

**ds**

**dnevi  
slovenske  
informatike**

**9. in 10. maj 2023**

Kongresni center Bernardin, Portorož

**30. konferenca Dnevi slovenske informatike**  
**Soustvarjamo digitrajno Slovenijo**

# **NAPOVEDOVANJE PLAČILNE (NE)DISCIPLINE Z ORODJI UMETNE INTELIGENCE**

Špela Ognjanovič, Tomaž Dular, Julija Lapuh Bele

9. 5. 2023

# Priprava podatkov

## CILJ: Napovedati denarne prilive podjetja

V začetno množico podatkov vzamemo vse zaprte račune od vključno leta 2016 do vključno leta 2021.

### Na nivoju računa obdržimo:

- Znesek računa,
- plačilni pogoji,
- kategorija računa.

Tabela 2: Razredi zamude.

Razred zamude	Opis
a	Brez zamude ali predplačila
b	Od 1 do vključno 5 dni zamude
c	Od 6 do vključno 15 dni zamude
d	Od 16 do vključno 30 dni zamude
e	Od 31 do vključno 60 dni zamude
f	Od 61 do vključno 90 dni zamude
g	Od 91 do vključno 120 dni zamude
h	Nad 120 dni zamude

Zgodovina poslovanja kupca je predstavljena kot lastnost računa.

Tabela 1: Osnovne in izračunane vrednosti računov.

Št.	Ime polja	Opis
1.	Znesek računa	Znesek izdanega računa.
2.	Dni za plačilo	Razlika med datumom dokumenta in datumom valute.
3.	Kategorija računa	Vrsta, tip računa.
4.	Število plačanih računov	Število plačanih računov pred računom, na katerega se nanaša zapis.
5.	Število izdanih računov, plačanih z zamudo	Število izdanih in z zamudo plačanih računov pred računom, na katerega se nanaša zapis.
6.	Znesek plačanih računov	Znesek plačanih računov pred računom, na katerega se nanaša zapis.
7.	Znesek izdanih računov, plačanih z zamudo	Znesek izdanih in z zamudo plačanih računov pred računom, na katerega se nanaša zapis.
8.	Povprečno število dni zamude	Povprečno število dni zamude z zamudo plačanih računov, pred računom, na katerega se nanaša zapis.
9.	Število odprtih računov	Število izdanih, neporavnanih računov pred računom, na katerega se nanaša zapis.
10.	Število zapadlih računov	Število izdanih, neporavnanih računov pred računom, na katerega se nanaša zapis in so že v zamudi.
11.	Znesek odprtih računov	Znesek izdanih, neporavnanih računov pred računom, na katerega se nanaša zapis.
12.	Znesek zapadlih računov	Znesek izdanih, neporavnanih računov pred računom, na katerega se nanaša zapis in so že v zamudi.
13.	Povprečno število dni zamude zapadlih računov	Povprečno število dni zamude izdanih, neporavnanih računov, pred računom, na katerega se nanaša zapis.
14.	Delež z zamudo plačanih računov	Vrstica 5 / vrstica 4.
15.	Delež zneska z zamudo plačanih računov	Vrstica 7 / vrstica 6.
16.	Delež zapadlih računov	Vrstica 10 / vrstica 9.
17.	Delež zneska zapadlih računov	Vrstica 12 / vrstica 11.

# Uporabljena programska oprema

## Microsoft Azure Machine Learning Studio

Omogoča pripravo in izvajanje poskusov avtomatiziranega strojnega učenja ter nadziranja njegovega okolja.

### Prednosti uporabe:

- Hitri rezultati (učenje več algoritmov hkrati),
- pregled podatkov (uravnoteženost, odvisnost, prazne vrednosti),
- ocena modela.

amiable\_grass\_htgqh3q7 Completed

Overview Data guardrails **Models** Outputs + logs Child jobs

Refresh Edit and submit (preview) Register model Cancel Delete | Deploy Download Explain model # View generated code ...

Created on All filters Clear all

Showing 1-10 of 60 models Page size:

Algorithm name	Explained	Accur... ↓	Sampling	Crea...	Dur...	Hyperparameter
<a href="#">StandardScalerWrapper, XGBoostClassifier</a>	<a href="#">View explanation</a>	0.71071	100.00 %	Apr ...	4s	booster : gbtree colsample_bylevel : 0.9 colsample_bytree : 1 eta : 0.4
<a href="#">StandardScalerWrapper, XGBoostClassifier</a>		0.70585	100.00 %	Apr ...	3m 59s	booster : gbtree colsample_bytree : 1 eta : 0.05 gamma max_depth
<a href="#">StandardScalerWrapper, XGBoostClassifier</a>		0.69929	100.00 %	Apr ...	1m 33s	booster : gbtree colsample_bytree : 1 eta : 0.1 gamma : 0.1 max_de
<a href="#">StandardScalerWrapper, XGBoostClassifier</a>		0.69890	100.00 %	Apr ...	4s	booster : gbtree colsample_bytree : 0.9 eta : 0.3 gamma max_depth
<a href="#">StandardScalerWrapper, LightGBM</a>		0.67292	100.00 %	Apr ...	4s	boosting_type : gbdtd colsample_bytree : 0.99 learning_rate : 0.08947473
<a href="#">SparseNormalizer, XGBoostClassifier</a>		0.66308	100.00 %	Apr ...	4s	booster : gbtree colsample_bylevel : 0.5 colsample_bytree : 1 eta : 0.3
<a href="#">MaxAbsScaler, XGBoostClassifier</a>		0.63369	100.00 %	Apr ...	35s	tree_method : auto
<a href="#">StandardScalerWrapper, XGBoostClassifier</a>		0.62949	100.00 %	Apr ...	38s	booster : gbtree colsample_bytree : 0.9 eta : 0.01 gamma max_de
<a href="#">StandardScalerWrapper, XGBoostClassifier</a>		0.62162	100.00 %	Apr ...	19s	booster : gbtree colsample_bytree : 0.5 eta : 0.3 gamma max_depth

Page 1 of 6

# Ocena modela

## Matrika razvrščanja

- Ponazarja podrobnejši opis rezultatov postopka razvrščanja.
- Rezultat razvrščanja v primeru 2 razredov razvrščanja (slika)
- Natančnosti =  $(RP + RN) / (RP + LN + LP + RN)$ .

		Predviden razred	
		Da	Ne
Dejanski razred	Da	Resnično Pozitivno	Lažno Negativno
	Ne	Lažno Pozitivno	Resnično Negativno

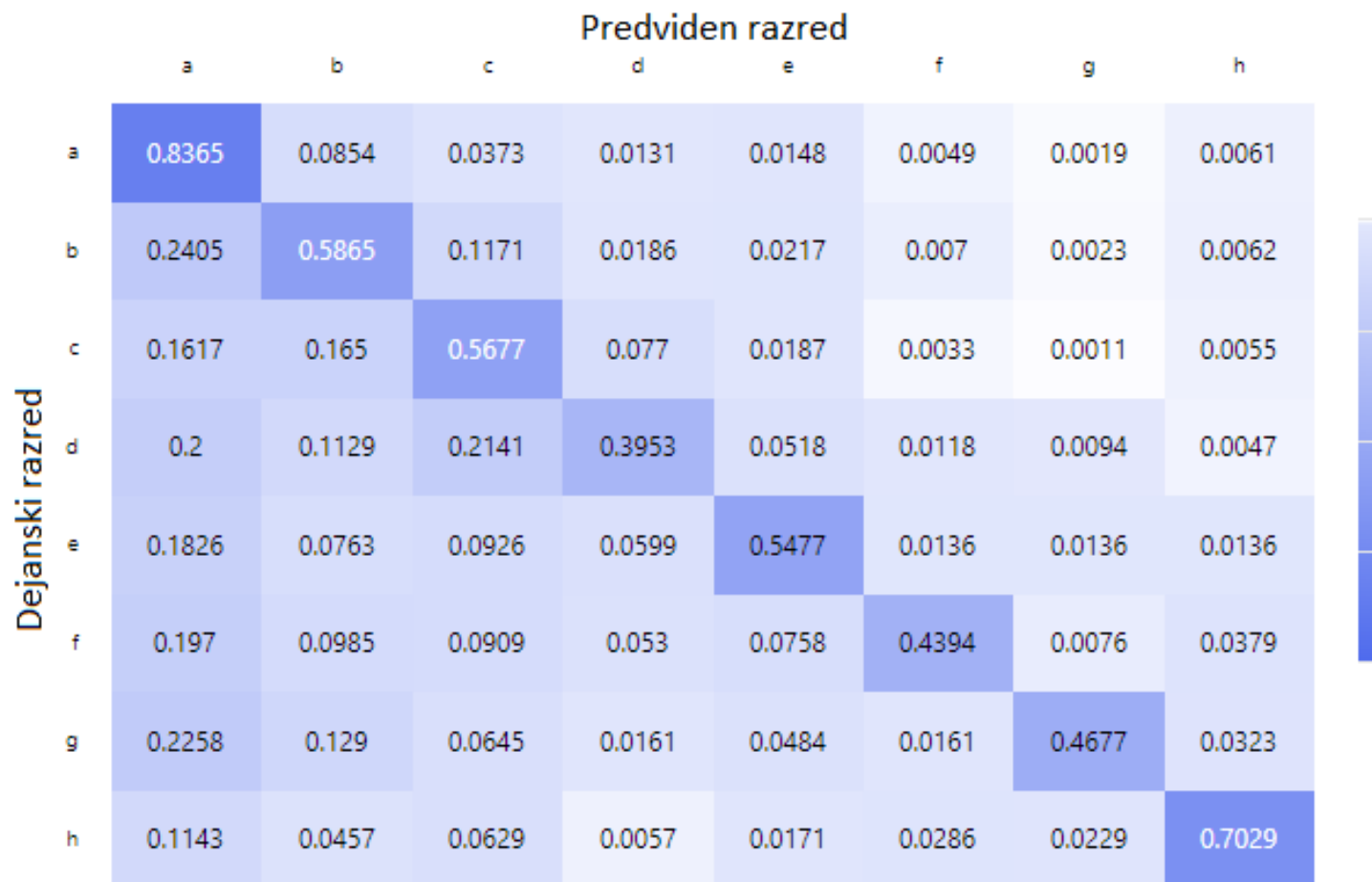
# Ocena modela

## Normalizirana matrika razvrščanja modela

Natančnost modela = 71 %

Tabela 2: Razredi zamude.

Razred zamude	Opis
a	Brez zamude ali predplačila
b	Od 1 do vključno 5 dni zamude
c	Od 6 do vključno 15 dni zamude
d	Od 16 do vključno 30 dni zamude
e	Od 31 do vključno 60 dni zamude
f	Od 61 do vključno 90 dni zamude
g	Od 91 do vključno 120 dni zamude
h	Nad 120 dni zamude



# Testni podatki / sodobni podatki

Natančnost modela (testni podatki) = 71 %

Dejanski razred	Predviden razred							
	a	b	c	d	e	f	g	h
a	0.8365	0.0854	0.0373	0.0131	0.0148	0.0049	0.0019	0.0061
b	0.2405	0.5865	0.1171	0.0186	0.0217	0.007	0.0023	0.0062
c	0.1617	0.165	0.5677	0.077	0.0187	0.0033	0.0011	0.0055
d	0.2	0.1129	0.2141	0.3953	0.0518	0.0118	0.0094	0.0047
e	0.1826	0.0763	0.0926	0.0599	0.5477	0.0136	0.0136	0.0136
f	0.197	0.0985	0.0909	0.053	0.0758	0.4394	0.0076	0.0379
g	0.2258	0.129	0.0645	0.0161	0.0484	0.0161	0.4677	0.0323
h	0.1143	0.0457	0.0629	0.0057	0.0171	0.0286	0.0229	0.7029

Natančnost modela (sodobni podatki) = 61 %

Dejanski razred	Predviden razred							
	a	b	c	d	e	f	g	h
a	0.8776	0.0787	0.0292	0.0117	0	0	0	0.0029
b	0.4681	0.383	0.1064	0.0426	0	0	0	0
c	0.6	0.075	0.175	0.125	0.025	0	0	0
d	0.4444	0.1556	0.2	0.0889	0.0889	0.0222	0	0
e	0.6739	0.0217	0.1087	0.1087	0.0652	0	0	0.0217
f	0.6	0.1	0	0.2	0.1	0	0	0
g	0.8	0	0.2	0	0	0	0	0
h	0.7778	0	0.1111	0	0	0	0.1111	0

## Diskusija