Overview of the Talks of Other Students

Jana Cavojska

Freie Universität Berlin

Abstract This paper provides a summary of the talks given by students during the "Data Visualization and Mining Seminar" in WS 2015/16, as well as the opinions of the author of this summary on each of the talks.

Keywords: visualization, data mining, talk summary

1   Talk 1

Author: **Alexandros Papanikos**
Topic: **Spatial Data Mining**

Summary:
This talk introduced the area of spatial mining as "discovering non-trivial, interesting and useful patterns from large datasets". Alexandros presented the reason why this task is difficult, such as

- complexity of data
- growth of the spatial data collection
- need for high efficiency of the algorithms being used

Some of the most common problems were mentioned, such as clustering.
A couple of applications were described as well, such as

- GIS (geographic information systems)
- medical imaging
- robot navigation
- public health
- transportation
- environmental science (climate change...)
- computer cartography
- ...

Some of the tasks that spacial data mining aims to accomplish:

- classification
- finding association rules (What belongs together?)
- discriminate rules (finding differences between parts of database)
- trend detection (finding temporal patterns in data)
- ...

Remarks:
Pros:
Alexandros gave a good overview of the topic, the problems-applications-goals structure was helpful, making it possible to create a mental picture of the field even if one has not dealt with the topic before.
He managed to make many general examples, mentioning areas where his topic is relevant, as well as introducing some technical terms from his field, but not too many, which would have overwhelmed the audience.

Cons:
The talk was a bit too short (only 15 minutes). Maybe to fill the gap, some of the content of the 2. half of the talk could have been outlined, or some specific algorithms could have been mentioned.
Also, a few specific examples could have helped to create a more persistent mental picture of the topic in the heads of the audience, like describing one scenario in which spatial data mining is used.
Also missing was a clear and simple definition of the term "spatial data".

2   Talk 2

Author: **Dimitris Bourgiotis**
Topic: **Data Mining in Medical Data**

Summary:
The author, Dimitris, explained that medical data in this context is all information patients provide during their care. Since it was not at the centre of this talk, the information gathering process was not further elaborated.
He described some of the problems that arise when handling this sort of data, such as medical errors (which are both harmful and expensive). Generally, medical data is very heterogeneous and has different levels of quality.

Aims of medical data mining:

- better disease prevention
- improvement of treatment effectiveness
- better categorizing of patients in order to provide care where it is needed
- pharmaceutic industry can improve their products
- medical device industry can improve their products

An example was provided to outline how exactly medical data mining could be useful in real life, when Dimitris described a situation in which a doctor is confronted with a patient with heart disease and is able to back his decision about further treatment by consulting a database of similar cases.

Remarks:
Pros:
Dimitris gave a good overview of the topic, explained well why the topic is important and how it is relevant in different parts of the medical field. The specific example of heart disease and how statistical data can provide guidance for further treatment was very helpful for illustrating a situation in which this kind of research is relevant.

Cons:
The slides could have contained a bit more text instead of just bullet points with one word per point and no further explanation. Some of the subtopics were not well integrated into the slides - for example, "heart disease" was mentioned under "Goals" of medical data mining. Also, maybe a couple of more examples could have been mentioned. This talk, too, was a few minutes shorter than it should have been.