Strojno učenje
Uvod

Tomislav Šmuc

Sadržaj – Uvod u Strojno Učenje (SU)

- O Strojnom učenju
- Oblici Strojnog Učenja: Tipovi, podjele, primjene
- Osnovni pojmovi

"Doba računarstva"

(1950 - ....)

- Fokus na informacijskom (algoritamskom) objašnjavanju/emulaciji prirodnih fenomena i ponašanja
- Procesima koje podržavaju (inteligentna) bića: skupljanje, procesiranje i korištenje podataka/informacija/znanja
- **Strojevi koji uče (AI/IA)**

Strojno učenje

•Strojno učenje je područje koje se bavi sustavima koji mogu učiti, bez da budu eksplicitno programirani za neki zadatak (A. Samuel, 1959).

• Algoritmi strojnog učenja uče iz podataka i stvaraju nove algoritme/modele.

Zašto (proučavati) strojno učenje?

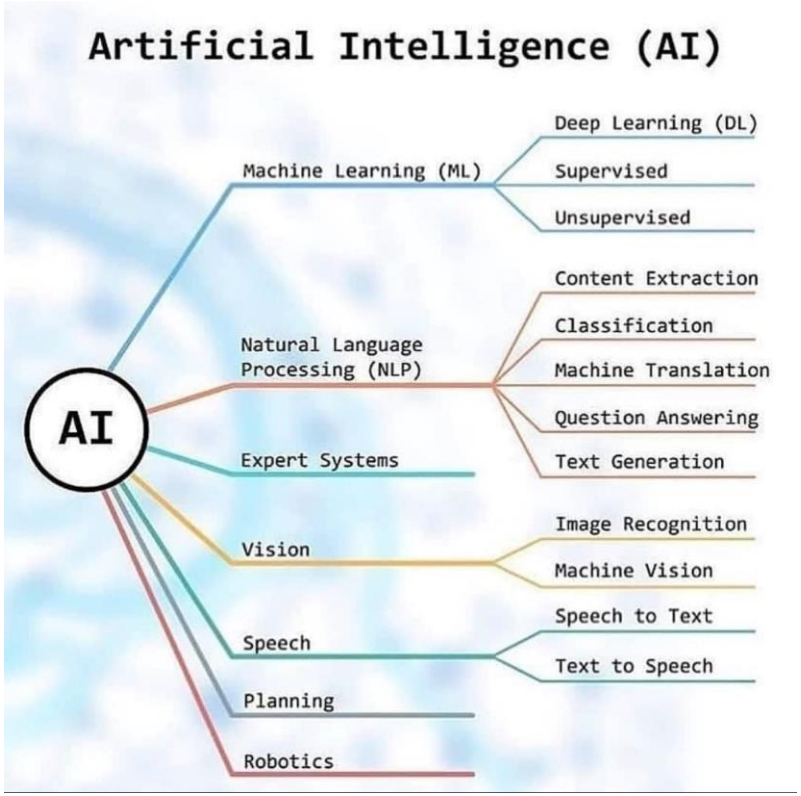
Znanstveni izazovi:

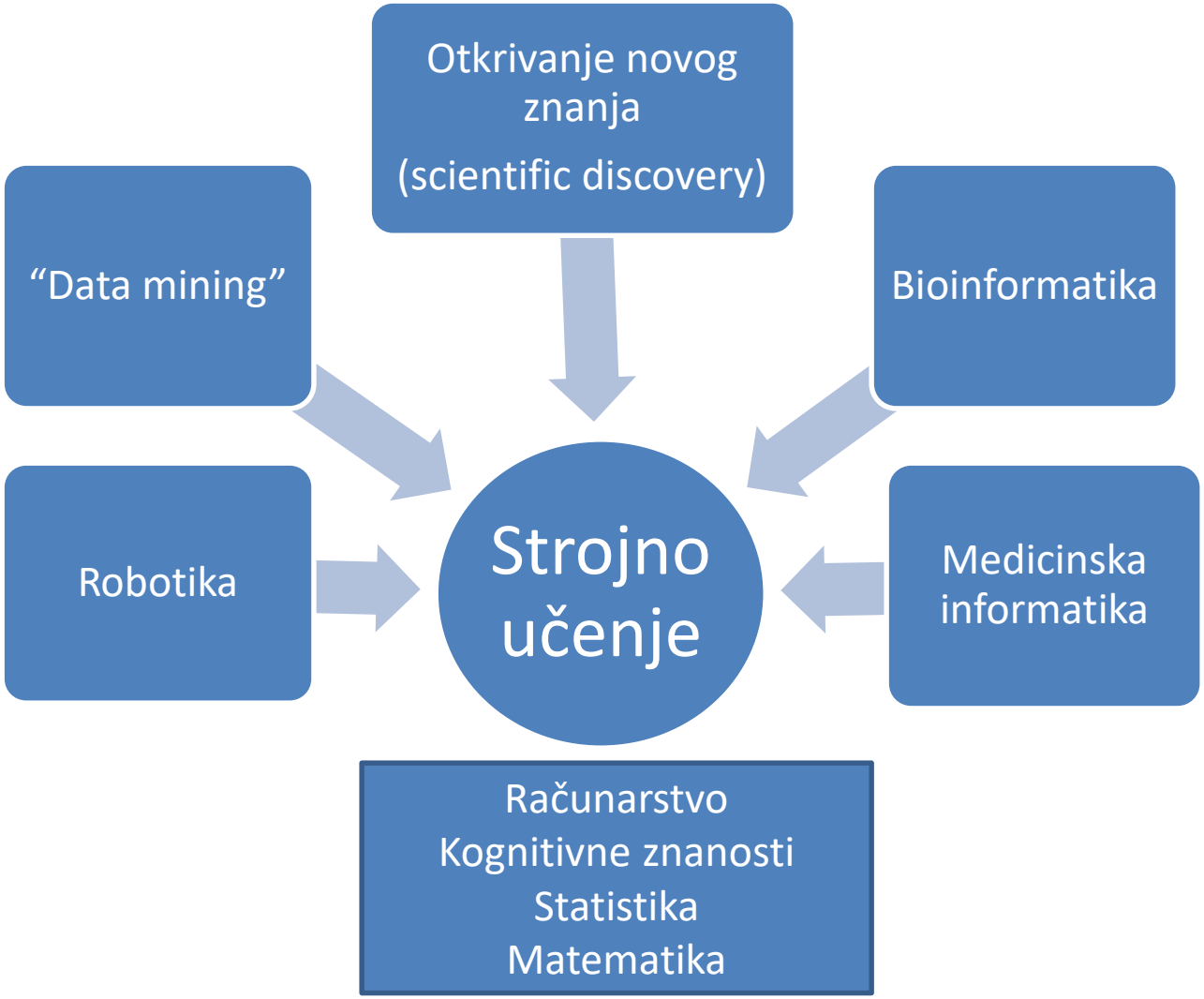
- Kako uče životinje ili ljudi
- Zahtjevi da bi se nešto moglo naučiti: Precizni uvjeti pod kojima su neki ciljevi učenja ostvarivi
- Kako poboljšati učenje – aktivno i pasivno učenje
- Računalne arhitekture strojnog učenja

Korisne primjene:

- Medicinska dijagnostika (EKG, EEG, mamografija)
- Otkrivanje novih zakonitosti u znanosti
- "Spam Filtering", prijevare (npr. kreditne kartice), upadi u rač.sustave
- sustavi za preporučivanje (google; amazon)
- automatsko prevođenje, Question & Answering
- praćenje objekata u slikama & videu
- igranje igara, vožnja automobila

Strojno učenje je (važan) dio AI




Otkrivanje novog znanja
(scientific discovery)
"Data mining"
Bioinformatika
Robotika
Strojno učenje
Medicinska informatika
Računarstvo
Kognitivne znanosti
Statistika
Matematika

Strojno učenje: usko vezane/preklapajuće discipline

- Data mining – rudarenje podataka;
- Data science
- Adaptivno procesiranje signala
- Probabilističko zaključivanje (Bayesove mreže)

Razlika između statistike i strojnog učenja

- **Primijenjena statistika:**
  - obično primjenjujemo na ”manjim” skupovima podataka
  - uloga statističara je velika – računalo je pomoćni alat
  - želimo nešto

- **Strojno učenje - naglasak je:**
  - na automatiziranju otkrivanja i korištenja pravilnosti u podacima;
  - kreiranju iskoristivog (prediktivnog?) modela (automatizacija, AI)
  - metodama koje garantiraju kvalitetu naučenih modela;
  - karakteriziranju što je „naučljivo” i pod kojim uvjetima;

Zaključivanje (en. Reasoning) je proces korištenja postojećeg znanja (ili podataka!) da bi se došlo do zaključaka, predikcija ili objašnjenja

Tri oblika:

Dedukcija → koristi strukturu znanja kojom se generiraju istiniti zaključci /zaključak slijedi iz premisa/ (ekspertski sustavi – GOFAI !)

Indukcija → zaključivanje na osnovu primjera ind. zaključivanje => nema garancije da su zaključci 100% točni (određena vjerojatnost da su točne) (strojno učenje => model i zaključivanje na osnovu primjera)

Abdukcija → Zaključivanje vodi ka najvjerojatnijem mogućem objašnjenju (XAI – objašnjiva AI?)

Dedukcija

- Pravilo (X)=> Y: Svi kolači iz Perine slastičarnice su dobri
- Primjer (X): Ovaj kolač je iz Perine slastičarnice
- Rezultat: Y Ovaj kolač je dobar

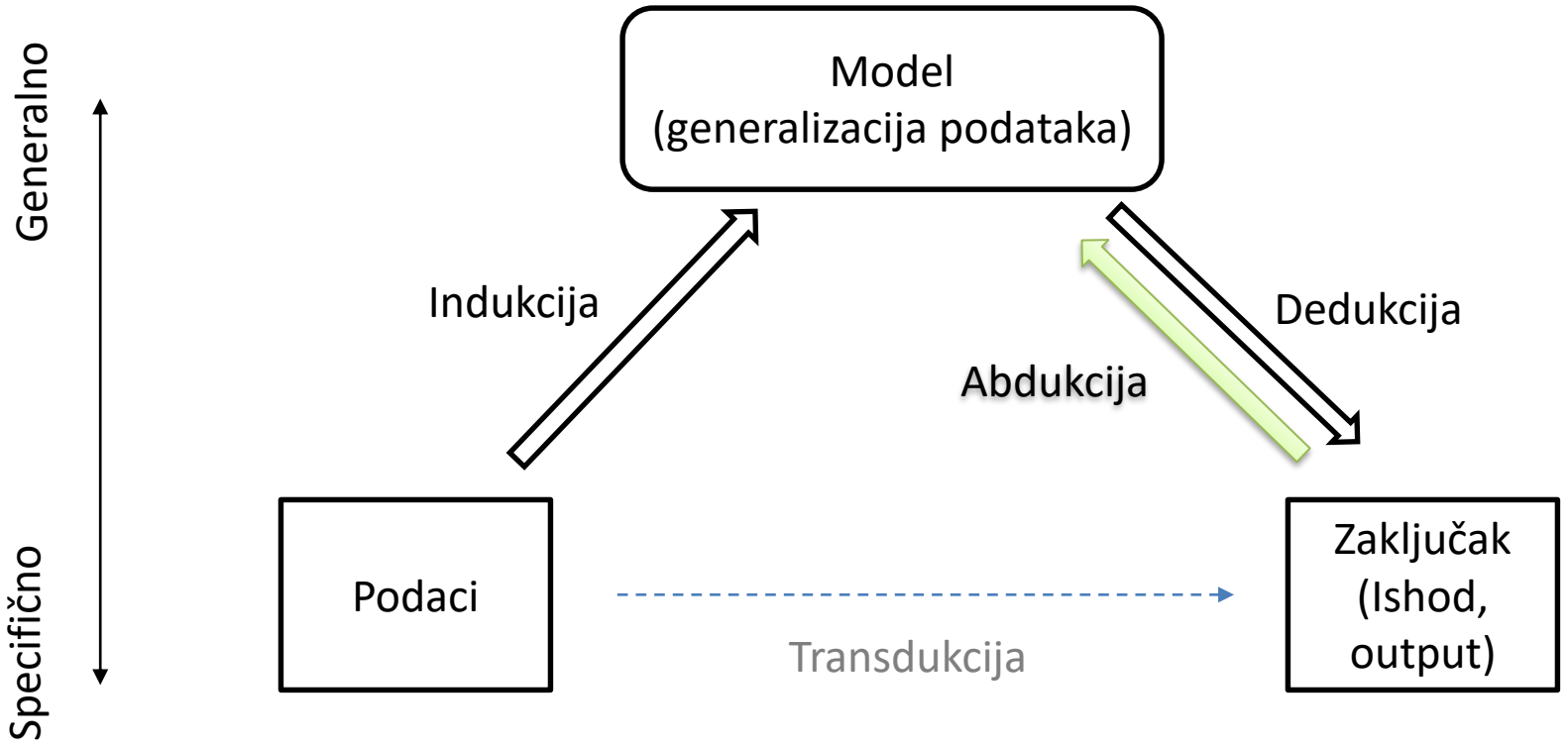
Indukcija

- Primjer: (X) Ovaj kolač je iz Perine slastičarnice...
- Rezultat: Y Ovaj kolač je dobar....
- Pravilo (X)=> Y: Svi kolači iz Perine slastičarnice su dobri

Abdukcija

- Pravilo (X)=> Y: Svi kolači iz Perine slastičarnice su dobri
- Rezultat: Y Ovaj kolač je dobar
- Primjer: (X) Ovaj kolač je iz Perine slastičarnice

Strojno učenje, indukcija i dedukcija



Abdukcija – rijetko korištena u strojnom učenju
(...ali često u znanosti!)

Mogu li strojevi učiti bolje i/ili brže ?

- Neki zadaci su najbolje definirani velikim brojem primjera (npr. dijagnostika)
- Ako imamo velike količine podataka – one mogu kriti vrlo korisne pravilnosti ili prediktivne relacije (data mining)
- U realnim okruženjima karakteristike kompleksnih sustava (npr. promet u gradu) se brzo mijenjaju – software koji ima mogućnost brzog adaptiranja može povećati korisnost/sigurnost ....

Oblici strojnog učenja – I podjela

Nadzirano učenje (en. Supervised Learning)
- eksplicitna informacija o primjerima i vrijednosti njihove ciljne varijable (en. label)
- cilj: napraviti model koji će raditi predikcije na još neviđenim (novim) primjerima (klasificiranje objekata, predikcija prodaje, cijene ...)
  - Klasifikacija
  - Regresija
  - Predikcija – predviđanje (en. Forecasting) vremenske serije

Nenadzirano učenje - Unsupervised Learning:
- imamo samo primjere, bez ikakve anotacije ili povratne informacije o njihovoj kategorizaciji
- cilj : grupirati primjere, otkriti neku strukturnu pravilnost u podacima, projekcija podataka u niže-dimenzionalne prostore

  - Grupiranje - Clustering
  - Otkrivanje-detekcija iznimaka (Outlier detection)
  - Kompresija podataka

Oblici strojnog učenja (nastavak)

Učenje s podrškom (povratnom vezom - en. reinforcement learning):

- dobivamo "nagradu" s odgodom ukoliko akcije predviđene našim modelom daju uspjeha
- nema eksplicitnog nadzora
- obično vezano uz učenje sekvenci akcija (igre, roboti!) **(AlphaGO!)**

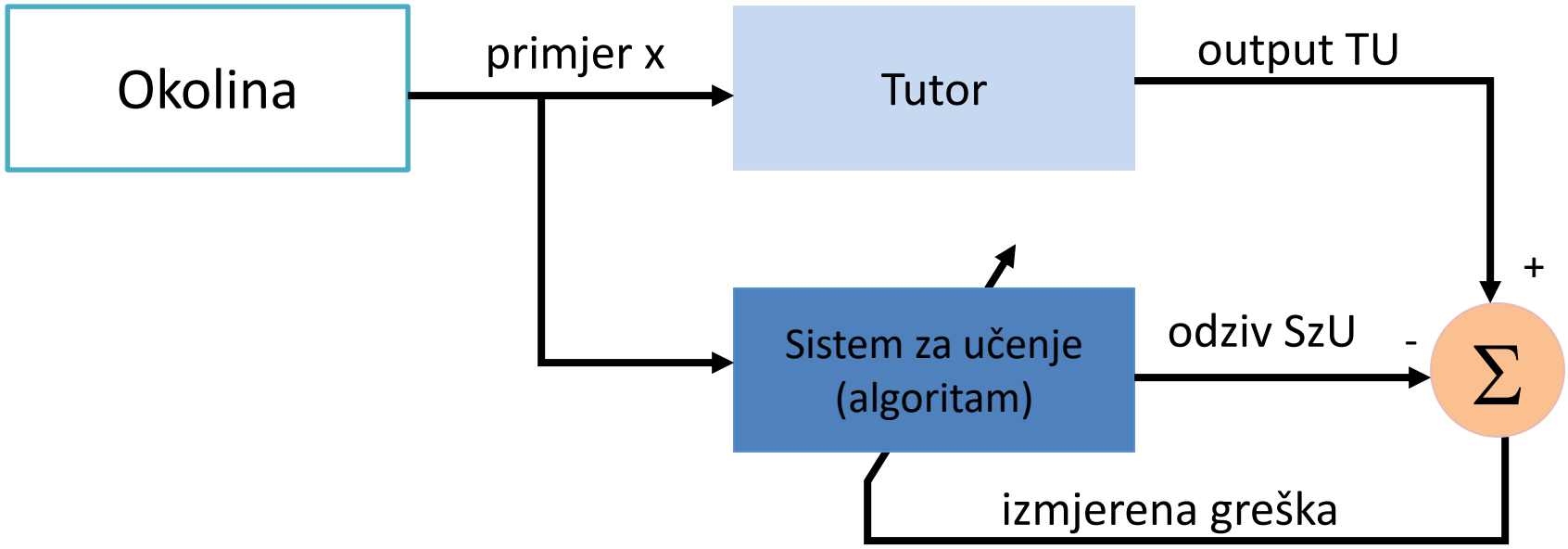
Učenje skupa pravila

- iz danog većeg skupa mjerenja ili transakcija, otkriti učestalo ponavljane uzorke mjerenja (npr. asocijativna pravila)

{Mlijeko, Jogurt} → Kruh *[support=5%, confidence=80%]*

Nadzirano učenje (učenje uz učitelja/tutora)

- Znanje – svodi se na primjere ($input_i$,$output_i$) = (x_i,y_i)
- Cilj: minimizirati grešku => razliku između stvarnog outputa (učenik) i željenog outputa (učitelj)



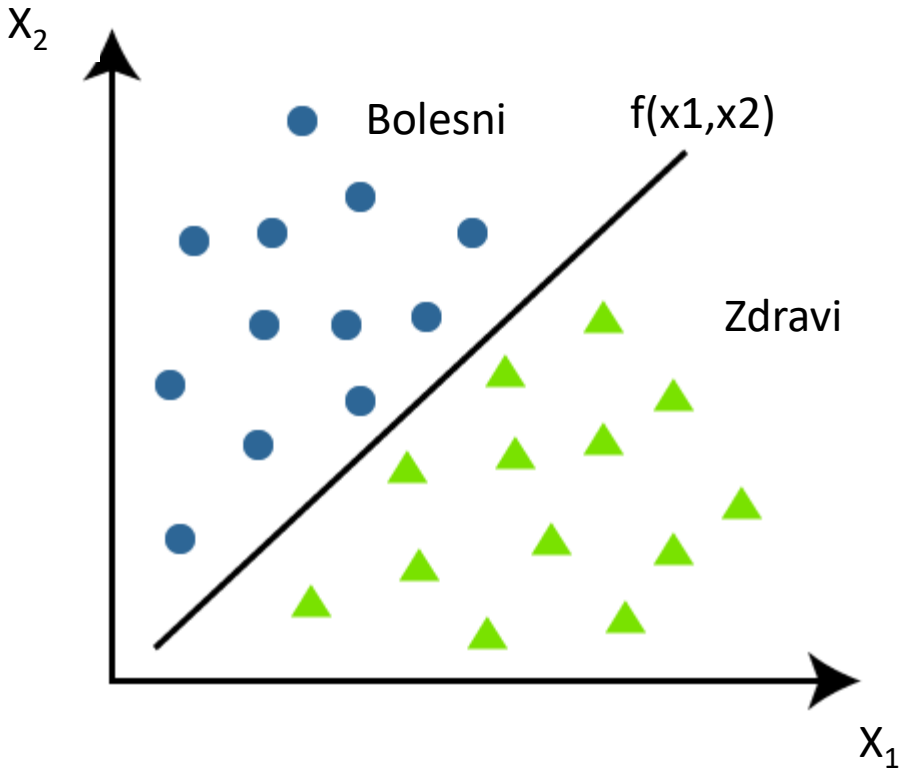
Nadzirano učenje (učenje uz učitelja/tutora)

Klasifikacija:

- Output je kategorička varijabla
- Input može biti bilo što
- Cilj - da naučeni model odabire korektnu klasuž
- Primjeri: klasifikacija slika, segmentiranje slika, klasifikacija tekstova, postavljanje dijagnoze

Klasifikacija

- Primjer: Dijagnostika
- Razlikovanje Bolesnih od zdravih ljudi na bazi izmjerenih vrijednosti X_1 i X_2



Diskriminant:

$y = w_0 + w_1 * x_1 + w_2 * x_2 > f(x_1, x_2) \Rightarrow$ Bolesni

$y = w_0 + w_1 * x_1 + w_2 * x_2 \leq f(x_1, x_2) \Rightarrow$ Zdravi

Klasifikacija: primjene

- **Održavanje strojeva**: prepoznavanje kvara monitoriranjem procesa rada (temperatura, zvuk, drugi fizički parametri koji opisuju ponašanje stroja...)
- **Medicinska dijagnostika:** simptomi => bolest
- **Biometrija:** prepoznavanje/autentikacija korištenjem fizičkih ili nekih drugih karakteristika
- **Prepoznavanje stila pisanja ili slikanja:** korištenjem posebnih reprezentacija za slike/tekst...

- ...

Regresija - aproksimacija funkcija

Cijlna varijabla je kontinuirana veličina
Cilj – aproksimirati nepoznatu funkciju $f_n(\mathbf{x})$

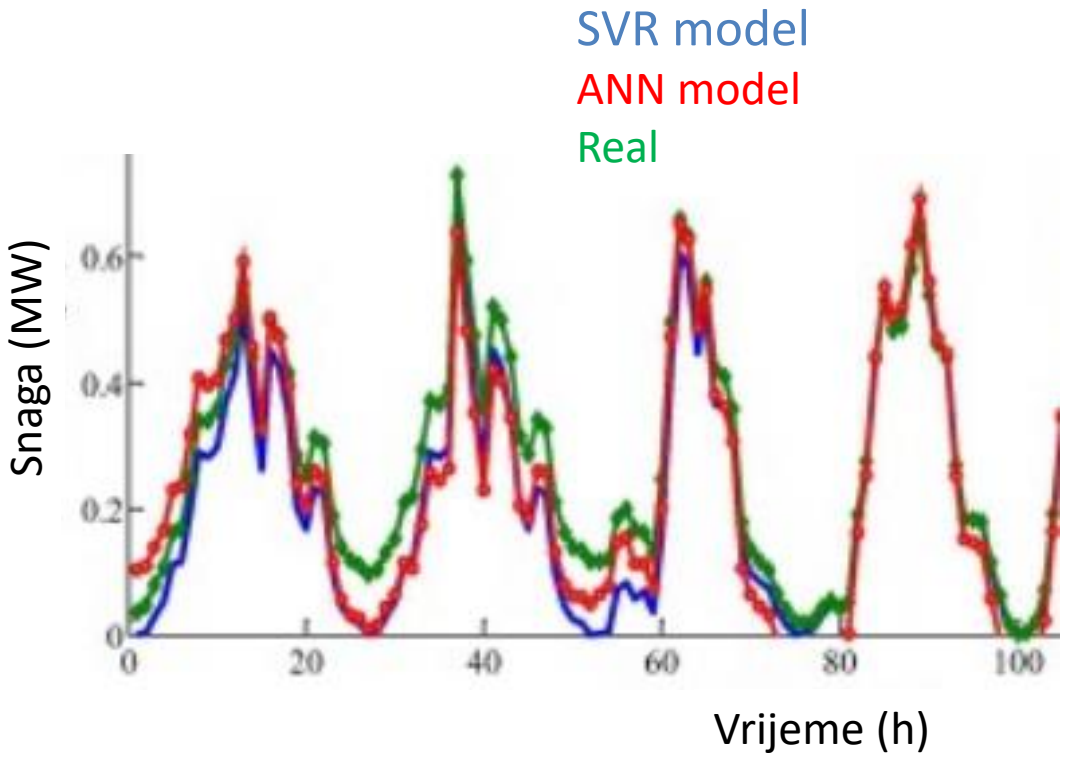
- tako da je preslikavanje $F_a(\mathbf{x})$ realizirano sistemom za učenje približno isto kao i ono nepoznate funkcije $f_n(\mathbf{x})$:

$$|F_a(\mathbf{x})-f_n(\mathbf{x})|<\varepsilon \quad \forall x$$

Primjeri:
Opis input-output odnosa nekog nekog složenog sustava kod kojeg je output kontinuirana veličina

Regresija

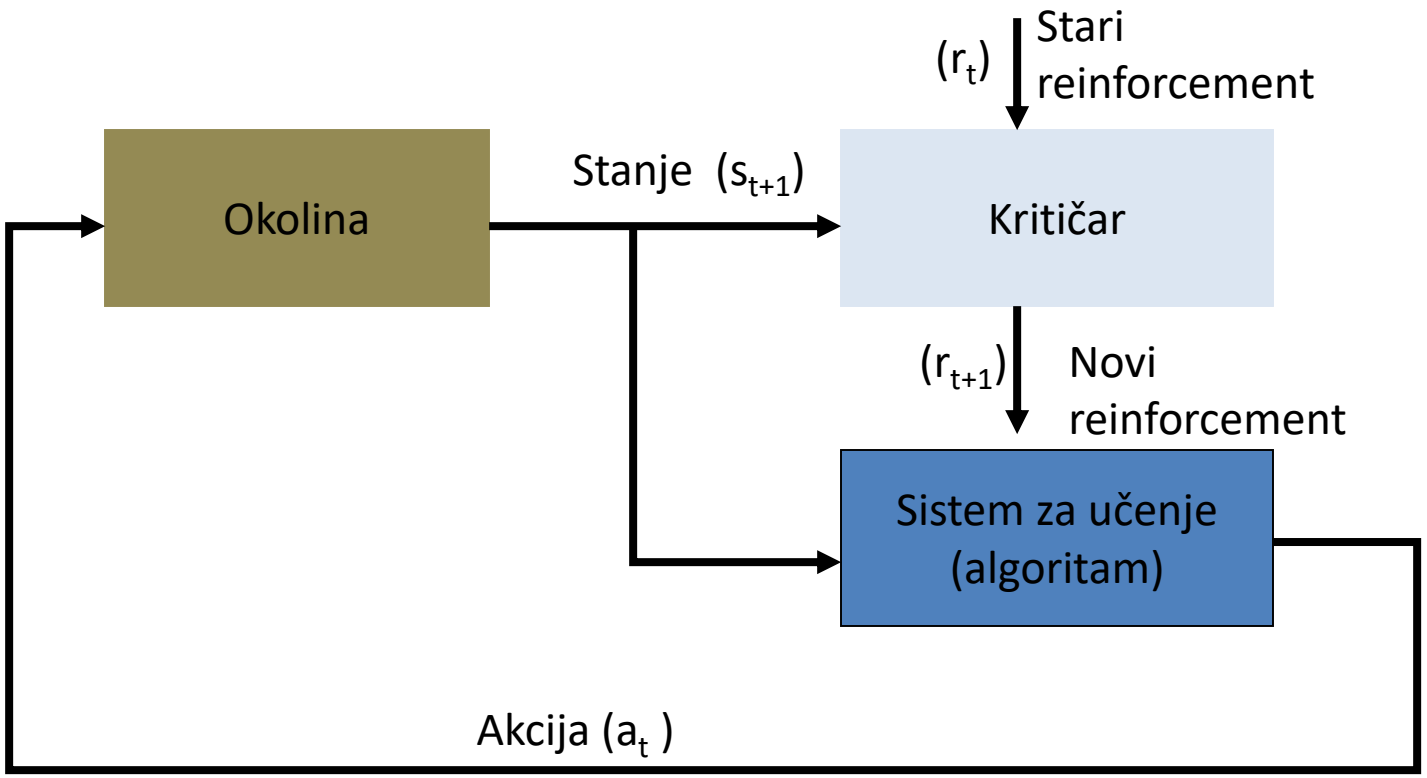
- Primjer: Predviđanje generirane snage solarne elektrane



y= f(vremenski uvjeti = T, p, UV zračenje, vjetar...)

Učenje podrškom - reinforcement learning – učenje niza akcija

- učenje – kroz interakciju s okolinom (ili preko simulacije interakcije)
- istraživanje odnosa stanja i akcija
- Cilj: maksimiziranje akumuliranih budućih “reinforcements” (rewards)



Primjer: kontrola procesa – učenje funkcije kontrole

- Podesiti parametre nekog kontrolnog sustava tako da na kraju vodi proces na optimalan način

Primjeri učenja kontrole (Reinforcement Learning)

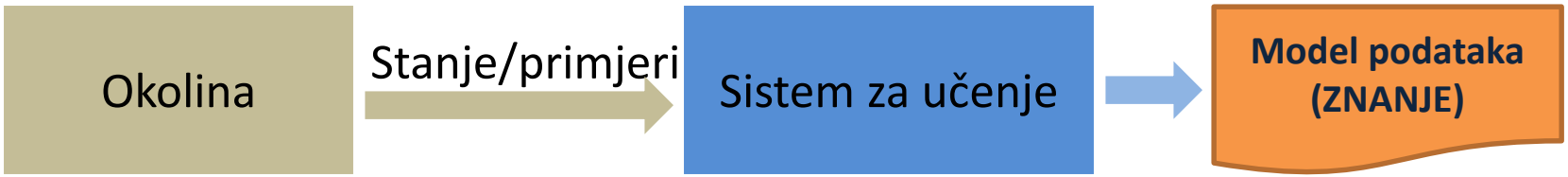
- Navigacija robota
- Učenje akcija koje maksimiziraju/optimiraju output nekog pogona (tvornice), procesa trgovanja dionicama na burzi….
- Učenje igranja igara (akcija/poteza u igrama kao šah, GO)

Karakteristike ovih problema:

- Odgođeno nagrađivanje umjesto instantnog - za dobre akcije (tzv. temporal credit assignment problem)
- Nema pravog nadzora – supervizije procesa učenja (primjeri su u obliku stanje, akcija)
- Postoji potreba za aktivnim istraživanjem prostora stanja i akcija

Nenadzirano učenje

- nema učitelja niti kritičara
- samo-organizirajuća svojstva
- mjera dobrote/kvalitete – neovisna o zadatku
- potrebno je naći pravilnosti u podacima – otkriti klase automatski – te otkriti što one znače ili predstavljaju!



Primjeri:
- segmentacija/grupiranje primjera
- otkrivanje izuzetaka (outlier detection) u podacima

Nenadzirano učenje

Primjene:

- Marketing: grupiranje potrošača prema uzorcima kupnje, demografskim karakteristikama...
- Astronomija: Grupiranje zvijezda prema fizikalnim svojstvima, sastavu..
- Osiguranje: identifikacija grupa korisnika prema policama, premijama i korištenju osiguranja
- Planiranje gradova: prometni uzorci, kućanstva, trošenje energije...

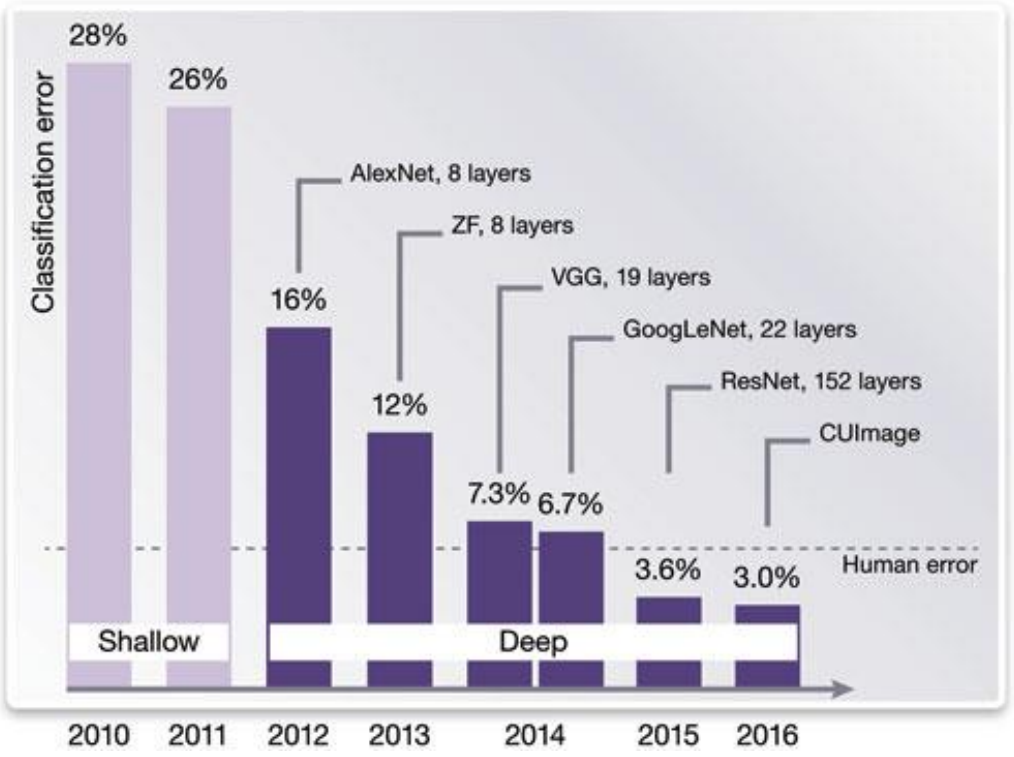
Novije primjene strojnog učenja (Duboke rekurentne mreže, transformeri, RL...)

- Text-to-speech synthesis
- Language translation
- Image caption generation
- Learning to program...

Novi modeli:

- AlphaFold2.0 – proteinska struktura
- GPT-3 – generator teksta (175*10^9 parametara)

ILSVRC – ImageNet Large Scale Visual Recognition Challenge



Vrlo poopćena definicija strojnog učenja

Poboljšati kvalitetu izvršavanja zadatka *t*,

***s* obzirom na definiranu mjeru uspješnosti *q*,**

na osnovu dostupnog iskustva (znanja i podataka) *e*

Primjeri:

t: prepoznavanje slova u tekstu
q: postotak točno prepoznatih slova
e: baza slika sa ručno napisanim tekstovima (pixel po pixel)

t: kome ponuditi novu policu osiguranja
q: postotak točno određenih kandidata iz baze poznatih primjera
e: baza podataka o klijentima osiguravajućeg društva

t: igranje šaha (povlačenje pravih poteza)
q: omjer dobivenih i izgubljenih igara
e: igranje igara protiv samog sebe

Stvaranje modela iz skupa primjera

1. **Pretpostavka**: postoji distribucija vjerojatnosti p(**x**, y) čiji pravi oblik ne znamo

   Obično na raspolaganju imamo tek (konačan) skup primjera (parova **x** i y) (obično uz određeni "šum" ili greške):

   $\{\boldsymbol{x}_1, y_1; \ldots; \boldsymbol{x}_n, y_n; \ldots; \boldsymbol{x}_N, y_N\}$

   Primjeri iz skupa su i.i.d. (independently and identically distributed)

2. Koristeći samo $\{\boldsymbol{x}_1, y_1; \ldots; \boldsymbol{x}_n, y_n; \ldots; \boldsymbol{x}_N, y_N\}$ generiramo model *M* koji za neki $\{\mathbf{x}_n\}$ daje $\{z_n\}$ (po mogućnosti tako da je što češće $z_n = y_n$, t.j. da model što manje griješi)

3. Zamislimo situaciju da je neki novi skup primjera $\{\mathbf{x}_k, y_k\}$ "izvučen" iz distribucije p(**x**, y). Naš model koristimo za određivanje $\{z_k\}$. (ukoliko znamo stvarni y_k možemo odrediti i grešku na novom, tzv. testnom skupu)

Osnovni problemi u ovom postupku:

- Koji algoritmi i u kojim uvjetima dobro aproksimiraju funkcije i pod kojim uvjetima?
- Kako broj primjera utječe na točnost modela?
- Kako kompleksnost reprezentacije mogućih modela/hipoteza utječe na točnost?
- Kako šum ili greške utječu na točnost?
- Koji su teoretski limiti “naučljivosti”?

Generalizacija

Osnovni problem train/test uzorkovanja i učenja:

- opasnost da ćemo dobiti model koji odlično "radi" samo na podacima na kojima je *istreniran*

  => **overfitting** („pretreniranosti" ili suprotno od generalizacije)

- učenje primjera napamet = savršen rezultat na skupu za učenje
  = slučajno pogađanje na novim primjerima

Generalizacija ~ sposobnost dobre predikcije na novim primjerima !

Kapacitet = kompleksnost (prostora) hipoteza

Osnovna pretpostavka induktivnog učenja

=> generalizacija je moguća

Ako naš model dobro radi na većini podataka na kojima je treniran, ako model nije prekompleksan – vjerojatno je da će dobro raditi i na novim podacima (napomena: ako su generirani iz iste distribucije kao i podaci za treniranje...)

Ova empirijska izjava je inače formalizirana u više desetaka godina istraživanja u području teorije strojnog učenja...
(SRM&VC-dimension, PAC learning, Occam's razor, MDL....)

Induktivna pristranost (inductive bias)

Posljedica osnovne hipoteze induktivnog učenja:

- generalizacija je moguća samo ako unaprijed radimo neke pretpostavke o konačnom izgledu hipoteze/modela odnosno ako je kompleksnost mogućih hipoteza/modela ograničena !
- **algoritam/učenik koji nema ugrađenu induktivnu pristranost ne može generalizirati !**

Induktivna pristranost – kako se postiže?

- Učenje napamet – nema induktivne pristranosti!

Induktivna pristranost u algoritmima strojnog učenja

- Linearna regresija – odnos između ulaznih varijabli i ciljne varijable je linearan
- Stabla odlučivanja – preferiranje plitkih stabala
- Naivni Bayes – ulazne varijable su međusobno nezavisne
- SVM – maksimiziranje margine izmedju različitih klasa primjera

- Regularizacija – penaliziranje složenosti modela (/hipoteza) u funkciji cilja (strojno učenje ~ optimizacijski problem)

Što čini metodu/algoritam strojnog učenja ?

- reprezentacija modela/ciljne funkcije

  (linearna funkcija, polinom, neuralna mreža....)

  (sa nepoznatim vrijednostima slobodnih parametara)
- algoritam pretraživanja/optimizacije

  (strojno učenje = pretraživanje/optimizacijski problem)
- metoda procjene greške na "neviđenim" primjerima – skalarna funkcija cilja kojom ćemo kvantificirati kako dobro radi/generalizira naš naučeni model

---

Cilj: podesiti vrijednosti parametara modela tako minimiziramo vrijednost funkcije cilja

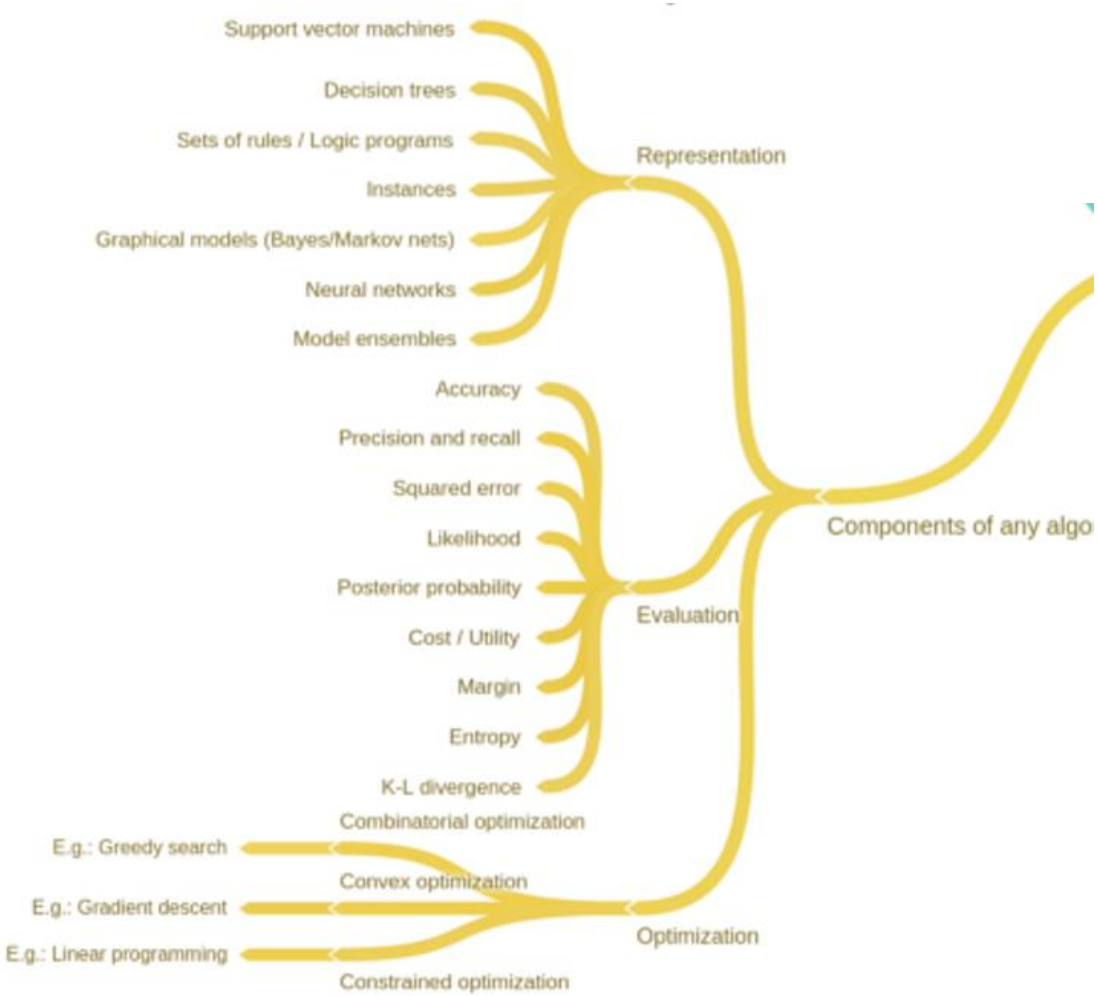
Funkcije cilja

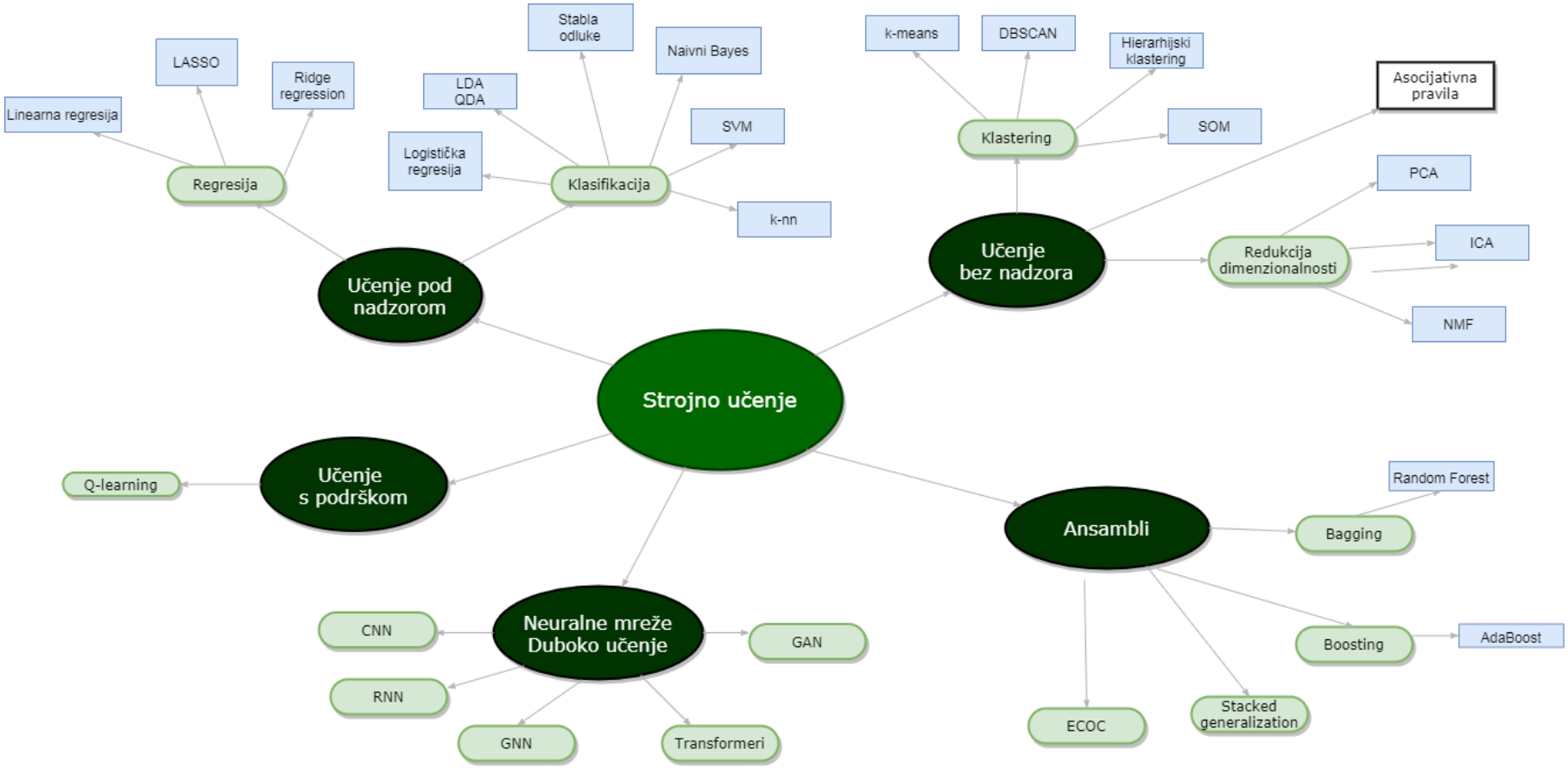
$$\phi(\mathbf{X},Y;\mathbf{w}) = L(\mathbf{X},Y|\mathbf{w}) + R(\mathbf{w})$$

w – parametri modela
L (**X**,Y|**w**) – mjera greške modela (loss function, npr RMSE, log-loss)
R (**w**) – funkcija kojom se penalizira kompleksnost modela – npr. $||\mathbf{w}||^2$

- oblik $\phi(\mathbf{X},Y;\mathbf{w})$ - opet govori o tome da želimo postići što manju grešku na skupu za učenje (L) ali isto tako da to želimo postići sa što jednostavnijim modelom **w** !

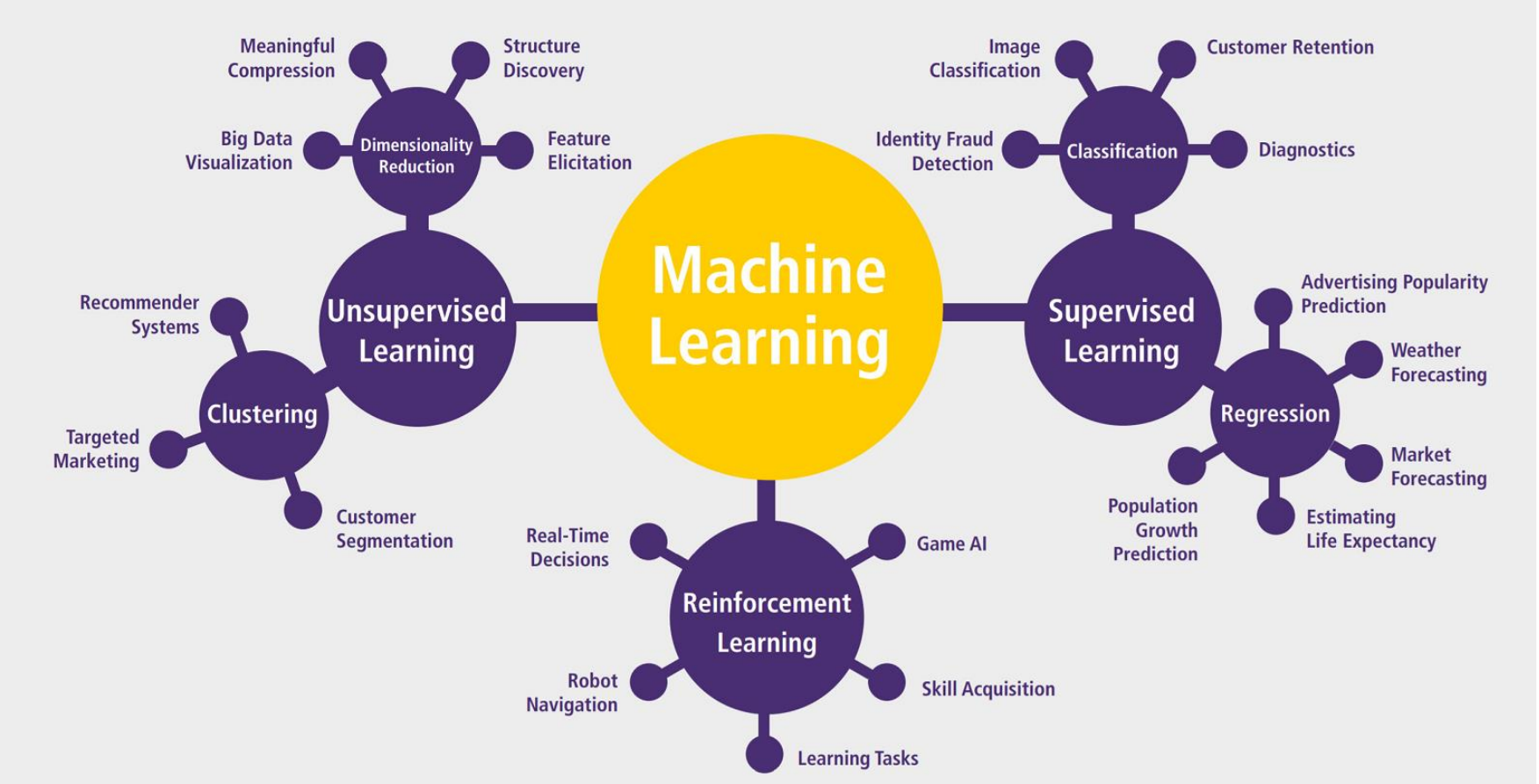
Strojno učenje @ math.pmf


Support vector machines
Decision trees
Sets of rules / Logic programs
Instances
Graphical models (Bayes/Markov nets)
Neural networks
Model ensembles
Representation
Accuracy
Precision and recall
Squared error
Likelihood
Posterior probability
Cost / Utility
Margin
Entropy
K-L divergence
Evaluation
Combinatorial optimization
E.g.: Greedy search
Convex optimization
E.g.: Gradient descent
E.g.: Linear programming
Constrained optimization
Optimization
Components of any algo



AI Primjene

Učenje reprezentacija | Distribuirane reprezentacije | Računalni vid | Jezični modeli (NLP) | Robotika | XAI - Objašnjiva umjetna inteligencija



Jha, V. An overview of machine learning techniques