

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Gregor Capuder

**Določitev priporočene cene na
spletni tržnici bolha.com**

DIPLOMSKO DELO
UNIVEZITETNI ŠTUDIJSKI PROGRAM
RAČUNALNIŠTVO IN INFORMATIKA

Mentor: doc. dr. Dejan Lavbič

Ljubljana, 2013

Rezultati diplomskega dela so intelektualna lastnina Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavljanje ali izkoriščanje rezultatov diplomskega dela je potrebno pisno soglasje Fakultete za računalništvo in informatiko ter mentorja.

Besedilo je oblikovano z urejevalnikom besedil \LaTeX .



Št. naloge: 01955/2013

Datum: 02.09.2013

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Kandidat: **GREGOR CAPUDER**

Naslov: **DOLOČITEV PRIPOROČENE CENE NA SPLETNI TRŽNICI
BOLHA.COM**
**DETERMINATION OF RECOMMENDED PRICE ON BOLHA.COM
ONLINE MARKET**

Vrsta naloge: Diplomsko delo univerzitetnega študija

Tematika naloge:

V času finančne krize uporabniki pogosto uporabljajo spletne tržnice, kjer se srečujejo kupci in prodajalci z namenom izmenjave izdelkov in tudi storitev. Omenjene tržnice ponujajo številne možnosti, a je nabor za navadnega uporabnika, ki bi želel brezplačno objavo oglasa, zelo omejen. Uporabnik lahko ponavadi le zapiše naziv in opis oglasa, skupaj s slikovnim gradivom, ter možnost opredelitve cene. Zanimiva funkcionalnost, ki jo prodajalci pogrešajo pa je priporočena cena. Uporabnik bi npr. želel čim prej prodati nek izdelek in ne želi izgubljati časa z raziskovanjem, kolikšna je vrednost tega izdelka, ampak bi želel priporočilo sistema, ki bi upošteval ceno novega izdelka in ceno iz oglasov primerljivih izdelkov. V okviru diplomske naloge naj študent pripravi prototip takšnega priporočilnega sistema, ki bi z različnimi strategijami in različnimi viri podatkov prodajalcu predlagal priporočeno ceno izdelka, ki ga želi prodati.

Mentor:

doc. dr. Dejan Lavbič



Dekan:


prof. dr. Nikolaj Zimic

IZJAVA O AVTORSTVU

diplomskega dela

Spodaj podpisani Gregor Capuder,

z vpisno številko 63070052,

sem avtor diplomskega dela z naslovom:

Določitev priporočene cene na spletni tržnici bolha.com

S svojim podpisom zagotavljam, da:

- sem diplomsko delo izdelal samostojno pod mentorstvom doc. dr. Dejana Lavbiča
- so elektronska oblika diplomskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko diplomskega dela
- soglašam z javno objavo elektronske oblike diplomskega dela v zbirki "Dela FRI".

V Ljubljani, dne 5.9.2013

Podpis avtorja/-ice:

Zahvalil bi se rad svojemu mentorju, doc. dr. Dejanu Lavbiču, za vodenje in pomoč pri izdelavi diplomske naloge. Zelo sem hvaležen tudi moji družini, puncu in vsem ostalim, ki ste me podpirali in spodbujali v času študija.

Kazalo

Seznam slik

Seznam tabel

Seznam uporabljenih kratic in simbolov

Uporabljene tehnologije in jeziki

Povzetek

Abstract

1	Uvod	1
2	Obstoječi priporočilni sistemi spletnih tržnic	3
2.1	Priporočilni sistemi za kupce	5
2.2	Priporočilni sistemi za prodajalce	5
3	Predstavitev sistema za priporočeno ceno	8
3.1	Namen razvoja priporočilnega sistema	9
4	Implementacija priporočilnega sistema	10
4.1	Potek raziskave s priporočilnim sistemom	10
4.1.1	Raziskava obstoječih izdelkov na spletni tržnici bolha.com	13

4.1.2	Raziskava na podlagi novega izdelka - mimovrste.com . .	19
4.2	Tehnične podrobnosti implementacije priporočilnega sistema . .	24
5	Analiza natančnosti razvitega priporočilnega sistema	34
5.1	Rezultati analize »iPad mini 16gb wifi«	35
5.2	Rezultati analize »Samsung galaxy S4 LTE«	37
5.3	Rezultati analize »HP 6735s«	38
5.4	Interpretacija rezultatov analize	39
6	Zaključek	41
	Literatura	43

Slike

2.1	Graf padca avkcij.	4
4.1	Začetni obrazec raziskave.	11
4.2	Diagram zaporedja celotnega postopka.	12
4.3	Okrnjen prikaz izdelkov.	16
4.4	Nastavitev parametrov - bolha.com	17
4.5	Prikaz rezultatov - bolha.com	18
4.6	Prikaz rezultatov - mimovrste.com	20
4.7	Prikaz parametrov - mimovrste.com	21
4.8	Priporočena cena - mimovrste.com	23
4.9	Zgodovina raziskav.	24
4.10	Seznam datotek projekta.	25
5.1	Graf natančnosti in priklica ob posameznih izločanjih pri analizi iPad mini 16gb wifi.	36
5.2	Graf natančnosti in priklica ob posameznih izločanjih pri analizi Samsung galaxy S4 LTE.	37
5.3	Graf natančnosti in priklica ob posameznih izločanjih pri analizi HP 6735s.	38

Tabele

4.1	Odbitki glede na starost izdelka.	22
4.2	Odbitki glede na poškodbo izdelka.	22
5.1	Tabela osnovnih podatkov analize iPad mini 16gb wifi.	35
5.2	Tabela natančnosti in priklica ob izločanjih pri analizi iPad mini 16gb wifi.	36
5.3	Tabela osnovnih podatkov analize Samsung galaxy S4 LTE. . .	37
5.4	Tabela natančnosti in priklica ob izločanjih pri analizi Samsung galaxy S4 LTE.	37
5.5	Tabela osnovnih podatkov analize HP 6735s.	38
5.6	Tabela natančnosti in priklica ob izločanjih pri analizi HP 6735s.	38
5.7	Tabela natančnosti ustreznih izdelkov na koncu izločanj.	39

Seznam uporabljenih kratic in simbolov

API – Application programming interface – programski vmesnik.

AJAX – Asynchronous JavaScript and XML – skupina povezanih spletnih tehnologij za ustvarjanje dinamičnih spletnih strani.

DOM – Document Object Model – jezikovno in platformsko neodvisen standard za predstavitev in interakcijo s objekti v HTML dokumentih.

HTML – HyperText Markup Language

CSS – Cascading Style Sheets

PHP – PHP: Hypertext Preprocessor

BIN – Buy It Now - oglas na eBay-u s fiksno odkupno ceno

Uporabljene tehnologije in jeziki

HTML – HyperText Markup Language – je zaporedje ukazov, ki povejo kako naj se spletna stran prikaže. Ukazi se imenujejo tudi značke. Spletni brskalniki so programi, ki berejo ukaze v jeziku HTML in prikažejo vsebino.

CSS – Cascading Style Sheets – so stili, predstavljeni v obliki preprostega slogovnega jezika, ki skrbijo za izgled spletnih strani. Najpogosteje se uporablja za oblikovanje spletnih strani v jeziku HTML. S pomočjo CSS stilov lahko popolnoma preoblikujemo izgled spletnih strani.

LESS – je JavaScript knjižnica, ki nadgradi CSS. Poleg osnovnega CSS-ja lahko uporabljamo tudi spremenljivke, gnezdenje, metode, funkcije in operacije.

JavaScript – je objektni skriptni programski jezik za ustvarjanje interaktivnih spletnih strani. Skupaj s HTML-jem poživi strani z dinamičnim izvajanjem.

jQuery – je knjižnica za skriptni jezik JavaScript, ki prinaša zbirko funkcij za hitrejši razvoj spletnih aplikacij. Glavna lastnost jQuery-ja je reševanje težav JavaScript-a v različnih brskalnikih, kjer je ista Ja-

JavaScript funkcija lahko implementirana drugače. S jQuery-jem pa vse funkcije delujejo enako na vseh podprtih brskalnikih.

PHP – PHP: Hypertext Preprocessor - je skriptni jezik, ki se izvaja na strežniku in služi za razvoj spletnih aplikacij.

Povzetek

V času ekonomske krize ljudje pogosto uporabljajo spletne tržnice kot npr. eBay.com ali bolha.com z namenom prihraniti ali pa zaslužiti nekaj dodatnega denarja. Pri objavi oglasa na spletni tržnici je navaden uporabnik zelo omejen, saj lahko vpiše le naziv, opis, ceno ter doda slikovno gradivo. V tem diplomskem delu smo razvili prototip sistema za priporočeno ceno na slovenski spletni tržnici – bolha.com. Namen tega sistema je, da uporabniku na intuitivni način pospeši čas raziskave ovrednotenja izdelka in mu predlaga priporočeno ceno. Sistem uporabi vnešene podatke ter informacije obstoječih izdelkov objavljenih na bolha.com ter mimovrste.com. Na začetku diplomskega dela smo pregledali sorodne rešitve razvitih priporočilnih sistemov, nato smo predstavili našo idejo ter podrobno implementacijo priporočilnega sistema, na koncu pa smo izvedli še analizo natančnosti in se prepričali, če je naš sistem uporaben.

Ključne besede:

priporočena cena, priporočilni sistem, bolha.com

Abstract

In time of economic crisis people very often use online auction websites like eBay.com or bolha.com, with the intent to save or earn some extra money. When a average user is posting an ad on bolha.com he's very limited to what he can do with his post. He can write the title, description, price and he can add some images of the product. In this thesis we developed a prototype system, which recommends a price to the user of bolha.com. The purpose of the system is to make the research process of the price evaluation intuitive and as fast as possible. The system uses users input and the information of existing ads on bolha.com and mimovrste.com. The thesis starts with an overview of related recommendation systems, then we introduce the idea behind our system and explain the details of the implementation. At the end we also made an analysis in which we tested how precise the system is and if we can trust it's recommendations.

Key words:

recommended price, recommendation systems, bolha.com

Poglavje 1

Uvod

Ne moremo zanikati popularnosti spletnih tržnic, saj so zelo uporabne, tako za prodajalca kot za kupca. Ena od največjih spletnih tržnic na svetu je eBay z dvema milijardama dolarjev prihodkov v samo prvi četrtini tega leta [1]. To je zelo velik znesek glede na to, da eBay-ev glavni prihodek predstavlja provizija ob objavi oglasa in ob prodaji izdelka.

V tem diplomskem delu pa se bomo osredotočili na najbolj popularno slovensko spletno tržnico - bolha.com. Na bolhi lahko prodajalec objavi oglas z nazivom, opisom, slikovnim gradivom in ceno. Funkcionalnost, ki bi bila zelo zanimiva pa je priporočena cena. Uporabnik bi npr. rad prodal svoj iPad in ne želi zapravljati časa z raziskovanjem vrednosti njegovega izdelka, ampak bi preprosto rad priporočeno ceno, ki bi bila izračunana na podlagi vpisanih podatkov (naziv, opis, itd.) in ostalih virov (cena novega izdelka ali že objavljenih izdelkov na spletni tržnici).

Verjetno ste se vprašali zakaj takšnega sistema še niso razvili? Kot boste videli v naslednjem poglavju takšen sistem že obstaja za priporočeno ceno izdelka na eBay-u. eBay ima javno dostopen API, s katerim so v podjetju

Terapeak [14] razvili sistem, ki s pomočjo zgodovine uspešno zaključenih oglasov priporoči ceno za določen izdelek. Zakaj takšnega sistema ni za bolho? Verjetno zaradi velikosti oz. majhnosti trga bolhe, ki je omejen na Slovenijo in sosednje države. Poleg tega pa bolha ne hrani zgodovine prodanih izdelkov in nima javno dostopnega API-ja, ki bi olajšal izdelavo takega sistema. V nadaljevanju boste videli kakšne strategije in tehnologije smo izbrali za pridobitev priporočene cene na bolhi.

Glavna naloga oz. cilj tega diplomskega dela je razviti intuitiven sistem, ki bo uporabniku bistveno skrajšal čas raziskave vrednotenja svojega izdelka in mu priporočil kar se da dobro ceno pri objavi oglasa na spletni tržnici bolha.com.

V sledečem besedilu si bomo ogledali obstoječe priporočilne sisteme, ki so bili razviti za pomoč kupcem in prodajalcem. Jedro diplomskega dela sestavljata predstavitev našega priporočilnega sistema in bolj podrobna implementacija le-tega. Ob zaključku pa si bomo ogledali tudi analizo, kjer bomo preverili uporabnost in natančnost našega sistema.

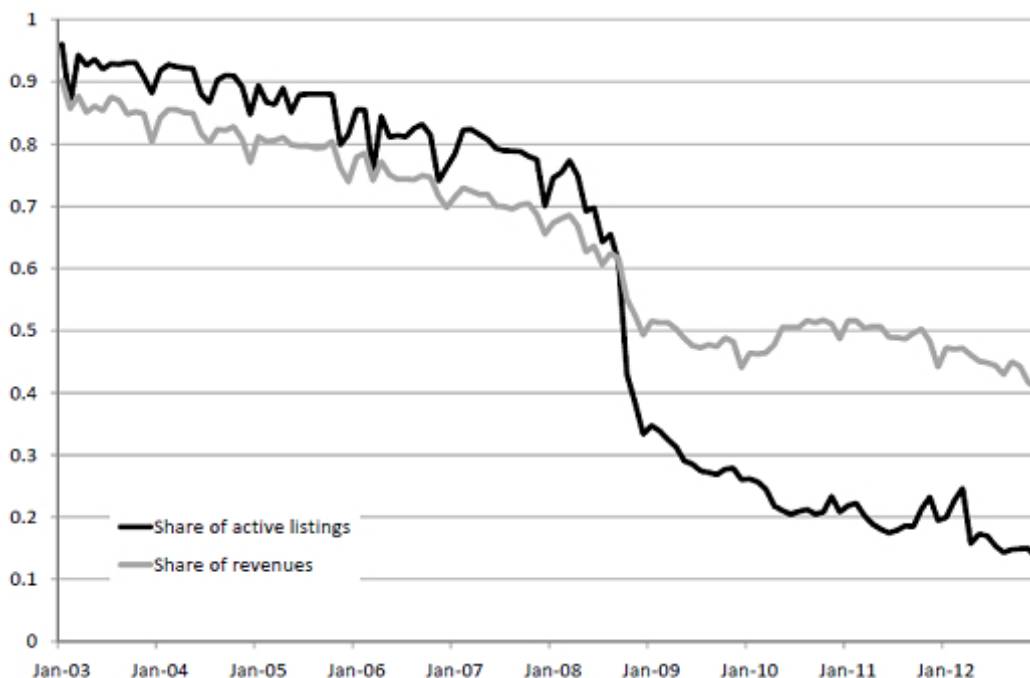
Poglavje 2

Obstoječi priporočilni sistemi spletnih tržnic

Največja spletna tržnica v Sloveniji je bolha.com, ki ima veliko uporabnikov in je popularna tudi v sosednjih državah (Hrvaška, Avstrija). Oglas lahko objavi fizična oseba ali podjetje. Mi se bomo osredotočili na pomoč fizični osebi oz. navadnemu uporabniku. Poleg tega pa bomo v tem poglavju delili uporabnika še na kupca in prodajalca. Ob iskanju obstoječih priporočilnih sistemov za spletno tržnico bolha.com smo našli samo pisna priporočila kako izboljšati prodajo izdelka z boljše izpolnjenim oglasom (dobro slikovno gradivo, dober opis, itd.). Hitro smo ugotovili, da za spletno tržnico bolha.com ni razvitih priporočilnih sistemov, mogoče tudi zato, ker bolha.com nima javno dostopnega API-ja, prav tako pa ne beležijo oz. nimajo možnosti beležiti zgodovine prodanih izdelkov, ker ni mogoče ugotoviti ali je bil izdelek prodan ali pa je uporabnik oglas izbrisal. Zato smo pogled, za raziskavo obstoječih priporočilnih sistemov, usmerili k največji spletni tržnici na svetu – eBay.com.

Razlika med bolho in eBay-om je ta, da mora prodajalec za objavo oglasa na

eBay.com plačati neko vsoto denarja, medtem ko je objava oglasa na bolha.com brezplačna. Prav tako se na eBay-u beleži zgodovina prodanih izdelkov. Na eBay-ju je bil v letu 2003 objavljen skoraj vsak oglas v obliki avkcij in zelo malo oglasov z »Buy It Now« (BIN) – prodaja s fiksno ceno. Ravno nasprotno pa je značilno za bolho, saj je skoraj vsak oglas objavljen s fiksno ceno. Vendar kot kažejo najnovejše raziskave [2] (slika 2.1) se vedno več uporabnikov eBay-a tudi odloča za objavo oglasa s fiksno ceno oz. BIN.



Slika 2.1: Graf padca avkcij: črn graf predstavlja procent objavljenih avkcij.

Ta trend kaže na spremembo obnašanja spletnih nakupovalcev. Pred desetimi leti so bile spletne avkcije zelo popularne, uporabniki so prebili veliko ur pred zaslonom in tekmovali za izbrane izdelke, to je predstavljalo neko vrsto zabave. Sedaj pa uporabnik ne posveti toliko časa za nakup željenega izdelka, ampak išče predvsem udobje in priročnost pri nakupu.

Poglejmo kakšne metode oz. sisteme so in se mogoče še vedno uporabljajo za pomoč kupcu in prodajalcu na spletnih tržnicah.

2.1 Priporočilni sistemi za kupce

Zelo priljubljen sistem za kupce na eBay-u je bil »Auction Sniping«, ena od aplikacij, ki izvaja to metodo avtomatično je Gixen [13]. Gixen je spletna aplikacija, kjer uporabnik vpiše produkt, ki ga želi kupiti na eBay-u ter maksimalno ceno, ki jo želi plačati za izbran izdelek. Aplikacija avtomatično poveča ponujen znesek ob zadnjih sekundah avkcije in tako v večini primerov uporabnik zmaga na avkciji. »Auction sniping« je bil zelo popularen ob začetku razvoja eBay-a.

Razvili so tudi sistem za napoved končne cene avkcije [3]. Cilj tega sistema je napovedati končno ceno avkcije in se s to informacijo odločiti ali bi se nakup izplačal. Sistem bazira na izluščanju podatkov iz preteklih uspešno zaključenih oglasov (opisa, slik, itd.) in uporabo »boosting« algoritmov na področju strojnega učenja.

Na hitro smo predstavili dve popularni metodi, ki koristita kupcu. Glavna tema tega diplomskega dela pa je pomagati prodajalcu določiti priporočeno ceno, zato si pogledjmo sisteme za prodajalce.

2.2 Priporočilni sistemi za prodajalce

V prejšnjem poglavju smo omenili sistem za napoved končne cene in kako lahko pomaga kupcu, Ilya Raykhel in Dan Ventura pa sta v letu 2009 [4] razvila sistem z napoved končne cene avkcije za korist preprodajalcev. Omejila sta se

samo na kategorijo prenosnikov in s pomočjo k-NN algoritma razvila sistem, ki ima 16% povprečno napako, sistem je skoraj popolnoma avtomatičen in poveča učinkovitost preprodajalca za kar 562% (merjeno v zaslužek/uro).

Pri avkcijah na eBay-ju lahko določimo začetno ceno in rezervacijsko ceno. Rezervacijska cena predstavlja minimalno vsoto denarja, za katero želimo prodati naš izdelek, če v avkciji rezervacijska cena ni dosežena nismo dolžni prodati našega izdelka. Veliko raziskav se je osredotočilo tudi na optimalno določitev začetne in rezervacijske cene [5] s ciljem največjega zaslужka.

Nobena od raziskav pa se ni osredotočila na priporočeno fiksno ceno. V času, ko je bil eBay zelo popularen se je razvilo veliko programov in aplikacij, ki so pomagale uporabniku s raziskavo izdelkov na eBay-u. Veliko takšnih podjetij je propadlo ali pa zastarelo. Eden od glavnih predstavnikov takšnih programov, ki še vedno uspešno deluje pa je »TeraPeak for eBay« [14]. Aplikacija ima veliko orodij, ki pomagajo prodajalcu:

- Analiza in statistika drugih prodajalcev – kaj prodajajo, kakšne strategije uporabljajo,
- Orodje za optimizacijo oglasa – predlog ključnih besed, kako besede vplivajo na prodajo,
- Analiza obnašanja kupcev – kaj kupci kupujejo, napovedi kaj se bo kupovalo ter
- Priporočila za ceno

»TeraPeak for eBay« uporablja zgodovino uspešno zaključenih oglasov v zadnjih 365 dneh in s pomočjo teh podatkov lahko predlaga zelo natančno priporočeno ceno. Poleg tega še predlaga kdaj naj oglas objavimo in kako ga objavimo (format oglasa, katere besede se uporabijo, itd.)

Ta aplikacija se še najbolj približa našemu cilju za določitev priporočene cene na spletni tržnici bolha.com. Problem je v tem, da bolha ne beleži prodanih izdelkov in tako ne moramo dostopati do takšnih podatkov kot jih uporablja »TeraPeak for eBay«. V naslednjem poglavju si oglejmo kakšne postopke in strategije smo si zamislili za razvoj našega priporočilnega sistema.

Poglavje 3

Predstavitev sistema za priporočeno ceno

Prodaja rabljenih ali novih izdelkov preko interneta je zelo priljubljena, saj nam omogoča hitro in enostavno prodajo izdelkov, ki jih ne potrebujemo več. Tako kot pravi znani pregovor: »za nekoga odpad, za drugega zaklad«, se vedno najde kupec, ki potrebuje tisto stvar, ki jo mi prodajamo.

Ponavadi pa ne znamo oceniti vrednosti izdelkov, ki jih želimo prodati in moramo zato narediti manjšo raziskavo. Da bi bila ta raziskava hitra, lahka in razumljiva smo se odločili za razvoj sistema za priporočeno ceno.

Uporabnik spletnih tržnic lahko pri objavi svojega oglasa zapiše naslov, opis in doda slikovno gradivo. Zelo zanimiva funkcija pa bi bila samodejna priporočena cena. V tem diplomskem delu smo se želeli kar se da približati avtomatskem priporočilu, vendar zaradi nestrukturiranih in pomanjkljivih podatkov spletne tržnice bolha.com avtomatizacija ni mogoča. Seveda pa smo običajno raziskavo za priporočeno ceno bistveno skrajšali in poenostavili.

3.1 Namen razvoja priporočilnega sistema

Namen razvoja sistema za priporočeno prodajno ceno je omogočiti navadnemu uporabniku hiter in intuitiven prikaz stanja na spletni tržnici bolha.com in mu priporočiti ceno.

Poglejmo si primer problema: Uporabnik želi prodati svoj mobilni telefon »HTC desire hd«. Ugotoviti mora koliko je njegov izdelek vreden oz. kakšno je stanje na slovenski spletni tržnici. Uporabnik bi moral, za natančno priporočeno ceno, pregledati vse rezultate za iskalni niz »HTC desire hd« in beležiti cene ustreznih izdelkov. Pregledati bi moral 178 izdelkov (število na bolha.com ob času pisanja - 14.7.2013) od katerih je približno 160 rezultatov nerelevantnih. Takšna raziskava je zamudna in nepraktična.

Uporabnik bi lahko z uporabo našega sistema prišel do rezultata bistveno hitreje, hkrati pa bi se vsi potrebni podatki beležili avtomatično. Torej bi do priporočene cene prišel z le nekaj kliki.

V prikazanem primeru je kar 90 % izdelkov nerelevantnih (podobno je tudi pri drugih iskanjih), ampak če želimo natančno priporočeno ceno, moramo pregledati vse vrnjene izdelke in izbrati le tiste, ki ustrezajo našemu kriteriju. Torej je glavna naloga našega priporočilnega sistema, da izloči nerelevantne izdelke in izračuna priporočeno ceno med preostalimi ustreznimi izdelki.

Poglavje 4

Implementacija priporočilnega sistema

4.1 Potek raziskave s priporočilnim sistemom

V tem poglavju bomo predstavili potek raziskave za priporočeno ceno z našim priporočilnim sistemom. Izdelek, ki ga želimo prodati je popularen tablični računalnik Apple iPad mini, bolj natančno bomo iskali priporočeno ceno za »ipad mini 16gb«.

Uporabnik vpiše iskalni niz v polje »Iskani izdelek« in izbere strategijo raziskave (slika 4.1). Izbere lahko med dvema strategijama. Prva oz. privzeta izbira je raziskava obstoječih izdelkov na spletni tržnici bolha.com. Če uporabnik izbere prvo strategijo, lahko v polje »Negativne besede« vpiše besede, ločene s vejico, ki ne smejo vplivati na priporočeno ceno . Z izbiro druge strategije pa se raziskava izvede na podlagi novih izdelkov. Novi izdelki se iščejo na spletni trgovini mimovrste.com. Ob nastavitvi teh parametrov uporabnik klikne na gumb »Razišči!«. V tej točki se potek raziskave loči na dve poti, od-

Podatki izdelka

Iskani izdelek

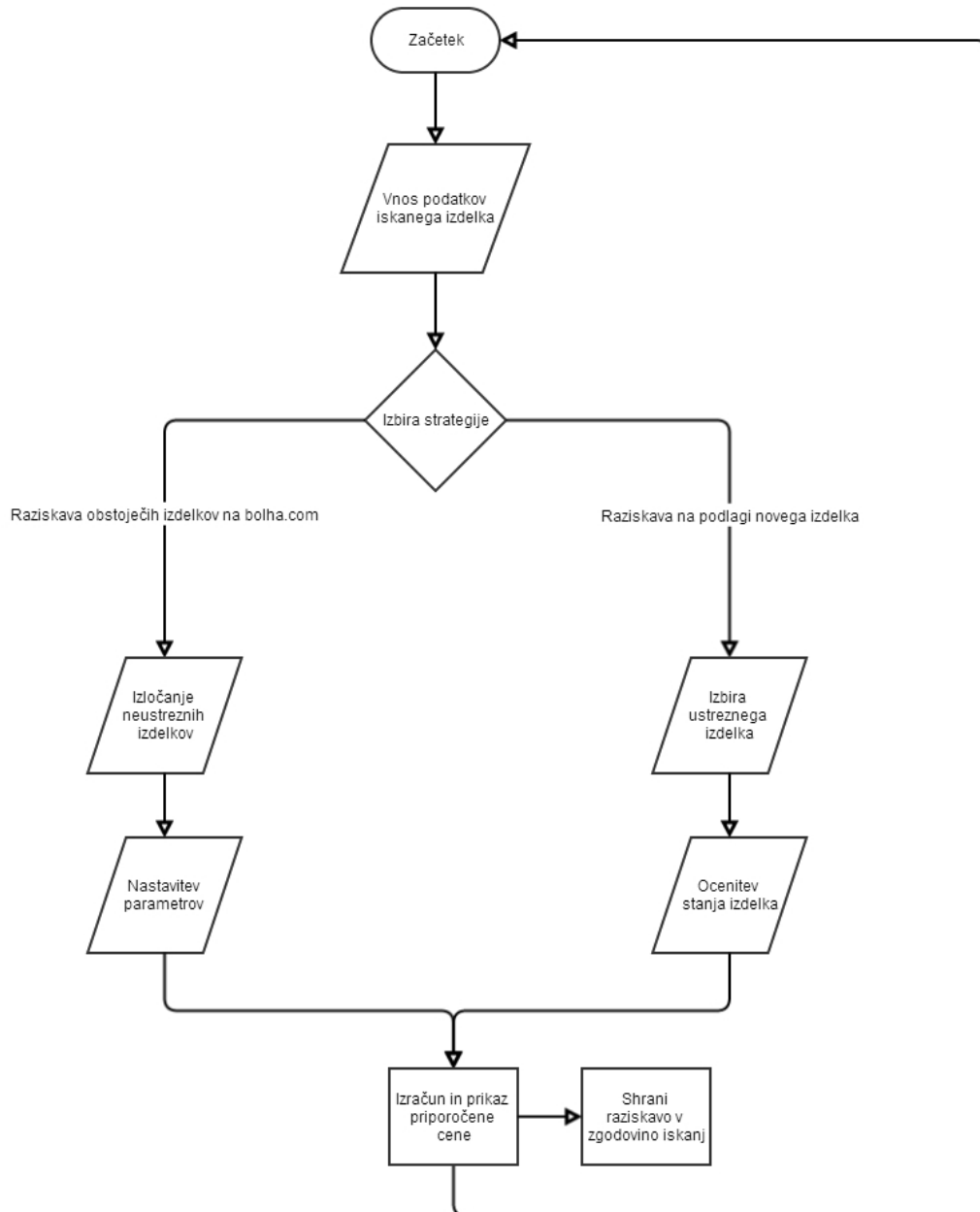
Negativne besede

Izberi strategijo

- Raziskava že obstoječih izdelkov na bolha.com
- Raziskava na podlagi novega izdelka (mimovrste.com)

Slika 4.1: Začetni obrazec raziskave.

visno od izbire strategije, zato sta obe poti opisani v naslednjih dveh poglavjih. Celotno delovanje sistema pa lahko vidite s pomočjo diagrama zaporedja na sliki 4.2.



Slika 4.2: Diagram zaporedja celotnega postopka.

4.1.1 Raziskava obstoječih izdelkov na spletni tržnici bolha.com

Kot smo že omenili je glavna naloga te strategije izločiti neustrezne izdelke, ki jih ponudi iskalnik bolha.com ter izračunati priporočeno ceno med preostalimi izdelki, ki so prestali naš kriterij.

Preden prikažemo podrobnosti sistema bomo pojasnili nekaj terminov, ki se bodo pojavljali v nadaljevanju:

- Negativne besede – besede, ki ne vplivajo na priporočeno ceno. Če se katera od negativnih besed pojavi v naslovu izdelka se ta izdelek izloči, razen če ni ta beseda tudi med dovoljenimi besedami.
- Dovoljene besede – besede, ki nasprotujejo negativnim. Dovoljene besede imajo višjo prioriteto kot negativne. Te besede se skirvajo v ozadju in uporabnik na njih ne mora neposredno vplivati. Izhodišče seznama dovoljenih besed so besede iskalnega niza. Npr.: če raziskujemo »ipad mini 16gb« in na seznam negativnih besed dodamo »Apple« in »iPhone«, na seznamu dovoljenih pa je samo »Apple« (poleg začetnega iskalnega niza). Sistem bi tedaj avtomatično pregledal naslednja izdelka:

A »Apple ipad mini 16gb« ter

B »Apple iPhone 5«

Izdelek A se ne bo izločil, kljub temu, da je »Apple« med negativnimi besedami. To je lep primer kako dovoljene besede nevtralizirajo negativne. Medtem pa bi se izdelek B izločil, zaradi negativne besede »iPhone«.

- Izločanje izdelka – ob izločitvi izdelka mislimo na obarvanje naslova izdelka. Za obarvanje namesto izbrisa izdelka smo se odločili zato, da vi-

dimo kaj sistem izloči in bomo tako kasneje tudi lažje analizirali učinkovitost sistema. Ob izločitvi na podlagi negativnih besed se naslov obarva rdeče, če pa se izloči na podlagi odstopanja od povprečne cene, se naslov obarva oranžno.

Za izločanje neustreznih izdelkov smo razvili dva postopka, ki sta med seboj prepletena:

1. Izločanje izdelkov na podlagi negativnih besed

Izhodišče za postopek so besede, ki jih uporabnik vnese ob začetku raziskave. Ob kliku na gumb »Razišči« se primerjajo vsi naslovi izdelkov z vnešenimi negativnimi besedami, če naslov izdelka vsebuje katero od negativnih besed se izdelek izloči (se obarva rdeče). Prav tako se izvede izločanje z negativnimi besedami vedno, ko izločimo izdelek ročno s klikom na »Izbriši!«. Ob kliku na ta gumb se poleg avtomatičnega dodajanja negativnih besed tega izdelka, pregleda vse izdelke levo od izbrisanega. Ob pregledu izdelkov na levi strani, torej od začetka do izbrisanega, se dodajo na seznam dovoljenih besed tiste besede, ki so v naslovu še ne izločenega (obarvanega) izdelka. Z novimi podatki (nove negativne in dovoljene besede), pa se preverijo vsi oglasi od izbrisanega naprej. Če katerega od izdelkov sistem avtomatično izloči se njegove besede zapišejo med negativne in se preverjanje nadaljuje.

2. Izločanje izdelkov na podlagi odstopanja od povprečne cene vzorca

V primeru, da ročno izločimo nek izdelek, ki ima levo od sebe vsaj pet ustreznih izdelkov, se izvede izločanje na podlagi odstopanja od povprečne cene. Med neizločenimi izdelki, ki predstavljajo vzorec ustreznih izdelkov se izračuna povprečje. Vsi izdelki desno od ročno izločenega

izdelka ne smejo odstopati za več kot 40% od izračunanega povprečja. Vsak izdelek, ki ne zadostuje temu pogoju se avtomatično izloči (obarva oranžno), prav tako pa se besede v naslovu zapišejo na seznam negativnih besed. Ta postopek se izvede samo enkrat (ko je pogoj za 5 ustreznih izdelkov izpolnjen) in je v določenih primerih zelo učinkovit, saj izloči večino nerelevantnih izdelkov.

Ob:

- izbiri strategije za raziskavo obstoječih izdelkov na spletni tržnici bolha.com,
- vnosu iskalnega niza,
- vnosu morebitnih negativnih besed ter
- kliku na gumb »Razišči!«

se prikažejo vsi izdelki, ki ustrezajo podanim parametrom. Vsak izdelek je predstavljen s sliko, naslovom, opisom, ter ceno izdelka. Vsi izdelki so razporejeni v mrežo za lažji in hitrejši pregled.

Ob vsakem izdelku je tudi gumb »Izbriši!« s katerim lahko izločimo neustrezen izdelek. Ko izločimo izdelek se v polje negativnih besed dodajo besede naslova izločenega izdelka in se z novimi negativnimi besedami pregledajo vsi preostali izdelki in se tako podobni izdelki samodejno izločijo. Prav tako se ob pogoju, da je pred izločenim izdelkom vsaj pet ustreznih izdelkov, zažene postopek za izločanje glede na odstopanje cen. Ta algoritem lahko pred začetkom izločanja tudi onemogočimo in tako sistem izloča izdelke samo na podlagi negativnih besed.

Pri naši raziskavi za »ipad mini 16gb« smo lahko izločili vse nerelevantne izdelke s samo tremi kliki, za vse ostalo je poskrbel sistem.


Izločite nerelevantne izdelke

Oranžna barva - Izločeni izdelki glede na odstopanje cen

Rdeča barva - Izločeni izdelki glede na negativne besede

Uporabi izključevanje na podlagi odstopanja cen

Število vseh oglasov: 05




iPad Mini 16GB WiFi

Prodajam iPad Mini 16GB WiFi, prodajam zato ker ga ne rabim, star je dva meseca s papirji in računom, znakov uporabe skoraj

230,00

[Izbrisi!](#)




IPAD MINI 16gb wifi+cellular

Prodajam NOV-ZAPAKIRAN APPLE IPAD MINI 16GB WiFi + Cellular črne barve, kupljen v Epi trgovini v Ljubljani. Cena novega v trgovini

360,00

[Izbrisi!](#)




iPad mini 16GB cellular, bele barve

Prodajam en mesec star iPad mini, s podporo cellular in kapaciteto 16GB. Menjave ne pridejo v poštev!

380,00

[Izbrisi!](#)

⋮




Orodje za popravilo Apple iPhone, iPad, iPod

(1Gb, 2Gb, 4Gb) Apple iPod Nano 2G 3G 4G 5G 6G 2Gb, 4Gb, 8Gb, **16Gb** Apple iPod Mini 1G 2G 4Gb, 8Gb Apple iPod Classic ... 64Gb Apple iPad WiFi **16Gb**, 32Gb, 64Gb Apple iPad 2 WiFi **16Gb**, 32Gb, 64Gb Apple iPad 2 WiFi + 3G **16Gb**, 32Gb, 64Gb Novi iPad (Apple iPad 3) WiFi **16Gb**, 32Gb, 64Gb Novi iPad (Apple iPad 3) WiFi + 4G

10,06

[Izbrisi!](#)




FM oddajnik z daljinskim upravljalnikom za Apple iPhone, iPad, iPod

Prilimo za: Apple iPod Nano 2G 3G 4G 5G 6G 2Gb, 4Gb, 8Gb, **16Gb** Apple iPod Mini 1G 2G 4Gb, 6Gb Apple iPod Classic 3G 4G ... 64Gb Apple iPad WiFi **16Gb**, 32Gb, 64Gb Apple iPad WiFi + 3G **16Gb**, 32Gb, 64Gb Apple iPad 2 WiFi **16Gb**, 32Gb, 64Gb ... Apple iPad 2 WiFi + 3G **16Gb**, 32Gb, 64Gb Novi iPad (Apple iPad 3) WiFi **16Gb**, 32Gb, 64Gb Novi iPad (Apple iPad 3) WiFi + 4G

15,76

[Izbrisi!](#)



Omrežni 220V polnilec za iPod in iPhone 3G 3GS 4 4S

2G 3G 4G 5G 6G 2Gb, 4Gb, 8Gb, **16Gb** Apple iPod Mini 1G 2G 4Gb, 6Gb Apple iPod Classic 3G 4G 30Gb, 60Gb, 80Gb, 120Gb ... 5,90 EUR Omrežni 220V polnilec za Apple iPod, iPhone, iPad ipadponudba Prilimo za: Apple iPod Nano ... 160Gb Apple iPod Touch 1G 2G 3G 4G 8Gb, **16Gb**, 32Gb, 64Gb Apple iPhone 2G 4Gb, 8Gb, **16Gb** Apple iPhone 3G 8Gb, **16Gb**

5,90

[Izbrisi!](#)

[Končaj z izbiro](#)

Slika 4.3: Okrnjen prikaz izdelkov.

S pritiskom na gumb »Končaj z izbiro« zaključimo izključevanje izdelkov in nadaljujemo z raziskavo. Prikaže se razdelek parametrov, kjer lahko izbiramo med metodo prodaje ter dostave.

Na sliki 4.4 lahko vidite gumb »Prikaži izbor izdelkov«, ki prikaže ali skrije izbor izdelkov, ki smo jih izbrali v prejšnjem koraku. Tako si lahko še enkrat ogledamo, katere izdelke smo izločili.

Metodi prodaje sta dve:

- »Izdelek bi rad prodal čim hitreje« - ob tej izbiri bo kot osnova za pri-

Prikaži izbor izdelkov

Nastavite parametre

Hitra prodaja ali večji zaslužek?

Izdelek bi rad prodal čim hitreje

Zaslužek je bolj pomemben, kot hitrost prodaje

Metoda dostave:

Osebni prevzem

Paket bom poslal po pošti

Pridobi priporočeno ceno

Slika 4.4: Nastavitev parametrov pri raziskavi obstoječih izdelkov na bolha.com

poročeno ceno izbrana minimalna cena, saj bomo s tem zagotovili, da se bo izdelek prodal kar se da hitro.

- • »Zaslužek je bolj pomemben, kot hitrost prodaje« - osnova za priporočeno ceno je mediana ali sredinska vrednost. S tem se znebimo ekstremnih vrednosti.

$$\tilde{x} = \begin{cases} x_{\frac{n+1}{2}}, & n = \text{liho} \\ \frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}), & n = \text{sodo} \end{cases} \quad (4.1)$$

4.1: Enačba mediane.

Izbira metode dostave pa ponuja »Osebni prevzem«, ki ne vpliva na pri-

poročeno ceno ter »Paket bom poslal po pošti«. Druga izbira prikaže vnosno polje kamor vnesemo težo paketa v kilogramih. Glede na težo pa se izračuna dejanska cena pošiljke na Pošti Slovenije [15], ki se prišteje k osnovni priporočeni ceni.

Korak nastavitve parametrov zaključimo s klikom na gumb »Pridobi priporočeno ceno«. Prikaže se nov razdelek, kjer se izpišejo vsi podatki raziskave, kot tudi priporočena cena. Prav tako pa se pravkar zaključena raziskava shrani na seznam zgodovine iskanj in tako lahko opravimo več raziskav in preklapljamo med opravljenimi raziskavami.



Slika 4.5: Prikaz rezultatov in priporočene cene pri raziskavi obstoječih izdelkov na bolha.com

Med končnimi rezultati so naslednji podatki (vidni na sliki 4.5):

- Število izdelkov - število ustreznih izdelkov, ki so prestali naš kriterij,
- Najmanjša cena,
- Najvišja cena,
- Povprečje cen,
- Izbrana metoda prodaje,
- Cena poštnine (opcijsko),
- Priporočena cena – sestavljena iz osnovne priporočene cene, glede na izbiro metode prodaje in ceno poštnine,
- Diagram razporeditev cen

Stolpičasti diagram razporeditev cen daje uporabniku predstavo kako so cene razporejene, v katerem intervalu cen je uvrščenih največ izdelkov. S to informacijo lahko uporabnik sam prilagodi priporočeno ceno.

Tako smo s raziskavo obstoječih izdelkov na spletni tržnici bolha.com tudi zaključili. Kasneje si bomo ogledali analizo učinkovitosti naše strategije, še prej pa strategijo za raziskavo na podlagi novih izdelkov s spletne trgovine mimovrste.com.


4.1.2 Raziskava na podlagi novega izdelka - mimovrste.com

Uporabnost strategije pridobivanja priporočene cene na podlagi novega izdelka se pokaže, ko s prvo strategijo (raziskavo obstoječih izdelkov na bolha.com)

ne dobimo pravih rezultatov. Lahko gre za pomanjkanje ponudbe na spletni tržnici bolha.com ali pa samo hočemo preprosto oceno vrednosti našega izdelka.

Ob vnosu iskanega izdelka, izbiri druge možnosti strategije in zagonu priporočilnega sistema s klikom na gumb »Razišči!«, se podobno kot pri prvi strategiji prikažejo izdelki, le da so izdelki pridobljeni iz spletne trgovine mi-movrste.com (slika 4.6).

Izberite vaš izdelek




Apple iPad mini 16 GB (Wi-Fi), bel

- Wi-Fi (802.11a/b/g/n)
- iOS 6
- Three-axis gyro Accelerometer Ambient light sensor Digital compass AirPlay
- ne
- 7,9" - 17,8 cm
- 16 GB

339,00

Izberi!




Apple iPad mini 16 GB (Wi-Fi), črn

- Wi-Fi (802.11a/b/g/n)
- iOS 6
- Three-axis gyro Accelerometer Ambient light sensor Digital compass AirPlay
- ne
- 7,9" - 17,8 cm
- 16 GB

339,00

Izberi!




Apple iPad mini 16 GB (Wi-Fi + Cellular), črn

- Wi-Fi (802.11a/b/g/n), Cellular (GSM/EDGE (850, 900, 1800, 1900 MHz), UMTS/HSPA+DC-HSDPA (850, 900, 1900, 2100 MHz), LTE)
- iOS 6
- Three-axis gyro Accelerometer Ambient light sensor Digital compass AirPlay
- A-GPS, GLONASS
- 7,9" - 17,8 cm
- 16 GB

469,00

Izberi!



Apple iPad 16 GB (Wi-Fi + Cellular), bel

- Wi-Fi (802.11a/b/g/n), Cellular (GSM/EDGE (850, 900, 1800, 1900 MHz), UMTS/HSPA+DC-HSDPA (850, 900, 1900, 2100 MHz), LTE)
- iOS 6
- Three-axis gyro Accelerometer Ambient light sensor Digital compass AirPlay
- A-GPS, GLONASS
- 7,9" - 17,8 cm
- 16 GB

469,00

Izberi!




Tablični računalnik Prestigio MultiPad 2 Prime Duo 8.0 PMP5780_DUO, črn + darilo 16 GB micro

- Wi-Fi 802.11b/g/n
- Android 4.1
- Video Accelerator OpenGL ES 2.0 G-Sensor 3D Accelerometer USB OTG Support
- ne
- 8" - 20,3 cm
- 16 GB

169,00

Izberi!



Apple iPad 16 GB (WiFi) 4. generacija, črn

- Wi-Fi (802.11a/b/g/n)
- iOS 6
- senzorji: Accelerometer, ambient light sensor, magnetometer, giroskop
- ne
- 9,7" - 24,6 cm
- 16 GB


509,00

Izberi!

Slika 4.6: Prikaz izdelkov pri raziskavi na podlagi novega izdelka

V tem koraku izdelkov ne izločamo, ampak izberemo tistega, ki ga želimo

prodati. Ko najdemo izdelek in ga izberemo s klikom na gumb »Izberi!« se vsi ostali izdelki izbrisejo, izbrani izdelek pa se postavi na vrh strani, prav tako pa se prikaže razdelek s parametri (slika 4.7).

Vaš izdelek	Ocenite stanje večšega izdelka
 <p>Apple iPad mini 16 GB (Wi-Fi + Cellular), črn</p> <ul style="list-style-type: none"> • Wi-Fi (802.11a/b/g/n), Cellular (GSM/EDGE (850, 900, 1800, 1900 MHz), UMTS/HSPA+/DC-HSDPA (850, 900, 1900, 2100 MHz), LTE) • iOS 6 • Three-axis gyro Accelerometer Ambient light sensor Digital compass AirPlay • A-GPS, GLONASS • 7,9" - 17,8 cm • 16 GB <p>469,00</p> <p>Vaš izbran izdelek</p>	<p>Starost izdelka:</p> <p>* vnesite starost v letih (če je izdelek star 6 mesecev vpišite 0.5) * Če je izdelek nov, pustite polje prazno ali pa vpišite 0</p> <input type="text" value=""/> let <p>Poškodba izdelka:</p> <p><input checked="" type="radio"/> Nepoškodovan <input type="radio"/> Rahlo poškodovan (praskе, lažje poškodbe) <input type="radio"/> Večje poškodbe (udolbine, večje praske) <input type="radio"/> Nedelujoč (uporabni deli)</p> <p>Metoda dostave:</p> <p><input checked="" type="radio"/> Osebni prevzem <input type="radio"/> Paket bom poslal po pošti</p> <p>Pridobi priporočeno ceno</p>

Slika 4.7: Izbran izdelek in ocenitev stanja.

Svoj izdelek ocenimo z naslednjimi parametri:

- Starost izdelka – starost izraženo v letih. Za izdelek star 6 mesecev se vpiše v polje »0.5«, če je izdelek nov pa »0« ali pa se polje pusti prazno,
- Poškodba izdelka – izbiramo lahko med: nepoškodovan, rahlo poškodovan, večje poškodbe in nedelujoč,
- Metoda dostave – identično kot pri prvi strategiji. Lahko izberemo pošiljko po pošti in vnesemo težo paketa s katero se izračuna cena poštnine.

Osnova za priporočeno ceno je cena izbranega izdelka. Glede na nastavitve parametrov pa se osnovna cena niža. Do vrednosti odbitkov v spodnji tabeli smo prišli s kratko raziskavo bolha.com in nekaterih elektronskih izdelkov. Zaradi prevelikih odstopanj in pomanjkljivih informacijah smo pri raziskavi odbitke rahlo zaokrožili. Rezultati so prikazani v tabelah 4.1 in 4.2.


Starost[v letih]	Odbitek [%]
0	0
$0 < x \leq 1$	10
$1 < x \leq 2$	20
$2 < x \leq 3$	30
$3 < x \leq 4$	40
$4 < x \leq 5$	50
$5 < x$	60

Tabela 4.1: Odbitki glede na starost izdelka.

Poškodba	Odbitek [%]
Nepoškodovan	0
Rahlo poškodovan	10
Večje poškodbe	40
Nedelujoč	70

Tabela 4.2: Odbitki glede na poškodbo izdelka.

Nastavitve parametrov zaključimo s klikom na gumb »Pridobi priporočeno ceno«. Prikaže se še tretji razdelek s končnimi rezultati prikazan na sliki 4.8.

Vaš izdelek	Ocenite stanje večega izdelka	Priporočena cena
 <p>Apple iPad mini 16 GB (Wi-Fi + Cellular), črn</p> <ul style="list-style-type: none"> • Wi-Fi (802.11a/b/g/n), Cellular (GSM/EDGE (850, 900, 1800, 1900 MHz), UMTS/HSPA+/DC-HSDPA (850, 900, 1900, 2100 MHz), LTE) • iOS 6 • Three-axis gyro Accelerometer Ambient light sensor Digital compass AirPlay • A-GPS, GLONASS • 7,9" - 17,8 cm • 16 GB <p>469,00</p> <p>Vaš izbran izdelek</p>	<p>Starost izdelka:</p> <p>* vnesite starost v letih (Če je izdelek star 6 mesecev vpišite 0.5) * Če je izdelek nov, pustite polje prazno ali pa vpišite 0</p> <p><input type="text" value="1"/> let</p> <p>Poškodba izdelka:</p> <p><input type="radio"/> Nepoškodovan <input checked="" type="radio"/> Rahlo poškodovan (praske, lažje poškodbe) <input type="radio"/> Večje poškodbe (udolbine, večje praske) <input type="radio"/> Nedelujoč (uporabni deli)</p> <p>Metoda dostave:</p> <p><input type="radio"/> Osebni prevzem <input checked="" type="radio"/> Paket bom poslal po pošti</p> <p>Teža paketa</p> <p><input type="text" value="2"/> kg</p> <p>Pridobi priporočeno ceno</p>	<p>Cena izdelka: 469€ Procent odbitka glede starosti: 10% Procent odbitka glede poškodb: 10% Cena poštnine: 4.47€</p> <p>Priporočena cena:</p> <p>384.36€</p>

Slika 4.8: Prikaz zaključene raziskave na podlagi novih izdelkov mimovrste.com

Za lažje razumevanje priporočene cene se izpišejo tudi naslednji podatki:

- Osnovna cena izdelka,
- Procent odbitka glede na starost,
- Procent odbitka glede na poškodbo,
- Morebitna cena poštnine

Hkrati pa se, podobno kot v prvi strategiji, končana raziskava shrani na seznam zgodovine iskanj (slika 4.9).

Med različnimi zaključenimi raziskavami se lahko po zgodovini iskanj preprosto sprehajamo in ponovno prikažemo vse podatke določene raziskave. Vsaka povezava v zgodovini se začne s črko »B« (raziskava obstoječih izdelkov Bolha.com) ali »M« (raziskava na podlagi novega izdelka Mimovrste.com).

Zgodovina iskanj

B: 1.iskanje-> ipad mini 16gb (priporočena cena=344.47€)

M: 2.iskanje-> ipad mini 16gb (priporočena cena=384.36€)

Slika 4.9: Zgodovina raziskav.

Predstavili smo delovanje našega sistema, v naslednjem poglavju pa si pogledimo kako je naš sistem zgrajen s tehnične strani.

4.2 Tehnične podrobnosti implementacije priporočilnega sistema

Potek in koncept našega sistema smo že predstavili, sedaj pa si pogledimo kako je sistem sestavljen in kakšno strukturo ima. Sistem je implementiran kot spletna aplikacija, ki je dostopna na računalniku, pametnem telefonu ali katerekoli napravi s sodobnim brskalnikom in povezavo z internetom.

Spletna aplikacija je spisana v jezikih:

- HTML5,
- CSS (oz. s predprocesno tehniko LESS),
- PHP
- Javascript s knjižnico jQuery

Osnovno ogrodje in podoba spletne strani je izdelana v HTML5 s pomočjo orodja Twitter Bootstrap, ki je namenjen hitrem spletnem razvoju z velikim številom prednastavljenih izgledov elementov, z lažjo postavitvijo ogrodja

```
Projekt
  assets
    css
      bootstrap-responsive.min.css
      bootstrap.min.css
      style.css
    img
      ajax-loader.gif
    js
      bootstrap.min.js
      Chart.min.js
      main.js
    less
      style.less
    scraper
      simple_html_dom.php
  getSearchResults.php
  index.php
```

Slika 4.10: Seznam datotek projekta.

strani in veliko dodatnih javascript funkcij, ki dodajo interaktivnost in funkcionalnost. Poleg tega pa Bootstrap omogoča tudi »responsive design«, ki spreminja izgled prikazane strani glede na širino zaslona, tako se uporabniku izboljša izkušnja uporabe spletne aplikacije. Za uporabo tega orodja je bila potrebna le vključitev določenih datotek (CSS in javascript) v element »head« glavne datoteke index.php. V tej datoteki je zapisana osnovna postavitev spletne strani in je hkrati edina stran za prikaz informacij uporabnikom, saj se vse poizvedbe za zunanje podatke vršijo preko AJAX klicev.

Del kode v datoteki index.php je tudi začetni obrazec, ki smo ga videli na

sliki 4.1:

```
<form id="searchForm" method="POST" action="getSearchResults.php">
  <fieldset>
    <legend>Podatki izdelka</legend>
    <label>Iskani izdelek</label>
    <input type="text" id="adTitle" name="adTitle">
    <div id="negativeWordsDiv">
      <label>Negativne besede</label>
      <input type="text" id="negativeWords" name="negativeWords">
    </div>
    <label>Izberi strategijo</label>
    <label class="radio">
      <input type="radio" name="strategy" id="strategy1" value="1"
        checked>
      Raziskava že obstoječih izdelkov na bolha.com
    </label>
    <label class="radio">
      <input type="radio" name="strategy" id="strategy2" value="2">
      Raziskava na podlagi novega izdelka (mimovrste.com)
    </label>
    <br>
    <button type="submit" class="btn btn-success"
      name="submitButton">Razišči!</button>
    <button id="clearBtn" class="btn btn-danger"
      name="clearBtn">Počisti!</button>
  </fieldset>
</form>
```

Ob kliku na gumb za oddajo obrazca se izvede javascript funkcija z AJAX klicem za pridobitev rezultatov iskanja na bolha.com ali pa na mimovrste.com, odvisno od izbrane strategije.

```
$("#form#searchForm").on("submit", function() {  
    $("#results").html(" ");  
    if ($('#strategy1').is(':checked')){  
        $('#resultsDiv legend').text("Izločite nerelevantne  
        izdelke");  
        $('#resultsDiv').removeClass('hide');  
    }  
    if ($('#strategy2').is(':checked')){  
        $('#resultsDiv legend').text("Izberite vaš izdelek");  
        $('#resultsDiv').removeClass('hide');  
    }  
    var myForm = $(this);  
    var url = myForm.attr("action");  
    var type = myForm.attr("method");  
    var data = myForm.serialize();  
    $("#loadingDiv").show();  
    $.ajax({  
        url: url,  
        type: type,  
        data: data,  
        success: function(response){  
            $("#results").html(response);  
            $("#loadingDiv").hide();  
        }  
    })  
})
```

```
    });  
    return false;  
});
```

Pri AJAX klicu se dostopa do strežnika, bolj natančno do PHP skripte »getSearchResults.php«. S pomočjo knjižnice »PHP Simple HTML DOM Parser« lahko iz spletne tržnice bolha.com in spletne trgovine mimovrste.com pridobimo potrebne podatke:

- Število vseh oglasov,
- Podatki o posameznem izdelku:
 - Spletni naslov slike izdelka,
 - Naslov izdelka,
 - Opis izdelka ter
 - Cena izdelka

Spletna tržnica bolha.com in spletna trgovina mimovrste.com nimata javno dostopnega API-ja, zato moramo vse te podatke pridobiti z metodo spletnega luščenja podatkov oz. web scraping, kar nam omogoča prej omenjena knjižnica. Spisali smo dve funkciji v odvisnosti od izbire strategije raziskave. Spodaj je prikazana koda za pridobitev podatkov iz spletne strani bolha.com.

```
function getResultatsFromBolha($adTitle){  
    $steviloOglasovNaStran = 20;  
    $steviloVsehOglasov = 0;  
    $steviloStrani = 0;
```

```

$html = file_get_html('http://www.bolha.com/iskanje?q='
.$adTitle."&adTypeH=00_Prodam%2FS+fiksno+ceno%2F");
$element = $html->find('div.allItemCount',0);
$temp = explode(':', $element->plaintext);
$steviloVsehOglasov = substr($temp[1], 1);
$temp = explode('.', $steviloVsehOglasov);
$steviloVsehOglasov = implode('', $temp);

if($steviloVsehOglasov==0){
    echo "<p class=\"noHits\">Za iskalni niz \"" . $adTitle . "\"
        trenutno ni zadetkov.</p>";
    exit(0);
}

echo "Število vseh oglasov: <span
    class=\"numberOfAllAds\">{$steviloVsehOglasov}</span> <br>";
$steviloStrani = ceil($steviloVsehOglasov/$steviloOglasovNaStran);
//echo "Število strani: " . $steviloStrani . "<br>";

$stevecOglasov = 0;
for ($i=1; $i <= $steviloStrani; $i++) {
    $html = file_get_html("http://www.bolha.com/iskanje?q=" .
        $adTitle .
        "&y0=Iskanje&nca=1&page=" . $i . "&adTypeH=00_Prodam%2FS+fiksno+ceno%2F");

    foreach($html->find('div.ad') as $element){
        $ad['imagelink'] = $element->find('img', 0)->getAttribute
            ('data-original');
    }
}

```

```

$ad['title'] = $element->find('div.content h3 a',
    0)->innertext;
$ad['description'] = $element->find('div.content',
    0)->innertext;
$temp=explode("</h3>", $ad['description']);
$ad['description'] = $temp[1];
$ad['price'] = $element->find('div.price', 0)->plaintext;
$temp=explode(" ", $ad['price']);
$ad['price'] = $temp[0];

echo "<div class=\"ad ad-{\$stevecOglasov}\">";
    echo "<div class=\"picture\">";
        echo "<img src=\"{\$ad['imagelink']}\" \>";
    echo "</div>";
    echo "<div class=\"title\">";
        echo "<h4> " . $ad['title'] . "</h4>";
    echo "</div>";
    echo "<div class=\"description\">";
        echo "<p> " . $ad['description'] . "</p>";
    echo "</div>";
    echo "<div class=\"price\">";
        echo $ad['price'];
    echo "</div>";
    echo "<div class=\"deleteLink\">";
        echo "<button class=\"btn btn-danger
            remove\">Izbriši!</button>";
    echo "</div>";
echo "<br></div>";

```

```
        $stevecOglasov++;  
    }  
}  
}
```

Ob uspešno zaključenem AJAX klicu se pridobljeni izdelki prikažejo uporabniku za nadaljevanje raziskave.

Ob zaključku raziskave obstoječih izdelkov na bolha.com se izriše zanimiv diagram razporeditev cen, ki smo ga izdelali s pomočjo knjižnice »Chart.js«. Zaradi dinamične narave izrisa diagrama je nastavitev le-tega malce drugačna.

```
var numberOfUnits = 10;  
var range = maxPrice - minPrice;  
var rangeUnit = Math.ceil(range/numberOfUnits);  
var labelsArray = new Array();  
var dataArray = new Array();  
  
for (var i = 0; i < numberOfUnits; i++) {  
    dataArray[i]=0;  
};  
  
for (var i = 0; i < numberOfUnits; i++) {  
    if(i == (numberOfUnits-1)){  
        labelsArray[i] = (minPrice+(i*rangeUnit)+1).toString() + " - " +  
            maxPrice;  
    }  
    else if(i == 0){
```

```
        labelsArray[i] = (minPrice+(i*rangeUnit)).toString() + " - " +
            (minPrice+((i+1)*rangeUnit)).toString();
    }
    else {
        labelsArray[i] = (minPrice+(i*rangeUnit)+1).toString() + " - " +
            (minPrice+((i+1)*rangeUnit)).toString();
    }
    var label = labelsArray[i].split(" - ");
    var min = label[0];
    var max = label[1];
    for (var j = 0; j < pricesArray.length; j++) {
        if(pricesArray[j]>=min && pricesArray[j]<=max){
            dataArray[i]++;
        }
    };
};

var barChartData = {
    labels : [],
    datasets : [{
        fillColor : "rgba(255,118,0,0.5)",
        strokeColor : "rgba(255,118,0,1)",
        data : []
    }]
}

for (var i = 0; i < labelsArray.length; i++) {
    barChartData.labels.push(labelsArray[i]);
}
```

```
    barChartData.datasets[0].data.push(dataArray[i]);  
  };  
  
var ctx = $("#graphCanvas").get(0).getContext("2d");  
var myLine = new Chart(ctx).Bar(barChartData);
```

Izris diagrama se sproži ob kliku na gumb za pridobitev priporočene cene.

Poleg prikazanih izsekov kode smo uporabili še nekaj drugih tehnik, med katerimi je tudi uporaba tehnologije LESS, ki nadgradi oz. razširi uporabo in zmožnosti navadnega jezika za oblikovanje spletnih strani CSS.

Poglavje 5

Analiza natančnosti razvitega priporočilnega sistema

Sedaj, ko smo predstavili zakaj smo priporočilni sistem razvili in kako le-ta deluje, je na vrsti analiza natančnosti. Preveriti moramo ali je sistem sploh uporaben oz. ali mu je moč zaupati.

Kot že omenjeno je bil glavni cilj razviti sistem za priporočeno ceno izdelkov, ki jih želimo prodati na slovenski spletni tržnici bolha.com. Pozornost analize bomo usmerili le v strategijo obstoječih izdelkov na bolhi, saj nam ravno ta poda priporočeno ceno, ki je najbolj uporabna. Največji problem pri tej strategiji pa je izločanje nerelevantnih izdelkov, zato bo število pravilno izločenih nerelevantnih izdelkov tudi merilo uspešnosti našega sistema.

Za analizo bomo uporabili metodo natančnosti in priklica (angleško: precision and recall [6]). Metodo smo izvajali ob vsaki izločitvi oz. kliku na gumb »Izbriši!«. Natančnost je bila izmerjena z enačbo 5.1.

Priklic pa z enačbo 5.2.

V sistem nismo ročno vpisovali negativnih besed in smo tako pustili sistemu,

$$\frac{\text{število ustreznih izločitev}}{\text{število vseh izločitev}} \quad (5.1)$$

5.1: Enačba natančnosti.

$$\frac{\text{število ustreznih izločitev}}{\text{število vseh izdelkov, ki jih je potrebno še izločiti}} \quad (5.2)$$

5.2: Enačba priklica.

da izloči izdelke avtomatično. Analizo smo opravili z večimi izdelki in izbrali 3, ki smo jih uvrstili v diplomsko delo:

- iPad mini 16gb wifi,
- Samsung galaxy S4 LTE,
- HP 6735s

Izbrali smo izdelke, ki se jih lahko kategorizira in se ob pregledu hitro vidi kateri izdelek na seznamu je ustrezen, hkrati pa imajo ob iskanju na bolhi relativno veliko število vrnjenih rezultatov.

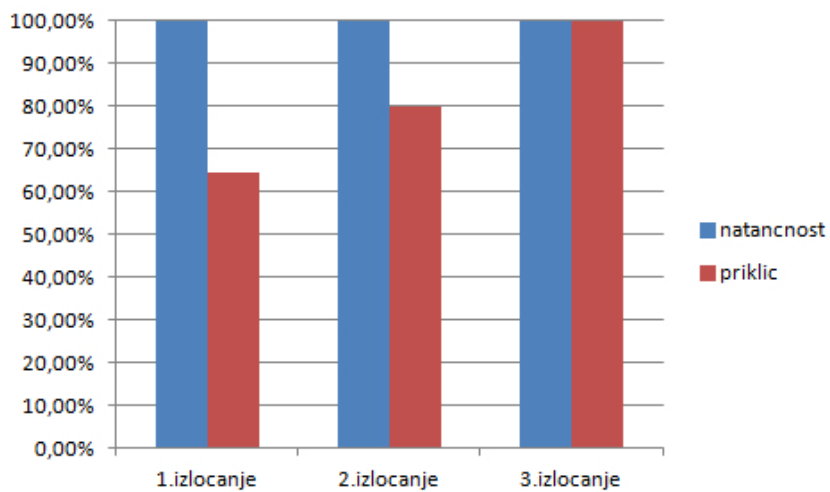
5.1 Rezultati analize »iPad mini 16gb wifi«

število vseh vrnjenih izdelkov	40
število ustreznih izdelkov	12
število izdelkov za izločitev	28

Tabela 5.1: Tabela osnovnih podatkov analize iPad mini 16gb wifi.

	natančnost	priklic
1. izločanje	18/18	18/28
2. izločanje	8/8	8/10
3. izločanje	2/2	2/2

Tabela 5.2: Tabela natančnosti in priklica ob izločanjih pri analizi iPad mini 16gb wifi.



Slika 5.1: Graf natančnosti in priklica ob posameznih izločanjih pri analizi iPad mini 16gb wifi.

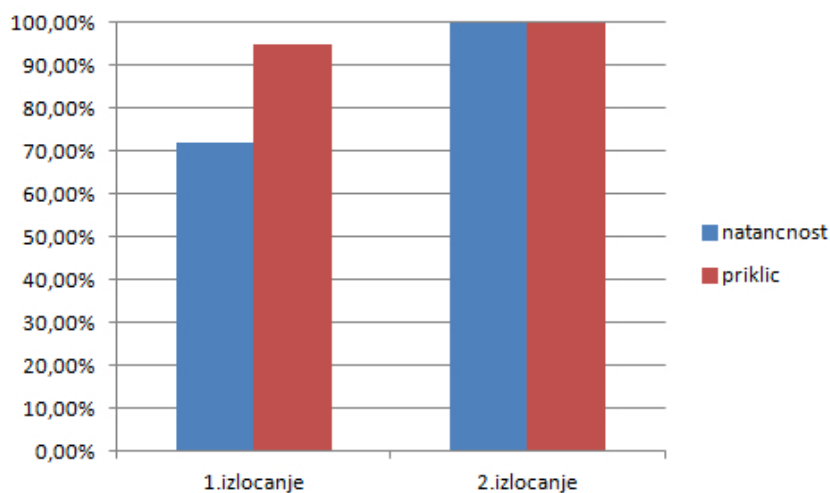
5.2 Rezultati analize »Samsung galaxy S4 LTE«

število vseh vrnjenih izdelkov	45
število ustreznih izdelkov	26
število izdelkov za izločitev	19

Tabela 5.3: Tabela osnovnih podatkov analize Samsung galaxy S4 LTE.

	natančnost	priklic
1. izločanje	18/25	18/19
2. izločanje	1/1	1/1

Tabela 5.4: Tabela natančnosti in priklica ob izločanjih pri analizi Samsung galaxy S4 LTE.



Slika 5.2: Graf natančnosti in priklica ob posameznih izločanjih pri analizi Samsung galaxy S4 LTE.

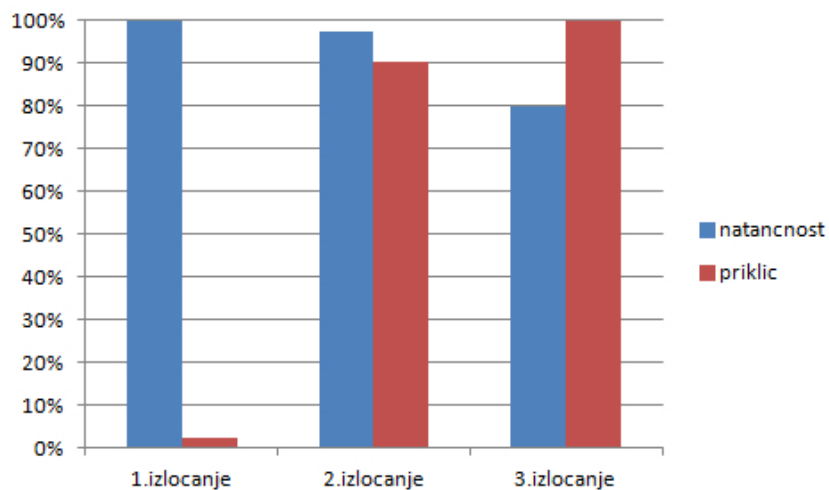
5.3 Rezultati analize »HP 6735s«

število vseh vrnjenih izdelkov	48
število ustreznih izdelkov	5
število izdelkov za izločitev	43

Tabela 5.5: Tabela osnovnih podatkov analize HP 6735s.

	natančnost	priklic
1. izločanje	1/1	1/43
2. izločanje	38/39	38/42
3. izločanje	4/5	4/4

Tabela 5.6: Tabela natančnosti in priklica ob izločanjih pri analizi HP 6735s.



Slika 5.3: Graf natančnosti in priklica ob posameznih izločanjih pri analizi HP 6735s.

5.4 Interpretacija rezultatov analize

Iz teh rezultatov analize lahko vidimo, da sistem v večini izločanj deluje natančno. Poleg natančnosti lahko razberemo, da sistem pride do končnega rezultata (izločitev neustreznih izdelkov) v zelo majhnem številu korakov. Zanimiv podatek in hkrati tudi ključen za določitev priporočene cene pa je natančnost števila ustreznih izdelkov na koncu izločanj.

	natančnost ustreznih izdelkov
iPad mini 16gb wifi	12/12 (100%)
Samsung galaxy S4 LTE	19/26 (73%)
HP 6735s	3/5 (60%)

Tabela 5.7: Tabela natančnosti ustreznih izdelkov na koncu izločanj.

V večini primerov analize na preostalih izdelkih, ki jih nismo uvrstili v to diplomsko delo, se je natančnost ustreznih izdelkov ob končanem izločanju gibala nad 60%.

Najslabši rezultat natančnosti ustreznih izdelkov na koncu izločanj smo dosegli pri iskanju priporočene cene za prenosni računalnik HP 6735s. Glavni razlog za slab rezultat sistema je pomanjkljiv ali prepodroben naziv objavljenih izdelkov. Npr. izdelek z nazivom: "HP 6753s, Dual core 2.0GHZ,320HDDD,3GB ram" je bil na začetku uvrščen med ustrezne izdelke, ker pa je bila večina oglasov na bolhi namenjena prodaji komponent tega prenosnika se je na podlagi negativnih besed (Dual core, 3GB, itd.) iz prej izločenih izdelkov izločil tudi ta izdelek.

Ob razvoju in analizi smo ugotovili, da je glavni problem razviti takšen

sistem za priporočeno ceno na bolhi, nestrukturiranost podatkov. Uporabnik, ki objavi oglas, lahko v naziv in opis oglasa vpiše poljubno vsebino, hkrati pa sta to tudi edina podatka iz katerih lahko črpamo informacije in na podlagi teh informacij izločamo izdelke. Zato mislimo, da naš sistem dobro izloča izdelke, določi priporočeno ceno in ostale informacije, ki so v pomoč uporabniku.

Poglavje 6

Zaključek

Cilj tega diplomskega dela je bil razviti intuitivni sistem, ki bi uporabniku bistveno skrajšal čas raziskave ovrednotenja svojega izdelka in priporočil ceno za objavo oglasa na spletni tržnici bolha.com. Najprej smo si ogledali sorodne rešitve sistemov, ki so jih razvili za svetovne spletne tržnice, nato pa smo vam predstavili namen in idejo za razvoj našega priporočilnega sistema. Sledila je predstavitev implementacije in analiza natančnosti našega sistema. Osnovni cilj, pohitriti ovrednotenje izdelka, smo gotovo dosegli, saj se pri večini raziskav z našim sistemom izloči nerelevantne izdelke s samo tremi kliki. Prav tako pa so priporočene cene v večini primerov dokaj natančne in uporabniku skupaj s drugimi uporabnimi informacijami pomagajo pri končni odločitvi. Seveda naš sistem oz. prototip ni popoln. Programska koda bi seveda lahko bila bolj optimizirana. Za drugo strategijo (priporočena cena na podlagi novega izdelka) smo se omejili samo na izdelke, ki so dosegljivi na spletni trgovini mimovrste.com, kaj če iskanega izdelka ni v ponudbi? Seveda bi našli še kakšno izboljšavo, ampak cilje tega diplomskega dela smo dosegli in ob naslednji prodaji izdelka na bolhi bomo lahko uporabili naš sistem.

Literatura

- [1] TechCrunch, Thanks To PayPal And Marketplace Growth, eBay's Q1 2013 Revenue Up 14 Percent To \$3.7B, Net Income Up 19 Percent, <http://techcrunch.com/2013/04/17/thanks-to-paypal-and-marketplace-growth-ebays-q1-2013-revenue-up-14-percent-to-3-7b/>
- [2] Liran Einav, Chiara Farronato, Jonathan D. Levin, Neel Sundaresan »Sales mechanisms in online markets: What happend to internet auctions?«, 2013
- [3] Dennis van Heijst, Rob Potharst, Michiel van Wezel »A Support System for Predicting eBay End Prices«, 2006
- [4] Ilya Raykhel, Dan Ventura »Real-time Automatic Price Prediction for eBay Online Trading«, 2009
- [5] Hai YU, Shouyang WANG, Chuangyin DANG »Optimal starting price for ebay-like online auctions«, 2005
- [6] Razlaga metode Natančnosti in priklica - Precision and recall, https://en.wikipedia.org/wiki/Precision_and_recall
- [7] LESS, <http://lesscss.org/>

- [8] jQuery,
<http://jquery.com/>
- [9] PHP - PHP: Hypertext Preprocessor,
<http://php.net/>
- [10] bolha.com,
<http://www.bolha.com/>
- [11] Mimovrste,
<http://www.mimovrste.com/>
- [12] eBay,
<http://www.ebay.com/>
- [13] Gixen,
<http://www.gixen.com/>
- [14] TeraPeak for eBay,
<http://www.terapeak.com/products/terapeak-for-ebay/>
- [15] Cenik paketov v notranjem promet - Pošta Slovenije,
Dostopno na: <http://www.posta.si/downloadfile.aspx?fileid=21341>
- [16] PHP Simple HTML DOM Parser,
<http://simplehtmldom.sourceforge.net/>
- [17] Twiter Bootstrap,
<http://getbootstrap.com/>
- [18] Chart.js,
<http://www.chartjs.org/>