



Kompetenčni
center za
umetno
inteligenco

Slovenian
National
AI Competence
Centre



Uvod v strojno učenje: Kako iz podatkov nastanejo napovedi

Miran Meža, Marko Meža, Andraž Sivec
Delavnica se izvaja v prostoru LUCAMI (klet nove stavbe)



REPUBLIKA SLOVENIJA
MINISTRSTVO ZA NOTRANJE ZADEVE
IN JAVNO UPRAVO



REPUBLIKA SLOVENIJA
MINISTRSTVO ZA IZOBRAŽEVANJE, ZNANOST
IN MLADINO



REPUBLIKA SLOVENIJA
MINISTRSTVO ZA GOSPODARSTVO
DELO IN ŠPORT



REPUBLIKA SLOVENIJA
MINISTRSTVO ZA OBRAMBO





— Temeljni koncepti strojnega učenja



UNIVERZA
V LJUBLJANI

FE

Fakulteta
za elektrotehniko



Povzetek

Uvod

- AI → Strojno učenje, Izrazi

Napovedovanje vrednosti

- Linearna regresija
- Multipla linearna regresija

Klasifikacija

- Logistična regresija
- Nevronske mreže

Nenadzorovano učenje

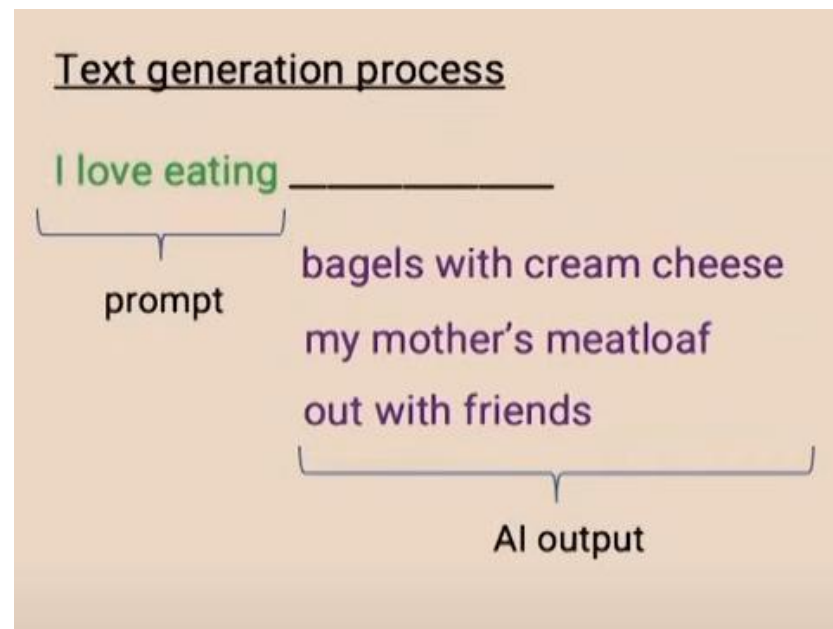
- Rojenje





Generativni modeli

Chat GPT, LLaVA,...
Klasifikacija, nadzorovano učenje
Predvidi naslednjo besedo v seriji



<https://www.youtube.com/watch?v=5p248yoa3oE>



Priporočilni sistem

PRIPOROČILA

Za vas

Najbolj gledano

Filmi

Šport

Nadaljevanke

Pompeji
Pompeii

Gladiatorju Milu so Rimljani pobili celotno pleme in družino, njega pa odpeljali v Pompeje, tik pred popolnim uničenjem mesta. Tega sta ob izbruhu zloglasnega Vezuva razdejala lava in pepel. Zaljubi se v Kasijo ter...

Cena: 3,5€ (24h) Ogled: sedaj

[OK Izposoja](#) [Podrobnosti](#) [Oceni](#) [Podaj mnenje](#) [Pomoč](#)

Storitev zagotavlja **Guru Cue**

TOR 15:28

[Pogačnik, Tasič., Meža, Košir. Personal content recommender based on a hierarchical user model for the selection of TV programmes. *User modeling and user-adapted interaction*]



UNIVERZA
V LJUBLJANI

FE

Fakulteta
za elektrotehniko



Klasifikacija autofloroscenčnih bronhoskopskih slik (maligno/vnetje)

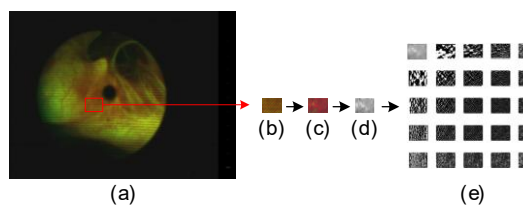
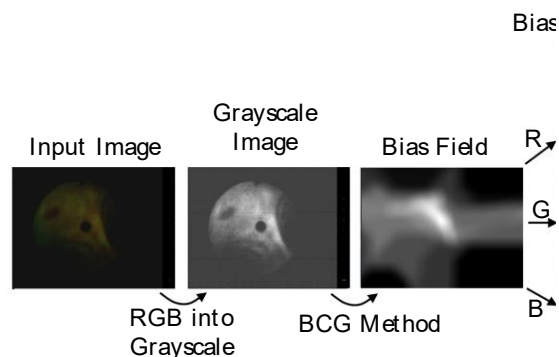
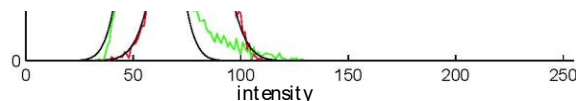


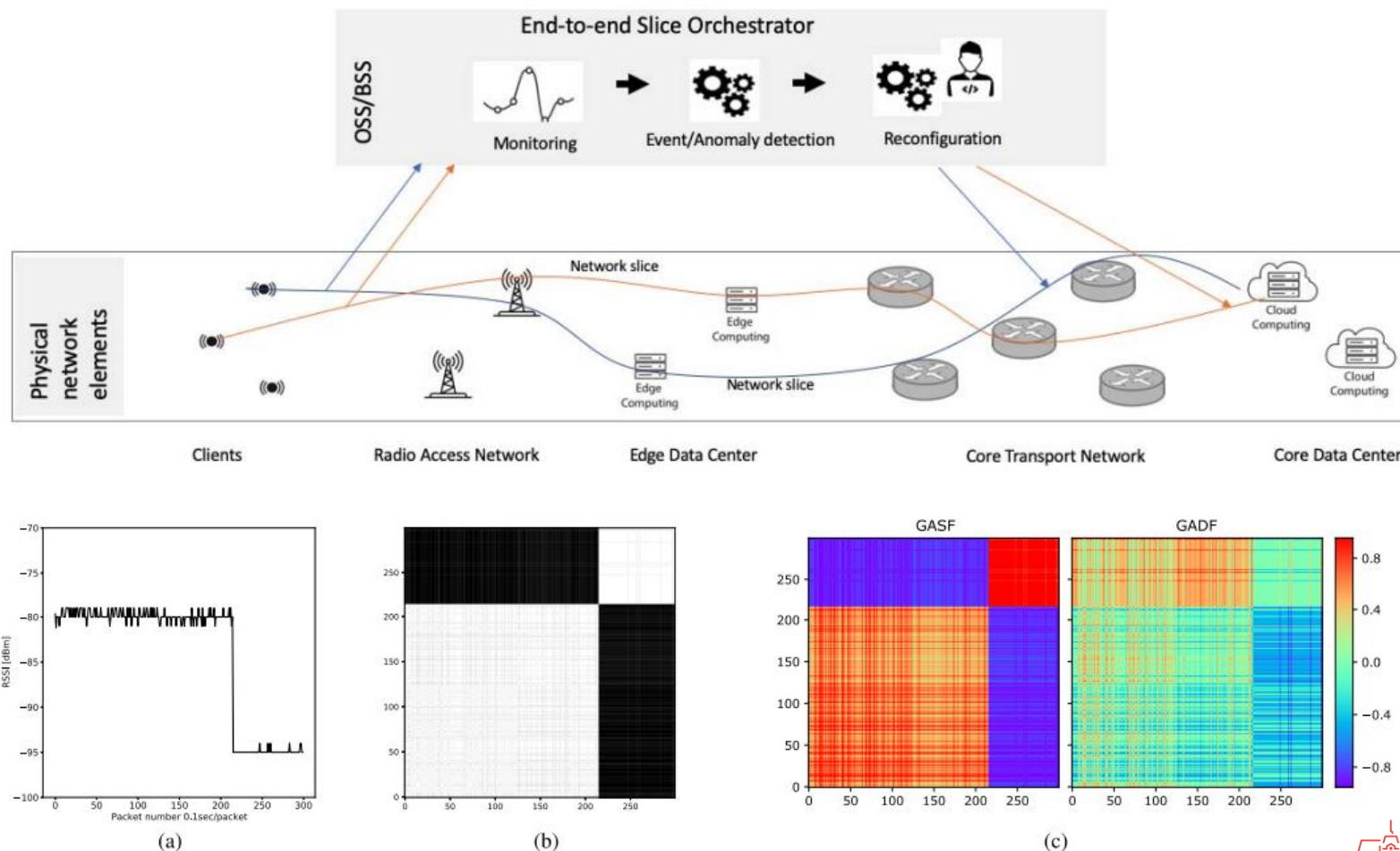
Table 3. Comparison of both methods

	<u>Bountris</u> method [18,19]	The new method
Classifier	Accuracy [%]	Accuracy [%]
k-NN	90,7	90,9
NB	90,4	91,5
SVM	92,1	95,8
Average accuracy	91,1	92,7





Zaznavanje anomalij v brezžičnih omrežjih



[Bertalanč, Meža, Fortuna. Resource-Aware Time Series Imaging Classification for Wireless Link Layer Anomalies. *IEEE TRANSACTIONS ON NEURAL NETWORKS AND*



UNIVERZA
V LJUBLJANI

Fakulteta
za elektrotehniko



Strojno učenje

Analiza podatkov v bazah (WEB click-stream, medicinski podatki, biologija, inženirski problemi)
Aplikacije, ki jih ni mogoče sprogramirati (prepoznavanje pisave, procesiranje naravnega jezika (NLP), računalniški vid)

Sistemi, ki se prilagajajo (priporočilni sistemi)

Razumevanje človeškega učenja (možgani, inteligenca)



UNIVERZA
V LJUBLJANI

FE

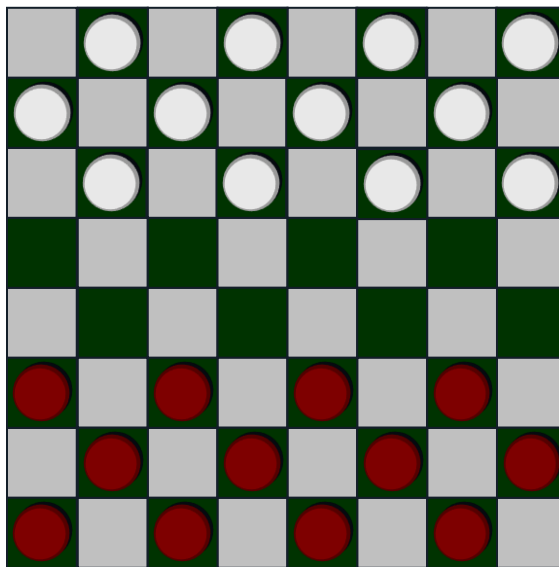
Fakulteta
za elektrotehniko



Definicija strojnega učenja

Arthur Samuel 1959

- Področje raziskav, ki strojem omogoča da se naučijo nalog, brez da jih neposredno programiramo
- Igranje dame



UNIVERZA
V LJUBLJANI

FE

Fakulteta
za elektrotehniko



Razdelitev algoritmov strojnega učenja

Nadzorovani (Supervised learning)

Nenadzorovani (Unsupervised learning)

Ostalo: Priporočilni sistemi, zaznavanje anomalij

Pomembno: Razumevanje kako uporabiti algoritme strojnega učenja



UNIVERZA
V LJUBLJANI

FE

Fakulteta
za elektrotehniko



Nadzorovano učenje



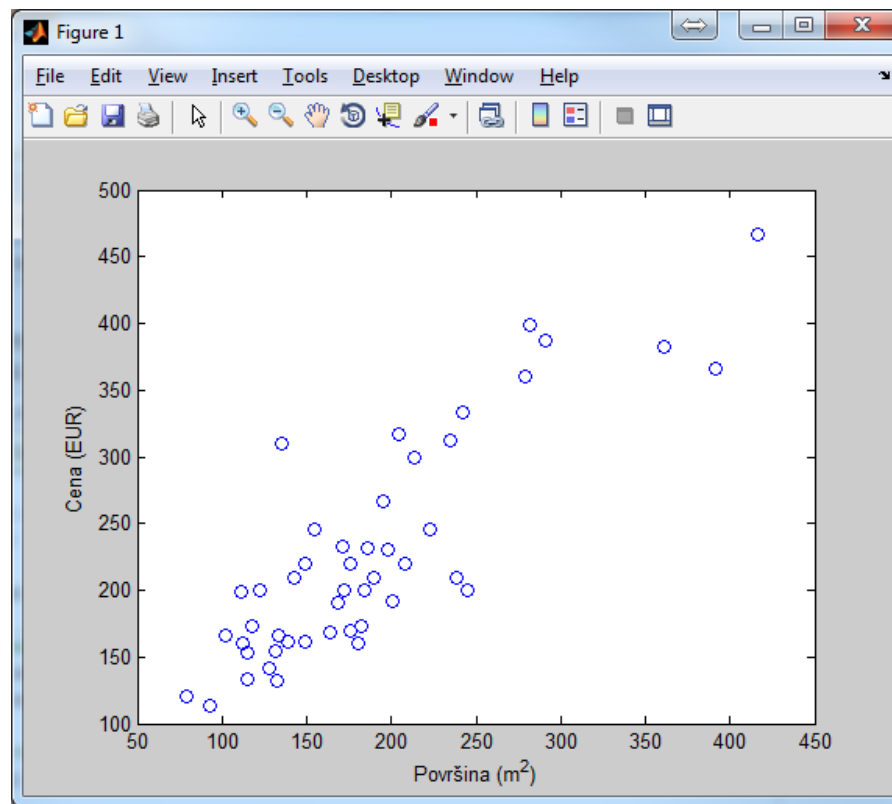
UNIVERZA
V LJUBLJANI

FE

Fakulteta
za elektrotehniko



Napoved cene hiš



Problem: Koliko stane 300m² veliko stanovanje?

Nadzorovano učenje – za podatke, ki so na voljo poznamo prave odgovore.

Regresija – Napovedovanje zvezne vrednosti rezultata (cene)



UNIVERZA
V LJUBLJANI

FE

Fakulteta
za elektrotehniko



Osnovni pojmi strojnega učenja

Primer: napoved cene hiše

Vzorec (sample)

En primer oziroma eno opazovanje v podatkovni zbirki.

Primer: ena hiša

Značilke (features)

Lastnosti vzorca, ki jih uporabimo kot vhod v model.

Primer: površina, št. sob, starost hiše

Ciljna spremenljivka (target / label)

Vrednost ali razred, ki ga želimo napovedati.

Primer: cena hiše

Model (model)

Iz podatkov naučena povezava med značilkami in ciljno spremenljivko.

Uporabimo ga za napoved pri novih vzorcih

površina • št. sob • starost → MODEL → napovedana cena

Pri nadzorovanem učenju v učnih podatkih poznamo ciljno spremenljivko.



UNIVERZA
V LJUBLJANI

FE

Fakulteta
za elektrotehniko



Tumor (maligen/benigen)



Problem: Ali je tumor dane velikosti benigni ali maligni?

Klasifikacija – Napovedovanje diskretne vrednosti (0 ali 1)

Nadzorovano učenje – za podatke, ki so na voljo poznamo prave odgovore.



UNIVERZA
V LJUBLJANI

FE

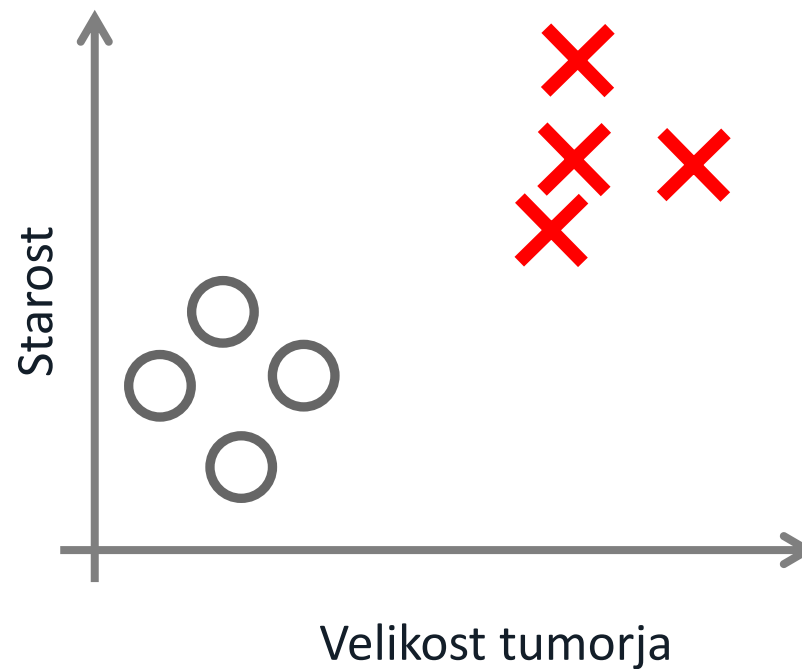
Fakulteta
za elektrotehniko



Tumor (maligen/benigen)

Druge značilke

- Uniformnost velikosti celic
- Uniformnost oblike skupkov celic
- ...





V podjetju želite uporabiti algoritme strojnega učenja za reševanje sledečih problemov:

- 1) Imate veliko skladišče identičnih izdelkov. Želite napovedati, koliko jih bo prodanih v naslednjih 3 mesecih
- 2) Za uporabniške račune želite ugotoviti ali je bil kompromitiran.
- 3) Kot zavarovalnica želite ugotoviti, ali je primer goljufija

Kakšen tip algoritma je potrebno uporabiti za posamezen problem (regresija ali klasifikacija)



Nenadzorovano učenje



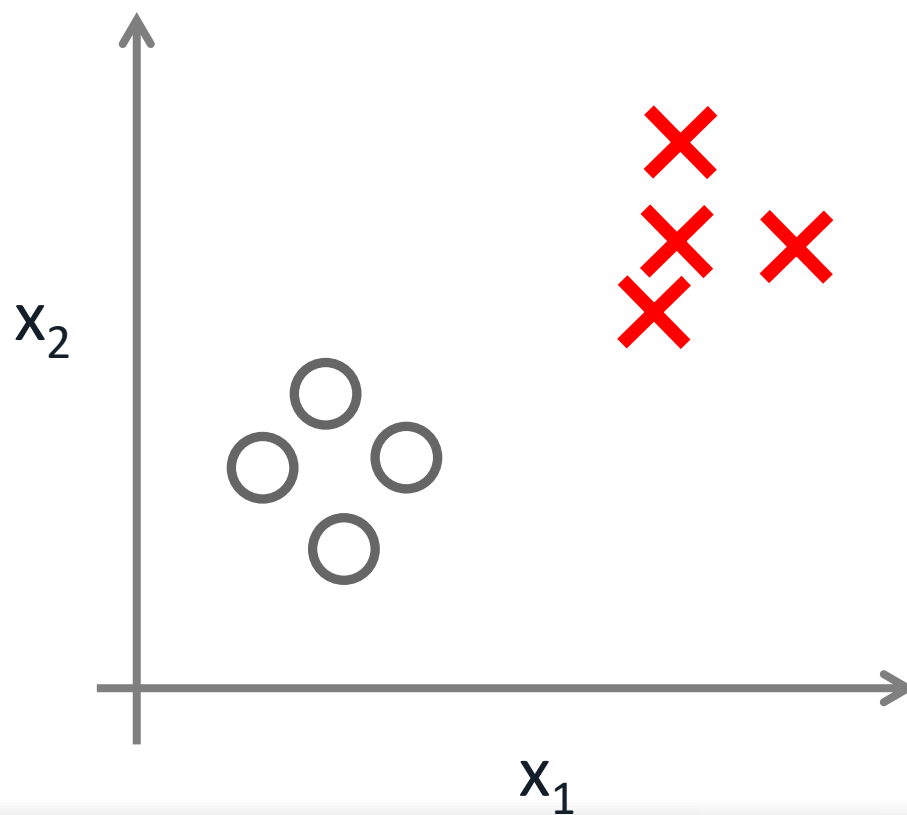
UNIVERZA
V LJUBLJANI

FE

Fakulteta
za elektrotehniko

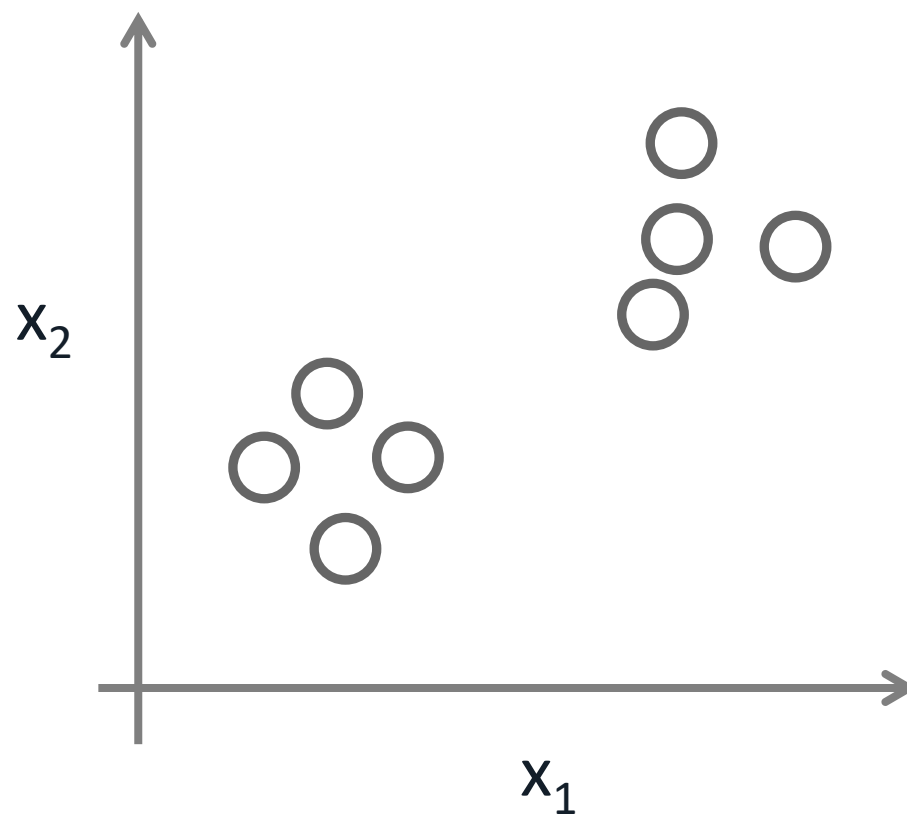


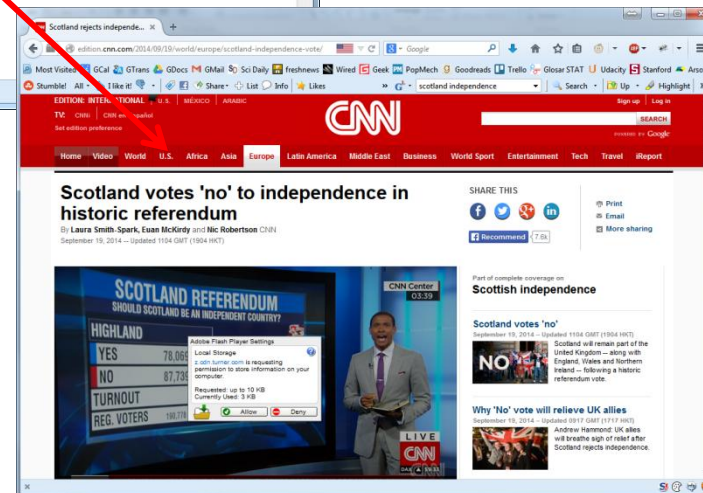
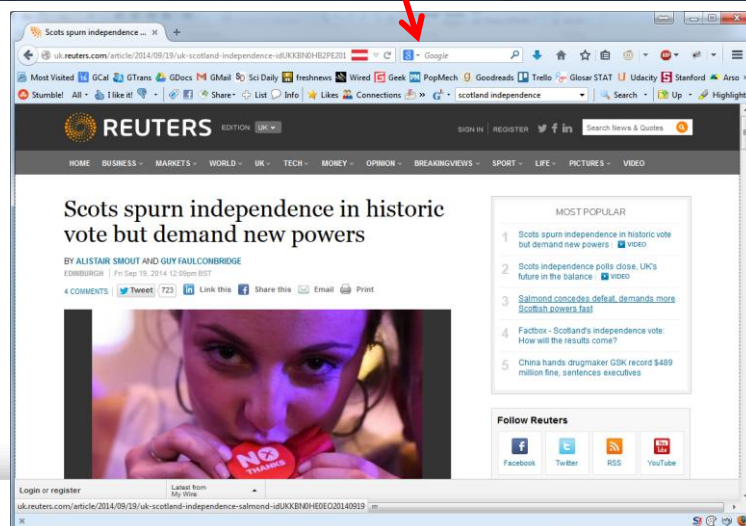
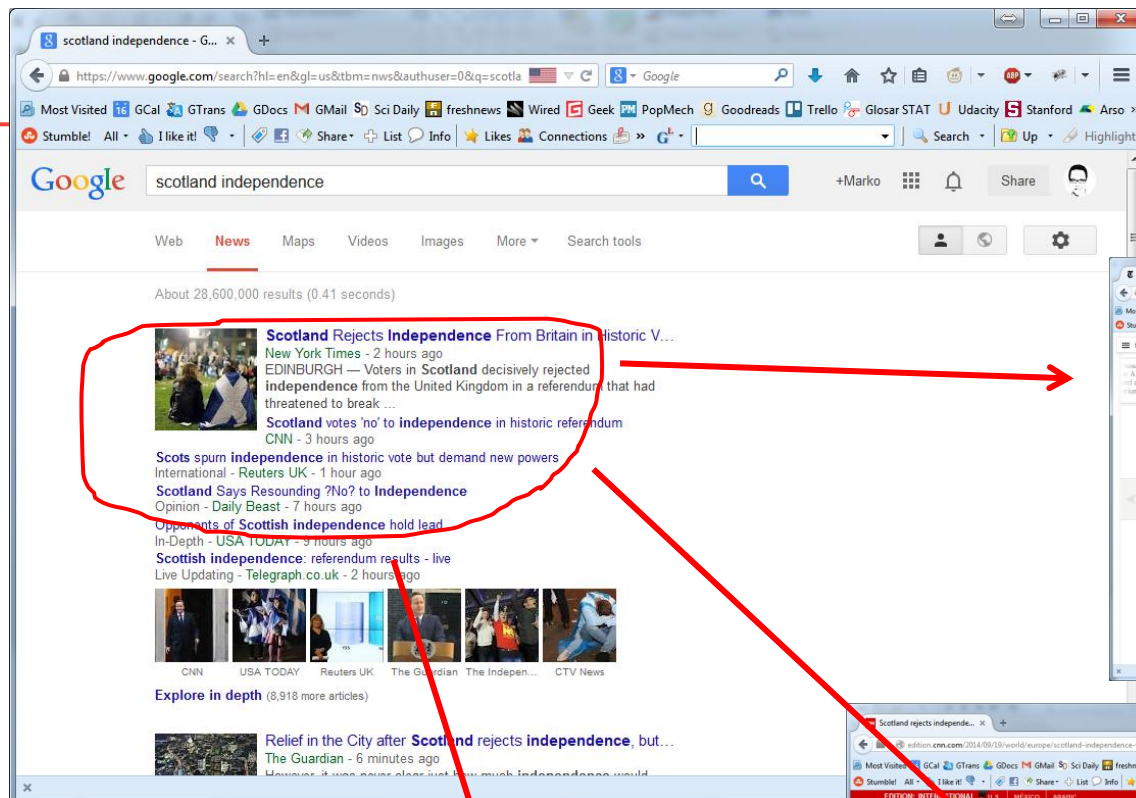
Nadzorovano učenje





Nenadzorovano učenje

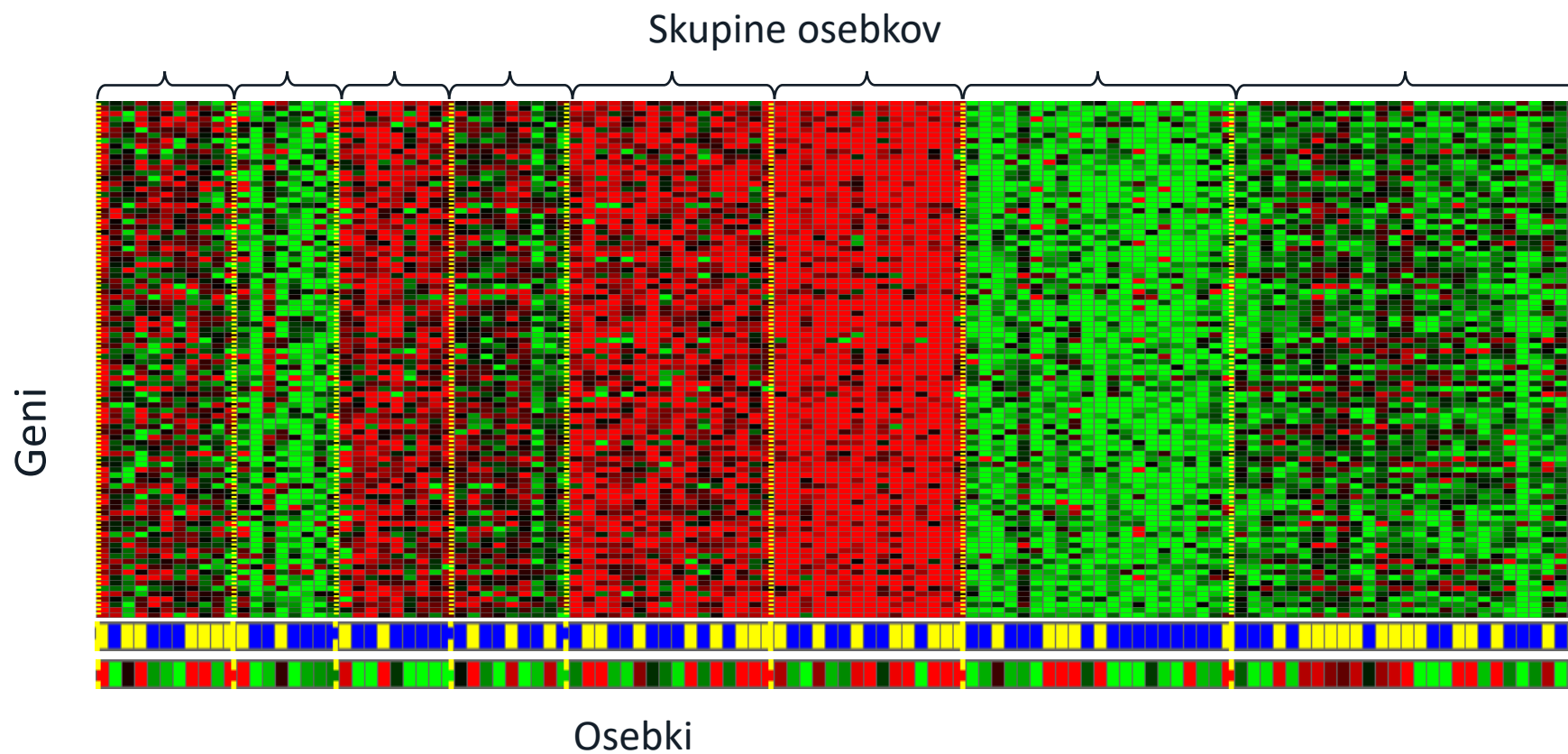




UNIVERZA
V LJUBLJANI

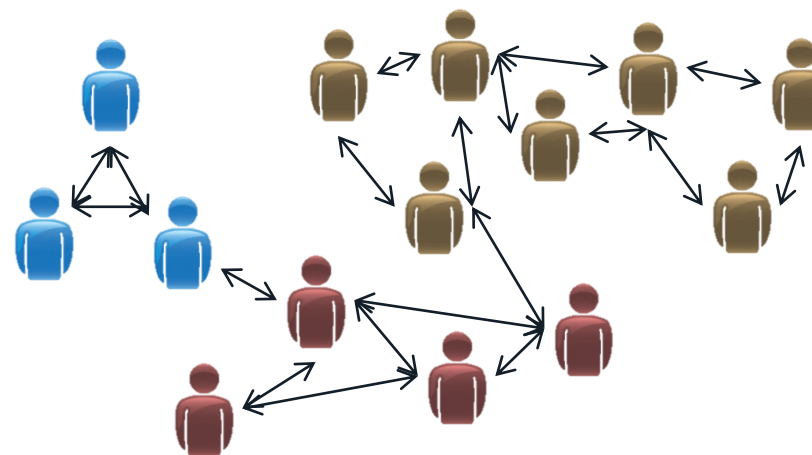
FE

Fakulteta
za elektrotehniko

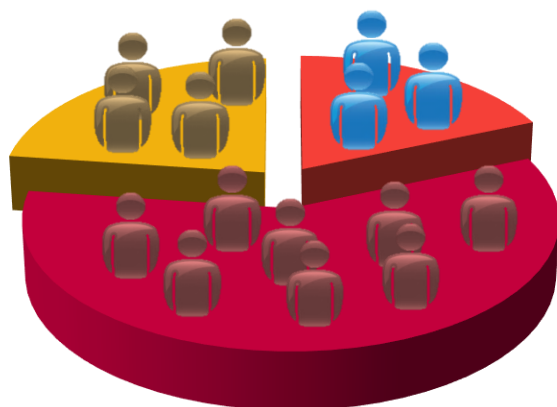




Organiziranje računskih centrov



Analiza socialnih omrežij



Segmentacija trga



Image credit: NASA/JPL-Caltech/E. Churchwell (Univ. of Wisconsin, Madison)

Analiza astronomskih podatkov



Primer

Za katere probleme bi uporabili nadzorovan/nenadzorovan algoritem strojnega učenja:

- Seznam e-mailov označen kot spam/nespam → učenje spam filtra
- Nabor člankov z novicami → združi jih v skupine glede na obravnavano tematiko
- Podatkovna baza kupcev → najdi tržne segmente in združi kupce v različne tržne segmente
- Podatkovna baza pacientov z diagnozo ima/nima diabetes → nauči se klasificirati ali ima pacient diabetes



Orodja

RapidMiner Studio

Orange

KNIME

WEKA

Python

R