Neurinfo: Anomáliadetekció

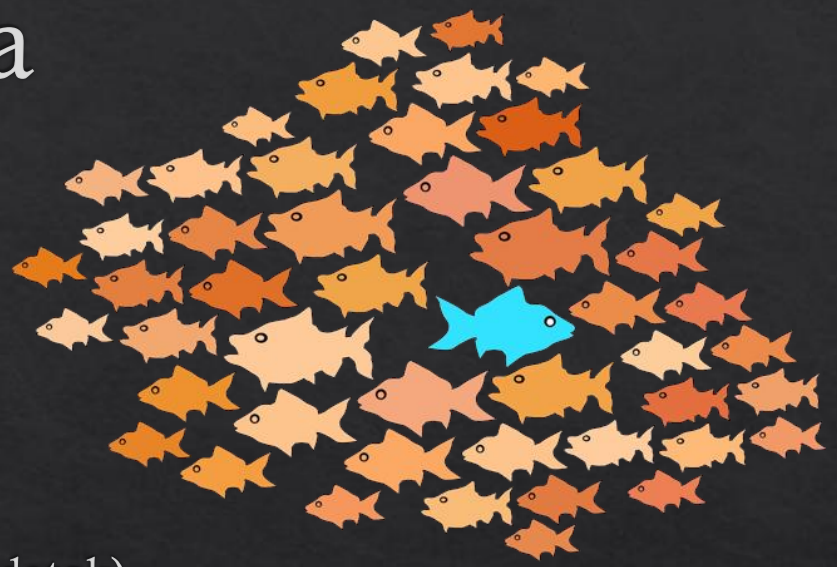
Stippinger Marcell
HUN-REN Wigner Fizikai Kutatóközpont
Komputációs Tudományok Osztálya

2024. december 17.

Rólam

- 2011: MSc fizikából (BME)
  MSc in Engineering (EC Lille, France, double degree programme)
- 2017: PhD fizikából (BME)
  - Statisztikus fizika, hálózatok és hibrid fázisátalakulás számítógépes szimulációja C++-ban
  - Gépi tanulás kurzus Python nyelven
- 2016-tól: HUN-REN Wigner FK, Komputációs Tudományok Osztálya
  - Elméleti idegtudomány, komplex rendszerek, gépi tanulás
  - Felhőalapú számítások, GPU programozás (Python-ban :D)
  - 2024. szeptember óra mesterséges intelligencia nagykövet

Az előadás tartalma



- Az anomália definíciója
- Az anomália típusai
  - Motiváció, példák
- Anomáliák detektálása:
  - Statisztika (kiszóró pont, szisztematikus hiba, modellillesztési feladatok)
  - Gépi tanulás (klasszifikációs algoritmusok, SVM, IF, *k*NN, LOF, TOF)
  - Mély mesterséges neurális hálók (VAE, CNN)
- Összefoglalás

Mi az anomália?

anomália főnév *..iát, ..iája*

- 1. (sajtónyelvi, rosszalló) **A szabályostól, a törvényszerűtől való eltérés; rendellenesség.** Közlekedési anomáliák. || a. (orvostudomány) Vmely szervnek a rendestől, szabályostól eltérő fejlődése, viselkedése, működése. Anomália mutatkozik a szívműködésben.
- 2. (sajtónyelvi, rosszalló) Visszaélés. Közigazgatási, választási anomáliák. Súlyos anomáliáknak jutottak nyomára. A szolgálati kocsit magáncélra igénybe venni anomália.
- 3. (fizika) Gravitációs anomália: eltérés a nehézségi erő fokának átlagos eloszlásától annak következtében, hogy az illető helyen a föld felülete alatt könnyebb v. nehezebb tömegek, talajfajták vannak-e. || a. (fizika) Mágneses anomália: a földmágnesség erős fokozódása a földkéreg vmely részén, ami nagy tömeg mágneses érc jelenlétére mutat.
- 4. (csillagászat) Vmely bolygó v. üstökös szögtávolsága a napközeltől.

Forrás: A magyar nyelv értelmező szótára

Anomália (természettudomány)

In the natural sciences, especially in atmospheric and Earth sciences involving applied statistics, an **anomaly** is a persisting deviation in a physical quantity from its expected value, e.g., the systematic difference between a measurement and a trend or a model prediction. (…)
A group of anomalies can be analyzed

- spatially, as a map, or
- temporally, as a time series.

It should not be confused for an isolated outlier. There are examples in atmospheric sciences and in geophysics.

Forrás: Wikipedia
(letöltve 2024.10.28.)

Anomália (természettudomány)

Anomaly refers to something that deviates from what is standard, normal, or expected. It can indicate an irregularity or an inconsistency in data, behavior, or patterns, often prompting further investigation to understand its cause or significance.

Forrás: ChatGPT
(model 4o-mini)

Pozitív definíciók anomáliára

Meglepetés (surprise, innovation)

- A *t* időben megfigyelt jel eltérése a *t*-nél korábbi információkra épülő optimális előrejelzéshez képest, ezt használhatjuk a modell javítására (vö. Kálmán szűrő).
  - Ha az agy ilyet észlel, az EEG-n eltérési negativitás (mismatch negativity) mérhető.
- A *t*-ig bezárólag elérhető összes információn alapuló optimális modelltől való eltérés a maradvány (reziduális), ami a véletlen hatásokból adódik.

Újdonság (novelty)

- Intelligens lény felismeri a korábban nem látott érzékszervi mintázatot
  - Világmodell frissítése (szemantikus memória), ellentétes tapasztalatok gyűjtése (epizodikus memória)
- Nagy pozitív vagy negatív hasznosság esetén további komputáció (Kahnemann-féle II. alrendszer, lassú gondolkodás)
- Ellentétes folyamata a megszokás (habituáció)

Miért detektáljunk anomáliát?

- Adatminőség javítása
  - Segíti az adatelemzést
  - Csökken a zaj, az adat jobban reprezentálja a valóságot
- Javul a gépi tanulás teljesítménye
  - Jobb döntések, jóslatok

Hogyan detektáljunk?

- Ábrázolás
- Statisztikai teszt
- Gépi tanulás

Az anomália típusai

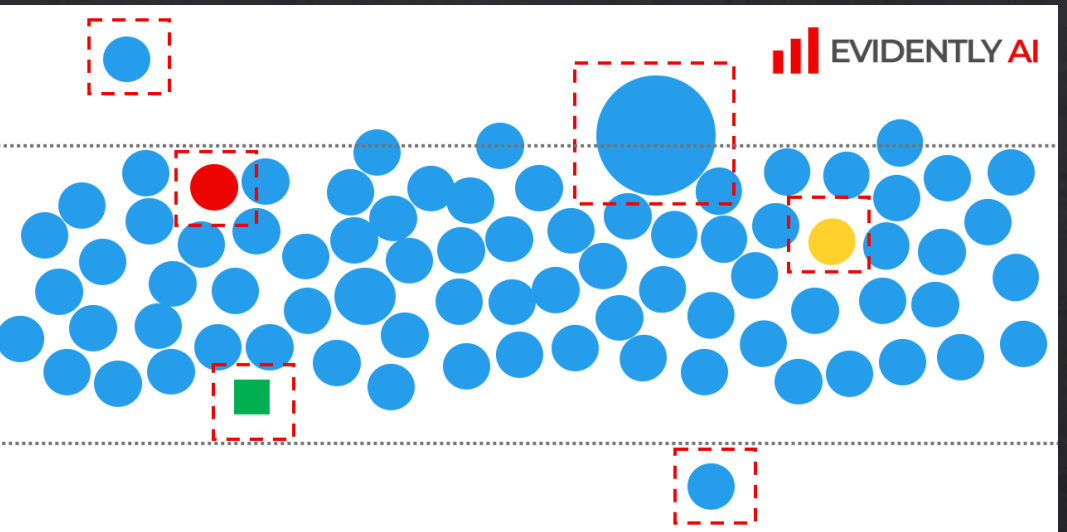
Az anomália oka

- Akaratlan: mérés, adatrögzítés hibája, zaj
  - Véletlen
  - Szisztematikus
- Szándékos/valódi: határozott cselekmények vagy események miatt
  - Egyedi előfordulások
  - Trend (pl. növekvő eladások karácsony közeledtével)

Érintett adatpontok kapcsolata

- Egyedi (kiszóró pont)
- Kontextuális: izolálva nem anomális, de az adott környezetből mégis kilóg
  - Pl. háztartási energia felhasználásában csúcs, amikor senki nincs otthon (az mért érték amúgy hasonlíthat a reggeli/esti csúcshoz)
- Kollektív: több dolog együttállása anomális
  - Pl. DoD vagy DDoS támadás (egyik támadó gép forgalma sem kirívó önmagában)
- Rendszerszintű: minden adatpont érintett

Kiszóró pont (outlier)

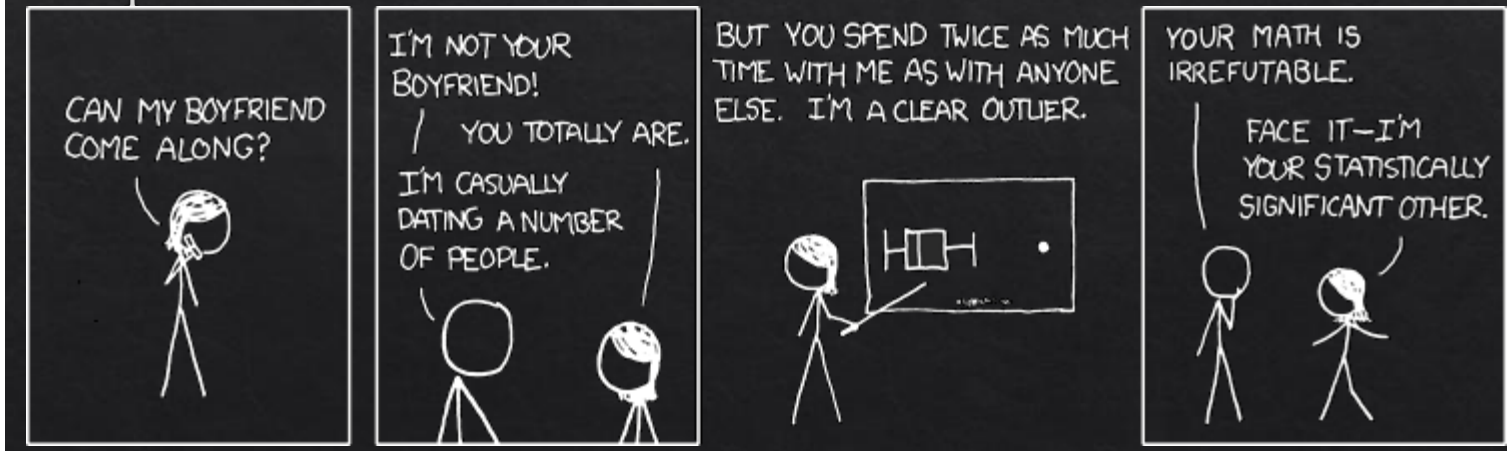


Példák kiszóró pontra.

“I know it when I see it.”

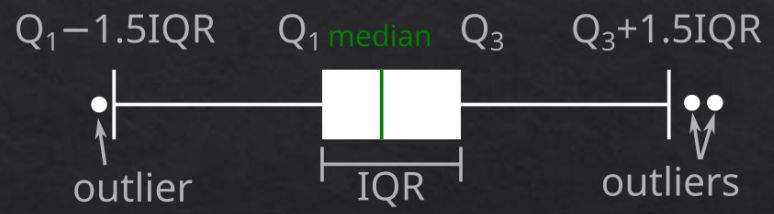
1,93
1,96
1,91
194
1,89
…

https://xkcd.com/539/



Kiszóró pont (outlier)

Definíció (kiszóró pont): olyan adatpont, ami szignifikánsan eltér a többi megfigyeléstől.

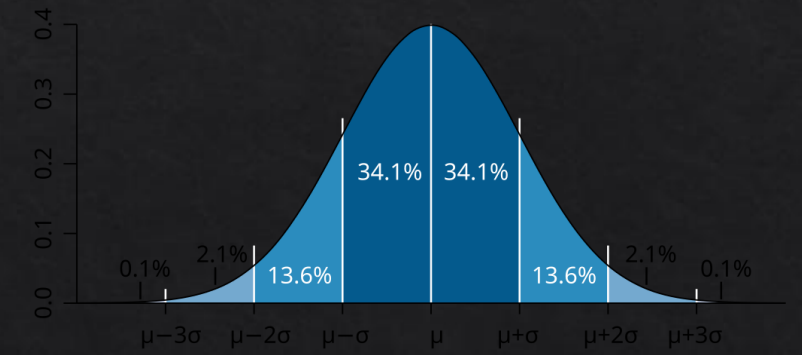


Forrása:

- Mérési, adatrögzítési hiba – felfogható keverék eloszlásként (helyes kísérlet vs. mérési hiba)
- Vastagfarkú eloszlás (heavy-tailed distribution) – ritka esemény

Hogyan dolgozzunk a kiszóró pontokkal:

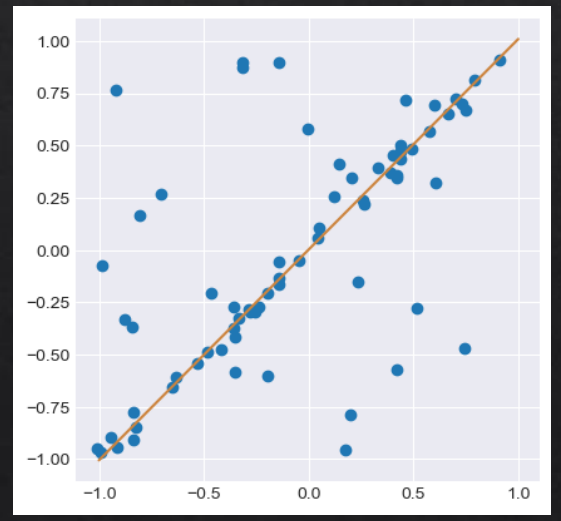
- Robusztus statisztika használata
  - Ne feltételezzünk normál eloszlást
- Kiszóró pontok (iteratív) eltávolítása
  - Tukey (boxplot), Chauvenet (Z-score), Peirce (likelihood ratio), Grubb (t)…
  - Diszkutáljuk (előny-hátrány)
  - Helyettesítés (csonkolás, Winsor-féle szomszéd módszer)



Random sample consensus (RANSAC)

Fischler és Bolles (1981)

- Cél: kiszóró pontoknak ne legyen hatása a modell paramétereire
- Szükséges: matematikai modell, kevés kiszóró pont
- Iteratív módszer
  1. Tanító halmazt erősen alulmintavételezve többnyire teljesen kimaradnak a kiszóró pontok
  2. Modellillesztés, megvizsgálja, mely pontokra illeszkedik a modell
  3. Döntés a modell paraméteriről és „konszenzus” a kiszóró pontok halmazáról
  4. A modell folyamatos finomítása az „konszenzusos” pontok alapján
  5. Modell elfogadása, ha elegendően sok pontra illeszkedik, különben újrakezdés
- Valószínűségi értelemben konvergál
- Példa kód Wikipédián (Python demó Jupyter notebook-ban)



Jupyter notebook

https://github.com/stippingerm/causality-course > Time Series Methods > Anomaly detection.ipynb

Futtatható online a Google Colab-ban!

Szisztematikus hibák

Kalibrációs hiba (tára vagy referencia szint, skála vagy erősítés, műszer torzítása)

Környezeti hatások (hőmérséklet, páratartalom, légnyomás miatti állandó hiba)

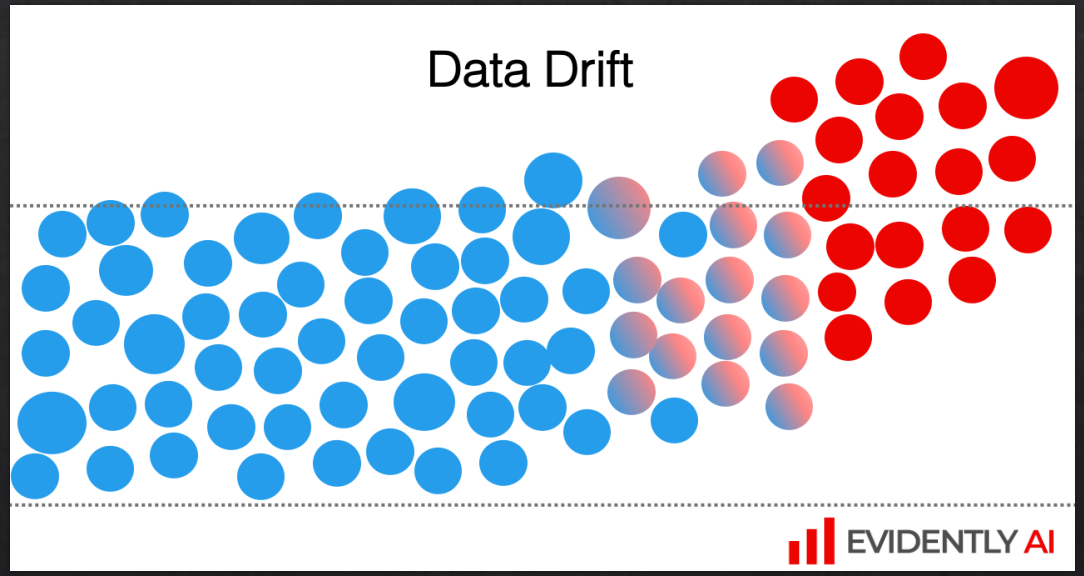
Megfigyelési torzítás
(observer bias – kísérletező elvárása;
sampling bias – mintavételi torzítás,
módszertani és protokollbeli hibák)

Válasz torzítás (kérdőív, önbevallás esetén)

adatokból észlelhető-e?

Időbeli hatások (drift, környezet változása)

Adattisztítás hibái (pl. kiszóró pontok)



Anomáliadetekció

Gépi tanulás típusai*

* Klasszifikációt tárgyaljuk, regressziót nem.

- Felügyelt anomáliadetekció:
  - Címkézett adat: „normális” vs. „abnormális”
  - Klasszifikátor tanítása
  - Adat ritkán elérhető, kiegyensúlyozatlan osztályok.
- Félig felügyelt (semi-supervised)
  - Az adat egy része címkézett
  - Gyakran modell illesztése normális adatra, likelihood
- Felügyelet nélkül (unsupervised)
  - Legnehezebb, leggyakoribb

Alkalmazási területek

- Kiberbiztonság (behatolás észlelése, IoT)
- Orvoslás, idegtudomány
- Gépi látás (konvolúciós neurális hálók)
- Bűnüldözés (videófelügyelet, csalás észlelés)
- Statisztika (előfeldolgozás)
- Ipar (gyártás, olajipar, csővezetékek monitorozása)

Gépi tanulási feladat

Adatszerkezet

	x1	x2	…	xd	Label
1	0.3	52	…	1	N
2	–0.2	71	…	1	N
3	0.7	30	…	0	A
4	0.9	27	…	1	A
5	–0.4	64	…	1	N
6	1.1	49	…	0	A
…	…	…	…	…	…
n	0.7	33	…	1	N

- Megfigyelések (observations, samples)
  - $i \in \{1, 2, \ldots, n\}$
- Jellemzők (X[i,j], features)
  - $j \in \{1, 2, \ldots, d\}$
  - Mért érték, előző napi érték, kép pixelek fényereje felsorolva, bináris jellemző, …
- Címkék (y[i], labels)
  - $c \in C$ osztályok (classes)
  - Összesen egy kategorikus oszlop, kétféle címke (bináris klasszifikáció)

Anomália vagy újdonság (gépi tanulás)

Anomália

- Egyetlen adatsor, meg kell találni, melyik pontok nem illeszkednek a mintába
- Szét kell válogatni a normális és az anomális pontokat címezetlen adaton
- Meg kell határozni azt a tartományt, ahol a normális pontok koncentrálódnak
- Általában sokkal kevesebb az anomális pont, mint a normális, kiegyensúlyozatlan osztályozási feladat (unbalanced classification)

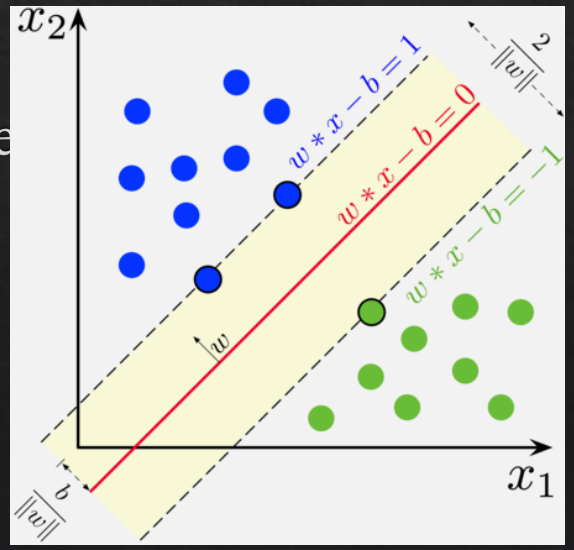
Újdonság

- A tanító halmaz elemei mind a „normális” halmazba tartoznak, azonos a címkéjük
- Keressük, hogy az új pontok közel vannak-e az eredeti halmazhoz
- A tanító halmaztól távol eső pontok számítanak újdonságnak
- Gondolhatunk-e rá úgy, mint teljesen kiegyensúlyozatlan (unbalanced) osztálymegjelenésre a tanító halmazban

Support Vector Machine

Support Vector Machine (SVM) klasszifikátor

- training vectors $x_i \in \mathbb{R}^d$, $i=1, \ldots, n$
- training labels $y_i \in \{-1,+1\}$, $i=1, \ldots, n$
- két osztályt legjobban elválasztó hipersík $w \in \mathbb{R}^d$ és $b \in \mathbb{R}$ paramétere



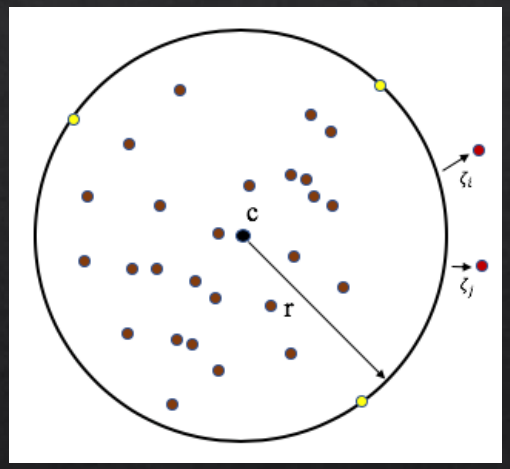
$$\min_{w,b,\zeta} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \zeta_i$$

$$\text{subject to } y_i(w^T \phi(x_i) + b) \geq 1 - \zeta_i,$$
$$\zeta_i \geq 0, i = 1, \ldots, n$$

Egy osztályos SVM

Schölkopf et al., Neural comp. 13.7 (2001)

- kernelfüggvény, hogy görbe felületeket is tudjon hipersík helyett



$$\min_{r,c,\zeta} r^2 + \frac{1}{\nu n} \sum_{i=1}^{n} \zeta_i$$

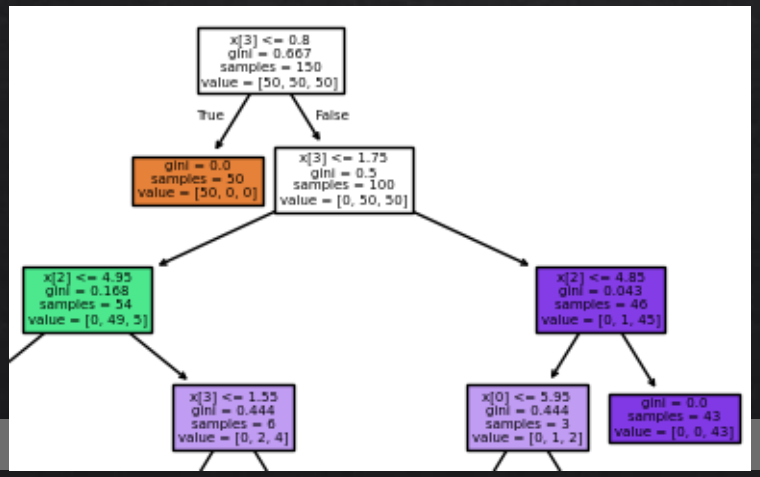
$$\text{subject to, } \|\Phi(x_i) - c\|^2 \leq r^2 + \zeta_i$$

Izolációs erdő

Véletlen erdő (random forest)

- Sokaság döntési fákból, többségi szavazás
- Fa építésénél mindig a jellemzők egy véletlen alhalmazát használjuk
- Vágásnál (fa elágazása) csökkenjen a címkék sokfélesége
- Alkalmas nagy adatmennyiségekre

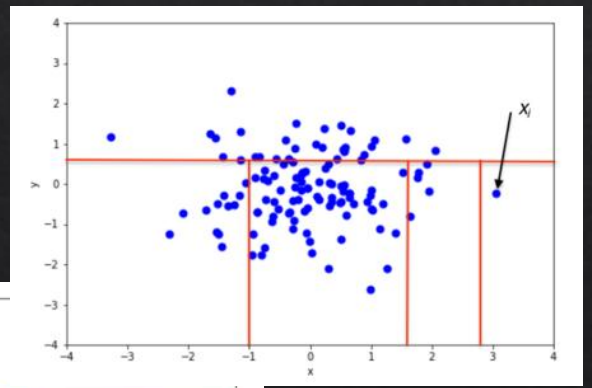
Példa döntési fa részlet az írisz adatok

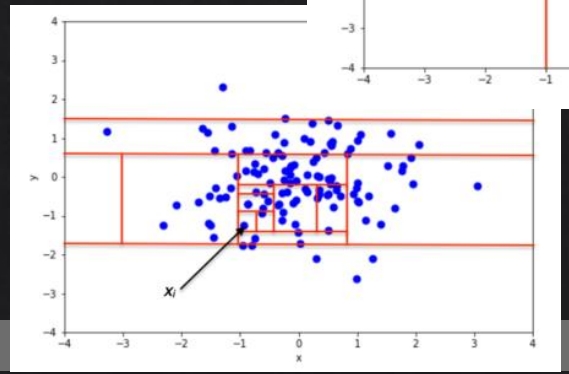


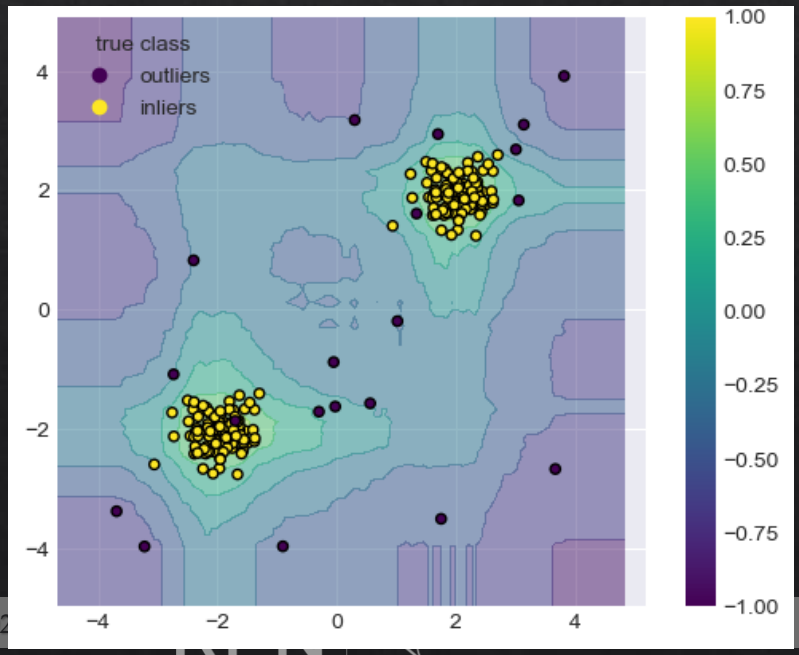
Izolációs erdő (isolation forest)

Liu et al., IEEE Int. Conf. Data Mining (2008)

- Anomális pontok könnyebben szeparálhatók
- Véletlen jellemzők véletlen megengedett értékénél vágunk
- Anomális pontok kevésbé mélyen lesznek a fában





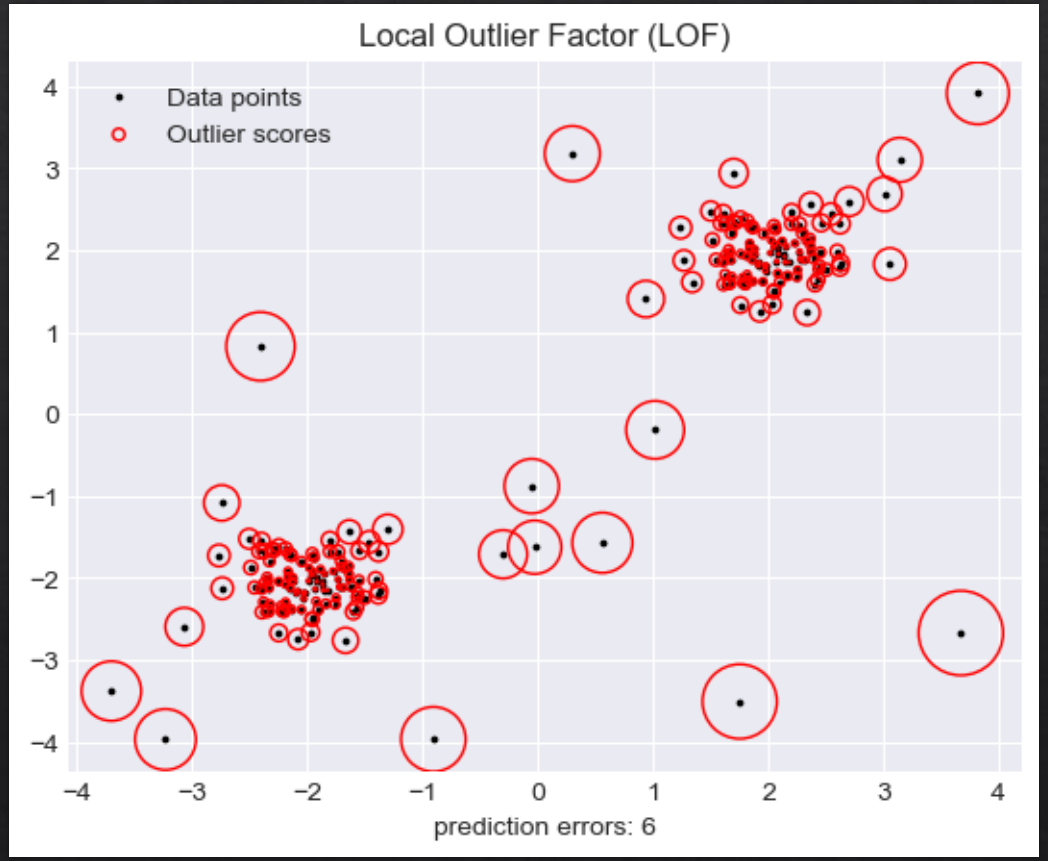


k legközelebbi szomszéd (kNN)

Helyi kiszóró faktor
(local outlier factor, LOF)

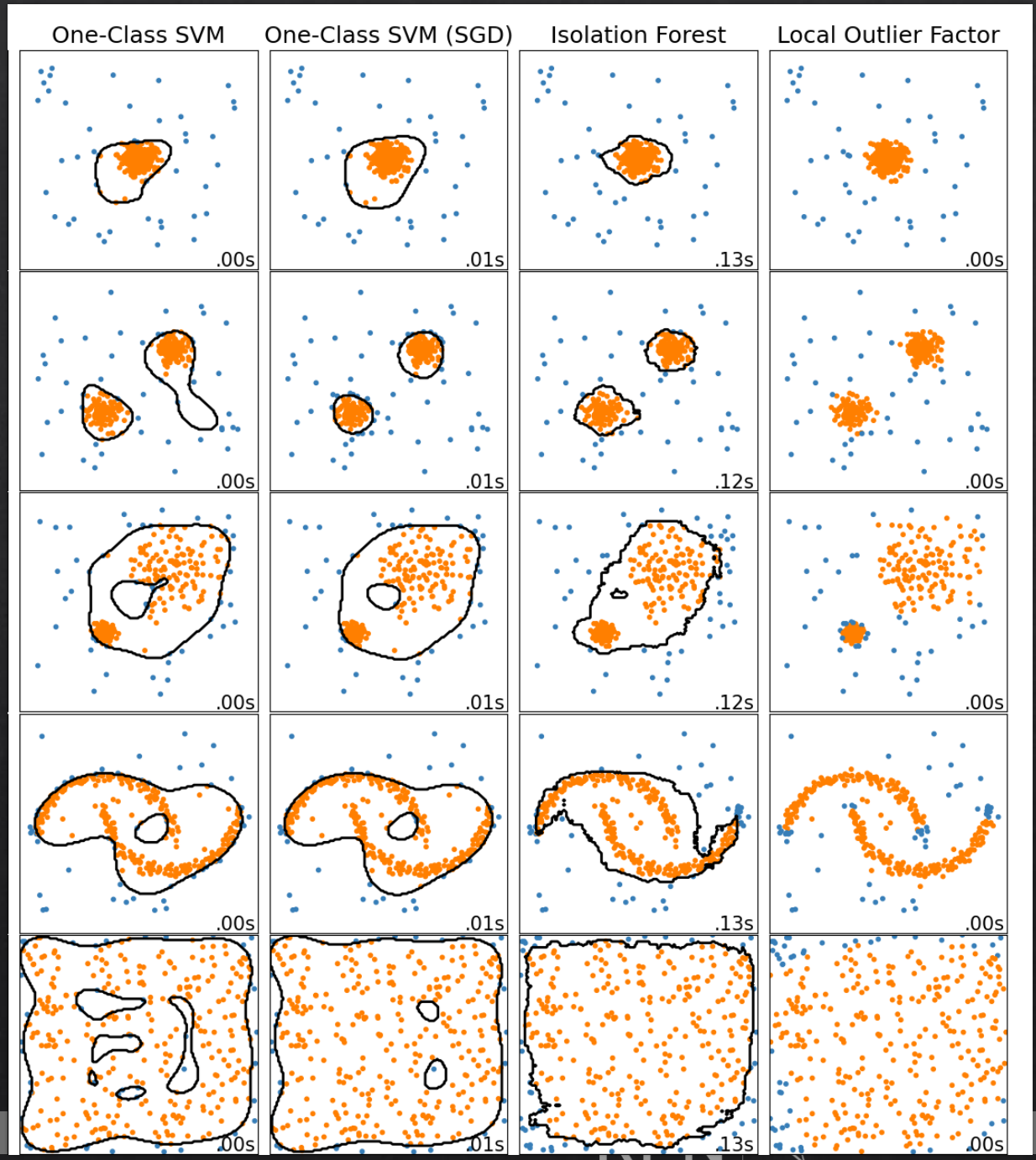
Breunig et al., Proc. ACM SIGMOD (2000)

- Kiszóró pontokban a térbeli sűrűség kisebb, mint az ő szomszédaiknál
- k-szomszédság miatt adaptív, az adatpontok inhomogén sűrűségéhez
- Legyen $k >$ minimum klaszterméret
- Legyen $k <$ #potenciális közeli kiszórók
- Tipikusan k $\approx$ 20



Példa kimenetelek

Forrás: https://scikit-learn.org/dev/modules/outlier_detection.html#overview-of-outlier-detection-methods



Idősorok vizsgálata

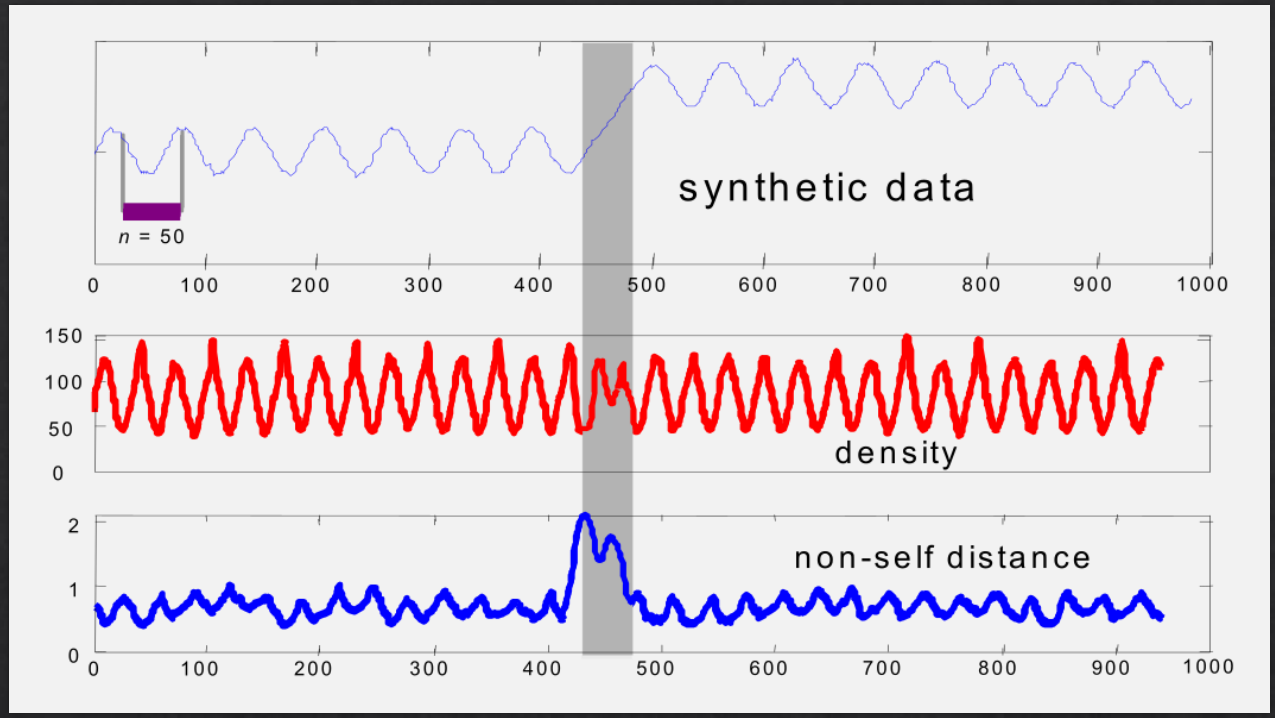
- Most nem tárgyaljuk: idősorok statisztikája, ARIMA, VAR modellek, véges állapotautomaták, rejtett Markov modell (HMM), több idősor összehasonlítása…
- Idősorok vizsgálata: rendszer, *d* szabadsági fokkal
- Trajektóriája az eredeti állapottérben
  $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_d(t)]$
  $d\mathbf{x}(t)/dt = F(\mathbf{x}) + \xi(t)$
- Időkéséses beágyazás (time-delay embedding, TDE)
  $X(t) = [x_1(t), x_1(t-\tau), x_1(t-2\tau), \dots, x_1(t-(k-1)\tau)]$
- TDE jelentősége determinisztikus dinamikai rendszerekben ($\xi=0$):
  $k=2d+1$ dimenzióban rekonstruálható topológiai ekvivalencia erejéig,
  pusztán egy változó megfigyeléséből. (Takens tétele, 1981)
- Más idősorokban: ablakos vizsgálat, mintaillesztés a korábban vizsgált módszerekkel

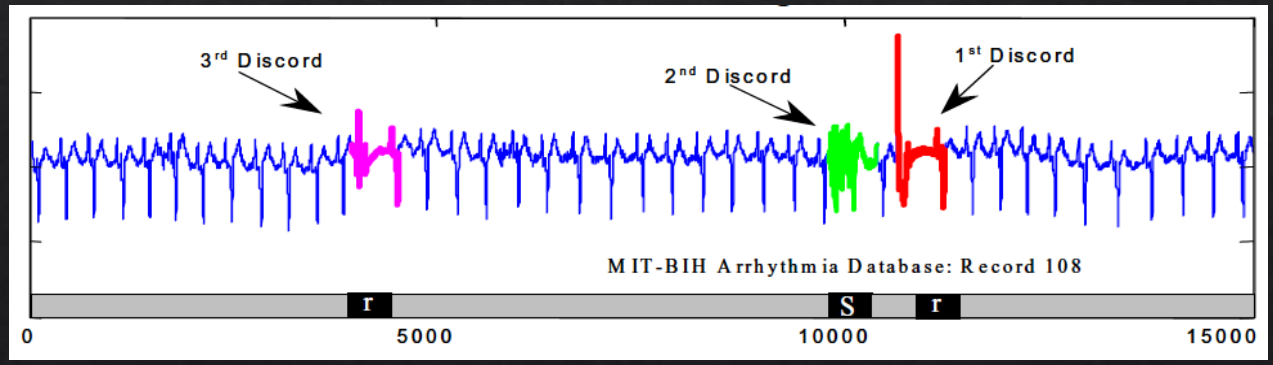
Idősor eltérés (time series discord)

Keogh et al., IEEE Intl. Conf. Data Mining (2005)

Beágyazott térben az a vektor (részsorozat), aminek a legközelebbi át nem fedő szomszédjához mért távolsága a legnagyobb.

$$\min(\mathrm{Dist}(D, M_D)) > \min(\mathrm{Dist}(C, M_C))$$





Időbeli kiszóró faktor (temporal outlier factor, TOF)

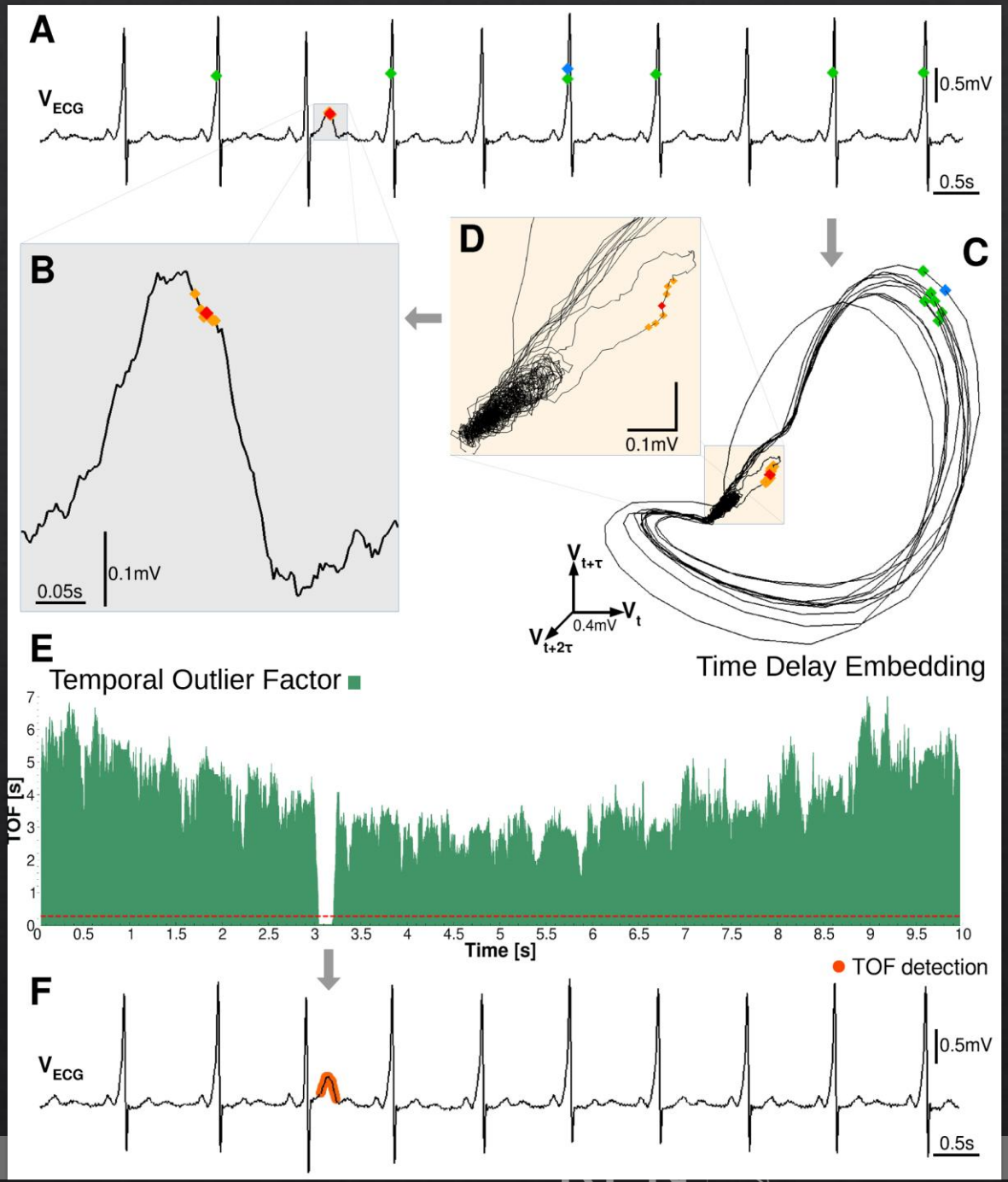
Benkő et al., Scientific Reports (2022)

Beágyazott térben elválnak az anomális időszakok a szokásos trajektóriáktól.

Beágyazott térbeli k legközelebbi szomszéd időindexeit vizsgáljuk, mennyire vannak az eredeti időpont közelében az időben:

- Az idősorban egyenletesen (normális)
- Egy helyre koncentrálódnak (anomália)

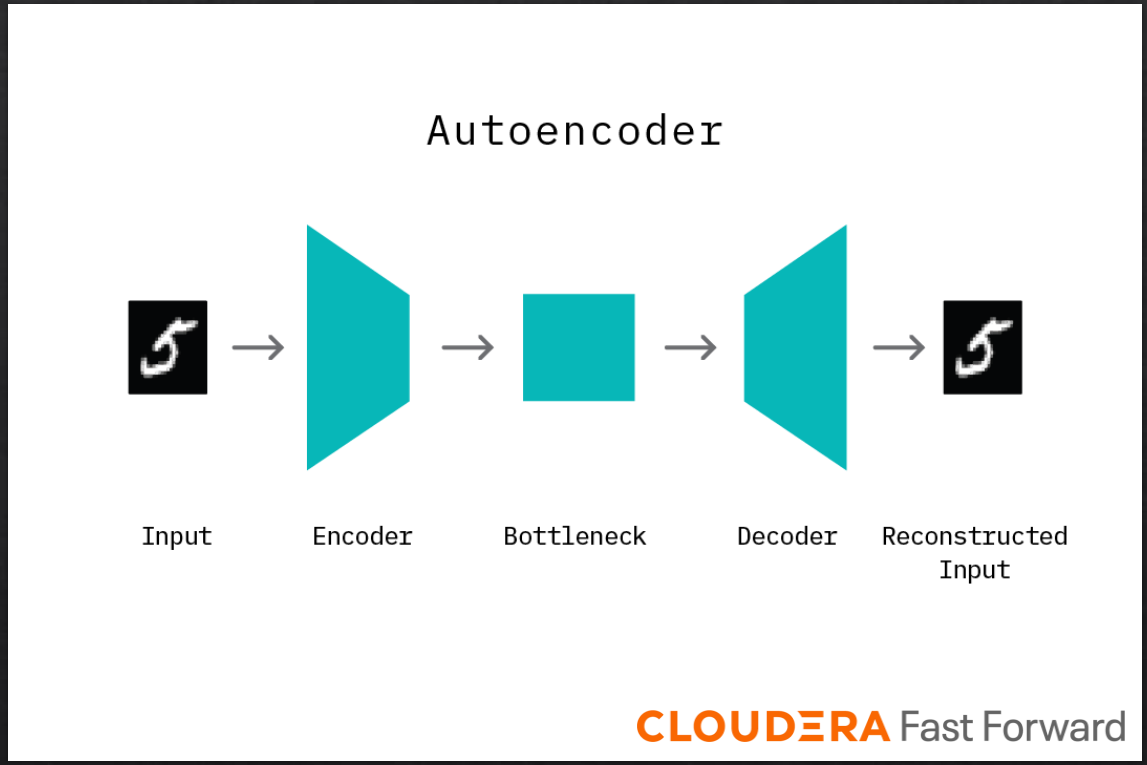
$$\mathrm{TOF}(t) = \sqrt[q]{\frac{\sum_{i=1}^{k} |t - t_i|^q}{k}}$$



Autoenkóder

Mély neurális háló enkóder–dekóder architektúrával, félig felügyelt tanítás

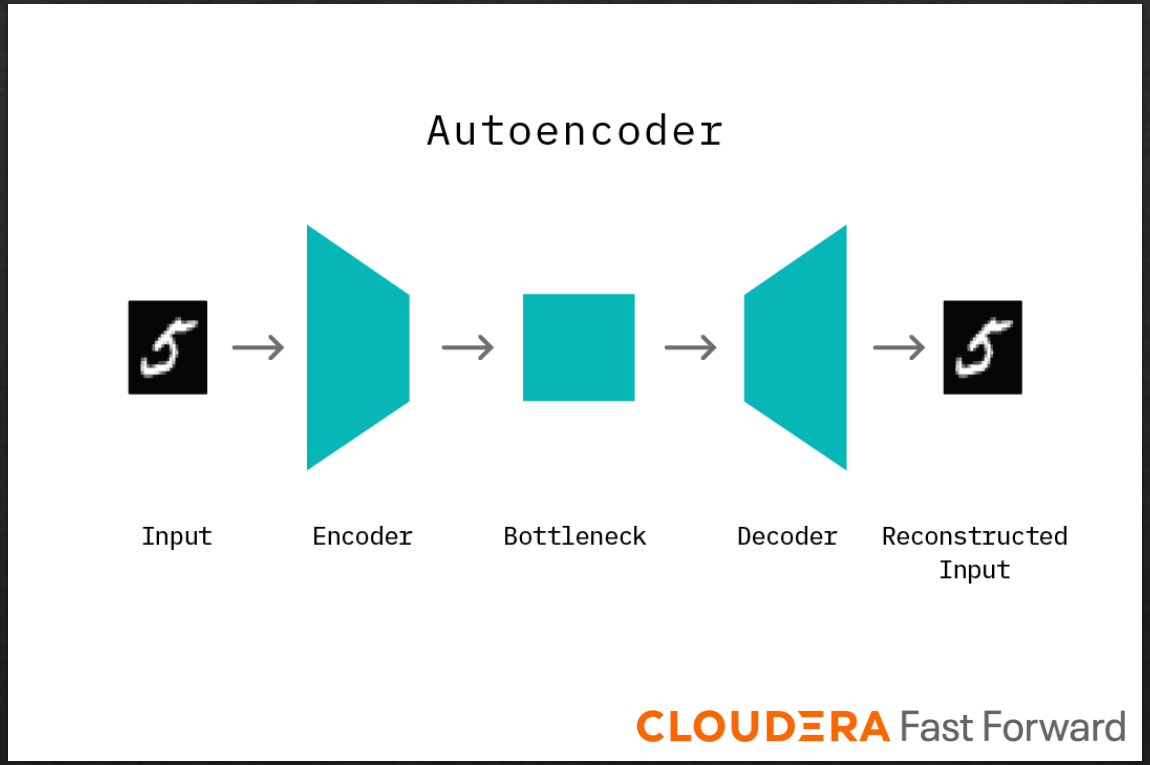
- Szűk keresztmetszetet:
  - biztosítja tömörítés hatékonyságát
  - alacsony dimenziós reprezentáció
- Hibafüggvény: rekonstrukció hibája (a kimenet és a bemenet eltérése) → minimalizálni
- Variációs autoenkóder: tanítás során sampling, jobb reprezentáció.
- Konvolúciós: eltolás-invariáns jellemzők



Autoenkóder implementáció

Minimalista Python programkód PyTorch segítségével

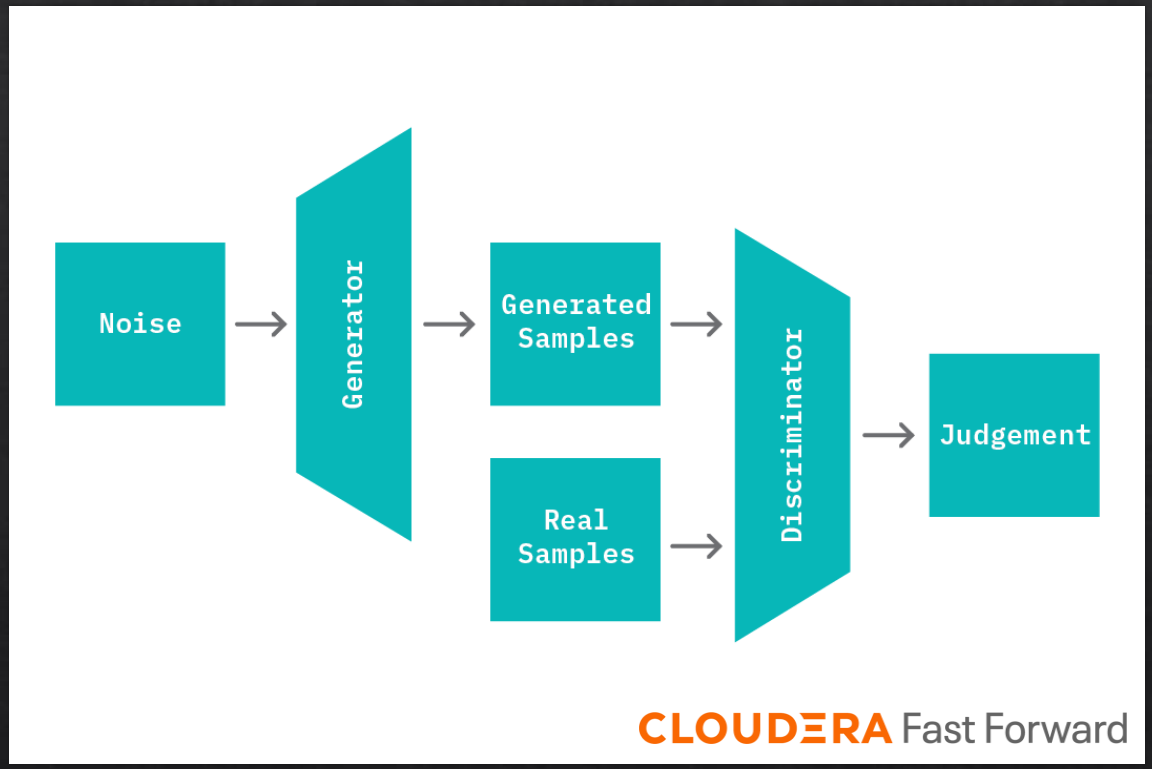
```
from torch import nn
class Autoencoder(nn.Module):
    def __init__(self):
        self.encoder = nn.Sequential(
            nn.Linear(256, 64),
            nn.ReLU(True),
            nn.Linear(64, 12),
            nn.ReLU(True),
            nn.Linear(12, 3)  # Latent space
        )
        self.decoder = nn.Sequential(
            nn.Linear(3, 12),
            nn.ReLU(True),
            nn.Linear(12, 64),
            nn.ReLU(True),
            nn.Linear(64, 256),
            nn.Sigmoid()  # Output layer
        )
    def forward(self, x):
        return self.decoder(self.encoder(x))
```



Versengő generatív neurális hálók

Generative adversarial networks

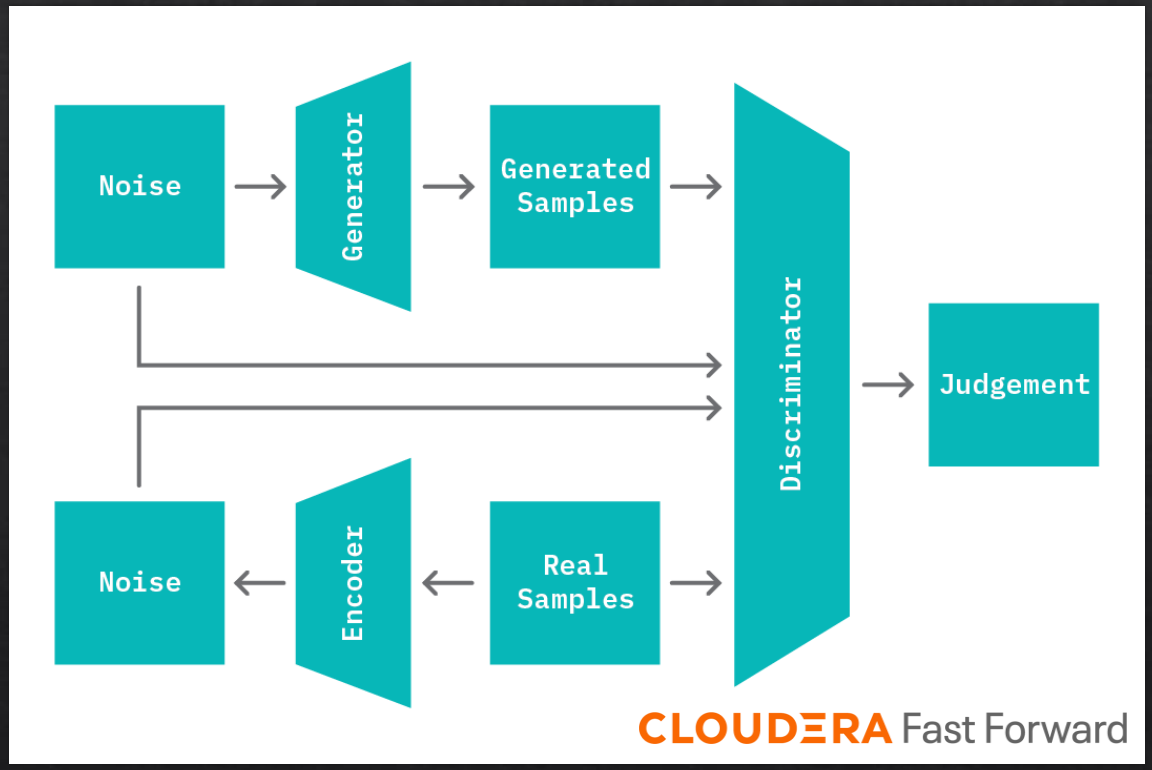
- Generátor G és diszkriminátor D
  - Versenyeznek miközben együtt tanulják a normális bemenet X eloszlását
  - G: véletlen zaj Z → imitált bemenet X_
  - D: X és X_ megkülönböztetése
- Anomáliadetekció
  - Tanítás során nem látott anomális mintát
  - D valódi G és különbségét tanulta meg…
  - Nagyon nehéz olyan Z-t találni, ami X-nek megfelel, ezért használnak BiGAN-t



BiGAN

BiGAN struktúra

- GAN kiegészítve E enkóderrel
  - E: X → látens reprezentáció Z_, együtt tanul a generátorral
  - G: véletlen zaj Z → imitált bemenet X_
  - D: X és X_ ill. Z és Z_ megkülönböztetése
- Anomáliadetekció
  - A GAN a tanítóhalmaz jellemzőit tárolja
  - Teszt mintából Z_ és X_
  - Anomália mértéke a rekonstrukciós hiba, lásd a Jupyter notebook-ban



Anomáliadetekciós MI összefoglalása

MI anomália detektálására

- Mintánként: VAE, GAN, BiGAN
- Seq2Seq: LSTM, Transformer
  (ezeket most nem érintettük, ld. irodalom)
- Amíg az anomáliák ritka események
  - a neurális háló normális eseményekre vonatkozó modellje alig változik
  - nem szükséges külön címkézett adat

vs. Hagyományos módszerek

- MI lassabb, mint a lineáris modellek
- MI általában pontosabb
  (accuracy, precision, recall)
- Bizonytalanság klasszifikálása
  - Probabilisztikus/variációs módszerek

Összefoglalás

Láttuk:

- Anomáliák definíciója, osztályozása, adatok kezelése
- Detektálás: statisztikai módszerek, klasszikus gépi tanulás, mély neurális hálók

Beadandó, ha ezt a témát választod:

- Saját kutatáshoz kapcsolódó vagy egyéni kérdést tegyél fel
- Mindenképp küldd meg előzetesen a projektötletet
- Az ötletet a beadandó kidolgozása előtt átbeszéljük, jóváhagyom

Irodalom

- Wikipédia szócikkek (angolul) https://en.wikipedia.org/
- IBM Anomaly detection https://www.ibm.com/topics/anomaly-detection
- Scikit-learn dokumentáció https://scikit-learn.org/
- Manish Gupta, Jing Gao, Charu Aggarwal, and Jiawei Han: Outlier Detection for Temporal Data. (Morgan & Claypool, 2014)
- Cloudera Deep Learning for Anomaly Detection https://ff12.fastforwardlabs.com/
- Venujkvenk: Anomaly Detection Techniques: A Comprehensive Guide …
- További olvasnivaló: PyTorch, towardsdatascience.com, Kaggle, …

Hirdetések

- Csoportunk honlapja http://cneuro.rmki.kfki.hu/
  - Komplex rendszerek, okságvizsgálat, forráslokalizáció, hálózatelemzés
- Tudományos Számítások Intézete Egyesület https://scicomp.hu/
  - Numerikában jártas szakemberek
- Lectures on Modern Scientific Programming (általában minden év novemberben)
  - https://gpu.wigner.hu/
  - Jó gyakorlatok a programozásban, GPU programozás, mesterséges intelligencia
- Mesterséges Intelligencia Akcióterv a HUN-REN Wigner FK-ban https://ai.wigner.hu
  - Mesterséges intelligencia nagykövetek, projekt szakmai támogatása