

7. Vremenske serije podataka

Sunday, 20 February 2022 16:52

Gledamo promene promenljive u određenim vremenskim intervalima

Namene:

- Razumevanje - sta uzrokuje promenu (vremenska prognoza)
- Nadgledanje - gledamo masinu da radi ispravno u granicama
- Predviđanje - na osnovu podataka predviđamo buduće vrednosti (vremenska prognoza)

Za opisivanje:

- Opšti trend - $f(t)$
 - Dugoročna promena vrednosti (neki trend)
- Sezonska komponenta - $s(t)$
 - Promena u manjim intervalima sezonama (znamo trajanje sezone, fiksna je dužina)
- Ciklična komponenta - $c(t)$
 - Kao sezonska samo nije vreme fiksne dužine
- Slučajna komponenta - $\varepsilon(t)$
 - Desava se al greskom I ne desava se cesto ili ne ponavljaju se

Modeli vremenskih serija

- Aditivni model - $X_t = f(t) + s(t) + c(t) + \varepsilon(t)$
 - Sve vrednosti su nezavisne jedna od druge pa se samo saberu
- Multiplikativni model - $X_t = f(t) \cdot s(t) \cdot c(t) \cdot \varepsilon(t)$
 - Sve vrednosti su zavisne medjusobno
- Mešoviti model - $X_t = f(t) \cdot c(t) \cdot (s(t) + \varepsilon(t) - 1)$

Time series databases - baze podataka za rad sa vremenskim serijama

Mogu relacije I noSQL baze da se koriste I neke druge

Zahtevi vremenske serije podataka

- Ko-lokacija podataka
- Efikan pristup podacima u proizvoljnom vremenskom intervalu

- Vrednost u sukcesivnim vremenskim intervalima
- Merimo na isti vremenski period vrednosti 3,6,9,12...
- Podaci potrebni zajedno da su radi brzog pristupa na disku
- Gledaju se samo CR a ne CRUD operacije

Osobine koje gledamo

- Preformanse I skalabilnost
- Kompakcija podataka
- Nizi troškovi - jer su optimizovani jeftini su
- Bolje poslovne odluke

RRDTool

- Round robin database tool
- Kad se stigne do kraja memorije podaci se brisu sa pocetka I zamenjuju novim podacima

OpenTSDB

- Java
- Čuva velike količine podataka, brzina nije bitna

KairosDB

Koristi kasandru

Kljuc je timestamp

Kolone vrednost koju merimo