

# **Strojno učenje 1**

## **Sažetak**

Prof. dr. sc. Jan Šnajder

Ak. god. 2022./2023.

# Plan

① Osvrt

② Kako dalje?

# Plan

① Osvrt

② Kako dalje?

# Plan nastave

- Uvod u strojno učenje
- Osnovni koncepti, regresija
- Linearni diskriminativni modeli, logistička regresija
- Stroj potpornih vektora, jezgrene funkcije, neparametarski modeli
- Opcionalno: Ansambli

## Međuispit

- Procjena parametara
- Bayesov klasifikator i probabilistički grafički modeli
- Grupiranje
- Vrednovanje modela
- Opcionalno: Odabir značajki

## Završni ispit

# Ishodi učenja

- **definirati** osnovne pojmove strojnog učenja
- **objasniti** teorijske pretpostavke, matematičke osnove te prednosti i nedostatke temeljnih algoritama nadziranog i nenadziranog strojnog učenja
- **primijeniti** temeljne algoritme klasifikacije, regresije i grupiranja na jednostavnije probleme
- **primijeniti** postupak odabira modela i evaluacije točnosti naučenog modela
- **analizirati i usporediti** temeljne algoritme strojnog učenja s obzirom na njihove komponente i računalnu složenost
- **razlikovati** temeljne pristupe strojnom učenju (generativni/diskriminativni, parametarski/neparametarski, bayesovski/frekventistički) te objasniti poveznice
- **procijeniti** prikladnost temeljnih algoritama strojnog učenja za zadani zadatak
- **oblikovati i implementirati** jednostavnije algoritme za klasifikaciju, regresiju i grupiranje

## U konačnici, želimo da...

- 1 Razumijete **osnovne ideje** i **razine apstrakcije**
- 2 Razumijete tipičnu **matematiku** ispod toga
- 3 Shvatite **algoritme i metode** tako da ih implementirate ili ispitajte na jednostavnim primjerima
- 4 Izgradite svoju **kartu područja** i dobijete **putokaze** za samostalno učenje

Vaši (lanjski) dojmovi. . .



# Izvođači

## **Predavač:**

- Prof. dr. sc. Jan Šnajder

## **Asistenti:**

- Zoran Medić, mag. ing. (glavni asistent)
- Ana Barić, Ivana Bašljan Puceković, Ivan Bilić, Filip Karlo Došilović, David Dukić, Antea Hadviger, Fran Jelenić, Josip Jukić, Miha Keber, Josipa Lipovac, Domagoj Plušćec

## **Demosi:**

- Dario Deković, Mladen Džida, Ivan Furač, Josip Hrvatić, Helena Ladić, Ivan Martinović, Mirta Moslavac, Magdalena Mucić, Leon Novački, Luka Pavlović, Lovre Petešić, Ivan Rep, Matej Škrabić, Matea Vasilj, Janko Vidaković



# Plan

① Osvrt

② Kako dalje?

# Kako dalje?

- Položite završni :-)
- Imate vrlo dobre osnove iz SU, čak i ako zaboravite detalje!
- U praktičnom smjeru:
  - ▶ razmislite gdje u svojem radu (diplomski rad?) možete upotrijebiti metode strojnog učenja
  - ▶ implementirajte sustav (možete se osloniti na postojeće alate)
  - ▶ napravite ispravno vrednovanje sustava (unakrsna provjera!)
- U teorijskom smjeru:
  - ▶ naučite teme koje nismo stigli obraditi
  - ▶ kompletno pročitajte jedan udžbenik iz strojnog učenja
  - ▶ produbite svoje znanje učenjem naprednijih tema koje vas zanimaju
  - ▶ kada postanete suvereni u nekoj temi, počnite čitati znanstvene članke


# Pet škola strojnog učenja

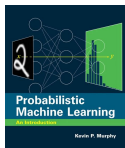
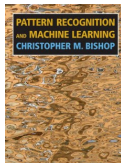
Pedro Domingos: The Master Algorithm



Škola	Temelji	Glavni algoritam
<b>Symbolists</b>	Logika, filozofija	Indukcija
<b>Connectionists</b>	Neuroznanost	Backpropagation
<b>Evolutionaries</b>	Evolucijska biologija	Genetičko programiranje
<b>Bayesians</b>	Statistika	Probabilističko zaključivanje
<b>Analogizers</b>	Psihologija	Jezgreni strojevi

# Literatura

- Ethem Alpaydin:  
*Introduction to Machine Learning*,  
3rd edition, MIT Press, 2015.
- Christopher Bishop:  
*Pattern Recognition and Machine Learning*,  
Springer, 2006.
-  Kevin P. Murphy: *Probabilistic Machine Learning: An Introduction*, MIT Press, 2022.



# Dodatna dodatna literatura

## Osnovno:

- Hastie, Tibshirani, Friedman: *Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, 2003.
- Deisenroth, Faisal, Ong: *Mathematics for Machine Learning*. Cambridge University Press, 2020.
- Marsland: *Machine Learning: An Algorithmic Perspective*, Chapman and Hall/CRC, 2009.
- Duda, Hart, Stork: *Pattern Classification*, Wiley-Interscience, 2000.

## Napredno:



- Murphy: *Probabilistic Machine Learning: Advanced Topics*, MIT Press, 2022.
- Mohri, Rostamizadeh, Talwalkar: *Foundations of Machine Learning*, MIT Press, 2012.
- Shalev-Shwartz, Ben-David: *Understanding machine learning: From theory to algorithms*, Cambridge University Press, 2014.

## Specifično:

- Koller, Friedman: *Probabilistic Graphical Models: Principles and Techniques*, MIT Press, 2009.
- Goodfellow, Bengio, Courville: *Deep Learning*, MIT Press, 2016.

# Korisne web-stranice

- Data Science StackExchange  
Podatkovna znanost, analiza podataka, strojno učenje  
<https://datascience.stackexchange.com/>
- CrossValidated QA  
Statistika, dubinska analiza i vizualizacija podataka  
<http://stats.stackexchange.com/>

## Vruće teme?

- Duboko učenje
- Pojačano učenje (reinforcement learning)
- Bayesovske metode
- Probabilističko programiranje
- Strukturno učenje
- Polunadzirano učenje
- Neparametarski bayesovski pristupi
- Strojno učenje za velike skupove podataka
- ...

- Videopredavanja na [videolectures.net](https://www.videolectures.net)
- Tutoriali na konferencijama [NeurIPS](#) i [ICML](#)
- Podcast: [Machine Learning Street Talk](#)
- Ljetne škole (MLSS, ESLLI, ...)

## Srodni predmeti na FER-u

- Analiza i pretraživanje teksta (2)
- Analiza velikih skupova podataka (2)
- Dubinska analiza podataka (2)
- Duboko učenje 1 (2)
- Duboko učenje 2 (3)
- Ekspertni sustavi (2)
- Meko računarstvo (1)
- Multivarijatna analiza podataka (2)
- Neuronske mreže (1, 3)
- Optimiranje evolucijskim računanje (1, 3)
- Poslovna inteligencija (2)
- Računarska statistika (2)
- Raspoznavanje uzoraka (1, 3)
- Strojno učenje 2 (3)
- Statistička analiza podataka (1)
- Uvod u znanost o podatcima (1)



# Analiza i pretraživanje teksta (TAR)

<http://www.fer.unizg.hr/predmet/apt>

Većina ljudskog znanja pohranjena je u **nestrukturiranom, tekstnom obliku**. Zbog ogromne i rastuće količine digitalnih tekstnih podataka, sustavi za **analizu i pretraživanje teksta** postali su neizostavan dio suvremene informacijsko-komunikacijske infrastrukture. Zbog složenosti i višeznačnosti prirodnog jezika, analiza teksta netrivialan je zadatak, koji se oslanja na postupke **obrade prirodnog jezika, računalne lingvistike i strojnog učenja**. Ovaj predmet daje cjelovit pregled tradicionalnih i naprednih postupaka za analizu i pretraživanje teksta, uključivo postupka temeljenih na statističkoj obradi jezika i strojnome učenju.

- Predmet se predaje isključivo na **engleskome jeziku**
- U drugom dijelu predmeta čitaju se i diskutiraju **znanstveni članci**
- Rade se **projekti** koji se prezentiraju ih u obliku znanstvenog rada

# Strojno učenje 2

...od ak. god. 2023./2024. ☺

- Statistical learning theory
- Kernel machines
- Bayesian machine learning
- Causal machine learning

- Teorijski predmet (teoremi i dokazi)
- Numerus clausus (odabir po ocjenama iz SU1 i motivacijskom pismu)
- Predavanja na engleskom jeziku
- Full-text pismeni ispiti + usmeni

Hvala!