



# **732A75 Advanced Data Mining**

## **TDDD41 Data Mining - Clustering and Association Analysis**

<http://www.ida.liu.se/~732A75>

<http://www.ida.liu.se/~TDDD41>



# Teachers

- Patrick Lambrix, examiner, lectures
  - Johan Alenlöv, lectures
  - Ying Li, labs (732A75)
  - Mina Abd Nikooie Pour, labs (TDDDD41)
- 
- Director of studies: Patrick Lambrix

## Course literature

- Jiawei Han and Micheline Kamber, *Data Mining: Concepts and Techniques*, 2nd edition, Morgan Kaufmann, 2006.

or

Jiawei Han, Micheline Kamber, Jian Pei, *Data Mining - Concepts and Techniques*, 3rd edition, Morgan-Kaufmann, 2011. ISBN: 978-0123814791 (*ROCK missing*)



# Course literature

- Articles.
- Lab descriptions.
- Slides related to book content: many slides adapted or taken from lecture slides by Jiawei Han and Micheline Kamber
- Note: slides are 'working material' – expectation to take notes

# Course organization

- Basic topics in association analysis and clustering
  - Lectures
  - Laboratory exercises (sign up in pairs)
  
- Credit requirements
  - Written examination
  - Laboratory exercises

# Connection to other courses

## ■ Companion course to Machine learning course

☐ Classification

☐ Prediction

☐ Association analysis

☐ Clustering

} ML Course

} TDDD41/732A75



## Changes wrt last year

- Some update of slides
- Former lecturer returns

# My own interest and research

- Modeling of data – Semantic Web
  - Ontologies (for Life sciences, animal health, materials design, crime scene investigation, sports analytics)
- Ontology engineering
  - Ontology alignment  
(Winner Anatomy track OAEI 2008 / Organizer OAEI tracks since 2013)
  - Ontology debugging and completion  
(Founder and organizer WoDOOM/CoDeS 2012-2016)
  - Ontology visualization  
(Founder and organizer VOILA since 2015)

→ Topics covered in TDDD43



# My own interest and research

- Sports Analytics

- Performance measures for players in ice hockey and football; player roles; sports data visualization; trajectory prediction in football; season prediction in basketball
- <https://www.ida.liu.se/research/sportsanalytics/>
- Founder of LINHAC – Linköping Hockey Analytics Conference

- Former work: knowledge representation, data integration, knowledge-based information retrieval, object-centered databases

- <http://www.ida.liu.se/~patla00/research.shtml>



# Sports Analytics

- TDDE64/753A01, 6 hp course in VT2
- MAX 20 participants
- Course page:  
<https://www.ida.liu.se/~753A01/index.en.shtml>



# Sports Analytics

- Lectures
  - Introduction to sports analytics
  - Lectures on research at sports analytics group
  - Guest lectures (e.g., Playmaker, LHC, other research groups)
- Schedule decided 2 weeks before course start
- Credits:
  - Presentations by students on topic of choice in sports analytics
  - Project OR LINHAC student competition

# LINHAC

## Linköping Hockey Analytics Conference - LINHAC 2024

June 3-5, 2024, Linköping, Sweden

[Home](#)[Calls](#)[Program](#)[Registration](#)[Student Competition](#)[Travel](#)[Venue](#)[Organization](#)[Industry packages](#)[Press](#)

### Registration

Previous conferences:

[LINHAC 2023](#)[LINHAC 2022](#)

For enquiries send mail to:

[LINHAC@groups.liu.se](mailto:LINHAC@groups.liu.se)

LINHAC 2024 is planned to take place June 3-5, 2024 in Linköping, Sweden, as a hybrid event.

LINHAC aims to bring together professionals and academics with an interest in hockey analytics. LINHAC will feature the latest research in hockey analytics in academia and companies, panel discussions with analysts, coaches, GMs and players, industry sessions with the latest hockey analytics products, and an analytics competition for students.

### Organized by



Tweets from

<https://www.ida.liu.se/research/sportsanalytics/LINHAC/LINHAC24/home.html>



# LINHAC student competition – open for **all** students

2022 and 2023:

- **Data set**

Data produced by Sportlogiq provided with permission of SHL, the Swedish Hockey League, representing event data from the 2020-2021 SHL season.

- **Task:**

Given the event data, generate findings/patterns related to sequences of events leading up to a particular outcome. You can choose the kind of outcome. For instance, find characteristics of sequences of events leading to, e.g., a goal or a successful zone entry.