Topics for Seminars

Prof. Myra Spiliopoulou
myra@iti.cs.uni-magdeburg.de
Timeline for seminars and teamprojects

- **REGISTRATION:**
  - You decide for a topic, *in agreement with the supervisor.*
  - You sign the form and submit to the KMD team by Monday 17.04.16

- **PRESENTATIONS for seminars:** 16.01.2017, 09:00–12:00
- **BACKUP presentations date:** 23.01.2017
- **REPORT for seminars:** 27.01.2017 till 13:00 s.t.
- **PRESENTATIONS/REPORT** for projects: Arrange with supervisor

**ALL MEETINGS FROM NOW ON:** KMD LAB (R 021)
Four seminar topics (Master level)

Polarity Learning in Micro-blocs or in News
(1) Polarity classification in opinionated streams

TASK: Write a literature overview of classifier methods (no ensembles) for sentiment analysis in opinionated streams of micro-blogs (e.g. tweets)

SUBTASKS:

1. Describe the formal problem solved by each method, focussing on commonalities among the methods.

2. Briefly describe how each method works, focussing on differences among the methods.

3. Specify at least two criteria that allow you to compare the methods.

4. For each criterion, explain which method is best and which methods are not performing well.

Your overview must contain at least 5 methods from different author teams, 2015 or later.

LITERATURE TO BEGIN WITH:

(2) Prediction of popularity and emerging trends in news

TASK: Write a literature overview of trend prediction and popularity prediction methods on streams of news

SUBTASKS:

1. Describe the formal problem solved by each method, focussing on commonalities among the methods.

2. Briefly describe how each method works, focussing on differences among the methods.

3. Specify at least two criteria that allow you to compare the methods.

4. For each criterion, explain which method is best and which methods are not performing well.

Your overview must contain at least 5 methods from different author teams, 2014 or later.

LITERATURE TO BEGIN WITH:

(3) Labeling microblog texts

TASK: Write a literature overview of (a) crowdsourcing approaches and of (b) active learning methods for the annotation of microblog texts

SUBTASKS:

1. Describe the formal problem solved by each method, focusing on commonalities among the methods.

2. Briefly describe how each method works, focusing on differences among the methods.

3. Specify at least two criteria that allow you to compare the methods.

4. For each criterion, explain which method is best and which methods are not performing well.

Your overview must contain at least 5+5 methods from different author teams, 2014 or later.

LITERATURE TO BEGIN WITH:

(4) Preprocessing high-dimensional trajectories

TASK: Write a literature overview of data preprocessing methods for high-dimensional trajectories, with focus on summarization, querying and intelligent matching

SUBTASKS:

1. Describe the formal problem solved by each method, focusing on commonalities among the methods.

2. Briefly describe how each method works, focusing on differences among the methods.

3. Specify at least two criteria that allow you to compare the methods.

4. For each criterion, explain which method is best and which methods are not performing well.

Your overview must contain at least 5 methods from different author teams, 2013 or later. You can include methods for multivariate timeseries in your overview.

LITERATURE TO BEGIN WITH:


- S. Gaffney and P. Smyth (1999). Trajectory clustering with mixtures of regression models. In Proc. of the 5th Int. Conf. on Knowledge Discovery and Data Mining (KDD), pages 63–72
Assigning the seminar topics to MDKE thematic areas

(1) Polarity classification in opinionated streams
(2) Prediction of popularity and emerging trends in news
(3a) Labeling microblog texts – crowdsourcing
(3b) Labeling microblog texts– active learning
(4) Preprocessing high-dimensional trajectories

EXCEPTION: If you focus solely on querying and efficient data engineering, then
Thank you very much!

Questions?