

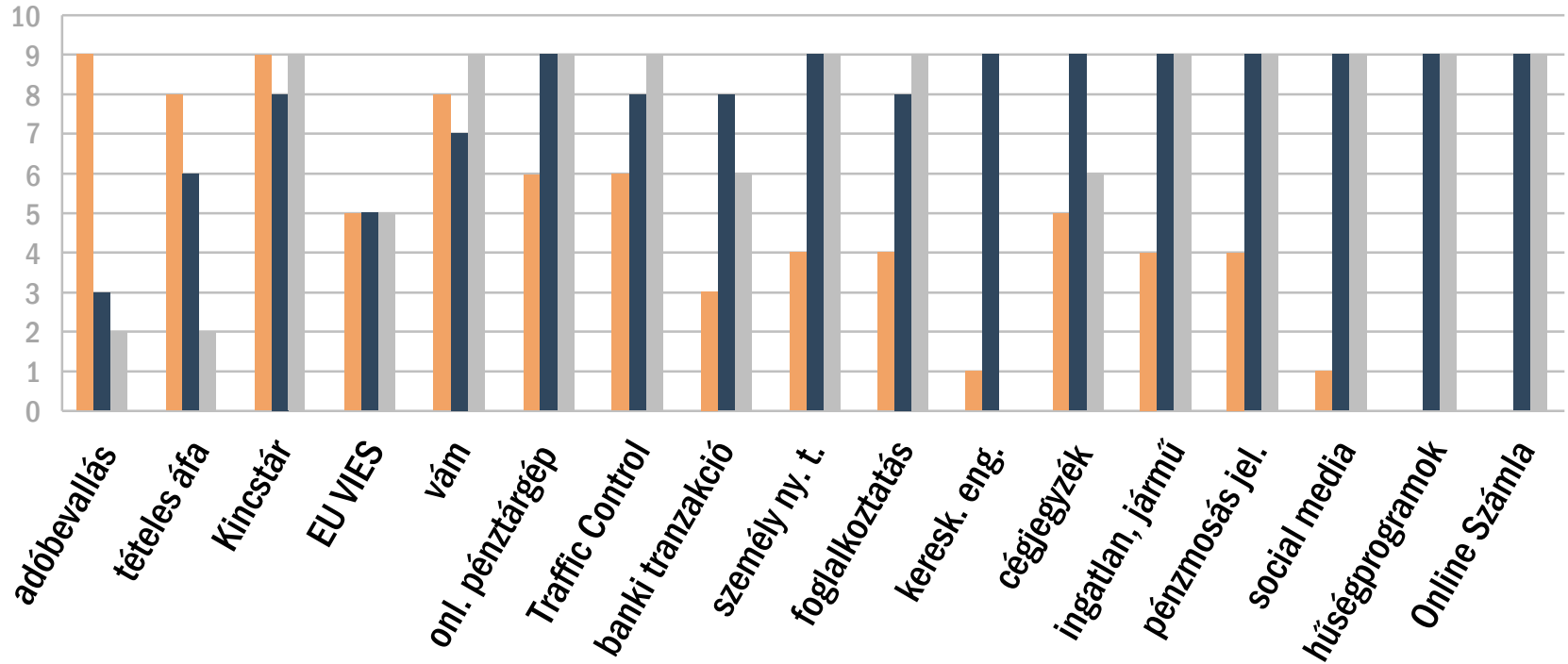
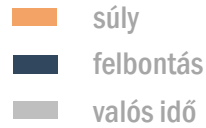
# Mélytanulási módszerek a közteherviselés szolgálatában

Vágújhelyi Ferenc  
államtitkár,  
a NAV elnöke



# Adatforrások

(relevancia, felbontás, időbeliség)



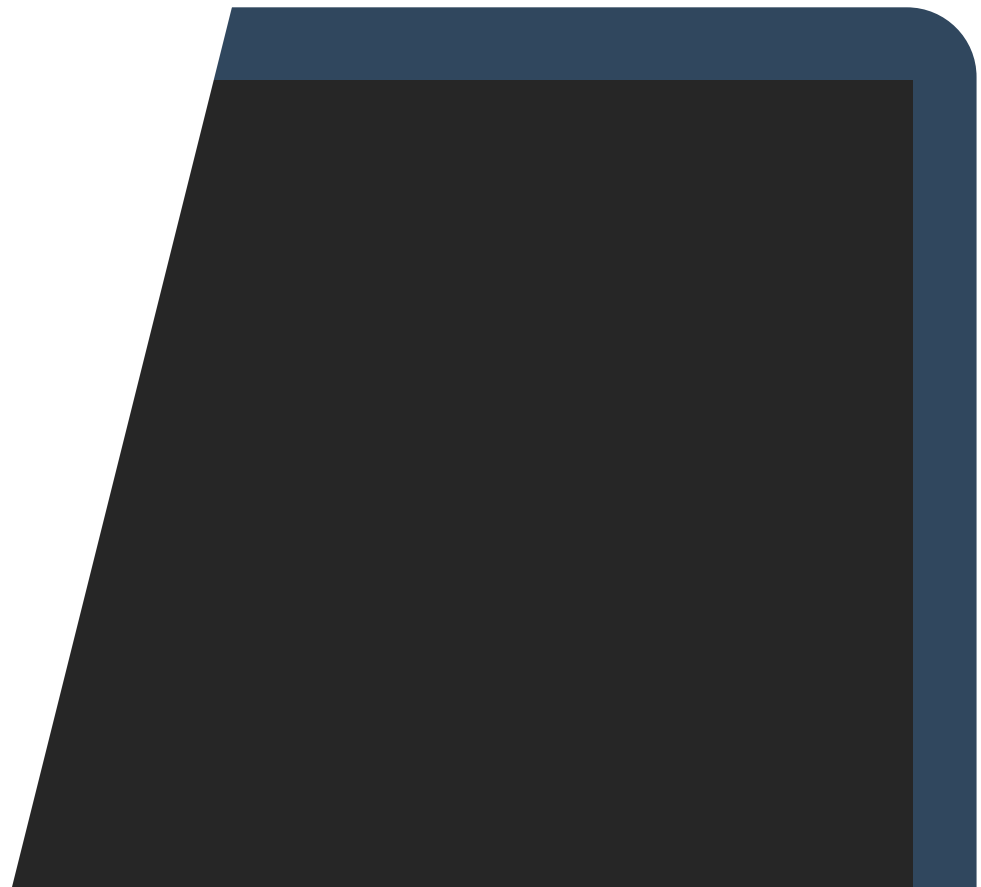
# Hogyan elemzünk?

```
{#define SPACE_METRICS}
```



# Hogyan elemzünk?

```
{#define SPACE_METRICS}
```



# Hogvan elemzünk?

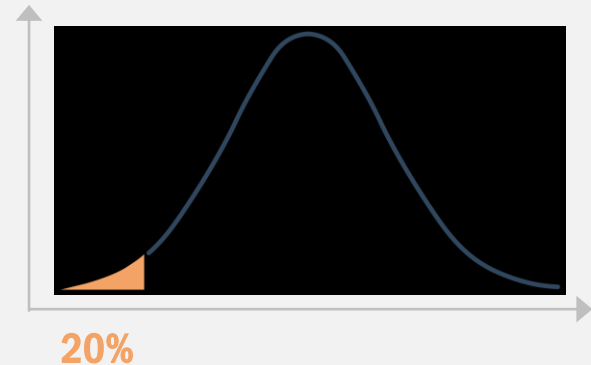
```
{#define SPACE_METRICS}
```

$$(1) \quad d_j(X_k, X_l) = \begin{cases} 0 & \text{if } X_k^j, X_l^j \text{ belong to the same class.} \\ 1 & \text{else} \end{cases}$$

$$(2) \quad d_j(X_k, X_l) = \frac{|X_k^j - X_l^j|^{w_j}}{r_j^{w_j}}$$

$$(3) \quad d(X_k, X_l) = \sum_j \alpha_j d_j(X_k, X_l)$$

$$(4) \quad A_{kl} = d(X_k, X_l)$$

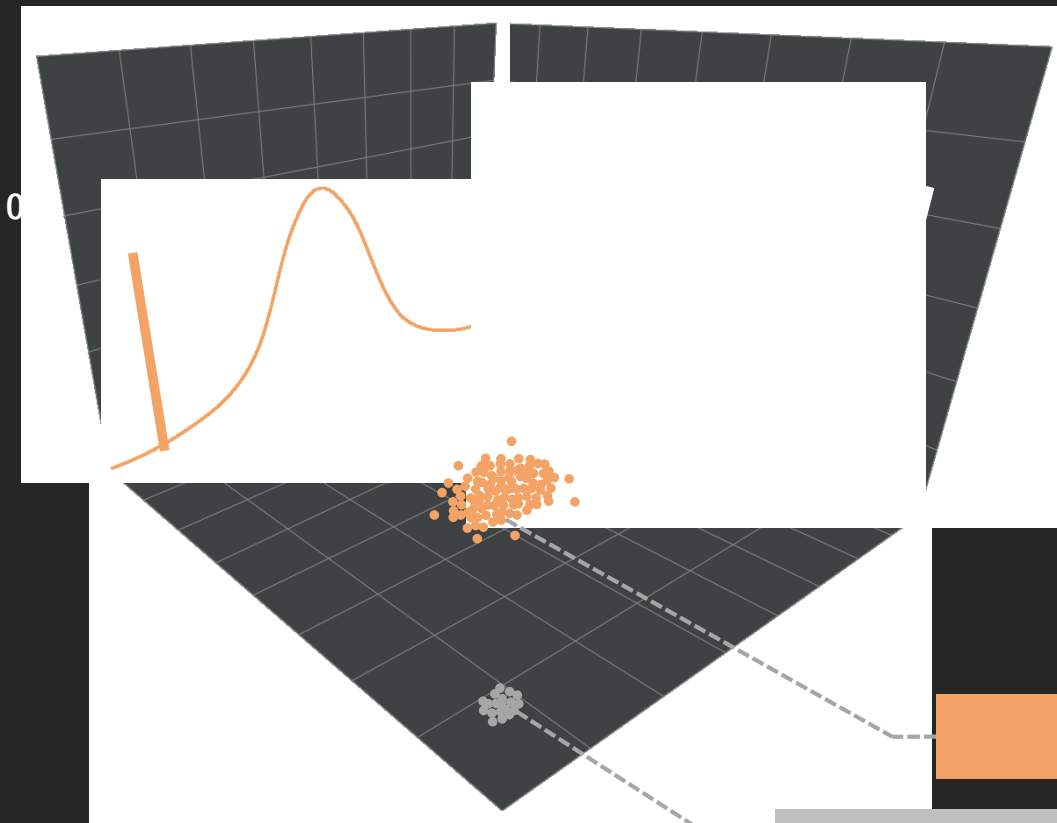


„Az ügyes” csaló feladata:

nem elég csalni, becsületesnek is  
kell látszani!

itt látszódni

itt lenni



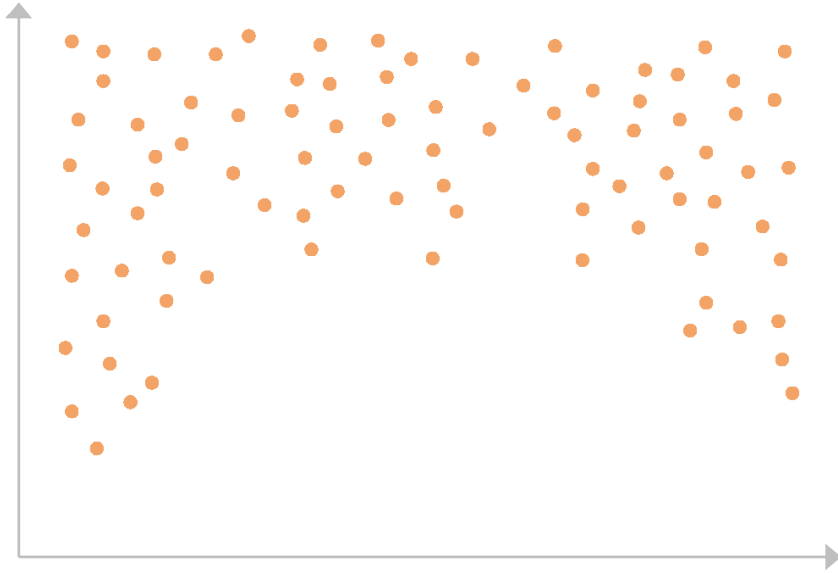
# Kontrollált gépi tanulás

**UNDERFITTING**



# Kontrollált gépi tanulás

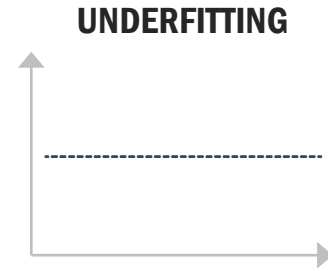
**UNDERFITTING**





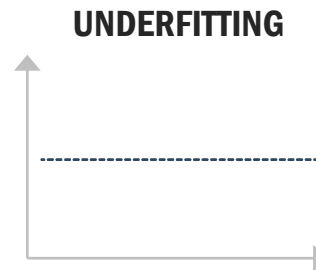
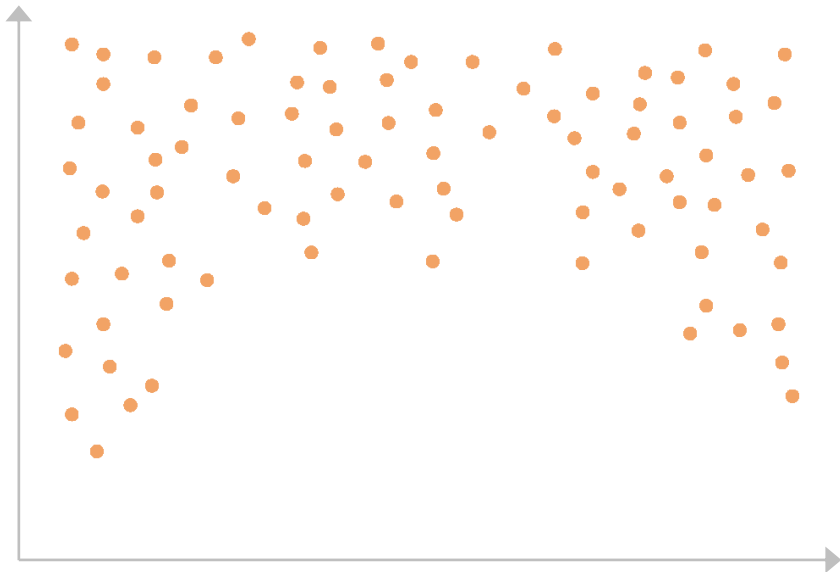
# Kontrollált gépi tanulás

**APPROPRIATE-FITTING**



# Kontrollált gépi tanulás

**APPROPRIATE-FITTING**



# Kontrollált gépi tanulás

**OVER-FITTING**



**UNDERFITTING**



**APPROPRIATE-FITTING**



# Kontrollált gépi tanulás

**OVER-FITTING**



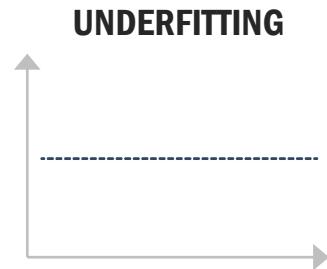
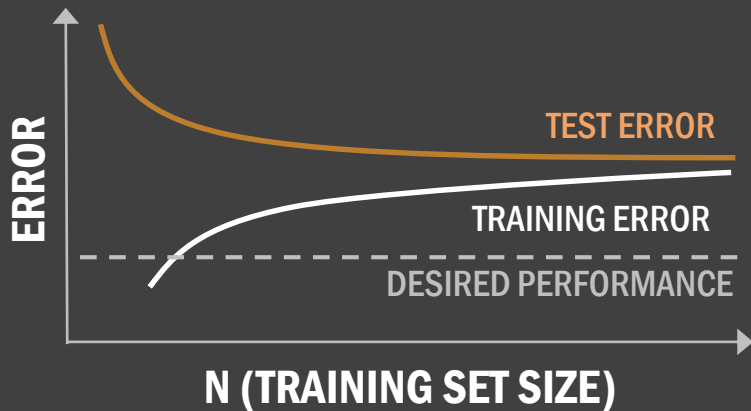
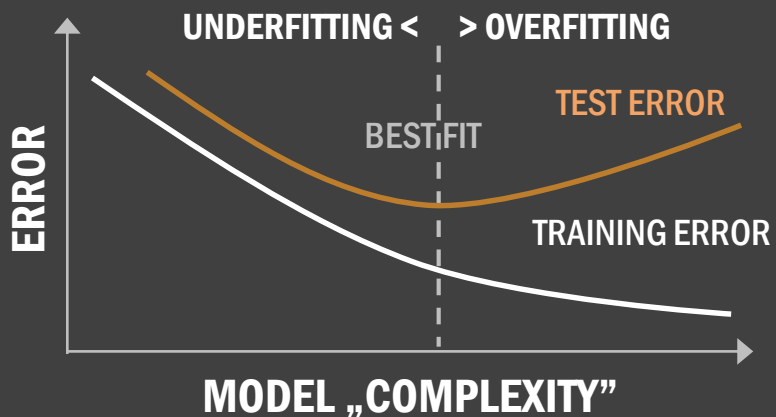
**UNDERFITTING**



**APPROPRIATE-FITTING**



# Kontrollált gépi tanulás



APPROPRIATE-FITTING



OVER-FITTING



# Megoldható a feladat?

Tetszőlegesen sok erőforrású bűnszervezet:



**bármennyi megbízható tisztségviselő, alkalmazott,  
különböző székhely, telephely,**



**bármennyi áru,**



**törvényes „fedő” tevékenységek,**



**hatósági elemzési módszerek  
ismerete**

akkor a bűnszervezet rövid ideig sikeresen működik. Minden ismert minta kikerülhető.  
A gépi tanulási módszerek is "nem deviáns" klasztert dobnak ki.

**A feladatot mégsem oldották meg!**



# A rejtőzködés korlátja

**Sok erőforrás:**

- veszteséges tevékenység





# A rejtőzködés korlátja

Az ideálisnál kevesebb erőforrás:

- felismerhető minta (mintaillesztés)

*de*tektálható



# Mi kell még?

Még több adat!

Adat külföldről is,  
legalább az EU-ból

Adatvédelem

- Céget
- Nemzetgazdaságot
- Szabályozás: fair anonimitás nemzeti és közösségi szinten is.





# Miért is? Konzisztenciaellenőrzés

Van adott VIES rekordhoz tartozó fuvar?

Stimmel a gépjármű és a súly?

Életszerű az időpont?

Az online pénztárgépes jelentéshez tartozik elég beszerzés, vagy a feketepiacról vásárol?

Hipermarketnél:  
a teherautómennyiség és a beszállítói lánc kockázatelemzése...

# Hol a határ?



## Új veszélyek jelentkeznek.

A teljes gazdaság elemi adataival még nincs tapasztalatunk.

A verseny majd abból áll, hogy „szerezz meg mindent” a versenytársról?

Ha a teljes adatbázis kompromittálódik, mik a következmények?

Egyes nagy cégek már rendelkeznek hasonló adatbázisokkal.



Nemzeti Adó-  
és Vámhivatal

[www.nav.gov.hu](http://www.nav.gov.hu)

