

Give answers in four (4) of the given five (5) questions according to your own choice.

Vastaa neljään (4) kysymykseen alla olevista viidestä (5) kysymyksestä oman valintasi mukaan.

1.

a) Which are the three most typical patterns in spatial data that you can identify by data mining methods?

Mitkä ovat kolme tyypillisintä spatiaalisen datan käyttäytymismuotoa/-kuviota, jotka voidaan tunnistaa datan louhintamenetelmillä?

b) Which are the four typical patterns in trajectory data?

Mitkä ovat neljä tyypillisintä liikerata-tyyppisen datan käyttäytymismuotoa/-kuviota?

In both answers / kummassakin vastauksessa:

- Mention the pattern by name /Mainitse käyttäytymismuodon nimi
- Explain the pattern behavior /Selitä käyttäytyminen
- Give an example of the pattern /Anna yksi esimerkki

2.

You have a database that contains information about a city in five tables:

- Incidents in the city (coordinates and type of the incident)
- Bars/restaurants in the city (center point coordinates of the building in which the bar/restaurant is located)
- Kindergartens (center point coordinates of the building in which the kindergarten is located)
- Offices (center point coordinates of the building in which the office is located)
- Small shops and warehouses (center point coordinates of the building and the type of the shop/warehouse)

Describe a spatial data mining method you can apply in order to identify the potential co-locations of the mentioned events/objects. Describe the method as a process (describe the essential steps of the mining process). Explain also what the result of the process is.

Käytettävissäsi on tietokanta, jossa on informaatiota kaupungista viidessä taulussa:

- Onnettomuudet (koordinaatit ja tyyppi)
- Baarit/ravintolat (kyseisen rakennuksen keskipisteen koordinaatit)
- Päiväkodit (kyseisen rakennuksen keskipisteen koordinaatit)
- Toimistot (kyseisen rakennuksen keskipisteen koordinaatit)
- Pikkukaupat ja tavaratalot ((kyseisen rakennuksen keskipisteen koordinaatit ja liikkeen tyyppi)

Kuvaa millaista spatiaalisen datan louhintamenetelmää voidaan käyttää mahdollisten co-lokaatioiden tunnistamiseksi tapahtumien/kohteiden välillä. Kuvaa menetelmä prosessina (kuvaa louhintaprosessin olennaiset vaiheet). Selitä myös, mitä saadaan tulokseksi.

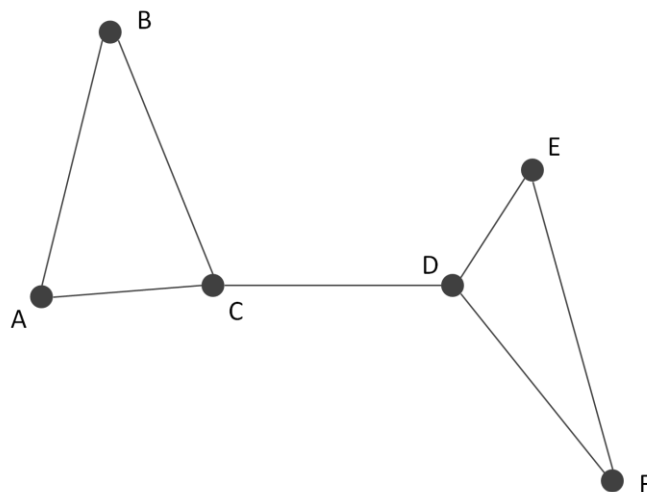
3.

What do the following terms mean in the context of networks/graphs:
Resilience and vulnerability?

Make the dual transformation (line graph) of the graph below and calculate betweenness values for the dual graph. Identify also the cut vertices of the dual graph. Convert the results into the original graph and explain their meaning in case the original graph was a small road network.

Mitä tarkoitetaan verkostojen / graafien yhteydessä termeillä:
Elpymiskyky ja haavoittuvuus?

Muodosta alla olevan graafin duaali (line graph) ja laske sille betweenness-arvot. Tunnista cut-solmut. Muunna tulokset alkuperäiseen graafiin ja selitä tulosten merkitys tapauksessa, että annettu graafi kuvaa pientä katuverkkoa.



4.

Explain the principles of the method, the main steps in using and the interpretation of the results in the use of the following methods:

- Geographically Weighted Regression
- SOM (Self-Organizing Map) that includes spatio-temporal data.

Kuvaa menetelmän pääperiaatteet, menetelmän käyttämisen vaiheet ja tulosten tulkinta seuraavien menetelmien käytössä:

- Maantieteellisesti painotettu regressio
- SOM (itseorganisoituva kartta) kun käytetään spatio-temporaalista dataa.

5.

- a) Describe the main steps of k-means clustering algorithm.
Kuvaa k-means -klusterointialgoritmin päävaiheet.

- b) Explain the use of spatio-temporal cube and 3D Kernel density method in trajectory data mining.
Selitä kuinka spatio-temporaalista kuution ideaa ja 3D Kernel tiheys -menetelmää voidaan käyttää liikeratatieiden louhintaan.