

MaxDiff TEST analiza (primjer)

Danijel Kopčinović, IT Market

MAXDIFF TEST

U upitniku "MAXDIFF TEST" postaviti ćemo Vam X pitanja s raznim varijantama, a Vi ćete u svakom pitanju odabrati Vama najbolju i najlošiju varijantu. Za ispunjavanje cijelog upitnika po našoj je procjeni dovoljno 15-20 minuta.

Kao znak zahvale za sudjelovanje u ovom upitniku, Y će Vam poslati mali poklon paket!

U ovom upitniku koristimo sljedeće osobne podatke:

- adresa e-pošte ispitanika
- IP adresa

Navedene osobne podatke koristit ćemo u svrhu identifikacije ispitanika i kreiranje profila ispitanika. Navedene podatke nećemo koristiti niti u koju drugu svrhu i nećemo ih dijeliti niti s kojom trećom stranom bez pismene privole ispitanika.

Više o politici zaštite osobnih podataka tvrtke Y pogledajte [ovdje](#).

Prihvaćam pravila o zaštiti osobnih podataka ☐

Uvod

U ovom dokumentu ćemo prezentirati testno istraživanje tržišta MaxDiff metodom. Istraživanje je bazirano na (pretpostavljenoj) želji lokalne pizzerije da ispita koju novu pizzu bi kupci najviše htjeli vidjeti u ponudi.

Nove pizze između kojih ispitanici biraju su:

```
## [1] miješana          povrtna          pikantna          4 vrste sira
## [5] s tunom           ribarska         4 godišnja doba calzone
## [9] slavonska        rukola/pršut     bolonjez          lovačka
## [13] losos            sa salamom
## 14 Levels: 4 godišnja doba 4 vrste sira bolonjez calzone losos ... slavonska
```

Osim glavnih proizvoda koji su nam u fokusu, zanima nas i mogući utjecaj “demografskih” faktora na rezultate:

```
## $lokacija
## $lokacija$naziv
## [1] "lokacija"
##
## $lokacija$vrijednosti
## [1] "blizu pizzerije"      "daleko od pizzerije"
```

U MaxDiff metodi ispitanici u svakom pitanju biraju najbolji i najlošiji proizvod između nekoliko alternativa. Pogledajmo neke kombinacije koje smo ispitivali, odn. postavili kao pitanja.

```
## task items best_choice worst_choice
## 1 1 4 vrste sira,slavonska,sa salamom ? ?
## 2 2 calzone,miješana,slavonska ? ?
## 3 3 4 godišnja doba,4 vrste sira,ribarska ? ?
## 4 4 bolonjez,calzone,ribarska ? ?
## 5 5 slavonska,ribarska,losos ? ?
```

Iz skupa svih 364 kombinacija proizvoda s 3 alternativa (broj alternativa u pitanju smo postavili kao parametar dizajna), odabrali smo 54 reprezentativnih. S odabranim kombinacijama moguće je izračunati (prosječnu ili

individualnu) vrijednost svakog proizvoda za ispitanike.

Sve kombinacije podijelili smo na 3 upitnika, od kojih svaki ima 6 pitanja, od kojih svako ima 3 opcija/alternativa.

Molimo odaberite Vama najbolju i najlošiju opciju među navedenim opcijama

★Opcije:

NAJBOLJA		NAJLOŠIJA
<input checked="" type="radio"/>	4 vrste sira	<input type="radio"/>
<input type="radio"/>	lovačka	<input checked="" type="radio"/>
<input type="radio"/>	calzone	<input type="radio"/>

🔗 Kliknite na kružić (radio dugme) za odabir

Da bismo bolje mogli procijeniti vrijednosti različitih opcija, dodali smo i pitanje u kojem ispitanik iskazuje općenitu pozitivnu ili negativnu tendenciju prema nekoliko odabranih proizvoda. Proizvodi u tom pitanju su odabrani s obzirom na prijašnje odabire, tako da sadrže raspon pizza - od pizza najčešće biranih kao najbolje do pizza najčešće biranih kao najlošije. U ovom pitanju se ne uspoređuju proizvodi jedan s drugim nego svaki u odnosu na tzv. graničnu opciju, koja u različitim kontekstima može značiti različite stvari, ali općenito znači točku između želim/ne želim, kupio bih/ne bih kupio, zanima me/ne zanima me i sl. Ovo pitanje nam omogućuje da bolje uskladimo odgovore svih ispitanika u odnosu na graničnu opciju pa time i međusobno, čime dobivamo relevantnije procjene vrijednosti.

★Kad biste imali mogućnost kupovine ovih opcija, da li biste ih kupili ili ne?

	KUPIO BIH	NE BIH KUPIO
miješana	<input checked="" type="radio"/>	<input type="radio"/>
lovačka	<input type="radio"/>	<input checked="" type="radio"/>
4 vrste sira	<input type="radio"/>	<input checked="" type="radio"/>
povrtna	<input checked="" type="radio"/>	<input type="radio"/>
calzone	<input type="radio"/>	<input checked="" type="radio"/>

🔗 Kliknite na kružić (radio dugme) za odabir

Rezultate upitnika smo generirali računalom, s određenim parametrima slučajnosti, za 200 ispitanika.

Jednostavna analiza - prebrojavanje

Postoji više načina na koje možemo analizirati podatke MaxDiff metodom. Najjednostavniji pristup je prebrajanje koliko je puta neki proizvod odabran kao najbolji, odn. najlošiji. Pogledajmo proizvode iz našeg istraživanja po broju odabira najbolji/najlošiji:

alt	chosen_as_best	chosen_as_worst
4 godišnja doba	161	20
4 vrste sira	164	36
bolonjez	7	175
calzone	18	148
losos	130	43
lovačka	24	154
miješana	70	19
pikantna	229	3
povrtna	115	61
ribarska	129	47
rukola/pršut	7	215
s tunom	52	68
sa salamom	56	59
slavonska	38	152

Možemo vidjeti da je pikantna najviše puta odabrano kao najbolja opcija, a rukola/pršut najviše puta kao najlošija opcija.

Analiza multinomnim logit modelom

Naprednija analiza uključuje modeliranje podataka (odgovora ispitanika) i zatim izračunavanje koeficijenata (“vrijednosti”) za svaki proizvod multinomnim logit modelom. Kad izvršimo analizu, dobijemo sljedeći sažetak:

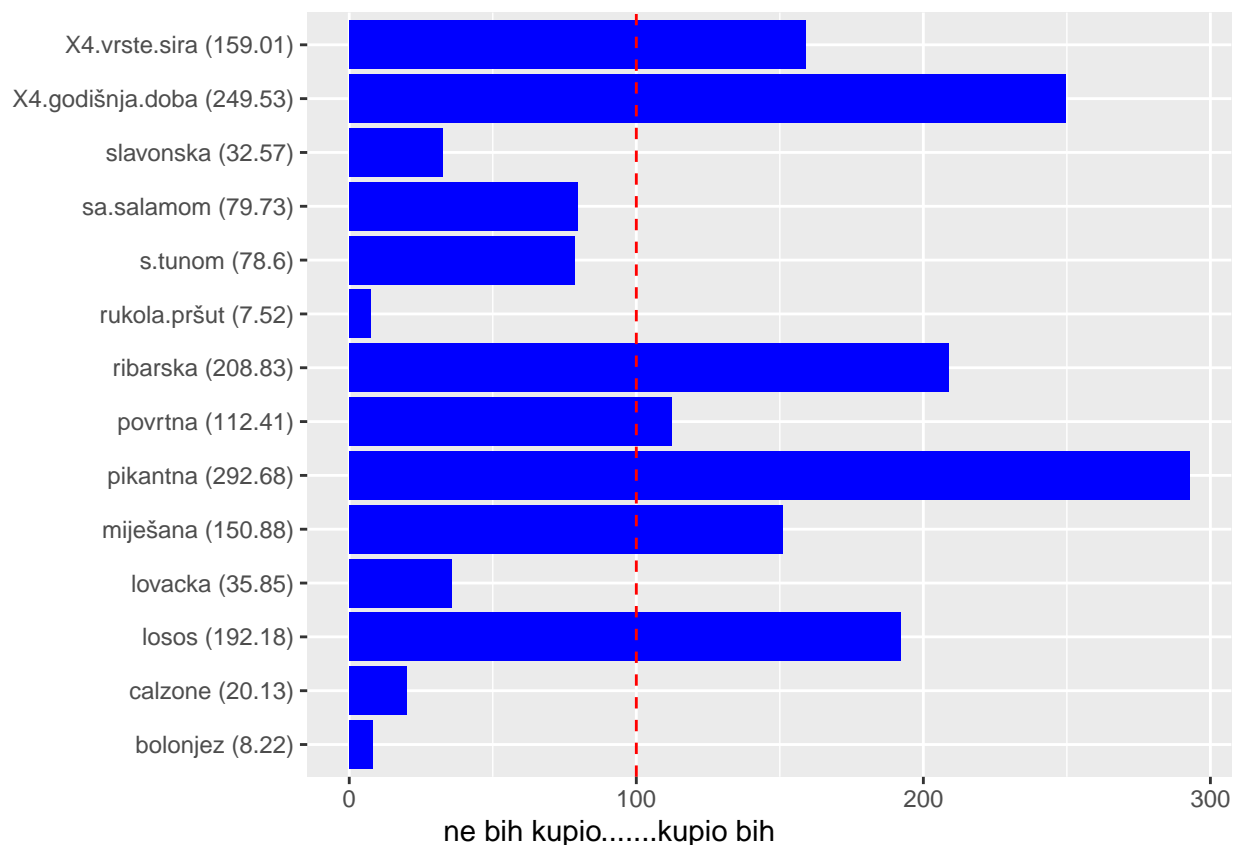
##	1st Qu.	Mean	3rd Qu.
## miješana	-0.888382307	0.2349791	1.3505137
## povrtna	-1.180309157	-0.2887543	0.6152998
## pikantna	2.700681397	3.9116654	5.0833724
## X4.vrste.sira	-0.599027521	0.3436022	1.2806869
## s.tunom	-1.906194196	-0.8122667	0.2935417
## ribarska	-0.009438742	1.0520189	2.1115724
## X4.godišnja.doba	0.596211851	1.8214363	3.0371697
## calzone	-3.448424673	-2.4087209	-1.3417371
## slavonska	-2.991635218	-1.8822479	-0.7353968
## rukola.pršut	-4.475042336	-3.4380242	-2.3552538
## bolonjez	-4.468180912	-3.3461902	-2.1785240
## lovačka	-2.751588330	-1.7738617	-0.7688110
## losos	-0.206722580	0.8012456	1.8071983
## sa.salamom	-1.744238546	-0.7930006	0.1748956

U prikazanom sažetku je dosta informacija koje nije tako lako direktno interpretirati. Ali da malo pojasnimo, evo nekoliko pojašnjenja koja slijede direktno iz sažetka:

- ‘calzone’ ima negativni (prosječni/mean) koeficijent -2.409, što znači da ga ispitanici slabije vrednuju nego osnovnu graničnu opciju koja označava točku kupio bih/ne bih kupio (usporedba koeficijenta za neku vrijednost je uvijek u odnosu na prvu/osnovnu deklariranu vrijednost, što je u ovom slučaju granična opcija) - jednostavnije rečeno ‘calzone’ ispitanici ne bi kupili, a zbog relativno velikog negativnog koeficijenta, ta je odluka prilično snažna,
- ‘pikantna’ i ‘ribarska’ obje imaju pozitivne (prosječne/mean) koeficijente (3.912, 1.052) što znači da ih ispitanici obje vrednuju više od osnovne granične opcije, ali ‘pikantna’ vrednuju više nego ‘ribarska’ - jednostavnije rečeno ispitanici bi kupili obje, ali ‘pikantna’ radije nego ‘ribarska’,

- prvi i treći kvantili označavaju raspon u kojem se (malo pojednostavljeno) nalazi svaki koeficijent sa 75% vjerojatnosti.

Ako sve koeficijente normaliziramo i skaliramo tako da granična opcija (kupio bih ili ne) bude na broju 100, lakše je interpretirati njihove vrijednosti i međusobne omjere:



Sad možemo vidjeti da je npr. najvredniji proizvod pikantna i da mu je vrijednost 1.84 puta veća od vrijednosti X4.vrste.sira.

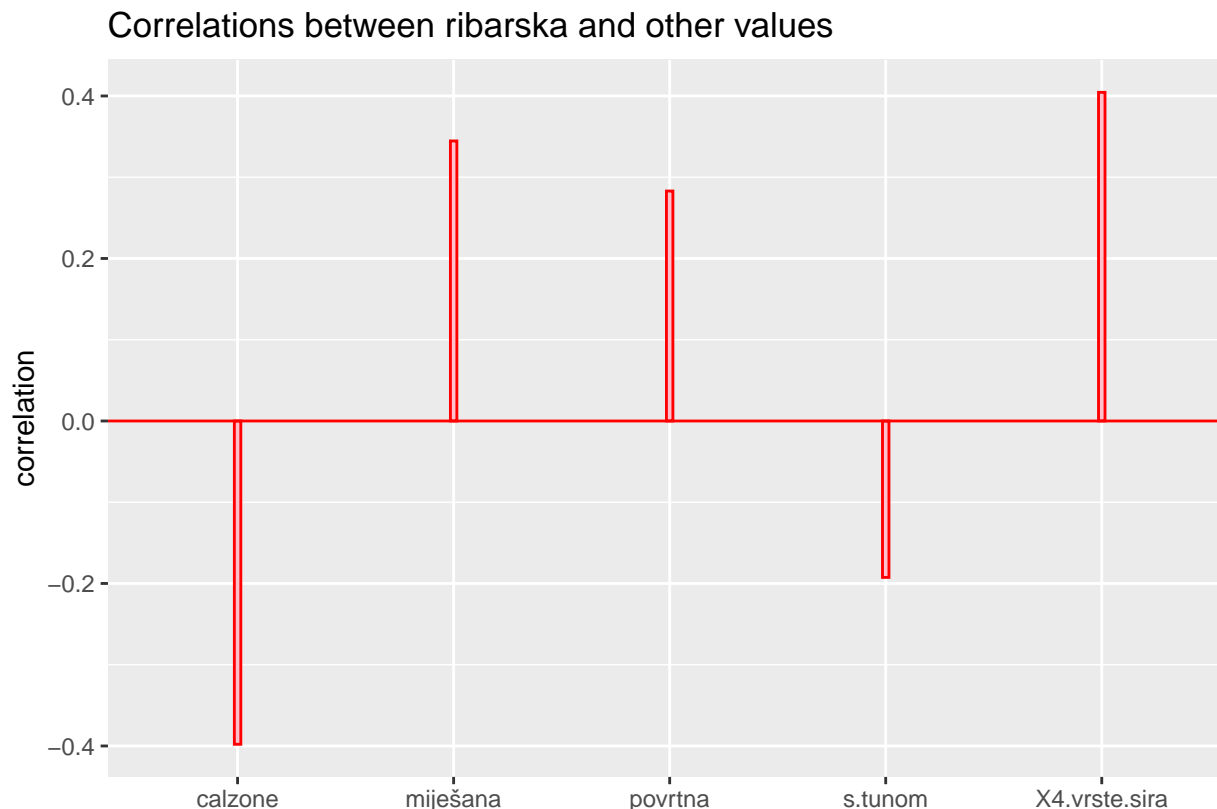
Ovi rezultati pokazuju i glavne prednosti MaxDiff analize u odnosu na jednostavno “redanje” proizvoda od najboljeg prema najlošijem koje se ponekad (prečesto :)) koristi za poslovne analize. MaxDiff nam pokazuje:

- *odnose među proizvodima* - koliko je koji vredniji ili manje vrijedan za ispitanike,
- *pozicije proizvoda u odnosu na (ključnu!) točku koja označava pozitivan ili negativan stav* - kupio bih ili ne, zanima me ili ne...

Time dobivamo kompletnu sliku svih proizvoda iz perspektive ispitanika i možemo donositi kvalitetne poslovne odluke.

Analiza koreliranosti proizvoda

Drugi element koji možemo analizirati je korelacija između vrijednosti pojedinih proizvoda.



Primjerice, u gore prikazanom grafu možemo vidjeti da je *ribarska* pozitivno korelirana s *miješana*, *povrtna*, *X4.vrste.sira*, a negativno korelirana s *s.tunom*, *calzone*. To znači da ako ispitanici daju veću vrijednost pizzi *ribarska*, onda daju veću vrijednost i *miješana*, *povrtna*, *X4.vrste.sira*, dok za *s.tunom*, *calzone* daju manju vrijednost.

Rezultate analize i procjene koeficijenata vrijednosti opcija najlakše je razumjeti preko “predviđanja udjela” (share prediction) što ćemo sada prikazati.

Predviđanje udjela (share prediction)

Jednom kada znamo vrijednosti koeficijenata svih proizvoda, možemo usporediti njihove različite kombinacije i predvidjeti udjele u prodaji. Predviđeni udjeli u prodaji su vjerojatnosti da će se određeni proizvod prodati u usporedbi s ostalim proizvodima koji su u kombinaciji koju testiramo.

Kao primjer uzet ćemo kombinaciju nekoliko slučajno odabranih proizvoda i izračunati njihove predviđene udjele u prodaji:

##	[,1]	[,2]	[,3]	[,4]	[,5]
## opcije	"povrtna"	"4 vrste sira"	"miješana"	"s tunom"	"pikantna"
## share %	"12.02"	"14.00"	"12.43"	"16.02"	"45.54"

U prethodnoj tablici možemo vidjeti kako bi ispitanici birali između 5 različitih opcija pizza. Svaka pizza dobiva predviđeni udio (postotak) u ukupnoj prodaji, prema procijenjenim vrijednostima pripadajućih koeficijenata.

Najčešće birana opcija (između navedenih, ne općenito) je *pikantna* s **predviđenim udjelom u prodaji 45.54%**.

Provjera modela na postojećim podacima

Korištenjem predviđanja udjela/vjerojatnosti prodaje za razne kombinacije proizvoda, možemo testirati i kako bi naš model predvidio izbor za (neke) kombinacije koje su dane u istraživanju:

```
##                      choice
## respondent/question true best predicted best true worst predicted worst
##          178/5          3          3          1          1
##          23/3          2          3          1          1
##          44/3          3          3          1          1
##          191/5         2          2          3          3
##          43/3          3          3          1          1
##          123/6         3          3          1          1
##          139/3         1          1          2          2
##          23/5          3          3          1          1
##          42/5          2          2          3          3
##          184/3         3          3          1          1
##          46/5          2          2          1          3
##          159/6         2          2          1          3
##          155/3         2          3          1          1
##          31/3          2          2          1          1
##          98/3          3          3          2          2
##          139/5         2          2          1          1
##          167/3         3          3          1          1
##          180/5         2          2          3          3
##          18/6          1          2          3          3
##          180/3         2          3          1          1

## [1] "Ukupno točno predviđanja:86.67%"
```

Ukupna točnost predviđanja daje nam sigurnost da naš model dobro funkcionira.

Segmentacija ispitanika

S obzirom na razne faktore (demografske ili sl.), ispitanike možemo podijeliti u segmente te zatim modelirati vrednovanje proizvoda za svaki segment (umjesto prosjeka ili individualno). Broj segmenata koji mislimo da postoji u skupu ispitanika treba zadati unaprijed kao parametar. Pretpostavimo da postoje 3 različita segmenta ispitanika (s tim parametrom ćemo napraviti analizu). Modeliranjem (multinomni logit model s LCA - latent class analysis) dobivamo sljedeći sažetak:

```
##
##          segment_1  segment_2  segment_3
## miješana          0.1148834  0.09101120 -0.3130916
## povrtna          -0.1178750 -0.28466952 -0.3599366
## pikantna          2.4660650  1.47187209  1.4795165
## X4.vrste.sira      0.2263574  0.06130955 -0.3630211
## s.tunom           -0.4368493  0.57051495  0.5078137
## ribarska          0.6289285  0.60133842  0.6285399
## X4.godišnja.doba   1.0681058  1.06660231  1.0389167
## calzone           -1.5812423 -1.12379359 -0.7151136
## slavonska         -1.1868617 -0.72700958 -1.0868358
## rukola.pršut      -2.3095288 -1.76848126 -1.3268809
## bolonjez          -2.1264829 -1.47366018 -2.2016732
## lovačka           -1.0616948 -0.99635986 -0.8848496
## losos             0.4854090  0.56463342  0.5819652
```

```
## sa.salamom -0.6502203 -0.02659289 -0.5247803
```

Vidimo da se segmenti razlikuju (za više od 0.5) u vrednovanju npr. *pikantna*, *s.tunom*, *rukola.pršut*, *bolonjez*, *sa.salamom*.

S obzirom na to da je cilj segmentacije i prepoznavanje koji korisnik (s najvećom vjerojatnošću) pripada kojem segmentu, pogledajmo i taj rezultat (navedeni su redni brojevi ispitanika koji pripadaju određenom segmentu):

```
## [1] "Segment 1"
## [1] 1 4 5 6 10 12 15 20 23 24 25 26 27 36 37 38 39 40
## [19] 41 42 44 46 48 50 51 52 53 55 57 58 59 62 67 68 69 70
## [37] 72 73 77 79 80 82 83 84 90 93 95 96 99 100 101 102 105 109
## [55] 110 111 116 117 119 124 126 128 129 130 131 135 137 138 139 140 142 145
## [73] 147 148 149 150 151 153 155 156 157 158 160 162 163 165 169 172 173 174
## [91] 178 179 181 182 183 185 186 189 190 191 194 196 197 199
## [1] "Segment 2"
## [1] 2 3 7 8 9 11 13 14 16 17 18 19 21 22 28 29 30 31 32
## [20] 33 34 35 43 45 47 49 54 56 60 61 63 64 65 66 71 74 75 76
## [39] 78 81 85 86 87 88 89 91 92 94 97 98 103 104 106 107 108 112 113
## [58] 114 115 118 120 121 122 123 125 127 132 133 134 136 141 143 144 146 152 154
## [77] 159 161 164 166 167 168 170 171 175 176 177 180 184 187 188 192 193 195 198
## [96] 200
## [1] "Segment 3"
## integer(0)
```

Kako je jedan segment prazan (niti jedan ispitanik mu ne pripada), zaključujemo da je broj segmenata koje s danim podacima i modelom možemo identificirati jednak 2.

S dobivenom segmentacijom možemo primjerice kreirati više različitih proizvoda od kojih će svaki biti najzanimljiviji određenom segmentu (s obzirom na vrednovanje). Tako ćemo povećati prodaju i zadovoljstvo kupaca.

TURF analiza

Ponudom u kojoj se nalazi više proizvoda povećat ćemo vjerojatnost da će barem neki proizvod biti kupljen od strane kupaca. Da bismo izračunali koji proizvodi bi trebali biti u ponudi da bismo postigli najveću prodaju, koristimo TURF (total unduplicated reach and frequency) analizu.

Pretpostavimo da u ponudu želimo staviti 3 proizvoda. Za svaku kombinaciju/ponudu koja sadrži 3 proizvoda i za svakog ispitanika iz našeg istraživanja izračunat ćemo da li bi kupio barem jedan proizvod iz ponude (reach - da li je kupac 'dosegnut' barem nekim proizvodom) i koliko (ako i jedan) bi proizvoda iz ponude kupio (frequency - koliko je puta kupac 'dosegnut' danom ponudom proizvoda). Odluka da li bi kupac kupio neki proizvod ili ne se donosi na osnovi vjerojatnosti prodaje proizvoda u usporedbi s ostalim proizvodima iz ponude i osnovnom (graničnom) opcijom koja označava granicu kupio bih/ne bih kupio.

Pogledajmo kako izgleda nekoliko redaka iz reach & frequency tablice za naše proizvode i ponude (poredanih silazno po reach):

```
## # A tibble: 6 x 3
## combination reach freq
## <chr> <dbl> <dbl>
## 1 pikantna, rukola/pršut, bolonjez 200 200
## 2 pikantna, calzone, slavonska 197 199
## 3 pikantna, slavonska, lovačka 196 197
## 4 pikantna, slavonska, rukola/pršut 196 197
## 5 pikantna, lovačka, sa salamom 195 200
```

6 pikantna, s tunom, slavonska 195 199

Možemo vidjeti da bi ponuda pikantna, rukola/pršut, bolonjez dosegla najviše ispitanika (kupili bi barem jedan proizvod iz ponude), njih 200, što je 100% od ukupnog broja ispitanika, s time što bi ukupno bilo 200 (potencijalnih) kupovina.

Na ovaj način možemo kreirati ponudu proizvoda kojom ćemo maksimizirati prodaju.

Zaključak

MaxDiff analiza je moćan alat za upoznavanje kupaca i njihovih vrijednosti i za kreiranje najboljih proizvoda za njih. Takav pristup uvijek vodi prema poboljšanju poslovanja, većim prihodima i većem zadovoljstvu kupaca.

Za više informacija kontaktirajte **danijel.kopcinovic@itmarket.hr**.