drevo</td>
</tr>
<tr>
<td>Description</td>
<td>opisovanje</td>
</tr>
<tr>
<td>Distance Function</td>
<td>funkcija razdalje</td>
</tr>
<tr>
<td>Estimation</td>
<td>ocenjevanje</td>
</tr>
<tr>
<td>Input</td>
<td>vhod</td>
</tr>
<tr>
<td>Interval Variable</td>
<td>intervalna spremenljivka</td>
</tr>
<tr>
<td>Machine Learning</td>
<td>strojno učenje</td>
</tr>
<tr>
<td>Market Basket Analysis</td>
<td>analiza nakupne košarice</td>
</tr>
<tr>
<td>Market Segmentation</td>
<td>segmentiranje trga</td>
</tr>
<tr>
<td>Nearest Neighbors</td>
<td>najbližji sosed</td>
</tr>
<tr>
<td>Nominal Variable</td>
<td>nominalna spremenljivka</td>
</tr>
<tr>
<td>Ordinal Variable</td>
<td>ordinalna spremenljivka</td>
</tr>
<tr>
<td>Predict Only</td>
<td>napovedna spremenljivka</td>
</tr>
<tr>
<td>Prediction</td>
<td>napovedovanje</td>
</tr>
<tr>
<td>Profile Analysis</td>
<td>izdelava profilov kupcev</td>
</tr>
<tr>
<td>Ratio Variable</td>
<td>razmernostna spremenljivka</td>
</tr>
<tr>
<td>Response Model</td>
<td>model reakcije strank</td>
</tr>
<tr>
<td>Risk Management</td>
<td>upravljanje s tveganji</td>
</tr>
<tr>
<td>Support</td>
<td>podpora</td>
</tr>
<tr>
<td>Unary Attributes</td>
<td>enostavne spremenljivke</td>
</tr>
<tr>
<td>Up-Selling</td>
<td>več enakih izdelkov</td>
</tr>
<tr>
<td>Weighted Voting</td>
<td>tehtano glasovanje</td>
</tr>
</tbody>
</table>