

OSNOVE UMETNE INTELIGENCE

2021/22

*uvod v predmet
umetna inteligenca
uvod v strojno učenje*

Osnove umetne inteligence

Izvajalci

- **nosilec:**
Zoran Bosnić
2. nadstropje, R2.17 (kabinet)

- **asistenti:**
Jure Žabkar
3. nadstropje, R3.54 (LUI)

Aleš Papič
2. nadstropje, R2.25 (LKM)

Anita Valmarska
2. nadstropje, R2.25 (LKM)



Cilji predmeta



- kaj je **umetna inteligenca**?
 - kaj si prizadeva?
 - kakšna je definicija?
 - kako dobro nam uspeva doseči te cilje?
 - kakšne so posledice v filozofiji, psihologiji, etiki?
- kaj vse lahko delamo z **metodami umetne inteligence**?
 - vrste problemov
 - načini reševanja problemov
- kako razumeti medije, literaturo in objave s področja umetne inteligence?

Obveznosti predmeta



- sprotno delo:
domače naloge - kvizi na učilnici, vezani na predavanja in vaje. Potrebno je pozitivno ($\geq 50\%$) opraviti vsaj 4 od 5 domačih nalog. Za vsak kviz so na razpolago 3 poskusi.
- končni izpit:
pisni izpit (ali opravljena kolokvija). Pogoji za pristop k izpitu so opravljene DN.
- **KONČNA OCENA pri predmetu je enaka oceni končnega izpita.**

Dodatno, neobvezno:

- 2 kolokvija, ki lahko nadomestita izpit (če je vsak od njiju $\geq 50\%$). Izvedba bo na računalnikih.
- Kolokvijev se lahko udeležijo samo študenti, ki imajo uspešno opravljene ($\geq 50\%$) vse kvize - domače naloge do kolokvija.
- Pozor: Pristop h kolokvijem šteje kot pristop k rednemu izpitnemu roku.

Literatura

- **[AIMA]**
S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, Third edition, Pearson Education, Prentice-Hall 2010, ISBN: 0136042597.
- **[PPAI]**
I. Bratko, Prolog Programming for Artificial Intelligence, 4th edition, Pearson Education, Addison-Wesley 2011, ISBN: 0201403757.
- **[PUI]**
I. Bratko, Prolog in umetna inteligenca, Založba FE in FRI, ponatis 2011.
- **[SU]**
I. Kononenko, Strojno učenje, Založba FE in FRI, 2005.
- **[MLDM]**
Kononenko, Igor, and Matjaz Kukar. Machine learning and data mining. Horwood Publishing, 2007.
- **[IS]**
I. Kononenko, M. Robnik Šikonja, Inteligentni sistemi. Založba FE in FRI, 2010.



Umetna inteligencija?



Artificial intelligence (AI) promises to transform the **media** and entertainment business – impacting everything from content creation to the consumer experience. ... “Through techniques such as image recognition and speech-to-text transcription, metadata tagging is the most widespread application of AI so far. Sep 7, 2018



[How artificial intelligence is transforming the media industry](https://www.technologyrecord.com)
<https://www.technologyrecord.com> › [Article](#) › [how-artificial-intelligence-is-t...](#)

Google's **AI** Can Make Its Own **AI** Now. **Artificial intelligence** is advanced enough to do some pretty complicated things: read lips, mimic sounds, analyze photographs of food, and even design beer. ... The project is called AutoML, and it's designed to come up with better machine-learning software than humans can.

[Google's AI Can Make Its Own AI Now | Mental](https://mentalfloss.com)
mentalfloss.com › [article](#) › [googles-ai-can-make-its-own-a](#)

Facebook's experiment isn't the only time that **artificial intelligence** has invented new forms of **language**. Earlier this year, Google revealed that the **AI** it uses for its Translate tool had created its **own language**, which it would translate things into and then out of. Jul 31, 2017

[Facebook's artificial intelligence robots shut down after they ...](https://www.independent.co.uk)
<https://www.independent.co.uk> › [life-style](#) › [gadgets-and-tech](#) › [news](#) › [face...](#)

Umetna inteligenca... ..zares?

Cost function

Logistic regression:

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m [y^{(i)} \log h_{\theta}(x^{(i)}) + (1 - y^{(i)}) \log(1 - h_{\theta}(x^{(i)}))] + \frac{\lambda}{2m} \sum_{j=1}^n \theta_j^2$$

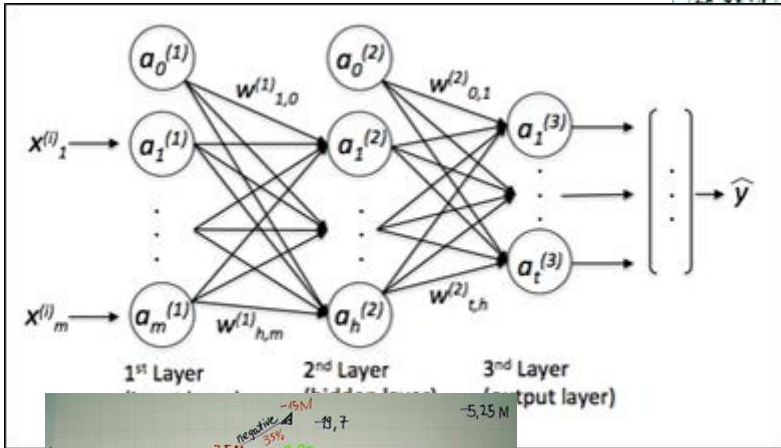
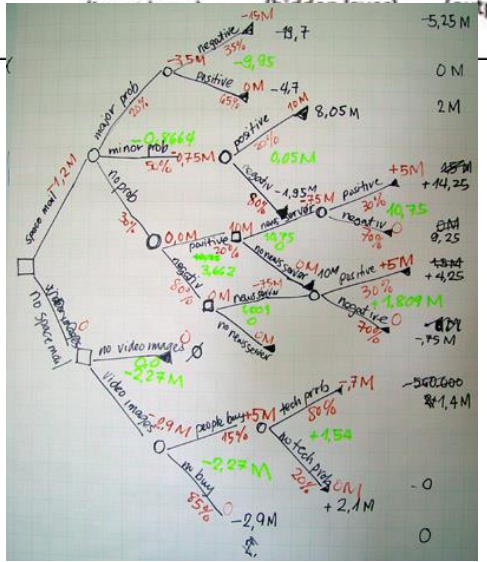
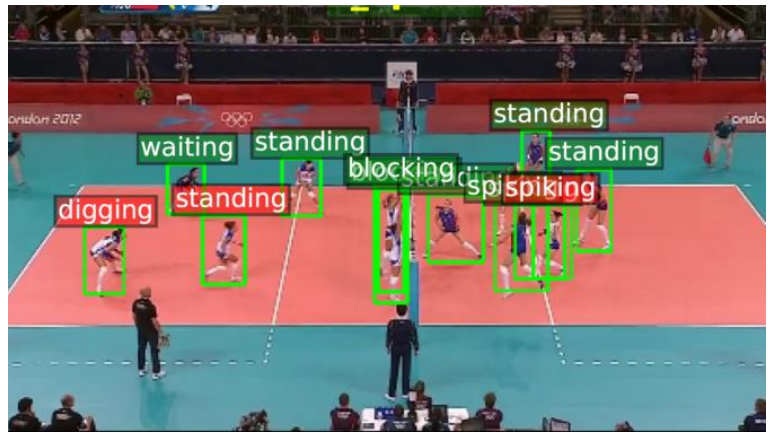
Neural network:

$\rightarrow h_{\Theta}(x) \in \mathbb{R}^K$ ($h_{\Theta}(x)_i = i^{th}$ output)

$$\rightarrow J(\Theta) = -\frac{1}{m} \sum_{i=1}^m \sum_{k=1}^K y_k^{(i)} \log(h_{\Theta}(x^{(i)}))_k + (1 - y_k^{(i)}) \log(1 - (h_{\Theta}(x^{(i)}))_k)$$

$$+ \frac{\lambda}{2m} \sum_{l=1}^{L-1} \sum_{i=1}^{s_l} \sum_{j=1}^{s_{l+1}} (\Theta_{ji}^{(l)})^2$$

(Handwritten notes include: $\Theta_{ji}^{(l)} = \omega_{ji}^{(l)} x_i + b_j^{(l)}$)

Back in 2000, **People Magazine** **PUBLISHER** highlighted **Prince Williams'** **PERSON** style who at the time was a little more fashion-conscious, even making fashion statements at times.

Now-a-days the prince mainly wears **navy** **COLOR** suits **ITEM** (sometimes **double-breasted** **DESIGN**), **light blue** **COLOR** button-ups **ITEM** with **classic** **LOOK** **pointed** **DESIGN** collars **PART**, and **burgundy** **COLOR** ties **ITEM**.

But who knows what the future holds ...

Duchess Kate **PERSON** did wear an **Alexander McQueen** **BRAND** dress **ITEM** to the **wedding** **OCCASION** in the **fall of 2017** **SEASON**.

Vsebina predmeta

- uvod, definicija umetne inteligence
- strojno učenje:
 - problemski prostor, hipoteze, ocenjevanje učenja
 - gradnja odločitvenih dreves
 - učenje iz šumnih podatkov, rezanje
 - manjkajoči atributi, regresija, naivni Bayes
- reševanje problemov kot preiskovanje grafov
 - neinformirani preiskovalni algoritmi
 - informirani preiskovalni algoritmi
 - lokalno preiskovanje
 - grafi AND/OR, nedeterministično okolje
- igranje iger
- planiranje, razporejanje opravil
- predstavitev negotovega znanja, Bayesovske mreže
- avtomatsko sklepanje



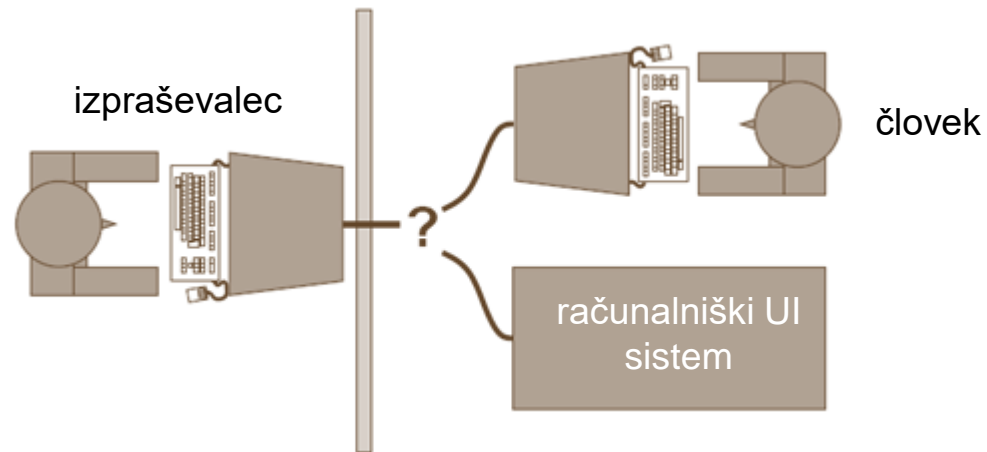
Kaj je umetna inteligenca?



- **cilj umetne intelligence:** razumeti in zgraditi *intelligentne* sisteme na osnovi razumevanja človeškega razmišljanja, sklepanja, učenja in komuniciranja (izvajanja dejanj, pojasnjevanja, komentiranja)
- **različne definicije**, ki izvirajo iz različnega pojmovanja človeške intelligence in cilja, ki ga pričakujemo od računalniškega sistema
- **težava:** definicija *intelligence*?
 - ali je vse, kar počne človek, inteligentno (reflexi, nekonsistentnost, napake, učinkovitost)?
 - ali zasledovati modeliranje človeka ali modeliranje ideala (princip **racionalnosti** – optimalnosti – "*doing the right thing*")
- **umetna inteligenca:**
 - sistemi, ki se vedejo/razmišljajo kot človek/racionalno
 - kako je s kreativnostjo, čustvi, zavestjo?
- velika medijska pokritost ("*AI will conquer the world*" (!@#!))

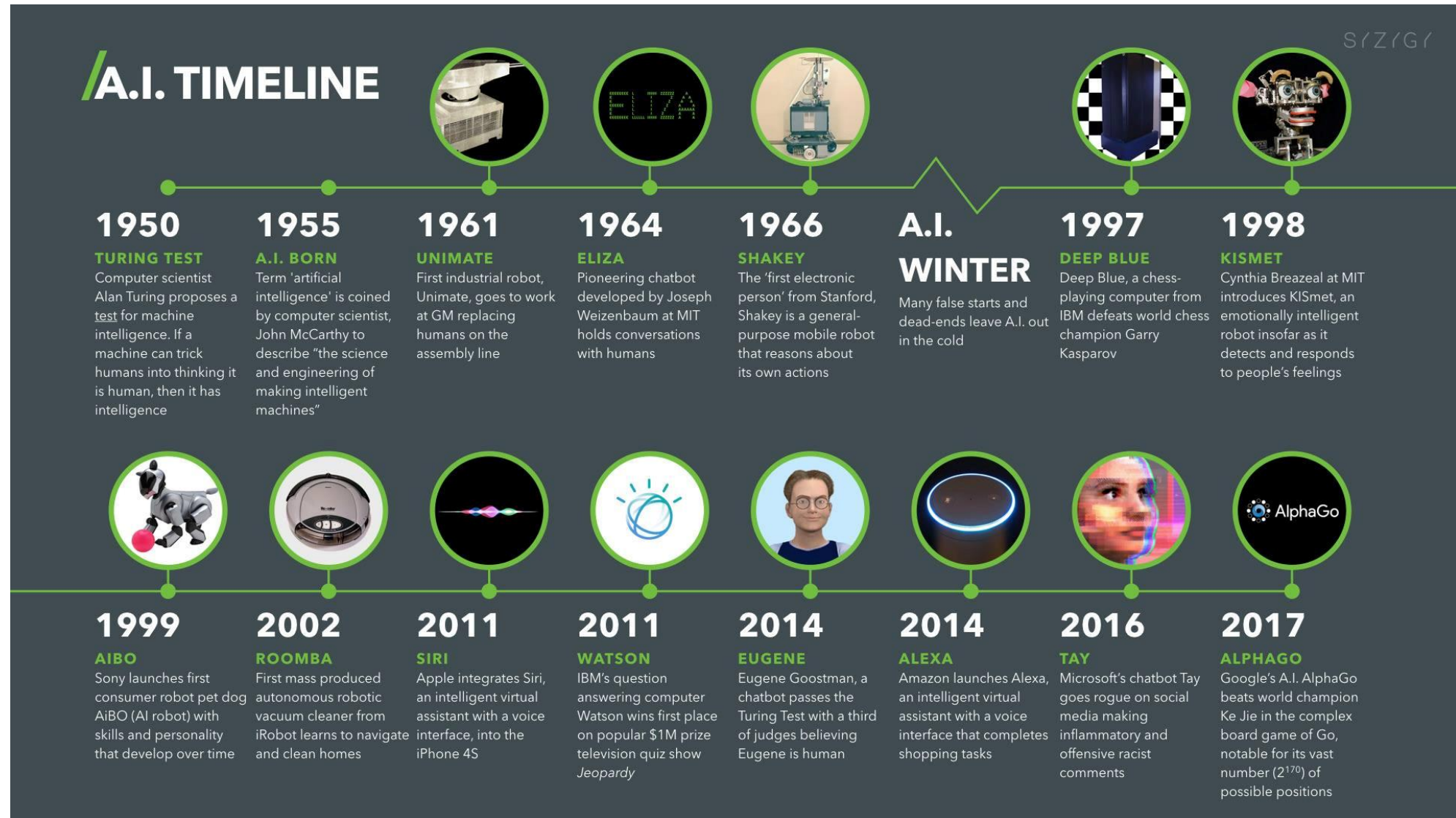
Turingov test

- sistem, ki se "vede kot človek"
- **praktični preizkus** (the imitation game), ki ga predlaga Turing (1950) za testiranje, ali je sistem dosegel stopnjo inteligence, primerljivo s človekom
- računalnik "opravi" preizkus, če človeški izpraševalec po računalnikovih odgovorih na zastavljena vprašanja ne more ugotoviti, ali odgovore podaja človek ali računalnik

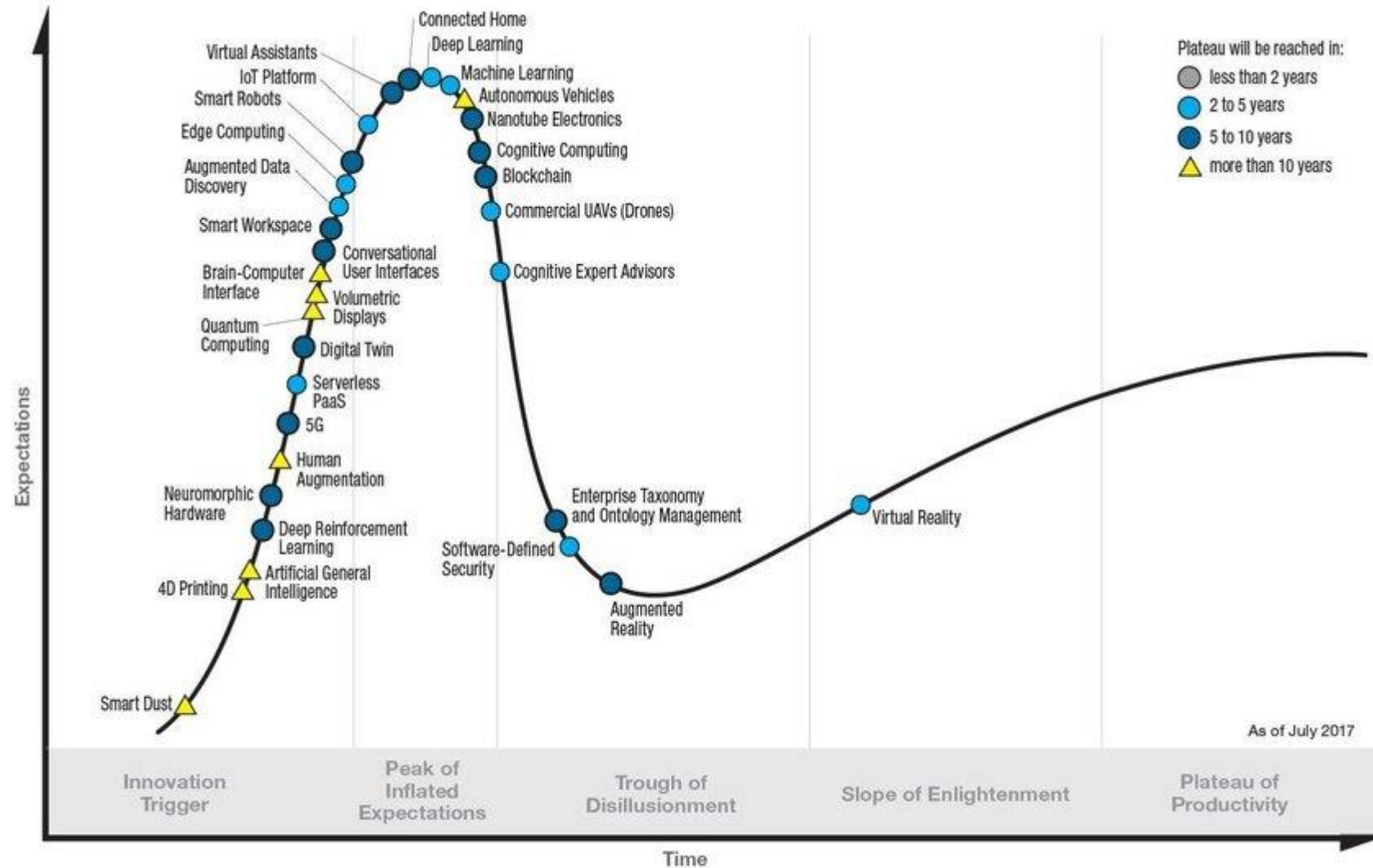


- ideja napeljuje na sposobnosti inteligentnih sistemov: obdelava naravnega jezika, predstavitev znanja, avtomatsko sklepanje, strojno učenje, računalniški vid, robotika
- problem: testa ni možno reproducirati ali podvreči matematični analizi

Zgodovina UI



Zgodovina UI

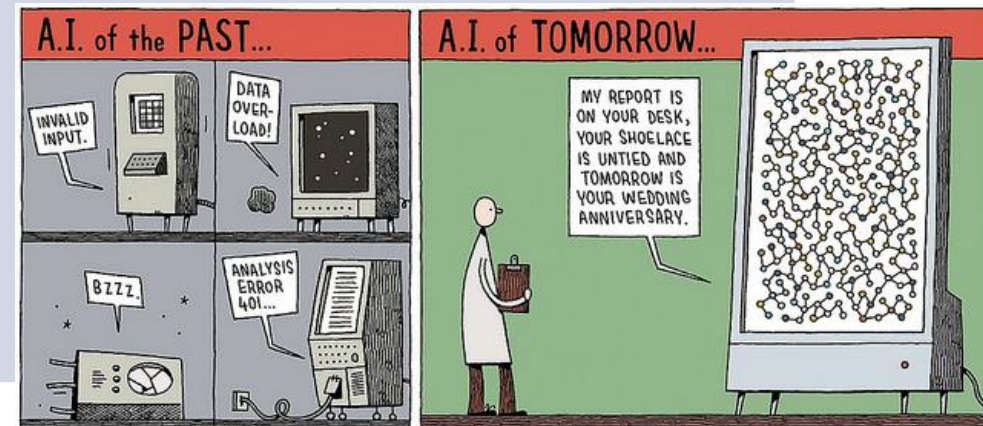


Cilj umetne inteligence



- razumeti in zgraditi "*intelligentne*" sisteme, na osnovi razumevanja človeškega razmišljanja, sklepanja, učenja, izvajanja dejanj, pojasnjevanja, komentiranja
- težava: definicija *intelligence*?

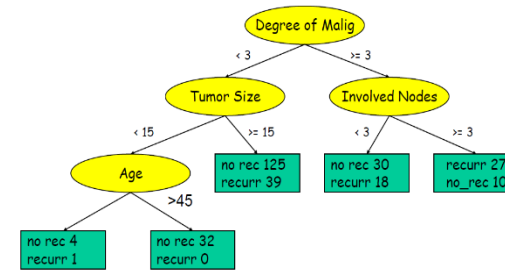
Kaj danes znamo z UI?	In česa še UI ne zna... ?
<ul style="list-style-type: none">• avtonomna (robotska) vozila• <u>robotska opravila</u> (pospravljanje pomivalnega stroja?)• razpoznavanje govora (klicni centri, pametni telefoni, ...)• <u>igranje iger</u> (šah - Deep Blue in Kasparov 1997), namizni tenis, karte)• <u>načrtovanje logistike</u> (porabe sredstev)• <u>detekcija neželene elektronske pošte</u>• robotika (Roomba, raziskovanje vesolja)• strojno prevajanje• nakupovanje živil za naslednji teden• iskanje in izpeljava novih matematičnih dokazov• <u>podajanje ustreznih nasvetov</u> na pravnem področju• prevajanje med dvema jezika v realnem času• izvajanje pogovora z osebo• analiza in sinteza jezika• ...	<ul style="list-style-type: none">• razlaga odločitev• odločanje po moralnih načelih• občutenje empatije, simpatije in drugih čustev• kreativnost• popolno nadomeščanje človeških delavcev• zavedanje (izdelava profila na FB?)• ...



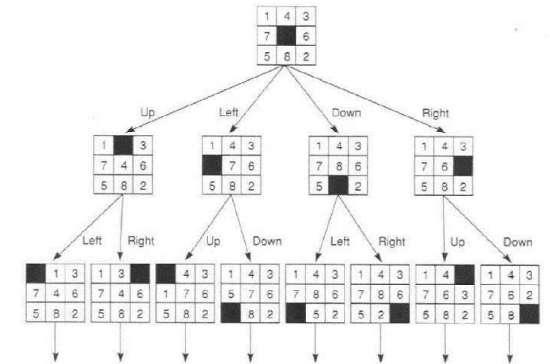
Osnove umetne inteligence

I. STROJNO UČENJE

(*machine learning*) in odkrivanje zakonitosti v podatkih (podatkovno rudarjenje, *data mining*, *knowledge discovery in databases*)



strojno učenje – odločitveno drevo za napovedovanje ponovitve raka na dojki



preiskovanja prostora stanj igre s ploščicami

II. REŠEVANJE PROBLEMOV

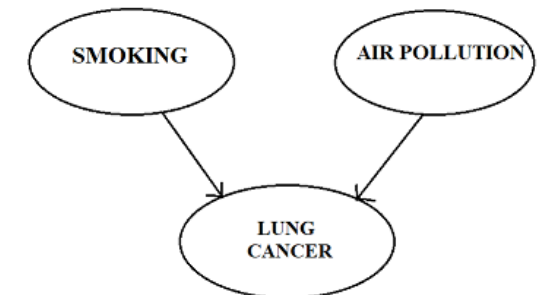
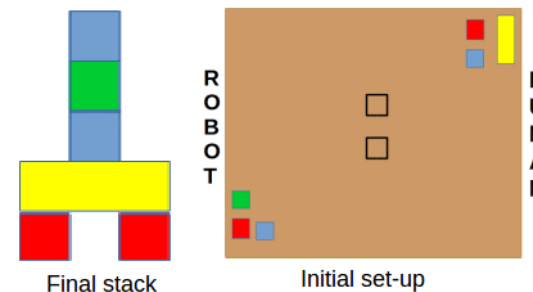
iskanje rešitev v problemskem prostoru, igranje iger

III. PLANIRANJE, RAZPOREJANJE OPRAVIL

izdelava načrta (plana) akcij za doseganje cilja

IV. SKLEPANJE

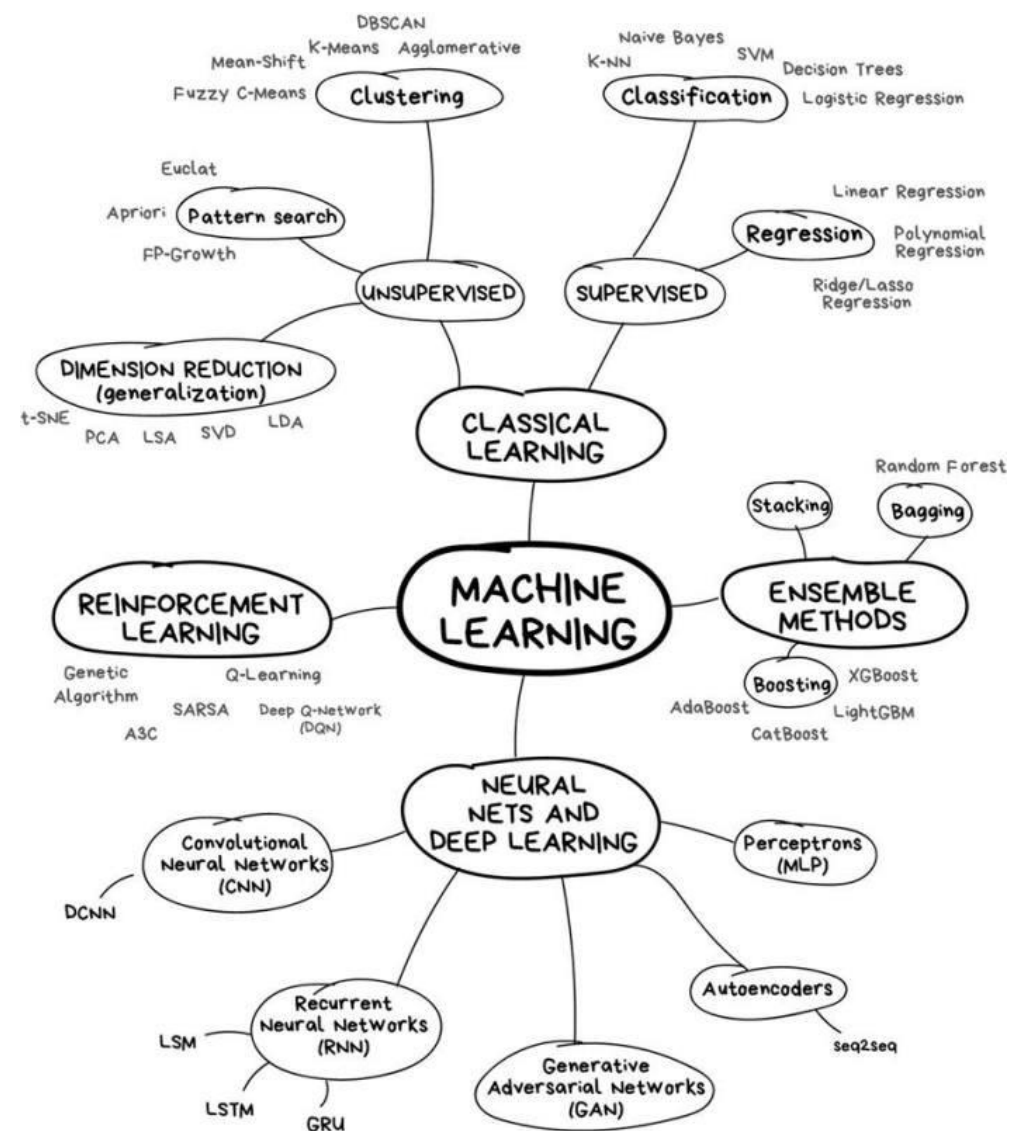
avtomatsko sklepanje (*reasoning*, *inference*), uporaba predstavitev znanj



I. STROJNO UČENJE

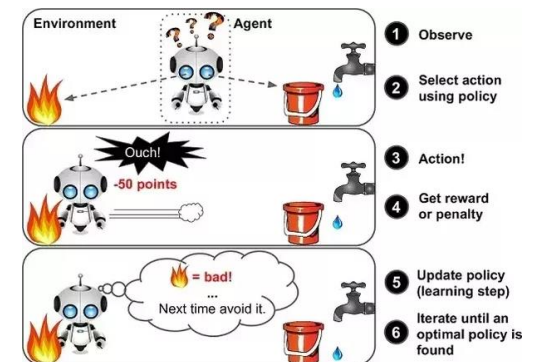
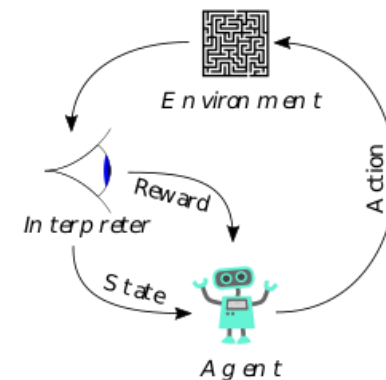
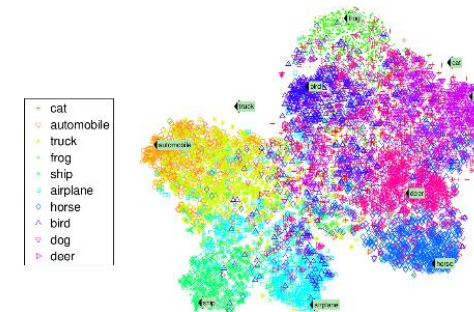
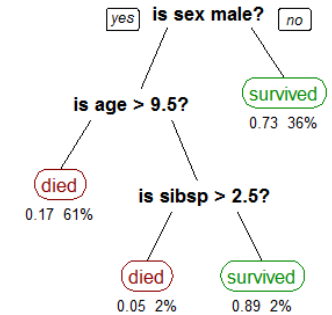
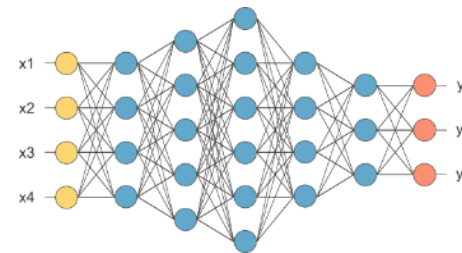
Strojno učenje

- angl. *machine learning*
- je **področje umetne inteligence**, ki raziskuje, kako se lahko algoritmi **samodejno izboljšujejo** ob pridobivanju izkušenj
- naloge: analiza učnih podatkov, gradnja modela, (napovedovanje)
- zakaj takoj ne vključiti vsega znanja v program?
 - razvijalci programske opreme ne morejo predvideti vseh možnih **problemских situacij**,
 - razvijalci ne morejo predvideti **sprememb** okolja skozi čas (prilagodljivost)
 - razvijalci **ne znajo** sprogramirati agenta z znanjem (npr. razpoznava obrazov?)



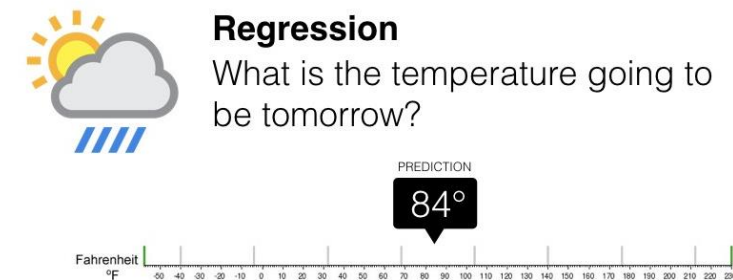
Vrste učenja

- **nadzorovano učenje** (angl. *supervised learning*): učni primeri so podani kot vrednosti vhodov in izhodov (učni primeri so označeni); učimo se funkcije, ki preslika vhode v izhode (npr. odločitveno drevo)
- **nenadzorovano učenje** (angl. *unsupervised learning*): učni primeri niso označeni (nimajo ciljne spremenljivke); učimo se vzorcev v podatkih (npr. gručenje)
- **spodbujevano učenje** (angl. *reinforcement learning*): inteligentni agent se uči iz zaporedja nagrad in kazni



Nadzorovano učenje

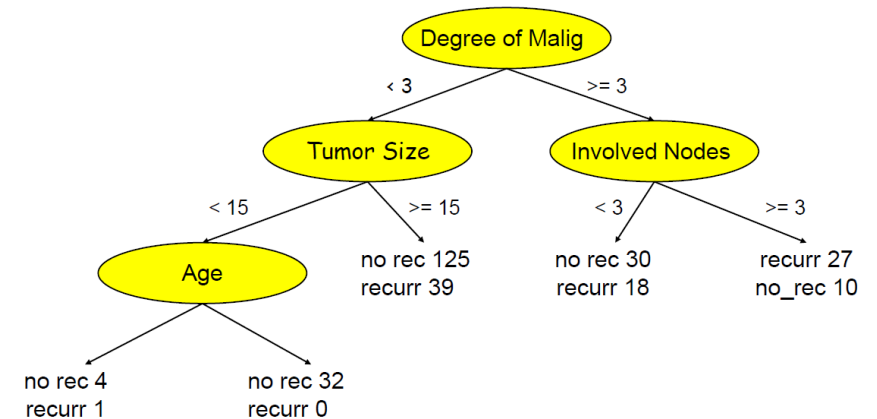
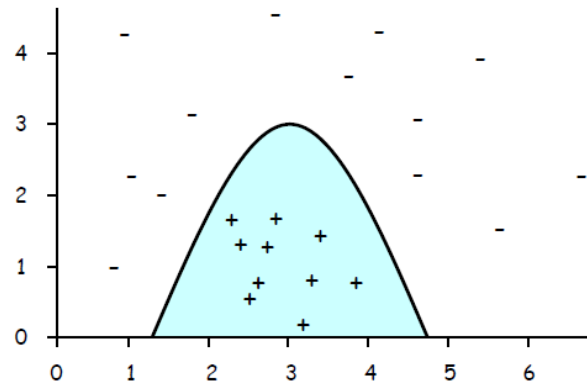
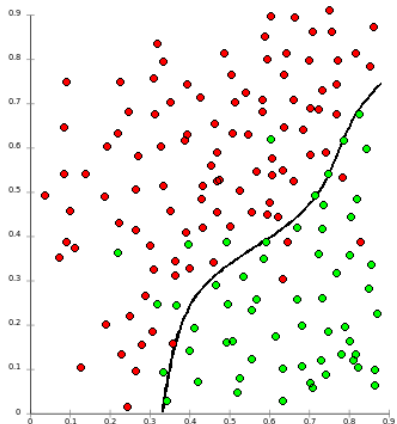
- **podana:** množica **učnih primerov**
 $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$,
kjer je vsak y_j vrednost neznane funkcije $y = f(x)$
- **naloga:** najdi funkcijo h , ki je najboljši približek funkciji f
- x_j so **atributi** (vrednost ali vektor)
- funkcijo h imenujemo **hipoteza**
- ločimo dve vrsti problemov:
 - če je y_j diskretna (kategorična) spremenljivka
→ **klasifikacijski** problem
 - če je y_j zvezna spremenljivka
→ **regresijski** problem



Vrste problemov

- **klasifikacija:**

- y pripada **končnem naboru vrednosti** (je diskretna spremenljivka)
- npr. $y \in \{užitna, strupena\}$, $y \in \{sonce, oblačno, dež\}$, $y \in \{zdrav, bolan\}$
- y imenujemo **razred** (*angl. class*)
- primeri:
 - napovedovanje vremena iz podatkov prejšnjih let
 - diagnosticiranje novih pacientov na osnovi znanih diagnoz za stare paciente
 - klasifikacija neželene elektronske pošte
 - napovedovanje vračila kredita

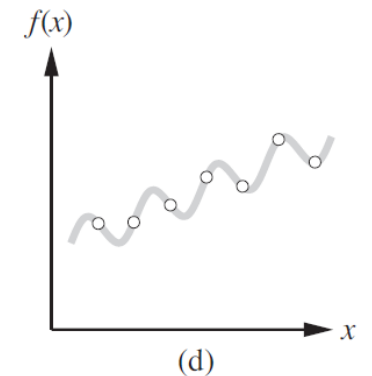
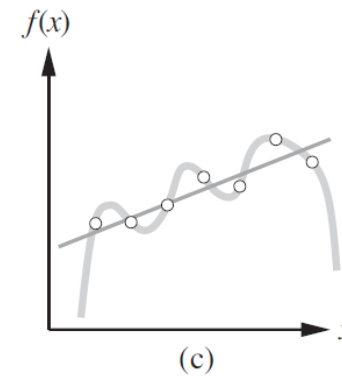
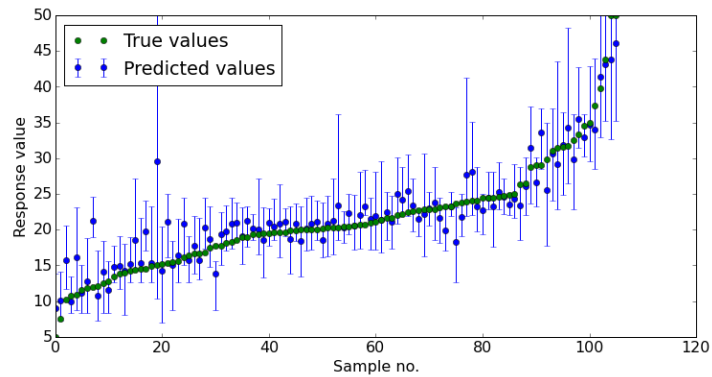
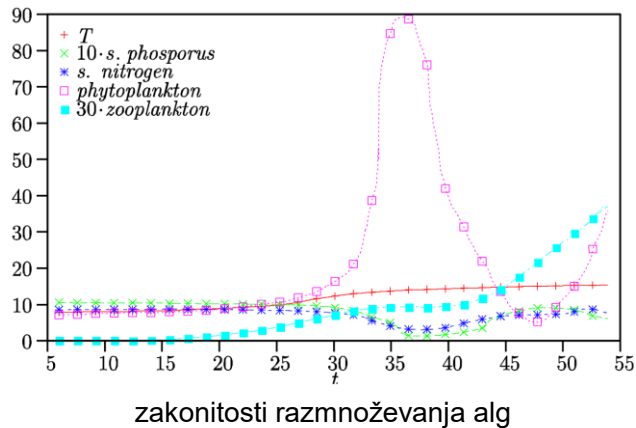


drevo, zgrajeno iz 139 učnih primerov; višja
klasifikacijska točnost kot zdravniška

Vrste problemov

- **regresija:**
 - y je število (običajno $y \in \mathbb{R}$, je zvezna spremenljivka)
 - npr. y je temperatura,
 - y imenujemo **označba** (angl. *label*)
 - primeri:
 - napovedovanje razmnoževanja alg
 - medicinska diagnostika
 - napovedovanje vremena
 - napovedovanje koncentracije ozona
 - napovedovanje gibanja cen delnic

Mid 1980s, Danish lake Glumso



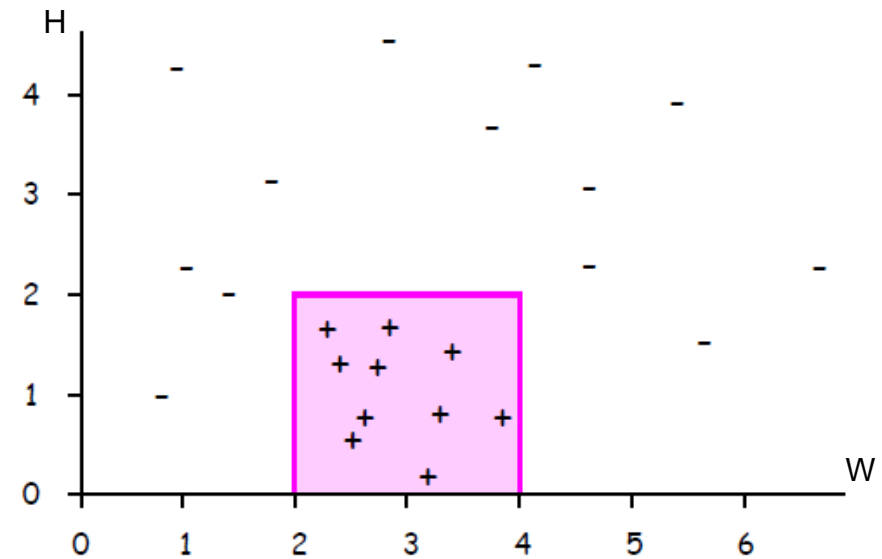
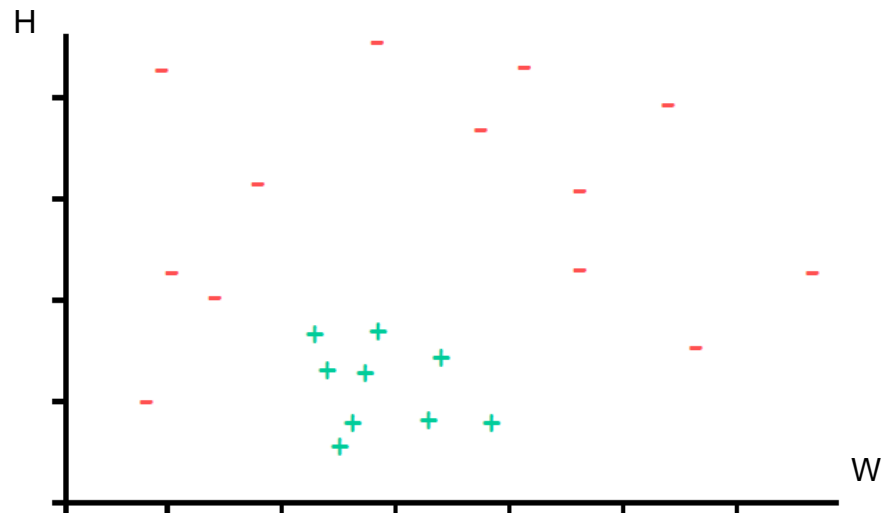
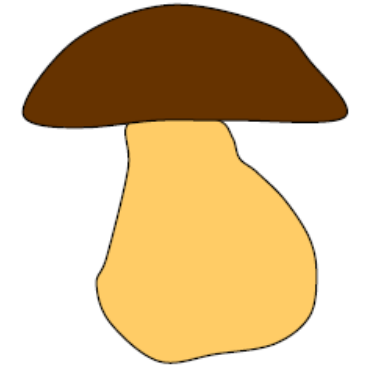
Atributna predstavitev podatkov

- učna množica: čakanje na prosto mesto v restavraciji
- ciljna spremenljivka: čakamo (T) ali ne čakamo (F)

Example	Attributes										Target
	<i>Alt</i>	<i>Bar</i>	<i>Fri</i>	<i>Hun</i>	<i>Pat</i>	<i>Price</i>	<i>Rain</i>	<i>Res</i>	<i>Type</i>	<i>Est</i>	<i>WillWait</i>
X_1	<i>T</i>	<i>F</i>	<i>F</i>	<i>T</i>	<i>Some</i>	<i>\$\$\$</i>	<i>F</i>	<i>T</i>	<i>French</i>	<i>0–10</i>	<i>T</i>
X_2	<i>T</i>	<i>F</i>	<i>F</i>	<i>T</i>	<i>Full</i>	<i>\$</i>	<i>F</i>	<i>F</i>	<i>Thai</i>	<i>30–60</i>	<i>F</i>
X_3	<i>F</i>	<i>T</i>	<i>F</i>	<i>F</i>	<i>Some</i>	<i>\$</i>	<i>F</i>	<i>F</i>	<i>Burger</i>	<i>0–10</i>	<i>T</i>
X_4	<i>T</i>	<i>F</i>	<i>T</i>	<i>T</i>	<i>Full</i>	<i>\$</i>	<i>F</i>	<i>F</i>	<i>Thai</i>	<i>10–30</i>	<i>T</i>
X_5	<i>T</i>	<i>F</i>	<i>T</i>	<i>F</i>	<i>Full</i>	<i>\$\$\$</i>	<i>F</i>	<i>T</i>	<i>French</i>	<i>>60</i>	<i>F</i>
X_6	<i>F</i>	<i>T</i>	<i>F</i>	<i>T</i>	<i>Some</i>	<i>\$\$</i>	<i>T</i>	<i>T</i>	<i>Italian</i>	<i>0–10</i>	<i>T</i>
X_7	<i>F</i>	<i>T</i>	<i>F</i>	<i>F</i>	<i>None</i>	<i>\$</i>	<i>T</i>	<i>F</i>	<i>Burger</i>	<i>0–10</i>	<i>F</i>
X_8	<i>F</i>	<i>F</i>	<i>F</i>	<i>T</i>	<i>Some</i>	<i>\$\$</i>	<i>T</i>	<i>T</i>	<i>Thai</i>	<i>0–10</i>	<i>T</i>
X_9	<i>F</i>	<i>T</i>	<i>T</i>	<i>F</i>	<i>Full</i>	<i>\$</i>	<i>T</i>	<i>F</i>	<i>Burger</i>	<i>>60</i>	<i>F</i>
X_{10}	<i>T</i>	<i>T</i>	<i>T</i>	<i>T</i>	<i>Full</i>	<i>\$\$\$</i>	<i>F</i>	<i>T</i>	<i>Italian</i>	<i>10–30</i>	<i>F</i>
X_{11}	<i>F</i>	<i>F</i>	<i>F</i>	<i>F</i>	<i>None</i>	<i>\$</i>	<i>F</i>	<i>F</i>	<i>Thai</i>	<i>0–10</i>	<i>F</i>
X_{12}	<i>T</i>	<i>T</i>	<i>T</i>	<i>T</i>	<i>Full</i>	<i>\$</i>	<i>F</i>	<i>F</i>	<i>Burger</i>	<i>30–60</i>	<i>T</i>

Primer: gobe

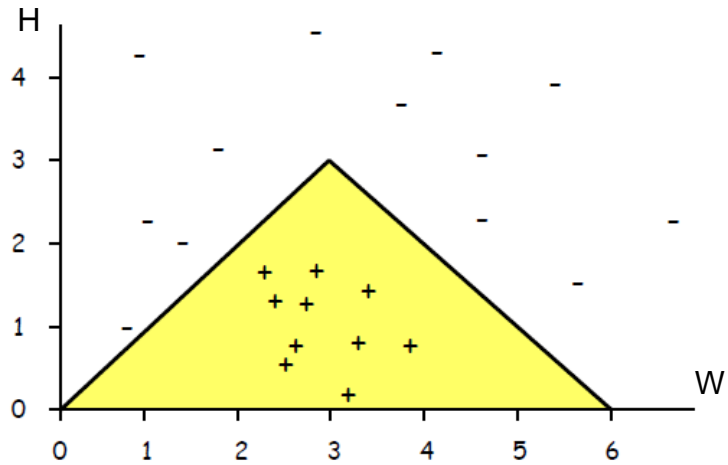
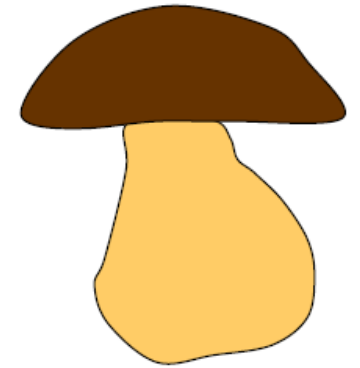
- razpoznavanje užitnih gob
- atributa (x): W (width) in H (height)
- razred (y): strupena (-), užitna (+)



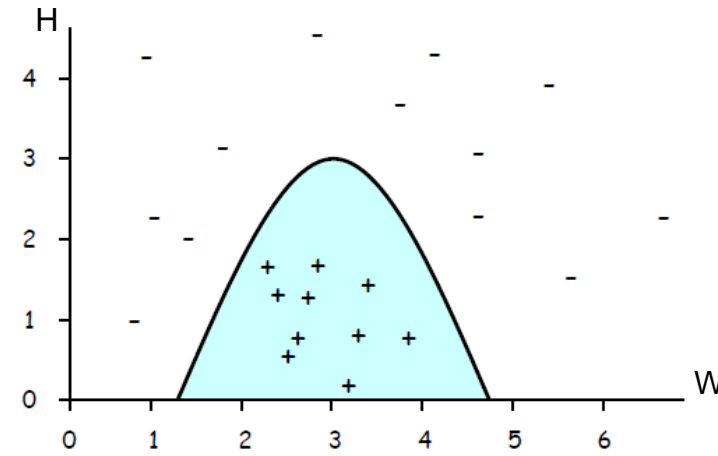
```
IF W>2 and W<4 and H<2  
THEN "edible" ELSE "poisonous"
```

Primer: gobe

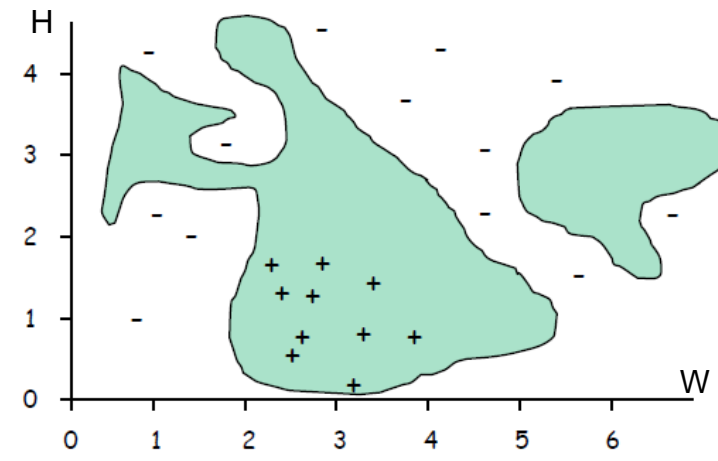
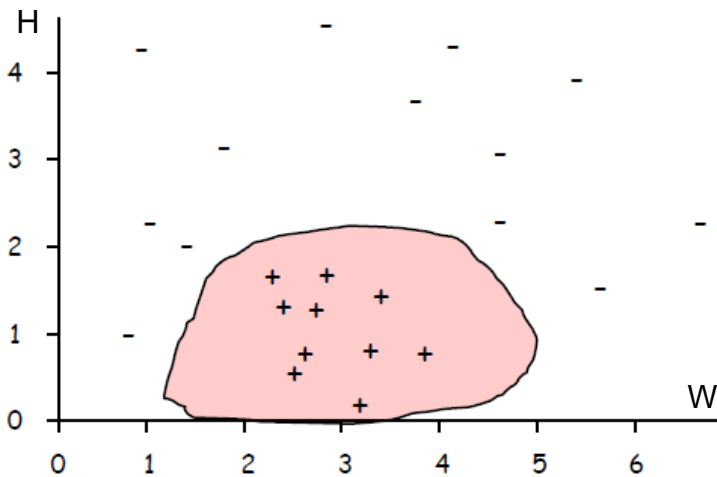
- ali pa ...



IF $H > W$ THEN "poisonous"
ELSE IF $H > 6 - W$ THEN "poisonous" ELSE "edible"

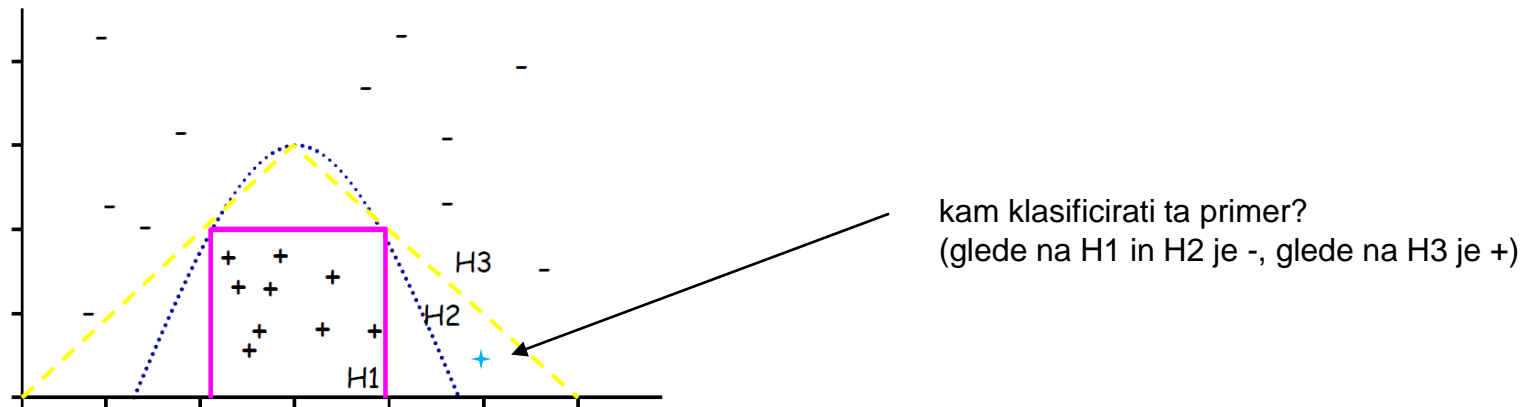


IF $H < 3 - (W-3)^2$ THEN "edible"
ELSE "poisonous"



Primer: gobe

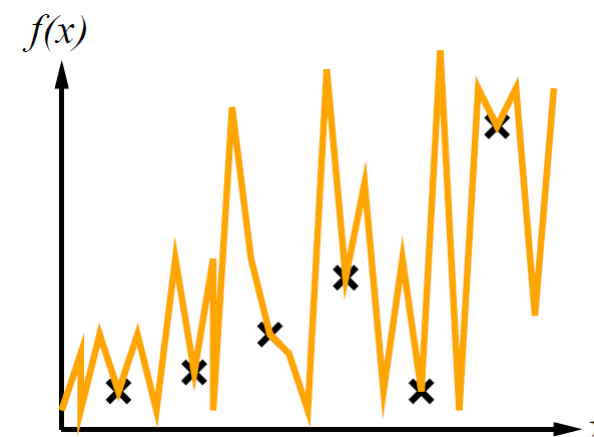
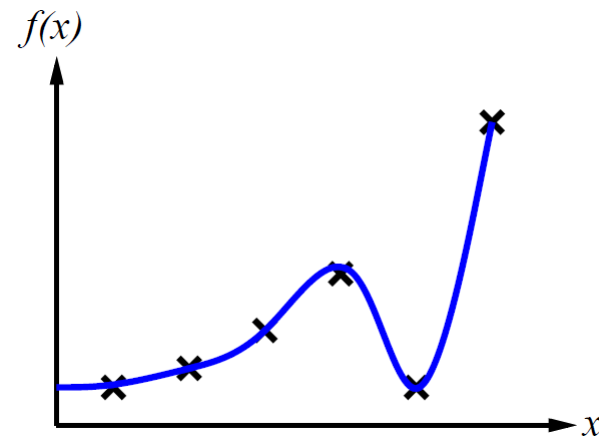
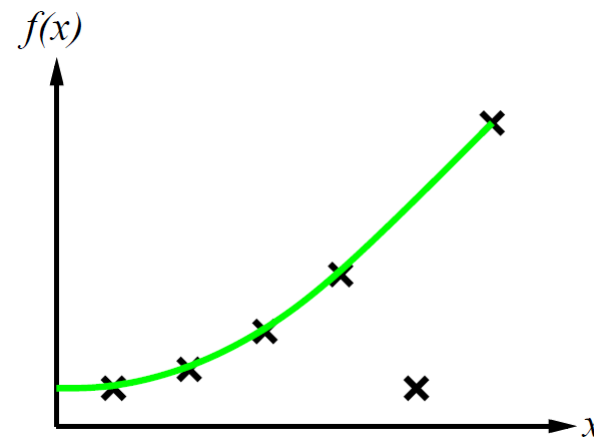
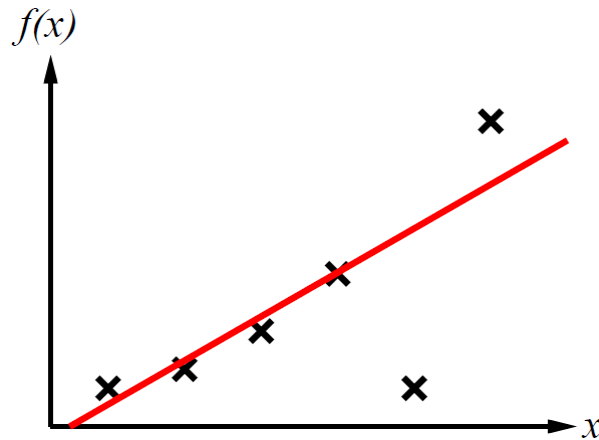
- prostor hipotez vsebuje več hipotez
- vse prikazane hipoteze so **konsistentne** z učno množico
- dobra hipoteza je dovolj **splošna** (angl. general), kar pomeni, da pravilno napoveduje vrednost y za nove (še nevidene) primere



- kako izbrati primerno hipotezo? Princip **Ockhamove britve** (*Ockham's razor*) (William o Ockham, 1320, angleški filozof):
 - prava hipoteza je najbolj preprosta hipoteza
 - *Entities should not be multiplied unnecessarily*
 - *Given two explanations of the data, all other things being equal, the simpler explanation is preferable.*

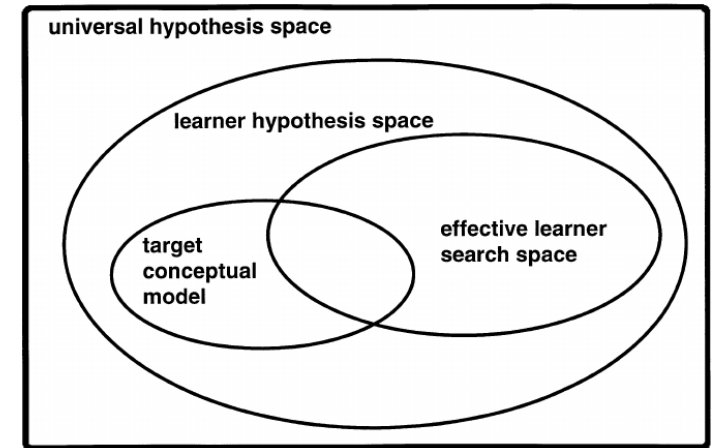
Primer

- podoben problem z izbiro hipoteze imamo tudi pri **regresijskih problemih** (iskanje funkcije, ki opisuje podane točke)



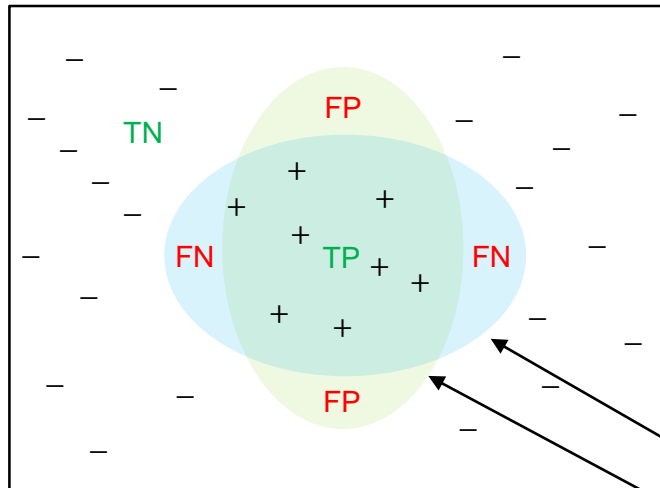
Prostor hipotez

- denimo, da imamo
 - binarno klasifikacijo
 - n binarnih atributov
- sledi:
 - 2^n različnih učnih primerov
 - 2^{2^n} hipotez (denimo, da lahko hipotezo opišemo s tabelo napovedi za vse primere)
- primer:
 - za 10 atributov izbiramo med 10^{308} možnimi hipotezami
 - za 20 atributov izbiramo med $10^{300.000}$ možnimi hipotezami
 - v resnici: hipotez je še več, izračunavajo lahko isto funkcijo
- potrebujemo:
 - **algoritme za gradnjo "dobrih" hipotez**
 - metode za **ocenjevanje hipotez** / ocenjevanje učenja
 - zavedanje o pristranosti hipotez



Evalviranje hipotez

- pomembni kriteriji:
 - **konsistentnost** hipotez s primeri (učnimi)
 - **splošnost** (točnost za nevidene primere)
 - **razumljivost** (interpretability, comprehensibility) hipotez
- ocenjevanje uspešnosti pri klasifikaciji na podlagi njihove **točnosti**:
 - točnost na učnih podatkih? (pristranost hipotez?)
 - točnost na testnih podatkih?
 - točnost na novih (še nevidenih) podatkih?



pravi (ciljni, neznani) pojem

naučena hipoteza

TP – pravilno pozitivno klasificirani primeri (angl. *true positive*)
TN – pravilno negativno klasificirani primeri (angl. *true negative*)
FP – napačno pozitivno klasificirani primeri (angl. *false positive*)
FN – napačno negativno klasificirani primeri (angl. *false negative*)

klasifikacijska točnost (angl. classification accuracy):

$$CA = \frac{TP + TN}{TP + TN + FP + FN} = \frac{TP + TN}{N}$$

Izpitna naloga

- 1. izpitni rok, 30. 1. 2018

3. NALOGA:

Podan je primer učenja iz primerov z atributoma A in B ter razredom C. Atribut A in razred sta binarna, atribut B pa lahko zavzame tri vrednosti. Učno množico primerov, ki smo jih zajeli in v katerih ni šuma, prikazuje tabela na desni. Denimo, da vemo, da pravo odvisnost med atributi in razredom izraža funkcija $C = IF (AB^2) < (A + B) THEN 1 ELSE 0$. Istočasno pa se z dvema različnima algoritmoma za učenje iz primerov naučimo naslednjih dveh hipotez:

$$H1: C = A + 1,5 \cdot B - AB - 0,5 \cdot B^2$$

$$H2: C = \min(A + B, 1) - \min(A, B)$$

Odgovori na naslednja vprašanja:

- a) Katera od podanih hipotez je bolj splošna? Kaj to pomeni?
- b) Katera od podanih hipotez ima višjo klasifikacijsko točnost na učni množici?
- c) Kakšna je razlika med nadzorovanim in nenadzorovanim učenjem? S katerim imamo opravka pri zgornji nalogi? Podaj primer problema nadzorovanega in nenadzorovanega učenja iz prakse.
- d) Kaj je to binarizacija atributa in zakaj je koristna? Podaj primer binarizacije atributa B.

A	B	C
0	0	0
0	1	1
0	2	1
1	0	1
1	2	0



Strojno učenje