

Similarity Search in Large Databases

Kickoff Meeting – Oct 7, 2021





Why this lab?

- The lecture gives you basics and zooms in on selected solutions
- Problem:
 - not feasible to cover all relevant topics
 - new and better solutions are continuously being developed
- Goal of the lab:
 - understand and assess technical, scientific publications
 - learn about clever ideas and solutions not covered in the lecture



Your job: 8 – 4 – 2 - 1

- 1. You will attend at least <u>8</u> talks given by your fellow students.
- 2. You will write <u>4</u> extended abstracts about selected talks.
- 3. Random: 2 / 4 abstracts will be marked.
- 4. You will give <u>1</u> talk (groups of 2-3).



Talks

- Topic: research result of scientific paper on similarity search
 - list of scientific papers will be made available online
 - groups pick topic: first come / first served (Slack #ssdb-ps-2021ws)
- Talks are prepared in groups of 2-3 students
- 30 minutes talk + 15 minutes discussion
- There will be more than 8 talks (at least 10) you choose!
- Attendance = audio + video! Please prepare the setup!



Extended abstracts

- Summarize motivation, key contribution, and limitations of work
- Two pages template will be available, unlimited references extra
- You pick the 4 talks you want to summarize
- High quality writing expected!
 - Should make sense for a reader that does not know the paper
 - Reader understands problem to solve and key ideas of solution
 - Language: concise; no typos, incomplete sentences, broken links ...



Preparing the talk 1/2

- Background / Motivation
 - give the context
 - intuition of the problem to solve / why is it important?
- Problem Definition
 - Short and precise: given ..., the goal is to ...
- State-of-the-art / research gap
 - Which solutions do exist? Why are they not good enough?



Preparing the talk 2/2

- Key Ideas and Solution
 - What are the new idea in this work?
 - How do these ideas help to solve the problem?
 - Describe the solution (long pseudo-code is not a good idea here!)
 - Explain with examples and figures running example if possible
- Experiments
 - How does the solution perform in practice?
 - Insights? Highlights and limitations?
- Main contribution
 - What does the paper add to the state of the art?
 - What can you do now that you could not do before?



Course organization – project groups

- Form project groups (2-3 students) within 1 week (by Wed Oct 13)
- Depending on the number of participants, a second course group will be available in PlusOnline
- Your course group assignment will not have an impact on schedules, talks that you can attend, topics you can choose, etc.



Schedule

- Will be made available when project groups and their topics are finalized
- Deadline group forming: Wed Oct 13
- Deadline topic: Wed Oct 20
- Talks will be 17:00 and / or 18:00
- First talk: earliest Nov 4



Information & Discussion

- Course Website: <u>https://dbresearch.uni-salzburg.at/teaching/</u>
- All discussions in Slack please do not send email
 - Slack channel and workspace \rightarrow see website
 - Also questions about your paper in Slack channel
 - Use a thread for each problem
 - Make clear in your message which paper you are talking about