



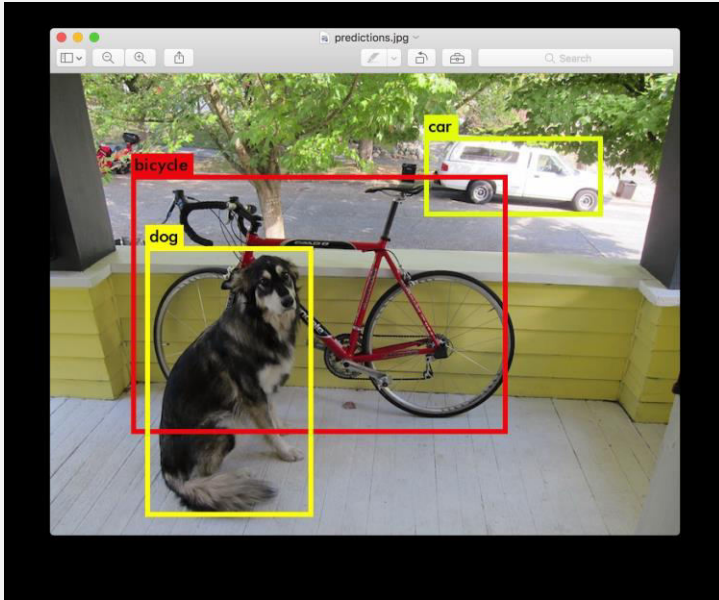
5. jubileumi
PROGRAMOZÓI
KONFERENCIA
2017

Gépi tanulás leleplezve

Remeli Mina

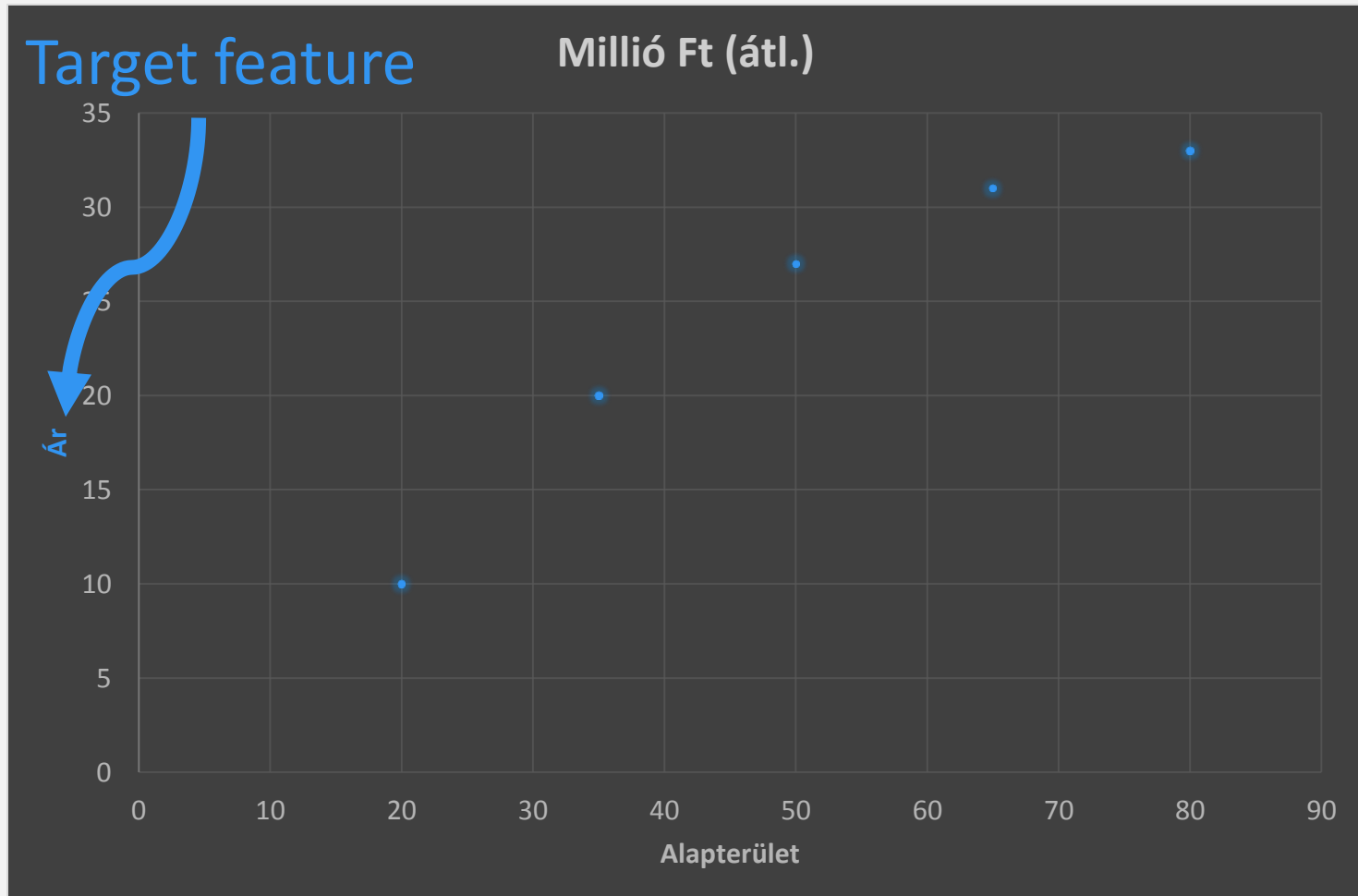


Machine Learning



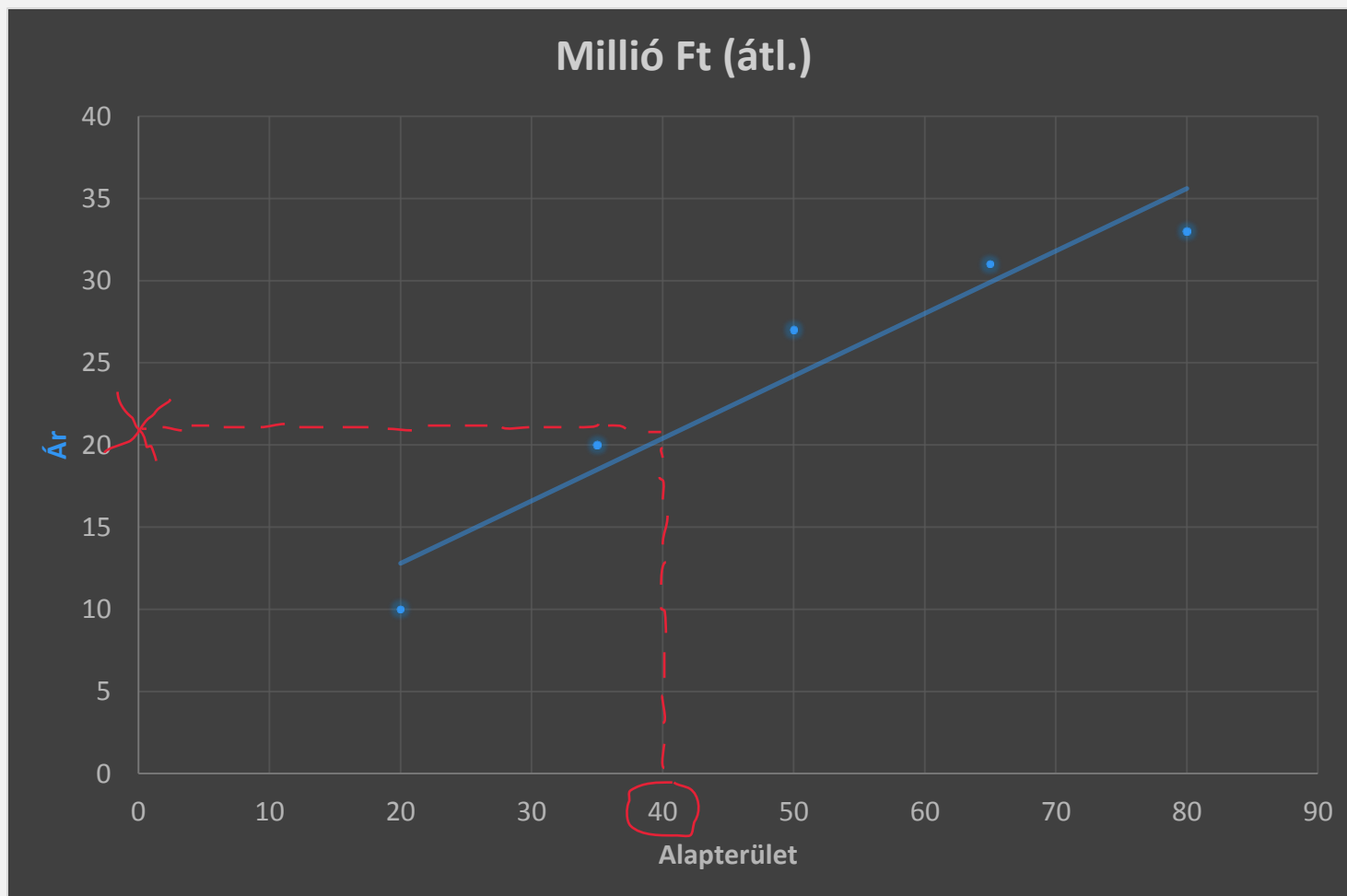


Példa: Budapesti lakásárak



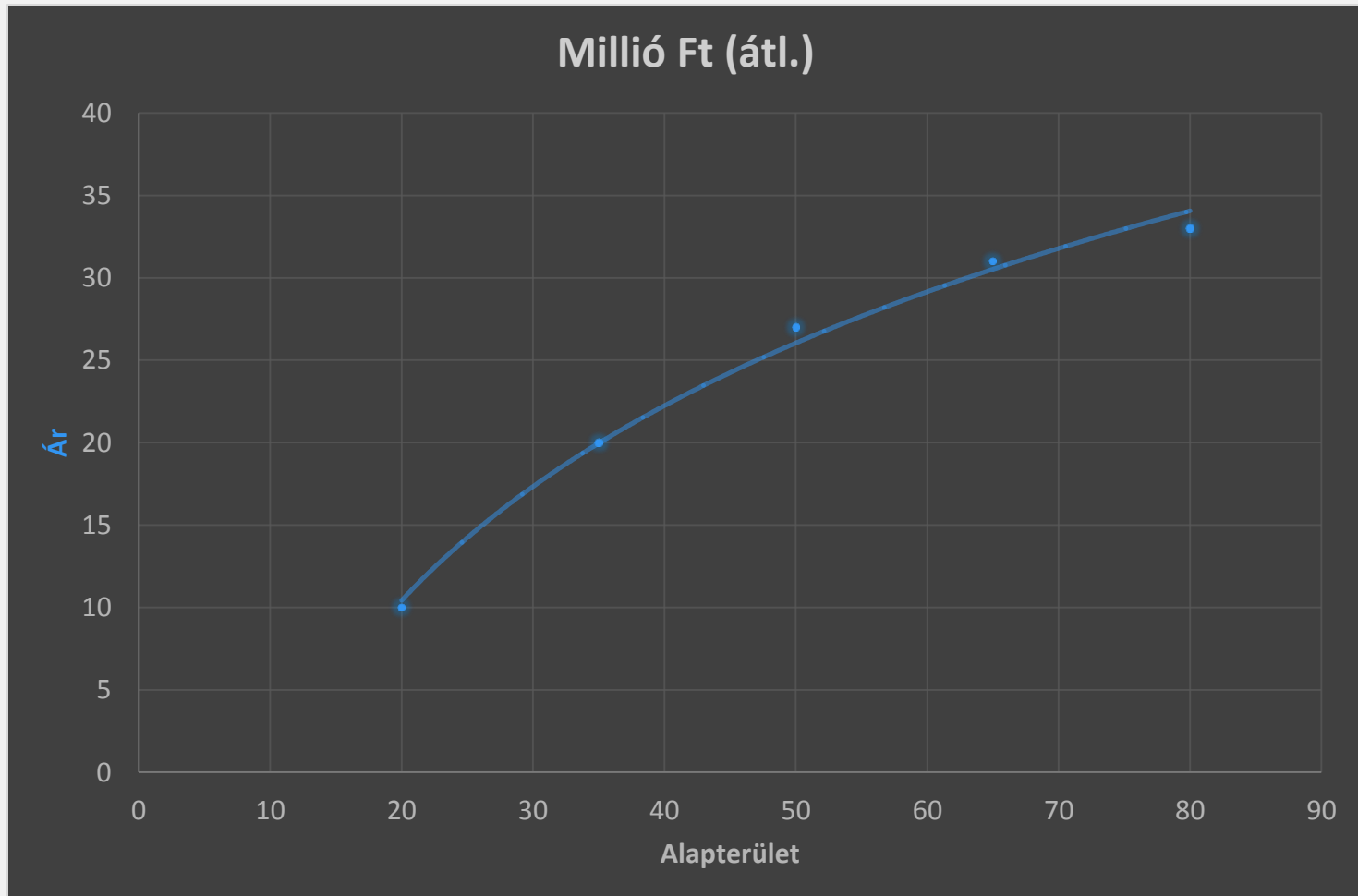


Példa: Budapesti lakásárak





Példa: Budapesti lakásárak



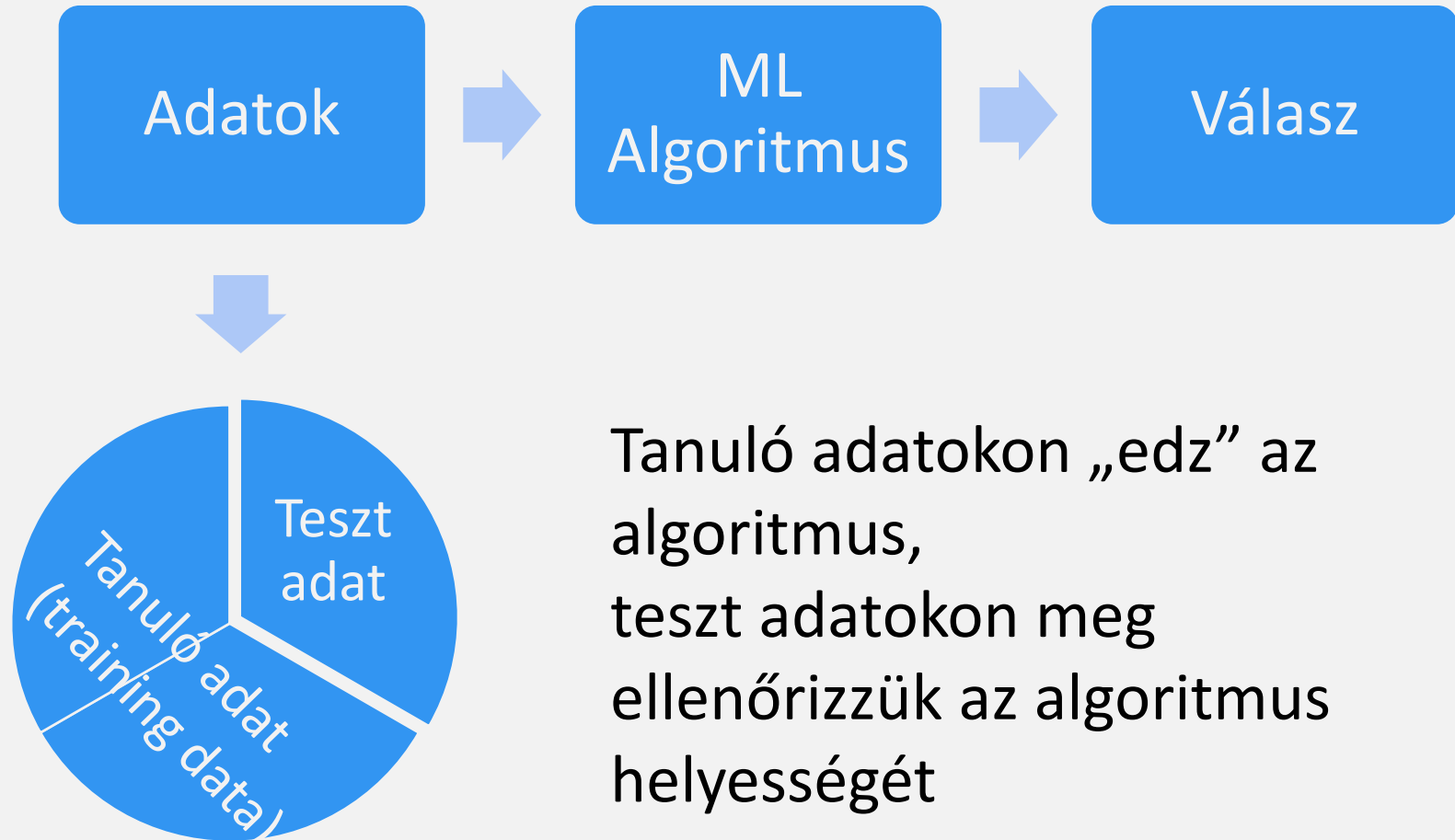


Az algoritmusok alapja matek





A folyamat





A folyamat



Honnan kapjuk?

Pl SciKit, Tensorflow,
Spark



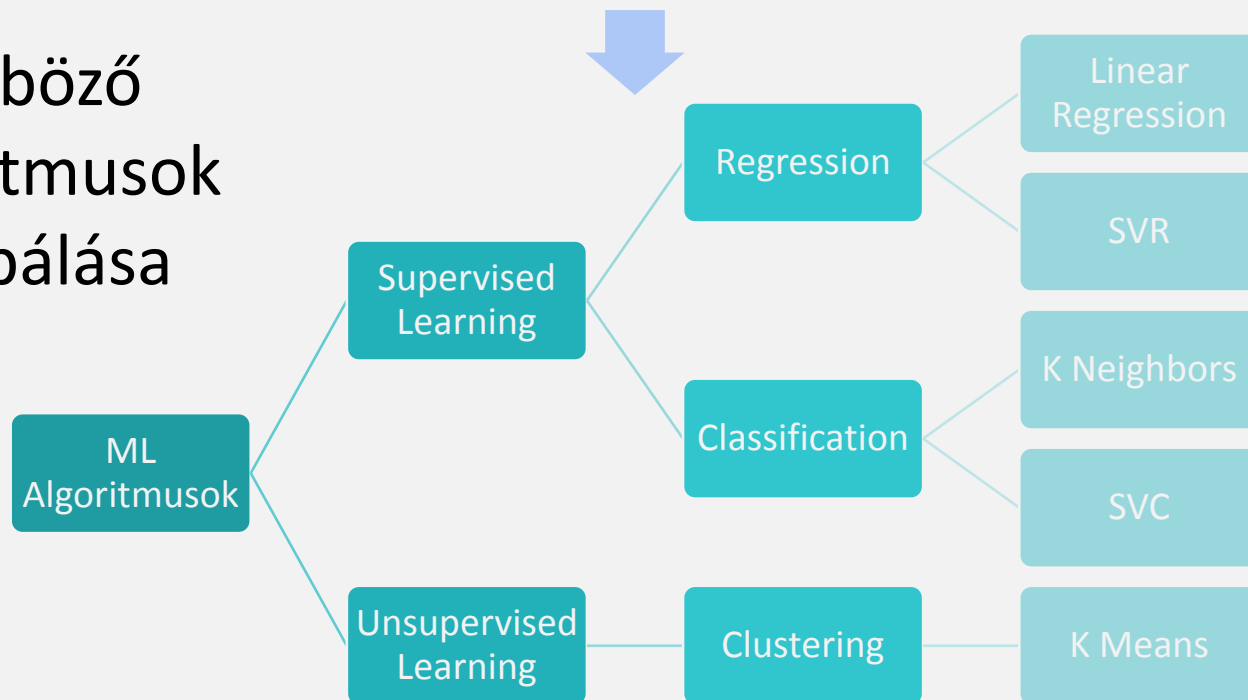
- ✓ Egyszerű használni
- ✓ Kevés hozzáértést igényel



A folyamat



Különböző
algoritmusok
kipróbálása





A gépi tanulás sötét oldala...

- **Feature engineering** - a 0. lépés
- Nem mindig állnak készen az adataink az algoritmushoz...
- Probléma, ha:
 - Nem számszerű az adat -> **mapping**
 - Nem megfelelő formátumú az adat (több forrás)
 - Nincs elég feature -> újakat kell konstruálni (**feature extraction**)

Feature extracting - példa

- Plusz információk kinyerése
 - Egy IP címből
 - Egy dátumból (+idő)

Feature extracting - példa

- Plusz információk kinyerése
 - Egy IP címből
 - ország
 - város
 - Egy dátumból (+idő)
 - Hétvége-e
 - Ünnepnep-e
 - A hét melyik napja ez



Feature extracting - példa

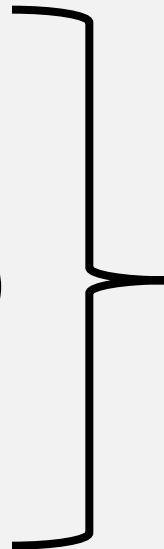
- Plusz információk kinyerése

- Egy IP címből

- ország
- város

- Egy dátumból (+idő)

- Hétvége-e
- Ünnepnep-e
- A hét melyik napja ez



- Helyi idő
- Reggel/délután/este
- Munkaidő-e

❓ Végezetül, egy kis kedvcsináló:

- Mindenki számára elérhető, feature engineering-et nem (vagy alig) igénylő adatok:
 - <https://www.kaggle.com/datasets>
- Ahol érdemes elkezdni:
 - Python, SciKit
- 10 Machine Learning based Products You MUST See
 - <https://www.youtube.com/watch?v=dcZvhP-lqY4>

