

Zakaj strojno učenje deluje?

Marko Grobelnik

Laboratorij za umetno inteligenco (AI Lab), Institut Jozef Stefan

Marko.Grobelnik@ijs.si



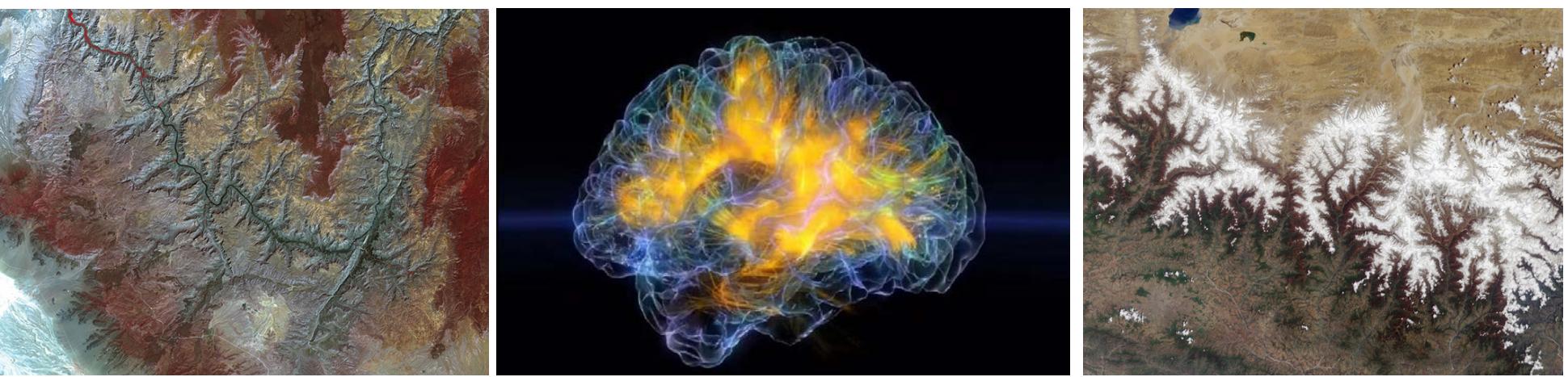
**Jožef Stefan
Institute
Ljubljana, Slovenia**

Podatkovna Znanost, Ljubljana, Mar 29th 2018
SINOG, Ljubljana, Apr 18th 2018

Kaj počnemo v znanost?

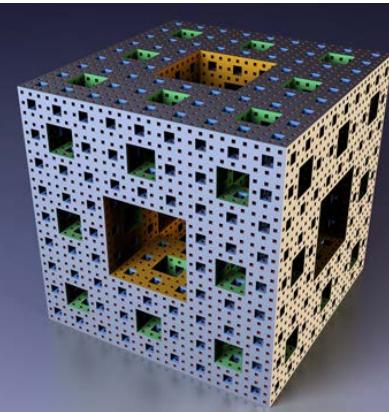
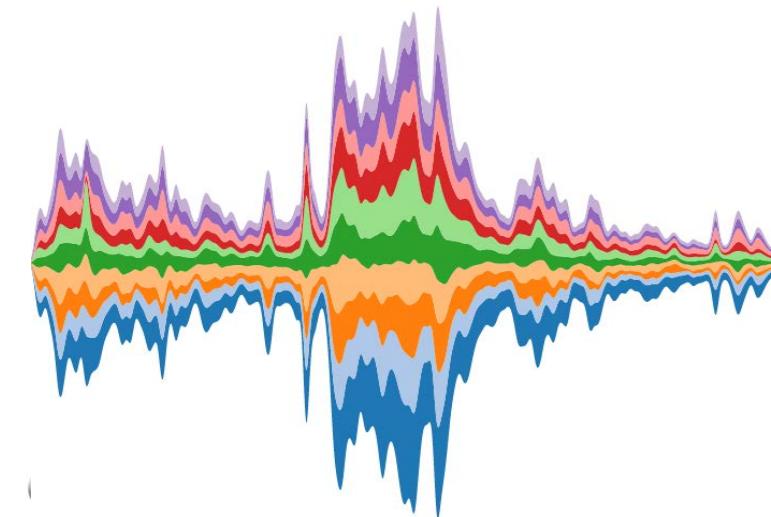
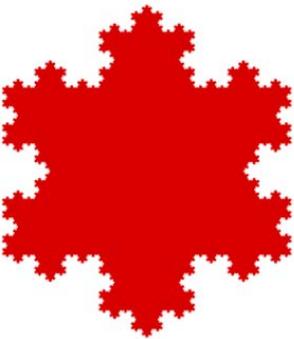
(in strojno učenje/AI je samo del znanosti)

- Narava je organizirana v strukture, ki so si zelo podobne
 - ...strukture izgledajo kot kristali
- Znanost skuša ugotoviti, kakšne so strukture v naravi
 - ...ne odkriva nič novega, samo skuša razumeti, kar že obstaja

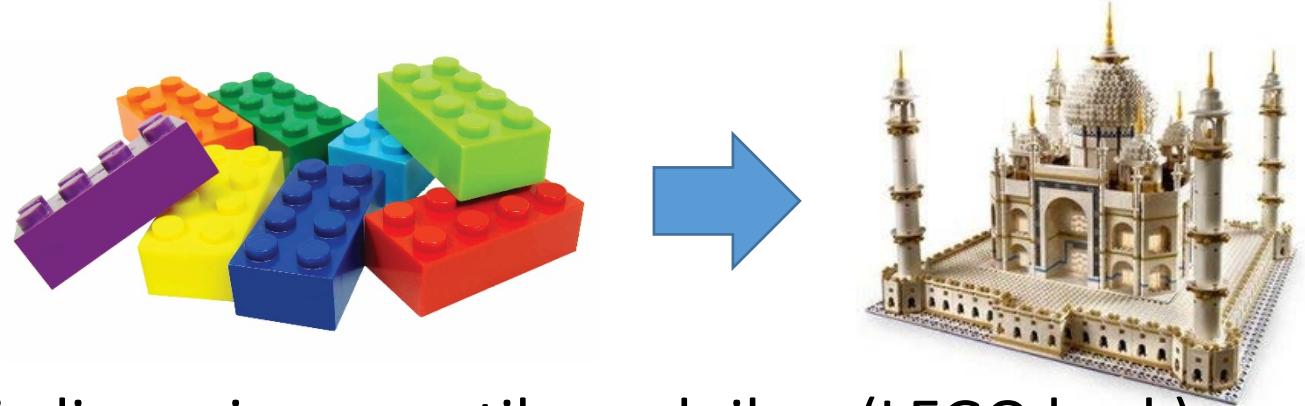


Kako v znanosti opisujemo naravo?

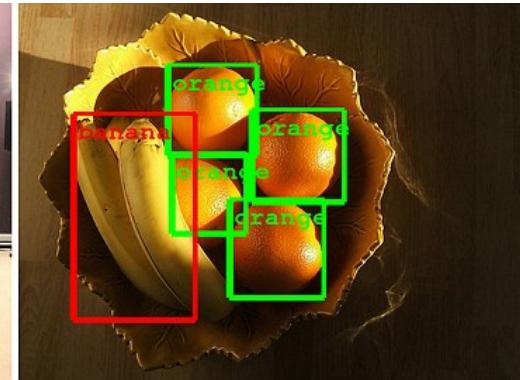
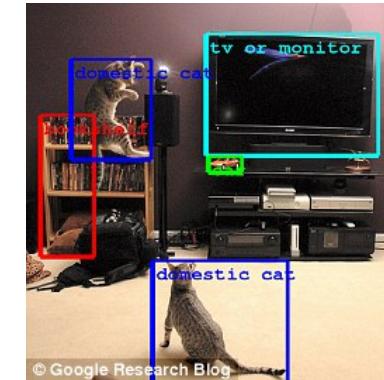
- V znanosti smo si izmislili jezike s katerimi bolj ali manj dobro opisujemo strukture, ki jih najdemo v naravi
 - ...običajno so to matematične formule (modeli)
 - ...če te matematične formule narišemo, izgledajo podobno kot strukture in pojavi v naravi
 - Najpogosteje v znanosti uporabljamo logiko in algebro, ki vsebujejo gradnike, s katerimi opišemo strukture iz narave



Kaj dandanes umetna inteligencia zmore?

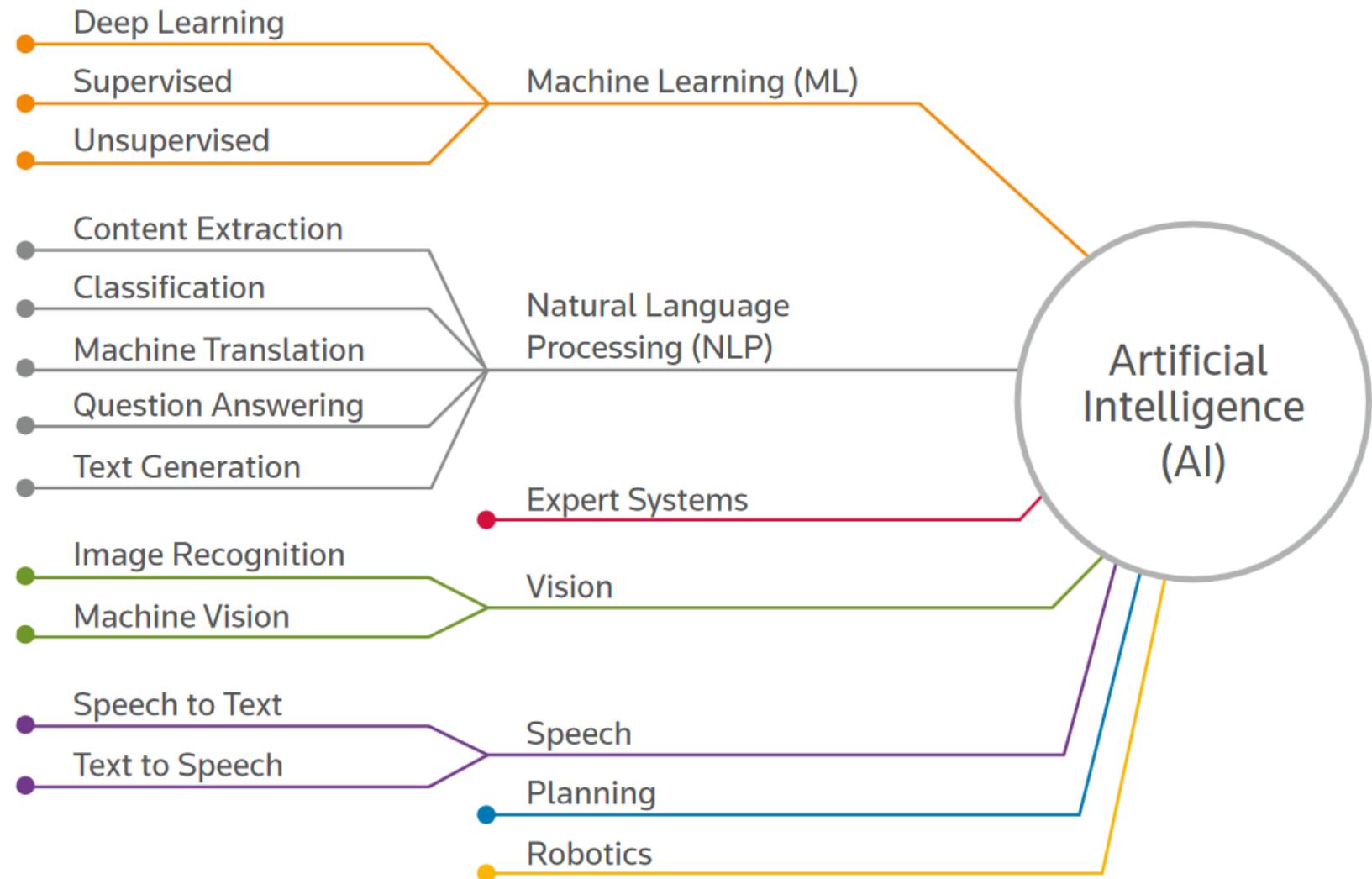


- Umetna inteligencia je zbirka bolj ali manj preprostih gradnikov (LEGO kock), ki jih zlagamo v bolj komplikirane strukture
 - ...take strukture rešujejo naše probleme
- Primeri problemov, ki jih znamo danes reševati, kjer z algoritmi ugotavljamo strukture v podatkih:
 - Prepoznavanje govora
 - Prepoznavanje objektov ali obrazov na sliki
 - Strojno prevajanje
 - Enostavno učenje iz podatkov
- Nikakor pa umetna inteligencia ne zna npr.:
 - Razumeti besedila, ker ima jezik prekompleksno strukturo
 - Računalniki še ne znajo opazovati in se učiti iz lastnih izkušenj



Kaj sestavlja umetno inteligenco?

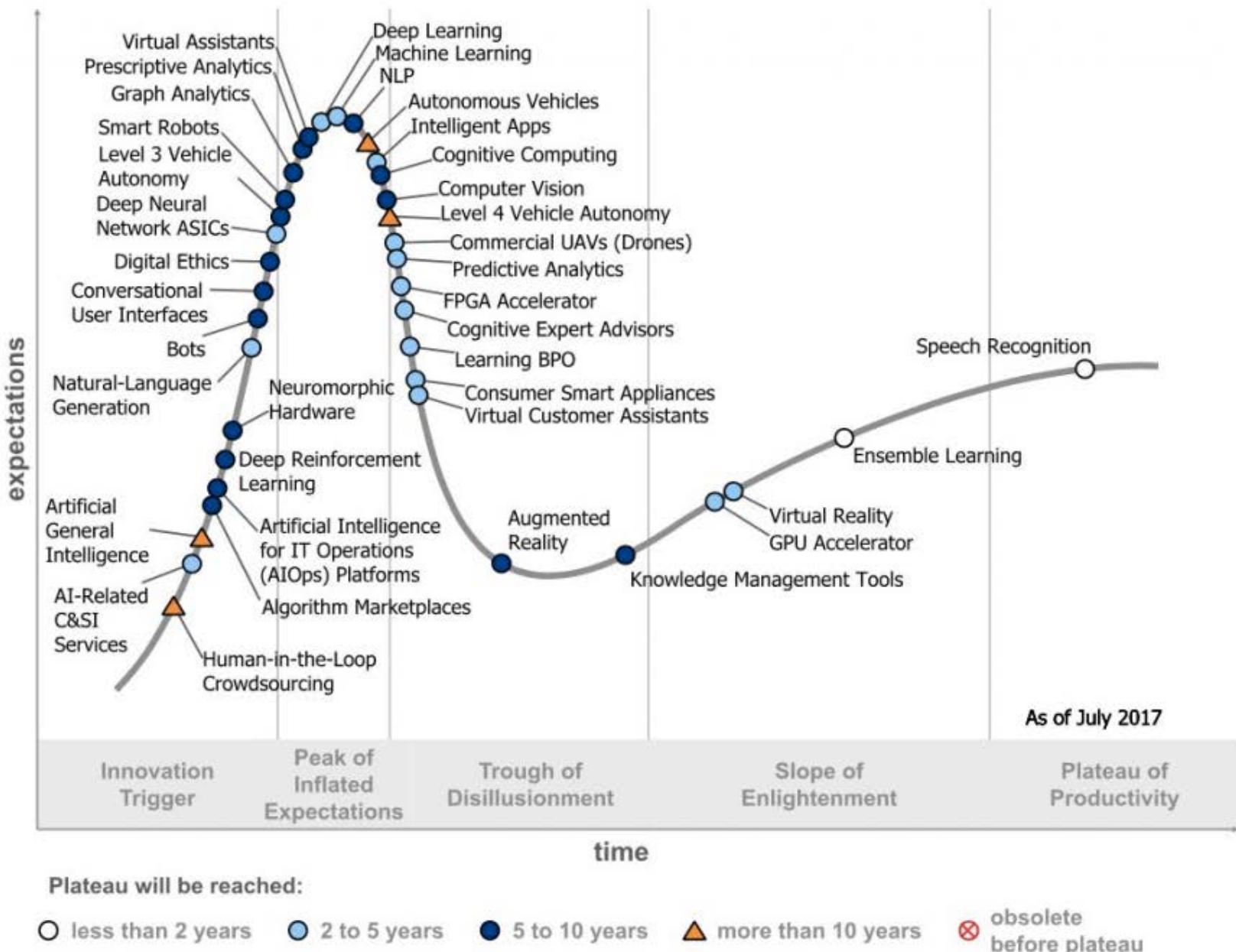
- UI sestavlja vrsta podpodročij, ki jih po vecini sestavljajo enaki ključni gradniki
- ...v zadnjem obdobju strojno učenje omogoča napredek v večini podpodročij umetne inteligence



Kaj sestavlja danasnjo umetno inteligenco?

Gartnerjev 'Hype Cycle' za področje umetne inteligence kaze katere tehnologije **prihajajo**, katere so **popularne** in katere so **zrele**

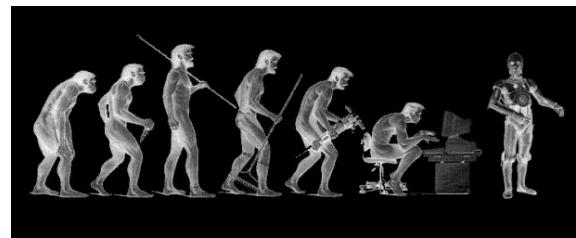
Figure 1. Hype Cycle for Artificial Intelligence, 2017



Zakaj je umetna inteligenca popularna danes?

Kaj je naredilo razliko?

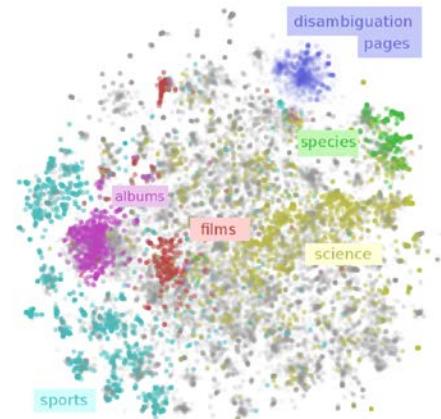
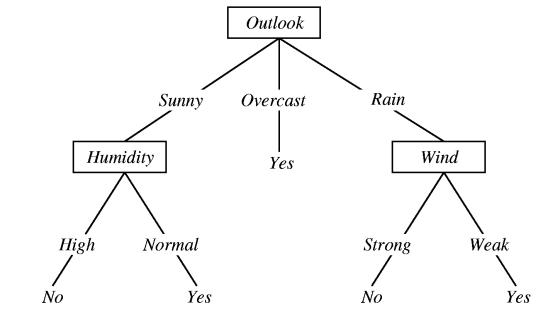
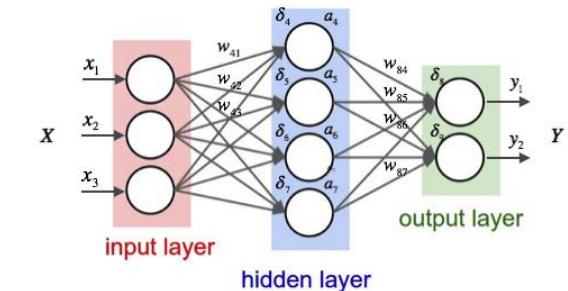
- Ključna razlika je naredil razvoj tehnologije
 - ...računska moč, kapaciteta shranjevanja podatkov, odprta koda
 - ...razvoj algoritmov za strojno ucenje po 2010 (posebej globokega učenja)
 - Razpolozljivost podatkov
 - Nekaj ključnih prebojev (računalniški vid, prepoznavava govora, ...)
- ...in, tehnologija je na razpolago vsakomor za dostopno ceno
 - Sposoben študent lahko danes naredi toliko kot pred 10+ leti cela znanstvena skupnost
 - (*...morda drzna trditev, vendar precej verjetno resnična*)



Kako lahko razumemo
strojno učenje?

Kaj so glavne sestavine strojnega učenja?

- Algoritem za učenje
 - ...ključen del je optimizacijski algoritem
- Jezik za opis modela
 - ...od enostavnih linearnih kombinacij spremenljivk do kompleksnih struktturnih formul
- Jezik za opis sveta (shema podatkov)
 - ...jezik, kako opišemo svet, ki ga skušamo modelirati
 - ...najpogosteje so to spremenljivke (značilke/featurji)
- Podatki o svetu
 - ...konkretna opazovanja iz sveta, ki ga modeliramo
 - ...več je opazovanj, bolj natančno lahko opišemo svet



Komentar o sestavinah strojnega učenja

- Algoritmi za učenje so presenetljivo malo pomembni
 - ...večina algoritmov je dovolj dobrih, da dokaj dobro modeliramo podatke oz. da rešimo nas analitski problem
- Jezik za opis modelov je za vecino tipičnih problemov lahko precej enostaven
 - pomembno je, da z jezikom lahko opisemo svet, ki ga modeliramo
- Ključna sestavina so podatki, se posebej shema podatkov
 - ...ce svet, ki ga skusamo modelirati ne predstavimo dovolj dobro, je algoritom za učenje oz modeliranje bolj ali manj slep
- Torej, morda najpomembnejša sestavina je predstavitev sveta skozi shemo podatkov in dejanske izmerjene podatke
 - ...če tu ne opravimo dobrega dela, potem težko rešimo problem
 - ...globoke nevronske mreže pomagajo pri situacijah, kjer analistik ni nujno najbolj siguren, kako predstaviti svet

Algoritem za učenje: Perceptron, kot morda najenostavnejši in dokaj učinkovit algoritem

- Perceptron je zelo enostaven algoritem, ki dokaj dobro in hitro rešuje lahko tudi dokaj zapletene analitske problem
- Ideja je v tem, da z modelom, ki je ravna črta (oz ravnina) razdelimo en del točk od drugih
- Prikaz delovanja Perceptrona:
<https://www.youtube.com/watch?v=GMDMm-Y65oA>

Algorithm 5 PERCEPTRONTRAIN($D, MaxIter$)

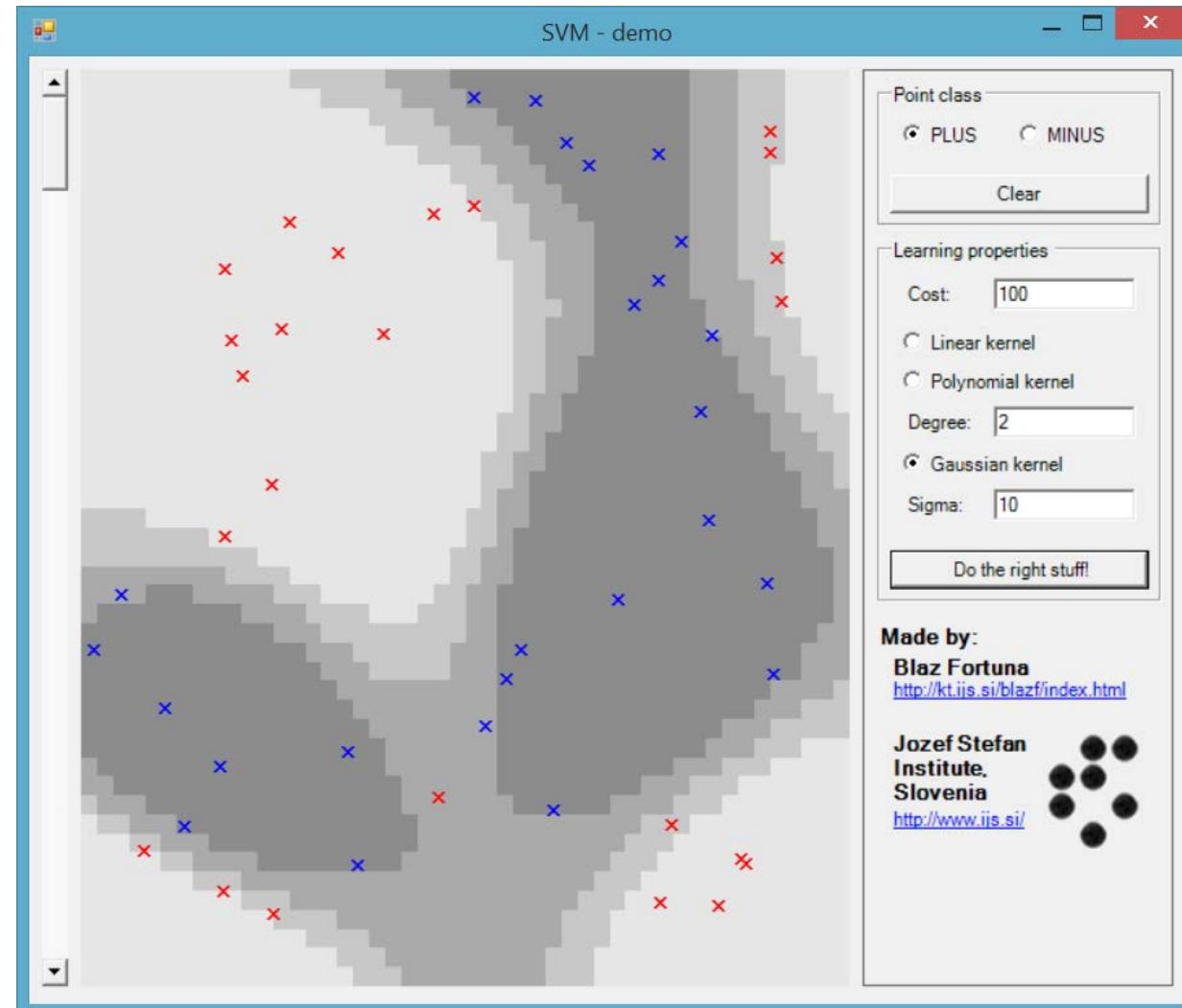
```
1:  $w_d \leftarrow 0$ , for all  $d = 1 \dots D$                                 // initialize weights
2:  $b \leftarrow 0$                                                        // initialize bias
3: for  $iter = 1 \dots MaxIter$  do
4:   for all  $(x,y) \in D$  do
5:      $a \leftarrow \sum_{d=1}^D w_d x_d + b$                                      // compute activation for this example
6:     if  $y_a \leq 0$  then
7:        $w_d \leftarrow w_d + y x_d$ , for all  $d = 1 \dots D$                   // update weights
8:        $b \leftarrow b + y$                                                  // update bias
9:     end if
10:   end for
11: end for
12: return  $w_0, w_1, \dots, w_D, b$ 
```

Algorithm 6 PERCEPTRONTEST($w_0, w_1, \dots, w_D, b, \hat{x}$)

```
1:  $a \leftarrow \sum_{d=1}^D w_d \hat{x}_d + b$                                      // compute activation for the test example
2: return SIGN( $a$ )
```

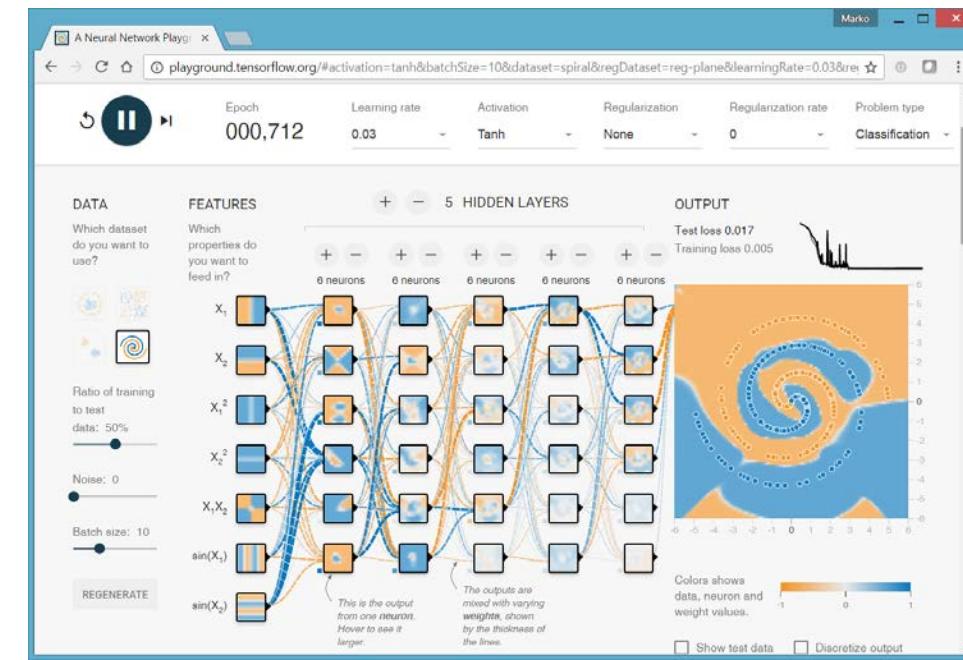
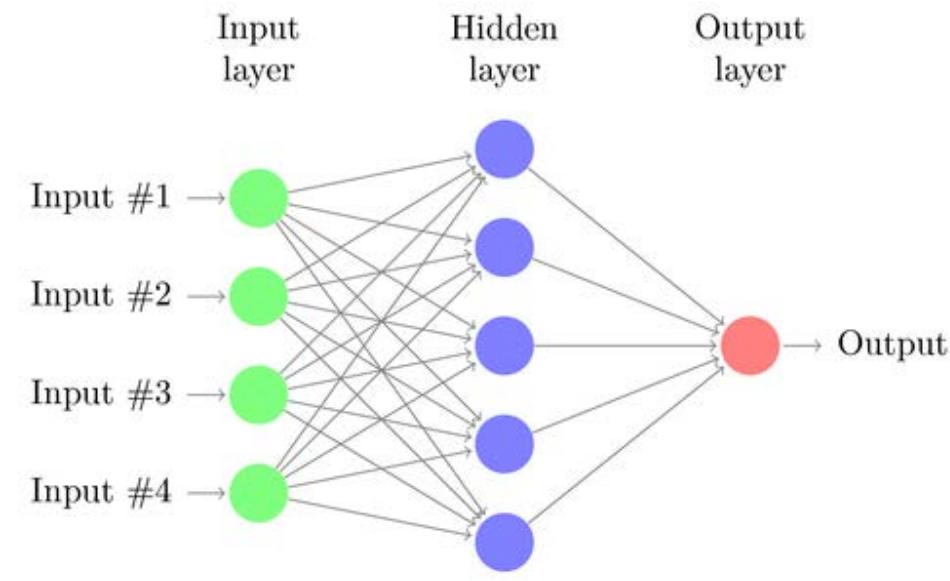
Povečevanje moči jezika: Support Vector Machine

- Metoda podpornih vektorjev (Support Vector Machine) je spremenila pogled na strojno učenje v 90tih
- Metoda je podobna Perceptronu, le da lahko reguliramo moč jezika za opis modela
- Prikaz delovanja SVMja:
 - <https://www.youtube.com/watch?v=3liCbRZPrZA>
 - <https://www.youtube.com/watch?v=9NrALgHFwTo>

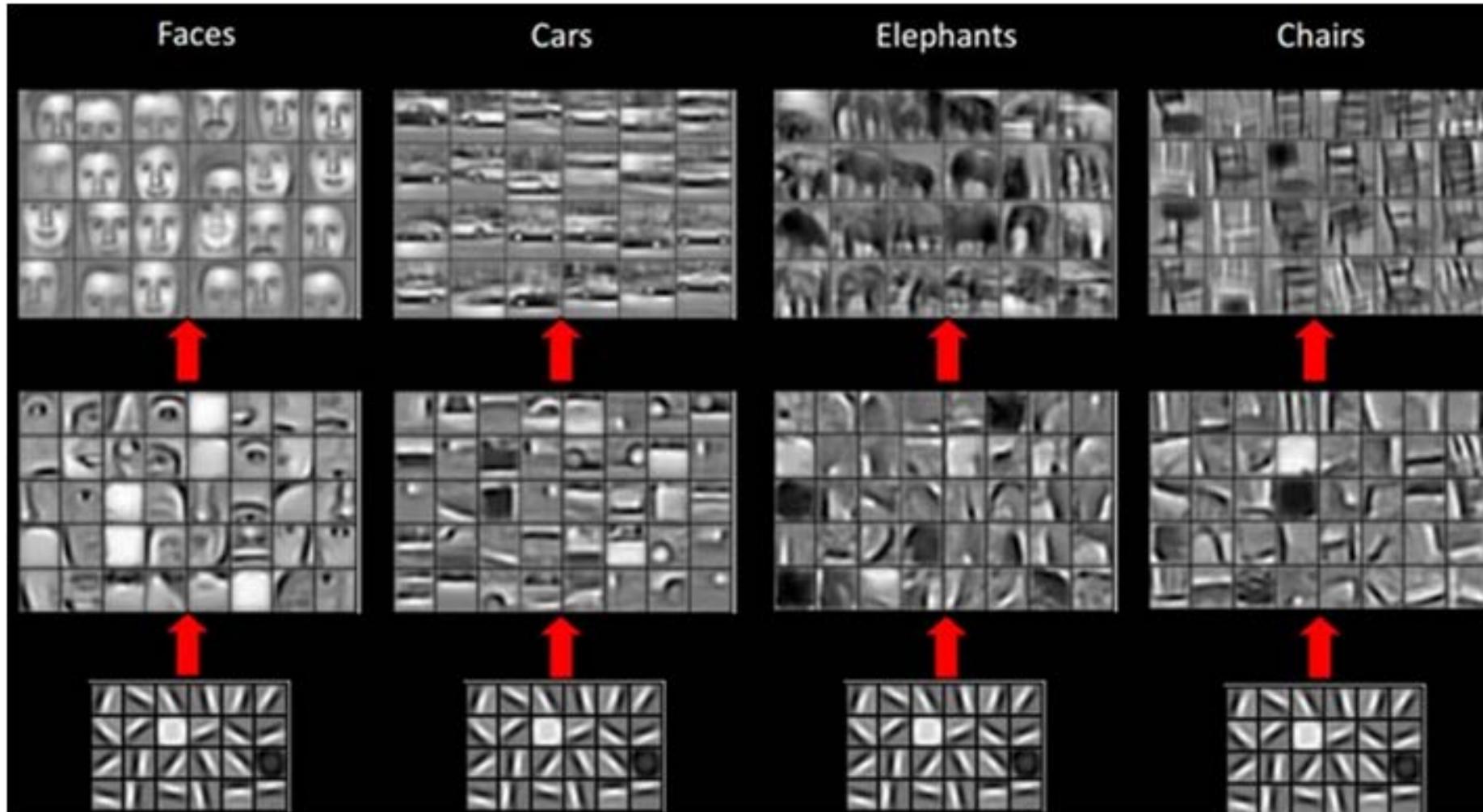


Nevronske mreže

- Nevronske mreže so se pojavile že v 50tih v Rusiji, postale popularne v 90, zares pa so naredile revolucijo po 2010
 - Danasne nevronske mreže so sinonim za "Globoko učenje" (deep learning), ker se je z njimi da resevati problem, ki so bili pred tem zelo težavno rešljivi
- Nevronske mreže so sestavljene iz veliko zelo enostavnih gradnikov (nevroni s funkcijami in sinapse z utežmi):
 - Ko mnogo takih enostavnih gradnikov povezemo v mrežo (ki predstavlja jezik za opis modela), lahko z njimi rešujemo zelo zahtevne probleme
- Ključna lastnost nevronskih mrež je, da si v času učenja razvijejo jezik za opis problema
 - (kar ni lastnost drugih algoritmov za strojno učenje)
- Googlov sistem **TensorFlow** za učenje nevronskih mrež:
 - <http://playground.tensorflow.org>
 - <http://scs.ryerson.ca/~aharley/vis/conv/flat.html>



Kako razvijajo nevronske mreže jezik za opis problema?



Nekaj zaključnih misli



- Umetna inteligenca je zaradi strojnega učenja naredila velik napredek po 2010, vendar je v absolutnem merilu se vedno zelo daleč od sposobnosti človeka pri sposobnostih sklepanja in upravljanja s kompleksnimi problemi (ki so po večine se nedotaknjeni)
 - ...ključni razlog je, da v Alju ne znamo predstaviti prave kompleksnosti sveta
- Umetna inteligenca napreduje, vendar verjetno počasneje kot se zdi
 - Nedavni uspehi višajo pričakovanja in v znanstveni skupnosti je začutiti inflacijo
 - ...vendar, uspehi so odprli široko polje za nove aplikacije
- Kaj so teme prihodnosti?
 - Vzročnost v sistemih (predvsem zaradi pridobivanja globjih vpogledov v kompleksnost opazovanega sveta)
 - Globoko razumevanje besedil (trenutno so besedila prekompleksna, da bi jih zares razumeli)