



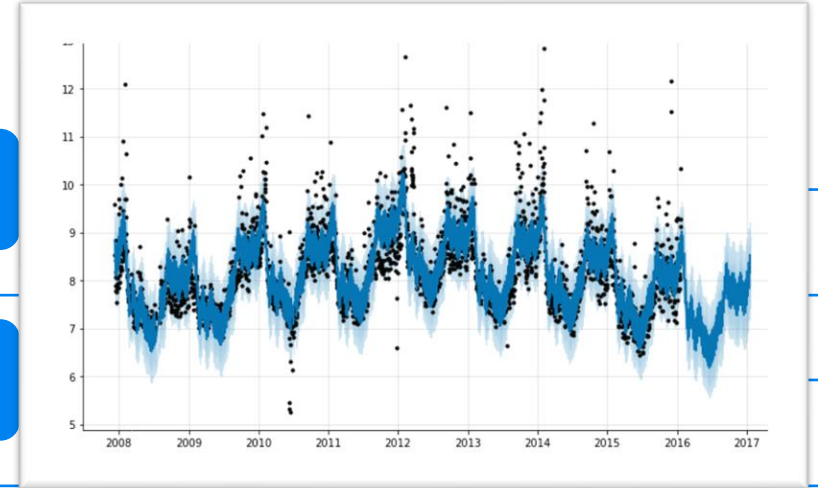
Online idősor monitorozás telekommunikációs hálózatokon

GAM – idősor modellezés

GAM = Generalised Additive Model

Nemlineáris idősorok modellezésére

Önmagában értelmezhető komponensek



$$y_t = T(t) + S(t) + H(t) + \varepsilon_t = \tilde{y}(t) + \varepsilon_t$$

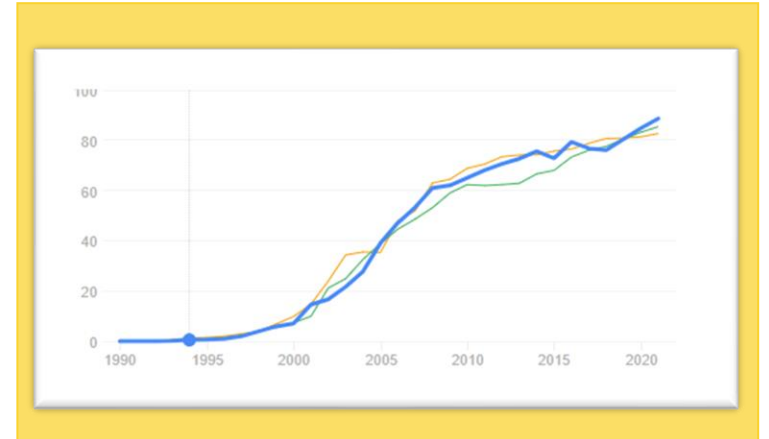
trend szezonális ünnepnap f.a.e. zaj

GAM – Trend

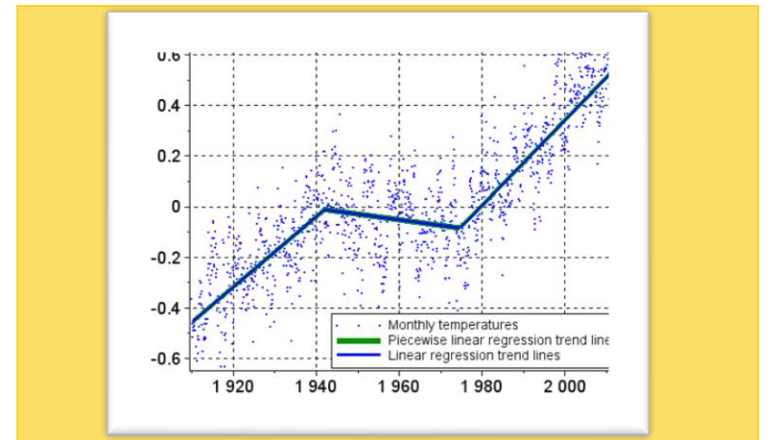
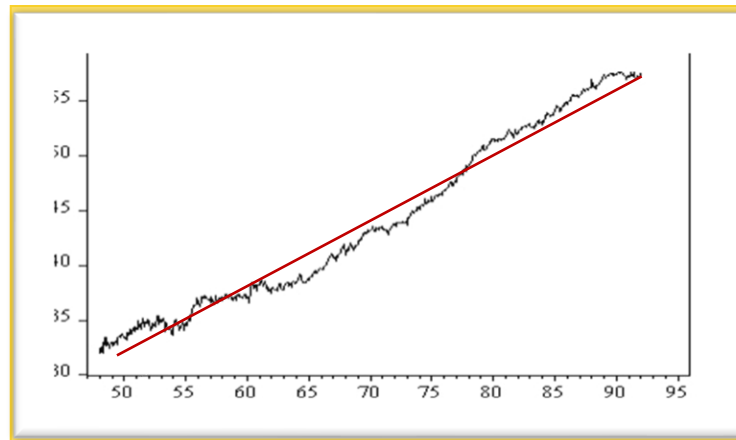


Hosszantartó változás

Logisztikus: gyors növekedés,
telítődés



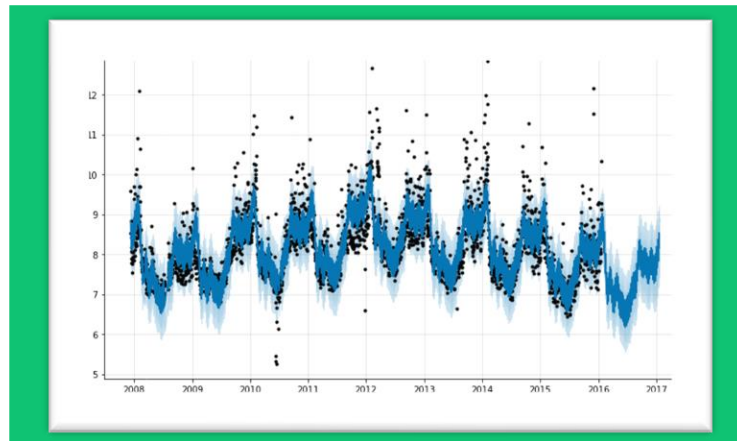
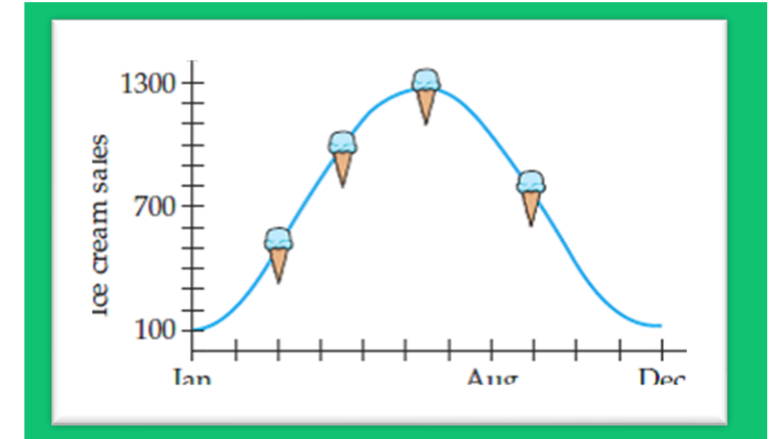
Lineáris: monoton növekedés



GAM – Szezonalitás

Naptárral összefüggő
ismétlődő változások

- Éves
- Havi
- Heti
- Napi
- Etc.



Hogyan illesztjük?

Fourier-sorokkal

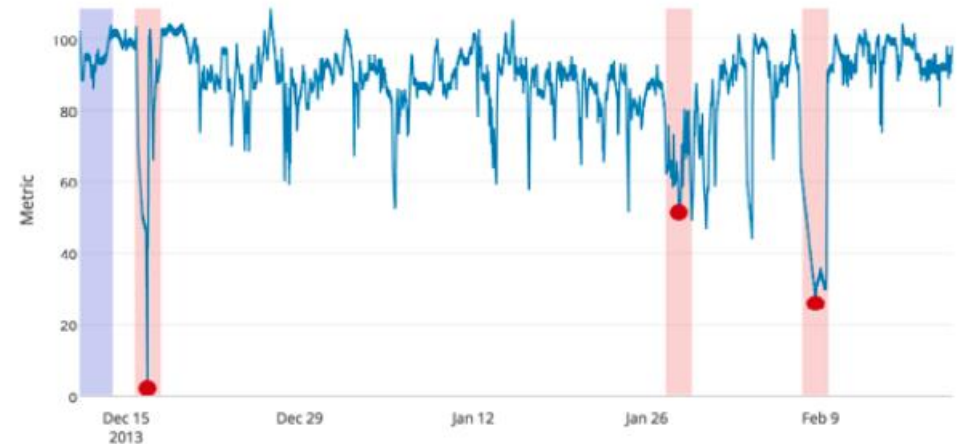
$$s(t) = \sum_{n=1}^N \left(a_n \cos\left(\frac{2\pi nt}{P}\right) + b_n \sin\left(\frac{2\pi nt}{P}\right) \right)$$

Anomália detektálás

- Anomália: kis időbeli kiterjedésű szokatlan esemény
- Kalibráció: mi számít váratlannak?

$$|y_t - \tilde{y}(t)| = \varepsilon_t \quad \text{nagy}$$

$$y_t = T(t) + S(t) + H(t) + \varepsilon_t = \tilde{y}(t) + \varepsilon_t$$

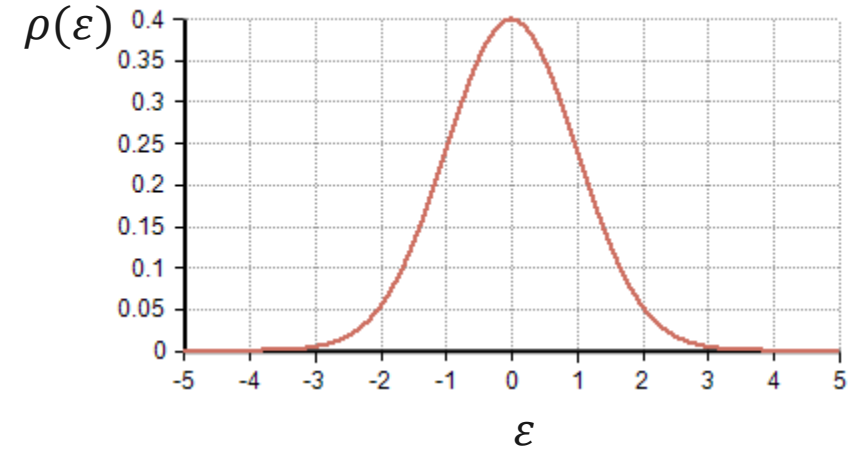


- y_t a múltban: tanítóhalmaz
- $\tilde{y}(t)$ a múltban: illesztett modell

Anomália detektálás 2.

- y_t a múltban: tanítóhalmaz
- $\tilde{y}(t)$ a múltban: illesztett modell
(Nem tökéletes az illesztés)

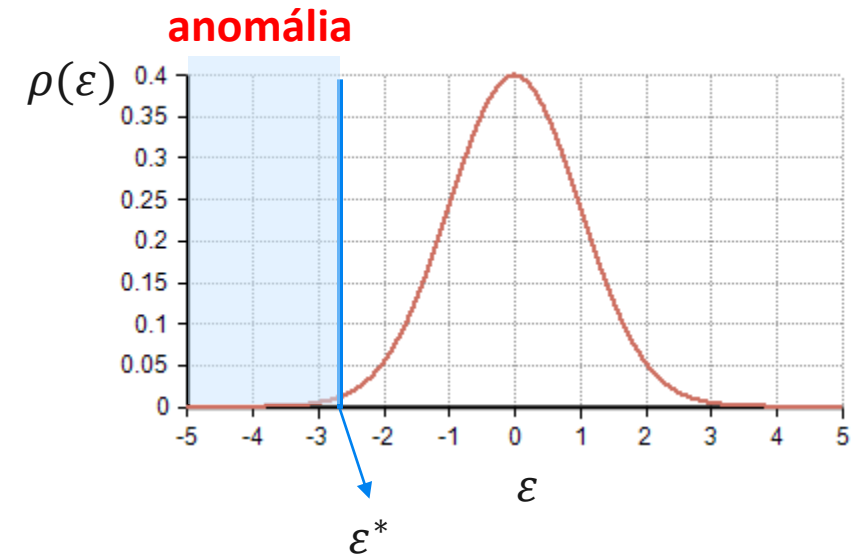
$$\left. \begin{array}{l} \end{array} \right\} \Theta \rightarrow \varepsilon_t$$



Anomália detektálás 2.

- y_t a múltban: tanítóhalmaz
- $\tilde{y}(t)$ a múltban: illesztett modell
(Nem tökéletes az illesztés)

$$\left. \begin{array}{l} \text{ } \end{array} \right\} \Theta \rightarrow \varepsilon_t$$

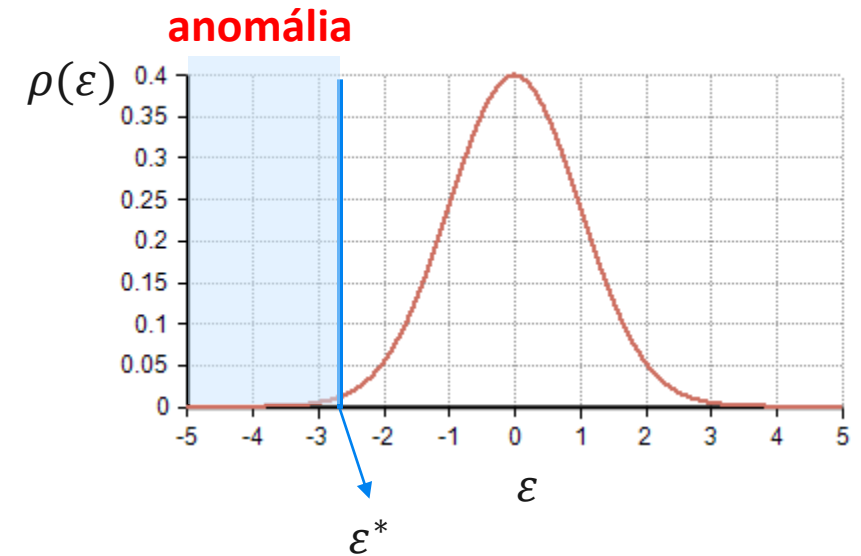


Hogy határozzuk meg a kritikus ε^* értéket?

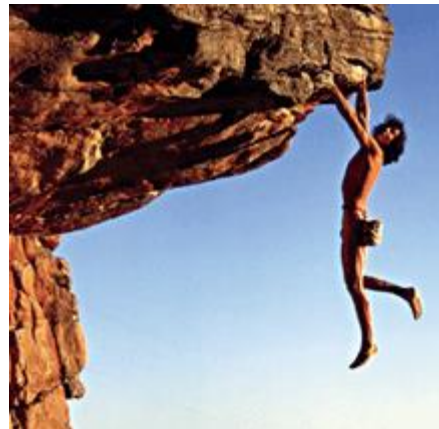
Anomália detektálás 2.

- y_t a múltban: tanítóhalmaz
- $\tilde{y}(t)$ a múltban: illesztett modell
(Nem tökéletes az illesztés)

$$\left. \begin{array}{l} y_t \\ \tilde{y}(t) \end{array} \right\} \Theta \rightarrow \varepsilon_t$$



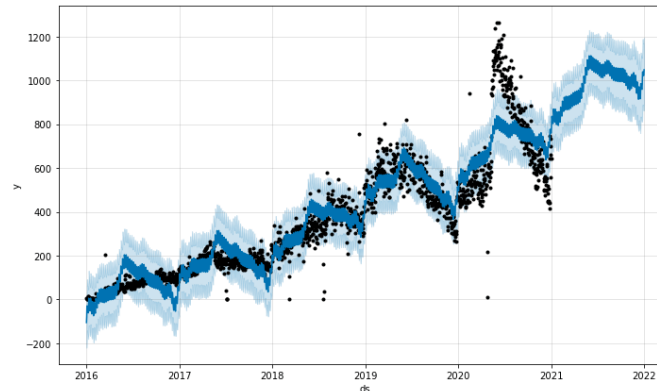
Ez titok ☺



Hogy határozzuk meg a kritikus ε^* értéket?

Modell degradáció

- Van egy problémánk...
- ... a modellek idővel elromlanak
- Minél több idő → annál pontatlanabb modell
- Idővel megváltozik a minták eloszlása

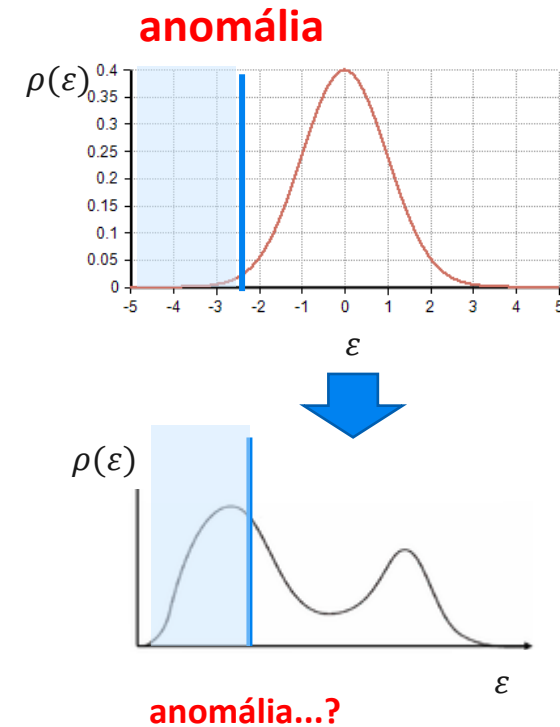


Modell degradáció 2.

- Idővel megváltozik, hogy mi számít normálisnak

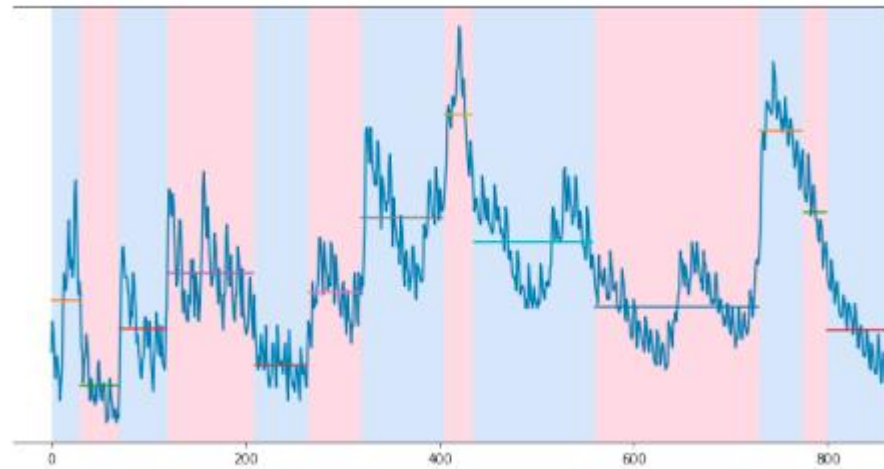
- Megoldások:

- Periodikusan újratanítunk (drága)
- Észrevesszük mikor romlik el a modell – Hogyan?



Changepoint detektálás

Adott y_t idősor, keressük azokat a $(t_1, \dots, t_k)$ időpontokat, ahol hosszútávú szignifikáns változások kezdődtek



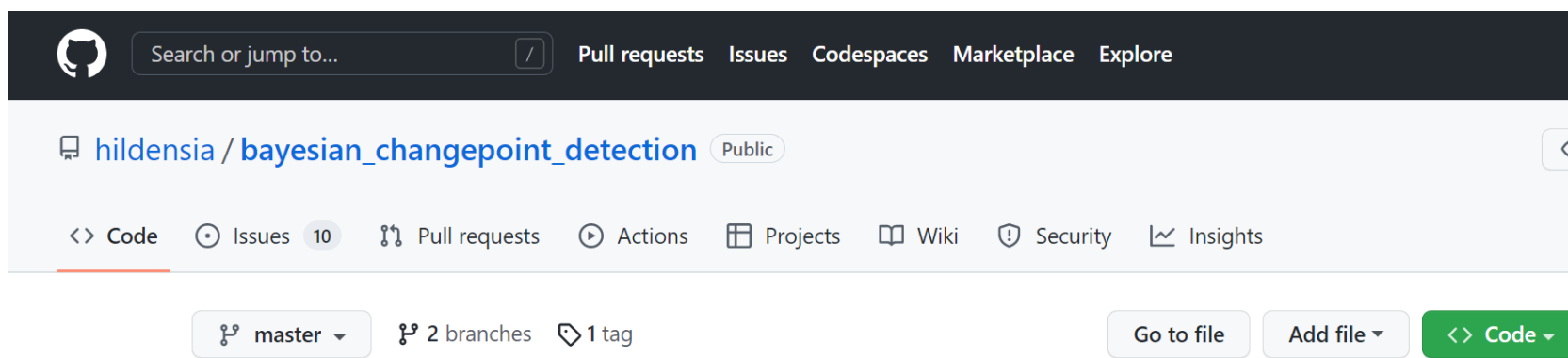
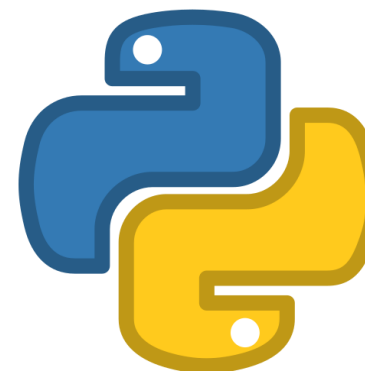


Changepoint detektálás

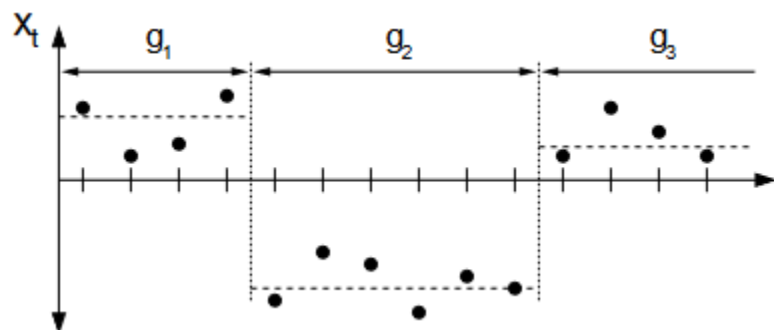
Bayesian Online Changepoint Detection

Ryan Prescott Adams
Cavendish Laboratory
Cambridge CB3 0HE
United Kingdom

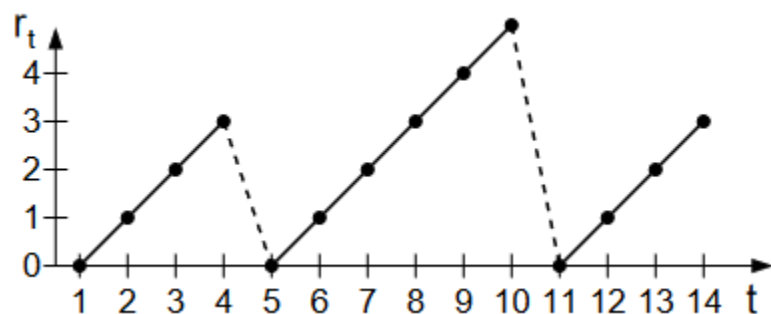
David J.C. MacKay
Cavendish Laboratory
Cambridge CB3 0HE
United Kingdom



Changepoint detektálás



Changepoint
detection



I.I.D. változók
minden ablakban

Nincs korlátozva
milyen eloszlás

Kiszámoljuk az
ablakhosszok
valószínűségét

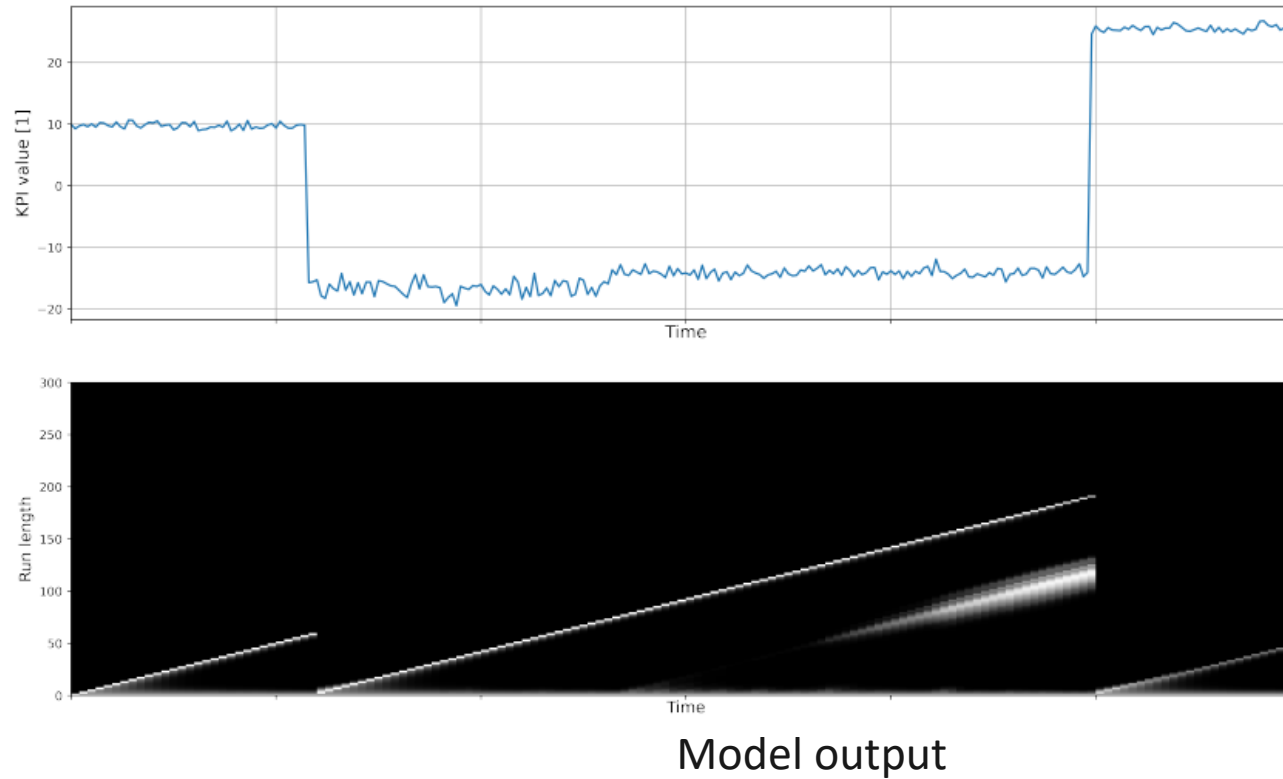
Becsüljük meg a
leghosszabb
ablakhosszt

Becsült
ablakhossz esés -
> changepoint



Changepoint detektálás

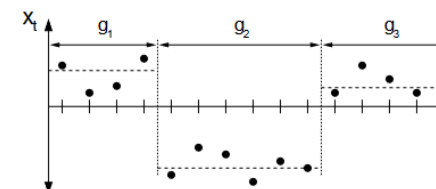
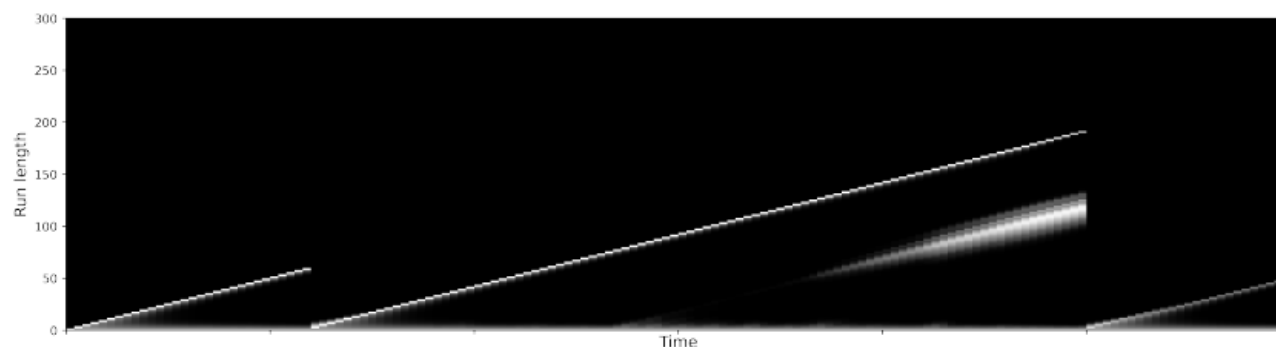
Kiszámoljuk az ablakhossz valószínűségeket





BOCPD működése

Ablakhossz valószínűségek



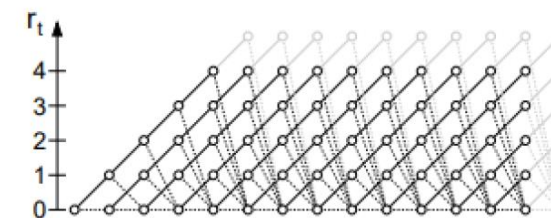
$$P(r_t, \mathbf{x}_{1:t}) = \sum_{r_{t-1}} P(r_t | r_{t-1}) P(x_t | r_{t-1}, \mathbf{x}^{(r_t)}) P(r_{t-1}, \mathbf{x}_{1:t-1})$$

Changepoint prior

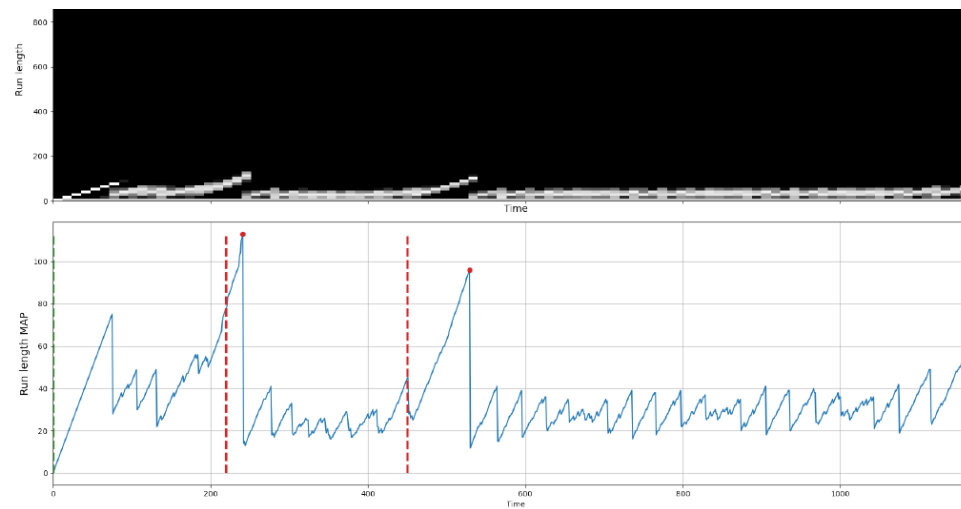
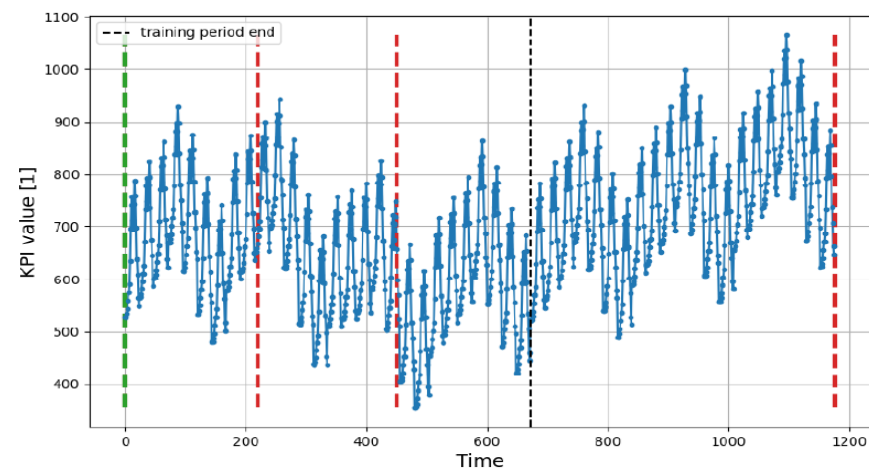
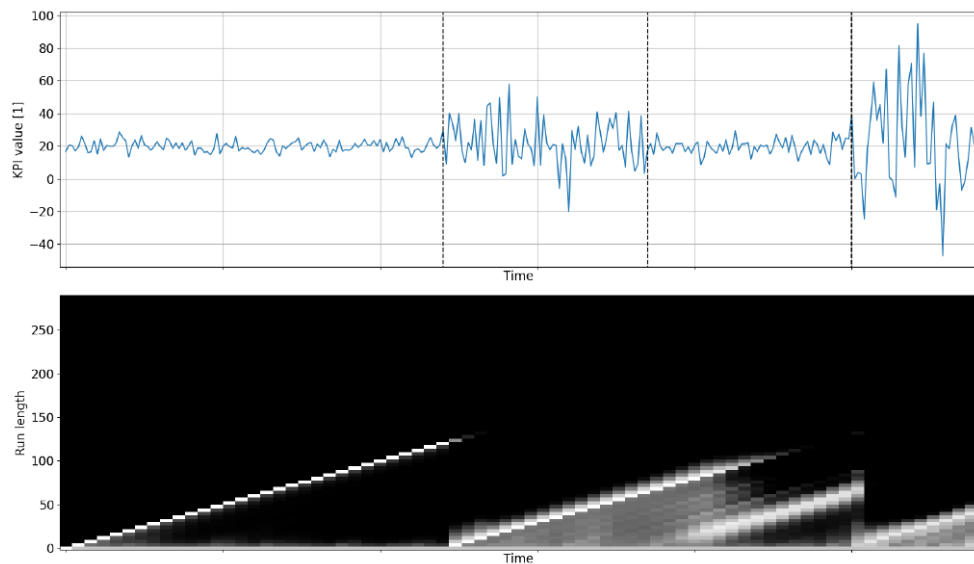
$$P(r_t | r_{t-1}) = \begin{cases} 1/\lambda, & \text{if } r_t = 0 \\ 1 - 1/\lambda, & \text{if } r_t = r_{t-1} + 1 \\ 0, & \text{otherwise.} \end{cases}$$

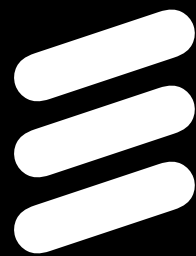
„Azonos eloszlásba tartozás” valsz.

Előző ablakhossz valószínűsége



Eredmények







<https://jobs.ericsson.com/job/Budapest-Data-Scientist-Intern-Buda/763203502/>

Balint.Hantos@ericsson.com