A Guide to Scientific Writing for (Young) Researchers: Handbook for DSBDA Template from https://tinyurl.com/dsbda-template

ANSGAR SCHERP, Ulm University, Germany

README.MD

Scientific writing is difficult work. This document shall help students understand how to write a scientific document, particularly in machine learning and artificial intelligence.
NOTE: This document is a work in progress. I am transferring content from the paper structure template to hear to clear up things.
For references to interesting papers, surveys etc. use this handbook now!

CCS Concepts: • **General and reference → Reference works**; **Computing standards, RFCs and guidelines**.

Additional Key Words and Phrases: scientific writing, machine learning, artificial intelligence, data science

Contents

Author's Contact Information: Ansgar Scherp, ansgar.scherp@uni-ulm.de, Ulm University, Ulm, Germany.

1 Data Science and Big Data Analytics (DSBDA) Group

1.1 Data Science Readings

We are running a reading club on Data Science on Wednesdays.

How it works: Idea of the reading club is to have a joined chat about recent research papers. Particular focus is text analytics and graph analytics, and general recent methods in deep learning.

Procedure is usually as follows:

- Someone proposes a paper/topic, which is well before the meeting disseminated.
- So everyone has time to read the paper and is actually also expected to have read the paper (otherwise discussions are not so much fun!)
- During the meeting, the proposer briefly summarizes the paper, including key strengths and weaknesses.
- Followed by a round-robin quick feedback from everyone.
- Discussion goes into the details ... :-)

How to subscribe: Interested? Go here to subscribe: https://imap.uni-ulm.de/lists/subscribe/data-science-readings

This is a mailing list on which you receive current information: mailto:data-science-readings@lists.uni-ulm.de

1.2 Lectures, Seminars, Project Groups, and Theses

Lectures: We offer a couple of different lectures for both BSc and MSc students. These are available for self-enrolment with all materials available for download. Please contact us to get information which lectures will be offered the next terms.

- "Graph Analytics and Deep Learning", Self-enrolment for slides (winter 2022/23): https://moodle.uni-ulm.de/course/view.php?id=36399
- "Text Analytics and Deep Learning", Self-enrolment for slides (winter 2021/22): https://moodle.uni-ulm.de/course/view.php?id=26119
- "Web Information Retrieval (and Deep Learning)", Self-enrolment for slides (summer 2021): https://moodle.uni-ulm.de/course/view.php?id=22260
- "Advanced Methods in) Data Mining and Machine Learning", Self-enrolment for slides (winter 2020/21): https://moodle.uni-ulm.de/course/view.php?id=16999
  There are also slides for the full 4 SWS module (same moodle course): https://moodle.uni-ulm.de/mod/folder/view.php?id=254324

My concept for research-based teaching: https://www.uni-ulm.de/fileadmin/website_uni_ulm/zle/Tag_der_Lehre/downloads/Scherp-TdL21-vortrag.pdf

Seminar and Projects: We also regularly offer seminars on data science (BSc/MSc), as well as the module "Project Data Science". For projects, please contact us.

Theses: If you are interested in a BSc or MSc thesis, please contact us. We have compiled a couple of topics here: https://docs.google.com/presentation/d/1k1aEZYX_UM8rWlojgGTV11O85Lu104e2K-CBDg-k-9A

1.3 Examples of Student Submissions

This folder contains examples of submissions from the last years (in PDF).

https://github.com/data-science-and-big-data-analytics/teaching-examples

Please refer to the corresponding sub-folders for an example relevant to a practical group project submitted in the context of a lecture, MSc project, seminar (written for MSc but also suitable for BSc), and MSc thesis.

1.4 Examples of Data Science Frameworks

This git repository explains how to use selected data science frameworks.

https://github.com/data-science-and-big-data-analytics/data-science-frameworks

A README explains how to use it. Furthermore, helpful tips and available infrastructure are stated (bwCloud, bwUniCluster, and Google Colab).

We have also added a slide deck explaining the frameworks a bit and how to use the cloud compute services available to you. Slides explaining this code (with comment function available):

https://docs.google.com/presentation/d/1v41r4zBfYMe7okcziThfDqt0vqsKrPPYjNDRQHZksRI

1.5 Examples of Peer-reviewed Publications from Student Submissions

Some selected publications from student submissions. Will be updated and completed shortly.

- MSc Thesis Fabian Singhofer [DocEng ‘21] (B ranked), **Best paper award!**, https://arxiv.org/abs/2105.08842
- Project STEREO [iiWAS’ 21] (C ranked), https://arxiv.org/abs/2103.14124
- Project Text Summarization [iiWAS’ 21] (C ranked), https://arxiv.org/abs/2105.11908
- MSc Thesis Ishwar Venugopal [IJCNN ‘21] (A ranked), https://arxiv.org/abs/2102.07838
- MSc Thesis Morten Jessen [DocEng ‘19] (B ranked), **Best student paper award!**, https://dl.acm.org/doi/10.1145/3342558.3345396
- MSc Thesis Florian Mai [JCDL ‘18] (**A* ranked**), https://arxiv.org/abs/1801.06717
- Project Quadflor: [KCAP ’17] (A ranked), https://arxiv.org/abs/1705.05311
- MSc Thesis Gregor Große-Bölting [KCAP ‘15, [4]]: **Best student paper nomination!**, https://dl.acm.org/doi/10.1145/2815833.2815838

2 Scientific Paper Writing Guidelines

2.1 General Tips

General writing guidelines.

- Write British English XOR American English, not both.
- Write in the present tense in your work, particularly the abstract, introduction, procedure, results, and discussion. Write in the *past tense* when describing prior work in the related work section.
- Use `seaborn` for data visualization, https://seaborn.pydata.org/index.html. It provides a set of pre-defined palettes, https://seaborn.pydata.org/tutorial/color_palettes.html.
- We use the notation from the Deep Learning book, see https://www.deeplearningbook.org/contents/notation.html.

2.2 Specific Tips

In the course of writing research papers, some discussions appear again and again that are worth mentioning.

2.2.1 Writing Your Proposal / To Begin With Your Resarch. A common question asked is "What is the initial set of items to write for a proposal"? This set of five items is like a guide through your paper. Write it at the beginning and iteratively refine it.

- Title
- Abstract (Jennifer Widom structure)
- Contributions list
- Datasets
- Procedure (including which models and baselines)

2.2.2 Scientific Writing Style.

- Use of `\mathcommands.tex` as much as possible. Unless it is not useful or it is uncommon in the community.
- A dataset D is a multiset of tuples of training samples and labels, but writing $\mathbb{D}$ is strange in the context of machine learning.
- Writing vector $\boldsymbol{z}$ instead of z is very helpful and needed (note the subtle bold font). It allows one to point to the i-th element z_i in the vector $\boldsymbol{z}$, which may be semantically the i-th dimension in cases of embeddings. Note that indexing of the vector starts at 1. So z_i is a convenient way to access elements in the vector $\boldsymbol{z} = (z_1, z_2, \ldots, z_d)$.
  As a convention, one can write $j : k$ to refer to a range of elements in the vector $z_{j:k}$ (note the spacing introduced before and after the colon). This range $j : k$ over elements in the vector $\boldsymbol{z}$ may be read as referring to the i-th elements over a *discrete* interval, i. e., $i \in [j, k] \cap \mathbb{N} = \{i \in \mathbb{N} \mid j \leq i \leq k\}$.
  (Experimental) As such, we may use a generalization of the range notation where the range is defined as a set. Particularly, we are interested in the subsets $S \in \mathcal{P}(\{1, \ldots, d\})$. We denote with z_S the selection of every i-th element in the vector z with $i \in S$.
  Further definitions of accessing elements in vectors, row/columns of matrices (=2-D tensors), and slices of 3-D tensors can be found in the Deep Learning book [3].
- When referring to multiple indices, should one use a comma or not? Consider this example for the cross-entropy loss, and response from ChatGPT (for entertainment). Which variant is better, i. e., more precise and better to read?
  Variant a):

$$L = -\frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{K} y_{ik} \log(\hat{y}_{ik}) \,,$$

  Variant b):

$$L = -\frac{1}{N} \sum_{i=1}^{N} \sum_{k=1}^{K} y_{i,k} \log(\hat{y}_{i,k}) \,,$$

  where N is the number of samples, K is the number of classes, $y_{i,j}$ is the true label (one-hot encoded), and $\hat{y}_{i,j}$ is the predicted probability for sample i and class k.
  "You're absolutely right about the clarity concern. To avoid confusion, especially when subscripts are involved, using a comma as a separator can make the indices clearer. The more explicit notation $y_{i,j}$ is indeed preferable because it clearly distinguishes the indices for each dimension of the matrix." In other words, use variant b).

2.2.3 *Missing Hyperparameter Details.*

- One often does not find all the necessary information about a model properly described in a paper. A lack of providing details about hyperparameter values, e. g., the learning rate, the train-test splits used, etc. happens quite often. This is comparably easy to spot. Sometimes, this is not so easy like this example. In any of those cases, it is generally necessary to consult the source code of the paper or even contact the authors themselves.
- For example, the authors of the ExaRanker paper [1] state about the choice of hyperparameters:

  > The model was finetuned for 30 epochs using the AdamW optimizer [24] with a learning rate of 3e-5, weight decay of 0.01, and **batch size of 128 examples** (64 positives and 64 negatives)

  The high batch size suggests that the authors make use of gradient accumulation as experiments with a batch size of 128 would result in out-of-memory on the GPU. Looking into the source code at https://github.com/unicamp-dl/ExaRanker/blob/main/monoT5-bin/main_trainer.py reveals that indeed gradient accumulation is used.

  ```
  [...]
  batch_n = 4
  [...]
  train_dataloader = torch.utils.data.DataLoader(train_dataset, batch_size=batch_n,
     num_workers = 0, shuffle=True)
  [...]
  accum_batch = 32
  [...]
  trainer = pl.Trainer(enable_checkpointing=False, log_every_n_steps=1,
      default_root_dir = 'monoT5-bin/chk', accumulate_grad_batches=accum_batch,
      gpus=num_gpus, max_epochs=num_epoch, logger=neptune_logger, callbacks=[lr_monitor])
  [...]
  ```

  The effective batch size is batch_n$\cdot$accum_batch $= 4 \cdot 32 = 128$. This information is crucial for reimplementing the model. It is advised to state in the hyperparameter section if gradient accumulation is used and how.

2.2.4 Datasets and Their Versions and Quality. When there are different versions of datasets around and/or if the datasets used in a are not properly reported/identified.

- For example, see the "The tale of two MS MARCO – and their unfair comparisons" [5]. There are "[...] two different corpora of MS MARCO are used in the literature, the official one and a second one where passages were augmented with titles, mostly due to the introduction of the Tevatron code base."
  It is more than just not knowing exactly what dataset is used. Sometimes, datasets are even harmful, as "the addition of titles actually leaks relevance information, while breaking the original guidelines of the MS MARCO-passage dataset." [5]
  "[...] if a paper does not properly report which version is used, reproducing fairly its results is basically impossible. Furthermore, given the current status of reviewing, where monitoring state-of-the-art results is of great importance, having two different versions of a dataset is a large problem." [5]
  So if you use MS MARCO, **you should use MS MARCO v1** and also write this version exactly into the paper. As test sets, one can use TREC DL 2019 and TREC DL 2020.

- This problem happens more than one might expect. Another example are the two graph datasets Chameleon and Squirrel as they contain a train-test leak [8]. *Should they be still used because other papers also still use it?* **No!** These datasets are faulty.

2.2.5 Prompt-hacking. "[...] we are in the midst of a "replication crisis" in AI research. Psychology and related social sciences have been experiencing a crisis in which a substantial number of published results do not replicate, often due to p-hacking to obtain statistically significant findings." [6]

"Much like the p-hacking crisis in the social sciences, prompt-hacking does not imply nefarious intent or active wrongdoing on the part of a researcher. Indeed, researchers may be entirely unaware they are engaging in this behavior." [6]

Prompt-hacking might include any of the following research practices [6]:

- "Carefully crafting dozens or even hundreds of prompts (manually, programmatically, or via generative AI tools) to obtain a desired result but not reporting in a paper the number of prompts tried that failed to produce desired results, and whether the prompt(s) that did produce desired results had any properties that systematically differentiated them from those that failed."
- "Not checking whether slight variations in a successful prompt alter the research results."
- "Not checking whether a prompt is robust across multiple models, multiple generations of the same model, or even the same model when repeated several times."

2.3 Paper Checklists

There are a couple of checklists that are employed at different conferences. Below are some examples.

- NeurIPS Paper Checklist Guidelines, https://neurips.cc/public/guides/PaperChecklist
- Guidelines for Answering Checklist Questions, https://aclrollingreview.org/responsibleNLPresearch/
- EMNLP 2021 Submission Guidelines, https://2021.emnlp.org/call-for-papers
- See also: resources → reproducibility-criteria

FROM EMMNLP Submission Call, https://2021.emnlp.org/call-for-papers ============================

Ethics / Impact Statement —————————- Tick below if your submission contains an ethics consideration / impact statement. Note that the impact statement is optional.. I/We have included an ethics / impact statement as part of our conference submission and understand that this will be taken into consideration during the review process.

Reproducibility Checklist —————————- Before you submit, please make sure that the following reproducibility checklist is filled.

For all reported experimental results: —————————————— A clear description of the mathematical setting, algorithm, and/or model (*) Submission of a zip file containing source code, with specification of all dependencies, including external libraries, or a link to such resources (while still anonymized) (*) Description of computing infrastructure used (*) The average runtime for each model or algorithm (e.g., training, inference, etc.), or estimated energy cost (*) Number of parameters in each model (*) Corresponding validation performance for each reported test result (*) Explanation of evaluation metrics used, with links to code (*)

For all experiments with hyperparameter search: ————————————————— The exact number of training and evaluation runs (*) Bounds for each hyperparameter (*) Hyperparameter configurations for best-performing models (*) Number of hyperparameter search trials (*) The method of choosing hyperparameter values (e.g., uniform sampling, manual tuning, etc.) and the criterion used to select among them (e.g., accuracy) (*) Summary statistics of the results (e.g., mean, variance, error bars, etc.) (*)

For all datasets used: ————————- Relevant details such as languages, and number of examples and label distributions (*) Details of train/validation/test splits (*) Explanation of any data that were excluded, and all pre-processing steps (*) A zip file containing data or link to a downloadable version of the data (*) For new data collected, a complete description of the data collection process, such as instructions to annotators and methods for quality control (*)

If the above items are not applicable or if you have any additional comments, please provide your feedback below.

Note: This list is based on Dodge et al, 2019 and Joelle Pineau's reproducibility checklist. Dodge: https://www.aclweb.org/anthology/D19-1224.pdf Pinaue https://www.cs.mcgill.ca/~jpineau/ReproducibilityChecklist.pdf

Further checklists for papers:

CoLLAs 2024, https://lifelong-ml.cc/reproducibility

NeurIPS 2021 Paper Checklist Guidelines, https://neurips.cc/Conferences/2021/PaperInformation/PaperChecklist

3 Administrative and Others

Structure of the Proposal. You may well use this template also for writing the proposal of your thesis. Please make sure to cover these topics.

- Motivation
- Problem statement (incl. assumptions!)
- Research questions (separate in mandatory / optional)
- Methods (you plan to apply and/or newly develop)
- Dataset(s) (possibly also: benchmarks)
- Related work (few, key papers only in the proposal)
- Schedule (how to use the 6 months of work; commonly we use 4 months for develop, 2 for evaluation; writing starts on day 1)

Proposal is typically short, few pages (e. g., 1-2 A4 pages) in this template.

Forms for registering a thesis at UULM. MSc Thesis: https://www.uni-ulm.de/fileadmin/website_uni_ulm/studium/Studienorganisation/Pruefungsanmeldung/Formulare/antrag_masterarbeit_WEB.pdf

BSc Thesis: https://www.uni-ulm.de/fileadmin/website_uni_ulm/studium/Studienorganisation/Pruefungsanmeldung/Formulare/antrag_bachelorarbeit_WEB.pdf

And do not forget to have your signature on the paper regarding the statement of originality, see following page.

4 About: Abstract

Information on how to write the abstract.

Abstract: How to write it

An abstract conveys in a summary of 150 words your research idea, experimental results, and their impact. It is an opportunity to directly communicate the key message of your proposal, which otherwise has to be collected from different places in the paper. With order words: *Not including an abstract in a proposal is a missed opportunity!*

This template is for papers, research-based group work reports, BSc and MSc theses, seminar works, etc. It is based on a common ACM style, which is both popular in the computer science research community as well as well maintained. For the author's information, create an ORCID and add it to your record, see the example of the first author. You can obtain an ORCID here: https://orcid.org/

For comments and feature requests, please email Ansgar at ansgar.scherp@uni-ulm.de.

Submission: *We pledge to make the source code and additional resources publicly available upon acceptance of the paper. An (anonymous) preview for the reviewers can be found at: http://anonoymo.us/me.*

Submission (if already available on arXiv): *An earlier version of this paper has been published on arXiv (add cite). We release the source code upon acceptance of the paper.*

Final: *The source code and additional resources are available at: http://anonoymo.us/me*

For the abstract, please follow the Jennifer Widom structure.

Note on the Use of Generative AI Tools

We are following the procedure of the German Research Foundation regarding the use of generative AI tools.

- Please carefully read the DFG's "Guidelines for Dealing with Generative Models for Text and Image Creation", which are available here: www.dfg.de/download/pdf/dfg_im_profil/geschaeftsstelle/publikationen/stellungnahmen_papiere/2023/230921_statement_executive_committee_ki_ai.pdf
- A very good "Artificial intelligence guidance" of what one can do and what not is also found here: https://www.essex.ac.uk/student/exams-and-coursework/artificial-intelligence
- This coincides with recent regulations at international conferences such as the International Conference on Machine Learning (ICML), which states: " The Large Language Model (LLM) policy for ICML 2023 prohibits text produced entirely by LLMs (i.e., "generated"). This does not prohibit authors from using LLMs for editing or polishing author-written text.". Source: https://icml.cc/Conferences/2023/llm-policy.

5 About: Introduction

What is Ego-less Research?

Define good research questions and run experiments that generate scientific insights, i. e., new knowledge. Do not aim to develop a new method and compare it to weak baselines, cherry-picked datasets, and experimental conditions that favor your model.

Have a throughline in your paper and maintain it!

A paper must be consistent and coherent in what it wants to convey to the reader. This means that you need to define and maintain a throughline in your paper.
Key place in the paper to check for coherence and consistency are

- title → does it contain the key message, which is then picked up in the abstract and elaborated in the introduction,
- abstract,
- introduction → contributions list and research questions, respectively,
- datasets,
- procedure → is there .

Whenever you make changes at one place, check and update the others, too!

Instructions: Write following this structure.

To organize the introduction, the proposed structure of Jennifer Widom should be used. Not using the structure may leave an introduction oftentimes meaningless, when it ends at the motivation and does not well explain the *why is it a problem* and *why is it not solved* parts.
Write explicit paragraphs for each of the questions. Furthermore, make sure that the introduction picks up every statement made by the abstract. The goal of the introduction is to extend the gist provided by the abstract by giving more detail, more context, explanations, and, very important, citations to definitions, related work, and methods.

This template is based on the official "Association for Computing Machinery (ACM) - SIG Proceedings Template" provided on Overleaf. A documentation is provided in this project. The template is taken from Overleaf: https://www.overleaf.com/latex/templates/association-for-computing-machinery-acm-sig-proceedings-template/bmvfhcdnxfty

The official URL to this Overleaf template is: https://www.overleaf.com/latex/templates/dsbda-templateforpaper-annotated/svwvwvqxfxtp You may also use the view link (ready only): https://www.overleaf.com/read/mpmsdhfcwdfk.
If you look for a template for presentations/slides, Fabian Singhofer is so kind to share his for DSBDA: https://www.overleaf.com/read/qxrdtnzrrpwc

Links are "read"-links, so one can copy it into a new project.

By default, the language is set to American English.

The concept of the teaching programme is also documented and available here: https://github.com/data-science-and-big-data-analytics/teaching-examples/blob/main/Scherp-TdL21-vortrag.pdf

Note that there are also new writing tools that support academic writing. For example, Grammarly: https://www.grammarly.com/blog/academic-writing/

Note: Yellow boxes provide background information, additional notes, recommendations, etc. and can later be removed.

Apply Jennifer Widom structure, which is encoded here in the yellow boxes.

What is the motivation?

Motivate your work.

What is the problem?

Describe in precise terms what the problem is that you address. This definition of the problem is used/referred to throughout the paper.

Why is it a problem?

Describe the relevancy of the problem.

Why is it not yet solved?

Describe why are existing solutions insufficient.

What is our solution approach?

Describe the method/algorithm that you propose to solve the problem.

What are the results?

Describe key results from your experiments. Mention datasets, measures, and observations. Reflect on the key insights by a brief discussion. Make the reader interested in your paper.

What are your contributions?

Instruction: Write down your list of contributions.

The introduction (and the structure of it) needs to match the bullet items of contributions at the end of the introduction. There is a clear disconnection and break in the paper if the introduction describes the motivation well, but the contributions list is about something else, see also comment below.
Your contributions list is a main point of discussion. It has to be done well.

Below, we summarize our contributions.

- Provide a bullet-itemized list of research questions that you address.
- Later, each research question will then be turned into a contribution, i. e., a brief answer to the question is given.

Introduction What is a contribution item and what not.

The bullet items of contributions need to be a precise description of research questions that are phrased as how they make a contribution beyond the state of the art. For example, “We compare our method X with three strong baselines A, B, and C to demonstrate the effectiveness of our approach on nine benchmark datasets. [...].” The contributions list may not be a description of implementation steps, e.g., we first pre-process data, we train the models, and we evaluate the models, etc.

The remainder of the paper is organized as follows. Below, we summarize the related works. Section 7 provides a problem statement and introduces our models/methods. The experimental apparatus is described in Section 8. An overview of the achieved results is reported in Section 9. Section 10 discusses the results, before we conclude.

6 About: Related Work

When reading the related work, we aim to understand the method(s), datasets used, results of the experiments, and what the results mean, i. e., how the authors argue about the results in the discussion.

Instructions

To check the trustworthiness of results, we always perform some checks (derived from [2]). Papers, where one has to tick one of the items below, do not allow for a fair comparison with the state of the art. Reasons include that they

- used different or non-standard benchmark datasets,
- modified the datasets to use a different number of classes (i. e., reducing the number of classes in the preprocessing),
- modified the datasets to use additional information (e. g., additional header metadata in the 20ng text dataset),
- employed different train-test splits (e. g., use more training samples than others),
- used a different, smaller number of training examples (e. g., run their methods only on 5% of the training data while using a benchmark dataset),
- not report the train-test splits (and thus the training data used remains unclear),
- do not report hyperparameter values (particularly the learning rate),
- do not report an average over multiple runs of the experiments together with the standard deviation (Avg. and SD will allow to assess the influence of random factors like the initialization of model weights),
- have not optimized or do not use optimal hyperparameter values (e. g., the learning rate strongly influences the results as demonstrated at the examples of BERT and RoBERTa by Galke et al. [2]),
- do unusual preprocessing on the datasets (e. g., apply preprocessing for models that do not require it like BERT, drop samples in a multi-labeling task that have 1 label and thus modify the datasets, etc.),
- are unclear about the measure(s) used (e. g.,, while writing “we use the F-score” most likely means the (harmonic) F1-score, it still does not detail if micro-averaging, macro-averaging, or samples-averaging F1 is reported),
  or
- it is not mentioned if the procedure applied considers training a (graph) neural network in an inductive versus transductive setting (transductive models are inherently performing better on graph tasks) .

IMPORTANT: See also, and tead the summary of dozens of practices in machine learning that may invalidate the results of a research paper. “Questionable practices in machine learning”, https://arxiv.org/abs/2407.12220

The rationales for not using benchmark datasets or employing other train-test splits are not always clear. Also, the papers often do not properly report hyperparameter values or miss reporting any other of the items above.

As a general rule when reading related work

Be suspicous and ask yourself: “Can I trust their results?” Keep in mind: A primary objective of the paper is to put their method in a good light.

And an important lesson when searching for literature.

Lesson learned (once) again!

If you search for literature and do not find anything. Likely you just did not search for the right keywords. For example, if you search for research on "(source) code segmentation", you will be disappointed (or happy) not to find any. But do not be a fool. There is work, it is "text segmentation" a classical area in natural language processing. You just have to think about source code being an (artificial) language that any modern tool will process in the same way as a natural language.
A good hint is also if the task is visible in the community. For text segmentation there exists its own category on Papers with Code, see https://paperswithcode.com/task/text-segmentation.

Writing hint: Use [**?**] or **?**].
But always put a tilde (̃) before the \cite.

6.1 Area 1

6.2 Area 2

6.3 Area ...

6.4 Summary/Reflection

What do we learn from the literature concerning your work? Where are their strengths, and where are their weaknesses? What is different in the related work compared to the proposed approach?

7 About: <MyMethod> *or* Methods *or* Models

Methods : Which methods do apply?

7.1 [Problem Statement/Problem Formalization]

(if not done as part of the introduction)

7.2 Assumptions

Assumptions: What are assumptions?

The assumptions describe explicitly what characteristics of the dataset, method, etc. are assumed when running the experiments. What assumptions you make are as different as the research questions. An example of an assumption in graph learning is "We assume to have access to unlabeled test nodes during training, i.e., we assume a transductive graph learning setting."

- What are the assumptions that you make?
Note: make sure there is an explicit section or subsection called "Assumptions" in your paper.

Example: A textbook example of what an assumption is

Our primary assumption [for bibliographic metadata extraction] is that all necessary information can be found within a one-hop crawl of the landing page associated with the DOI. This assumption is based on our observation that publishers present key bibliographic information on the landing page or pages directly linked to it e. g., the PDF of the publication.

Assumptions: Difference to research questions.

The assumptions are clearly not the same as the research questions (that are to be stated in the introduction). *Writing the research questions in the section on assumptions is not possible.*

7.3 Methods for Aspect 1

Point of Discussion: Provide a bullet-itemized list of the aspects that are considered by your research. For each aspect, provide a description of the methods/models used and proposed (own methods). Make sure it is consistent with the research questions/contributions describe in the introduction.
Example: Aspects are: a) clustering algorithms, b) embedding methods, c) similarity measures. Instances for a) are DBCAN, k-means, etc., b) TF-IDF, BERT, etc., c) cosine similarity.

- Method 1
- Method 2
- ...

7.4 Methods for Aspect 2

7.5 Methods for Aspect 3

7.6 Summary

8 About: Experimental Apparatus

Follow the description of the experimental apparatus given the structure below.

Make sure to cover the questions provided in the paper writing guidelines, see Appendix 2.

8.1 Datasets

Dataset: What needs to be included in the description?

The used datasets need to be described including a table showing relevant descriptive statistics. This includes the number of samples in the data set and the split of the dataset into the train, validation, and evaluation sets. Other information relevant to the experiment needs to be included such as the total number of classes and the average number of classes per sample (in case of multi-label classification), the average length of a document, etc. Commonly this information is provided in tabular form. What information is to be included depends on the research question. A good guide is to look it up from closely related papers. *Independent of what is reported on the datasets, it is always necessary to add for each average also the standard deviation.*

Datasets: Which datasets do you use? Provide descriptive statistics, usually in tabluar form.

Point of Discussion: Make sure that your datasets fit to the problem and research questions, respectively. Make sure that the datasets are available. Available means that you have a) the license obtained (if needed) and b) the datasets are actually on your disk (copied).

8.2 Preprocessing *or* Pre-processing

Describe the steps that are needed to prepare the datasets for the experiments. It is commonly about rather technical steps that are important for a good reproducibility of the work.

8.3 Procedure

Procedure: What needs to be described to understand the experiments.

The experimental procedure needs to be clearly described such that one can understand precisely which experiments are carried out and how. Do not mix in pre-processing (it is its own subsection above) nor implementation details (it is a subsection below). Focus on describing how the experiments are used to answer your research questions. So if there are three research questions in the order A, B, and C, one would expect that the procedure describes experiments corresponding to these research questions in exactly this order. If not already clear from the dataset description, include a clear statement about the dataset split including a rationale why this specific split is used. It can be as short as "We use a standard train/validate/test-split of 80, 10, and 10 percent of the dataset, following the literature (cite the papers)."

Point of Discussion: Describe which methods you use along the aspects defined in your research, on which datasets they are applied, etc. Make sure it reflect fully the experiments that you want to carry out according to your own plan defined in the research questions.

Procedure: How do you run your experiments?

8.4 Hyperparameter Optimization

Note: If space is limited, this can be moved to supplementary materials

Point of Discussion: What are the (critical) hyperparameters that you need to consider (beyond the learning rate)? How do you plan to optimize the hyperparameters with respect to the models and datasets? What is the hyperparameter search space?

8.5 Measures *or* Metrics

Measure: How do you measure the results?

Point of Discussion: Regarding the measurements and what to measure, i. e., to which level of detail, please carefully read: John Ousterhout's article on "*Always Measure One Level Deeper*" [7].

9 About: Results

- Report your results in tabular or otherwise structured form.
- Limit to objective results, no interpretation of results
- The results should be written up in the present tense. Not in the past tense.

9.1 RQ1 Results

9.2 RQ2 Results

9.3 ... Results

10 About: Discussion

- Now interpret and reflect on your results.

10.1 Key Scientific Insights [Gained from the Results]

- What is the key takeaway? Reflect on the results (what have we learned from them)?

- What are the key results of your research?
- What interesting insights could you obtain?
- Break down by research question.

10.2 Threat to Validity

- Why may your results be biased/not trustworthy? And why in fact are they trustworthy! How reliable are your analyses? Meaning, critically reflect on whether there may be errors / biases in your analyses. So: What (possible) threats exist that could have made the results unreliable, AND why are these not threats?

- Trick is to write down potential threats and explain why they don't hold true here!

- How reliable are your analyses? Meaning, critically reflect on whether there may be errors / biases in your analyses.

10.3 Generalization

- Will the results be transferable/generalize to other datasets, tasks, models, etc?

- Can one transfer the insights/results to other datasets? ... other scenarios? ... other algorithms? Why can we assume that the results generalize?

Why?

10.4 Future Work and Impact

What is future work?

What is the general impact of your work? — pick up arguments from introduction etc.

[- But also: What is the practical impact.]

11 About: Conclusion

Summarize the key results in an interesting and new way. For example by expanding it to a general broader scope of science, economics, impact to life, etc. :-)

Provide a brief outlook to future work! (If not described in the Section 10.4)

12 About: Limitations

- Reflect on the limitations of your work, so what conclusion cannot or should not be derived from the work.

See also EMNLP's **Mandatory Discussion of Limitations**.

> We believe that it is also important to discuss the limitations of your work, in addition to its strengths. EMNLP 2023 requires all papers to have a clear discussion of limitations, in a dedicated section titled "Limitations". This section will appear at the end of the paper, after the discussion/conclusions section and before the references, and will not count towards the page limit. Papers without a limitation section will be automatically rejected without review.

[...]

> While we are open to different types of limitations, just mentioning that a set of results have been shown for English only probably does not reflect what we expect. Mentioning that the method works mostly for languages with limited morphology, like English, is a much better alternative. In addition, limitations such as low scalability to long text, the requirement of large GPU resources, or other things that inspire crucial further investigation are welcome.

https://2023.emnlp.org/calls/main_conference_papers/#mandatory-discussion-of-limitations

13 About: Author Statement

Author statement based on CRediT (Contributor Roles Taxonomy), see: https://www.elsevier.com/authors/policies-and-guidelines/credit-author-statement

14 About: Ethical Statement

Write about ELSI, i. e., Ethical, Legal, and Social Implications of your research.

Instructions: How to write an ELSI statement?

If you have no idea what to write here, consult your favorite AI. Ask it for a checklist for ELSI considerations. Should you ask the AI? Is it sufficient to ask the AI?

15 About: Acknowledgements

Acknowledgments

Add this mandatory acknowledgment if you use the bwHPC.

The authors acknowledge support from the state of Baden-Württemberg through bwHPC.

This template is co-funded under the “2LIKE - Artificial Intelligence for Individualised Learning Path and Processes” (16DHBKI001) project by the German Federal Ministry of Education and Research (BMBF) and the Ministry of Science, Research and the Arts Baden-Württemberg within the funding line Artificial Intelligence in Higher Education.

The presented research is the result of a Master module “Project Data Science” taught at the University of Ulm in SEMESTER+YEAR. The last author is supervisor of the student group.[1]

The presented research is the result of a Master module “Project Data Science” taught at the University of Ulm in 2022. The last author is supervisor of the student group.

16 About: References

Use the BibTeX entries from DBLP.

Check for correct capitalization of the titles. The BibTeX styles usually ignore capitalization in the `title`-attribute of the entries.

For example, the language model “BERT” [?] may appear in the references as “Bert”. Since it is an acronym, one can force its capitalization by adding braces, i. e., by writing `{BERT}` in the title attribute.

If space is needed, you can abbreviate conference names. Remove information like “Proceedings of the 49th [...]”.

Furthermore, one can use common abbreviations or even full conference acronyms if the abbreviation is clear for the target venue. For example, “International Conference on Learning Representations” may become “Int. Conf. on Learning Representations” or just “ICLR”. Another nice example is “37th IEEE/ACM International Conference on Automated Software Engineering” becomes “Int. Conf. on Automated Software Engineering” or “Automated Software Engineering” (as people will understand that it is a conference).

[1] Author is contributing Conceptualization, Writing - Review & Editing, and Supervision. Statement is based on the Contributor Roles Taxonomy, see: http://credit.niso.org/

17 About: Supplementary Materials

Note: Backward references to main part of the paper is ok. But do not directly refer to figures or tables from body to here.

17.1 Extended Related Work

17.2 Extended Results

17.3 Hyperparameter Optimization

17.4 Detailed Discussions

17.5 ...

18 About: Declaration

References

[1] Fernando Ferraretto, Thiago Laitz, Roberto de Alencar Lotufo, and Rodrigo Frassetto Nogueira. 2023. ExaRanker: Synthetic Explanations Improve Neural Rankers. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2023, Taipei, Taiwan, July 23-27, 2023*, Hsin-Hsi Chen, Wei-Jou (Edward) Duh, Hen-Hsen Huang, Makoto P. Kato, Josiane Mothe, and Barbara Poblete (Eds.). ACM, 2409–2414. https://doi.org/10.1145/3539618.3592067

[2] Lukas Galke, Andor Diera, Bao Xin Lin, Bhakti Khera, Tim Meuser, Tushar Singhal, and Ansgar Scherp. 2023. Are We Really Making Much Progress in Text Classification? A Comparative Review. *CoRR* abs/2204.03954 (2023). https://doi.org/10.48550/ARXIV.2204.03954 arXiv:2204.03954

[3] Ian J. Goodfellow, Yoshua Bengio, and Aaron C. Courville. 2016. *Deep Learning.* MIT Press. http://www.deeplearningbook.org/

[4] Gregor Große-Bölting, Chifumi Nishioka, and Ansgar Scherp. 2015. A Comparison of Different Strategies for Automated Semantic Document Annotation. In *Proceedings of the 8th International Conference on Knowledge Capture, K-CAP 2015, Palisades, NY, USA, October 7-10, 2015*, Ken Barker and José Manuél Gómez-Pérez (Eds.). ACM, 8:1–8:8. https://doi.org/10.1145/2815833.2815838

[5] Carlos Lassance and Stéphane Clinchant. 2023. The Tale of Two MSMARCO - and Their Unfair Comparisons. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2023, Taipei, Taiwan, July 23-27, 2023*, Hsin-Hsi Chen, Wei-Jou (Edward) Duh, Hen-Hsen Huang, Makoto P. Kato, Josiane Mothe, and Barbara Poblete (Eds.). ACM, 2431–2435. https://doi.org/10.1145/3539618.3592071

[6] Meredith Ringel Morris. 2024. Prompting Considered Harmful. *Commun. ACM* 67, 12 (2024), 28. https://doi.org/10.1145/3673861

[7] John K. Ousterhout. 2018. Always measure one level deeper. *Commun. ACM* 61, 7 (2018), 74–83. https://doi.org/10.1145/3213770

[8] Oleg Platonov, Denis Kuznedelev, Michael Diskin, Artem Babenko, and Liudmila Prokhorenkova. 2023. A critical look at the evaluation of GNNs under heterophily: Are we really making progress?. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023.* OpenReview.net. https://openreview.net/forum?id=tJbbQfw-5wv

A Supplementary Materials

B Resources: Very Interesting Paper

Language Model "Alternatives".

- S4: Efficiently Modeling Long Sequences with Structured State Spaces, https://arxiv.org/abs/2111.00396
- Mamba: Linear-Time Sequence Modeling with Selective State Spaces, https://arxiv.org/abs/2312.00752
- xLSTM: Extended Long Short-Term Memory, https://arxiv.org/abs/2405.04517
- KAN: Kolmogorov-Arnold Networks, https://arxiv.org/abs/2404.19756
- gMLP: Pay Attention to MLPs, https://arxiv.org/abs/2105.08050
- Pretraining Without Attention, https://arxiv.org/abs/2212.10544

Language.

- Do Llamas Work in English? On the Latent Language of Multilingual Transformers, https://arxiv.org/abs/2402.10588
- Evolutionary Optimization of Model Merging Recipes, https://arxiv.org/abs/2403.13187
  TL;DR: Language model merging on the level of layers.
- When LLMs are Unfit Use FastFit: Fast and Effective Text Classification with Many Classes, https://arxiv.org/abs/2404.12365v1
- Leak, Cheat, Repeat: Data Contamination and Evaluation Malpractices in Closed-Source LLMs, https://aclanthology.org/2024.eacl-long.5/
  and
  Proving Test Set Contamination in Black-Box Language Models, https://openreview.net/forum?id=KS8mIvetg2
- LLM2Vec: Large Language Models Are Secretly Powerful Text Encoders https://arxiv.org/abs/2404.05961

Graphs.

- GraphAny: A Foundation Model for Node Classification on Any Graph, https://arxiv.org/abs/2405.20445
- Mirage: Model-Agnostic Graph Distillation for Graph Classification, https://arxiv.org/abs/2310.09486

C Resources: Interesting Paper

Language.

- MM1: Methods, Analysis & Insights from Multimodal LLM Pre-training, https://arxiv.org/abs/2403.09611
  TL;DR: Show among others that adding automatically generated image captions improves the image classification task.
- Better & Faster Large Language Models via Multi-token Prediction, https://arxiv.org/abs/2404.19737
  TL;DR: Learn to predict n tokens in causal language modeling rather than a single token is better.

Graphs.

- Graph Language Models, https://arxiv.org/abs/2401.07105
  TL;DR: Transfer of T5's model weights and modification of the attention mechanism to support graphs.
- Utilizing Description Logics for Global Explanations of Heterogeneous Graph Neural Networks, https://arxiv.org/abs/2405.12654
  TL;DR: Bridges the classical world of logic to fancy graph neural networks.

D Books

Machine Learning. Probabilistic Machine Learning: An Introduction https://probml.github.io/pml-book/book1.html
Probabilistic Machine Learning: Advanced Topics https://probml.github.io/pml-book/book2.html
Lifelong Machine Learning https://www.cs.uic.edu/ liub/lifelong-machine-learning.html
Lifelong Learning: https://www.cs.uic.edu/ liub/lifelong-machine-learning-draft.pdf
The Modern Mathematics of Deep Learning https://arxiv.org/abs/2105.04026
Probabilistic ML intro, https://probml.github.io/pml-book/book1.html,
Probabilistic ML Adv, https://probml.github.io/pml-book/book2.html,
Interpretable Machine Learning
A Guide for Making Black Box Models Explainable https://christophm.github.io/interpretable-ml-book/

Graphs. Deep Learning on Graphs, https://yaoma24.github.io/dlg_book/
Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges https://arxiv.org/abs/2104.13478
GeomDL (also videos etc): https://geometricdeeplearning.com/,
Graph Representation Learning, https://www.cs.mcgill.ca/~wlh/grl_book/

Natural Language Processing. * Eisenstein - Natural Language Processing, https://cseweb.ucsd.edu/ nnakashole/teaching/eisenstein-nov18.pdf (also: https://www.amazon.de/Jacob-Eisenstein/dp/0262042843/)
* Formal Aspects of Language Modeling, 2023, https://arxiv.org/abs/2311.04329
* Formal Aspects of Language Modeling, 2024, https://drive.google.com/file/d/1IYgjs0Vf8TPmVW6w4S125j3G5Asatn4f/view
* Training, Fine Tuning, Inference and Applications of Language Models, https://drive.google.com/file/d/1PtxuMe6JZyBXBuuGkgDnnD JEl5j/view
Modules
* Advanced Formal Language Theory, Spring 2023, https://rycolab.io/classes/aflt-s23/
* Large Language Models, Spring 2023, https://rycolab.io/classes/llm-s23/

Theory/Math.

- Mathematics for Machine Learning, https://mml-book.github.io/
- Mark de Berg, Otfried Cheong, Marc van Kreveld, Mark Overmars: *Computational Geometry*, https://www.cs.cmu.edu/afs/cs/academic/class/15456-s14/Handouts/BKOS.pdf
- Leonid Libkin: *Elements of Finite Model Theory*, https://homepages.inf.ed.ac.uk/libkin/fmt/fmt.pdf

E Surveys

E.1 General Machine Learning

Attention.

- A General Survey on Attention Mechanisms in Deep Learning, https://arxiv.org/abs/2203.14263

Continual Learning.

- A Wholistic View of Continual Learning with Deep Neural Networks: Forgotten Lessons and the Bridge to Active and Open World Learning, https://arxiv.org/abs/2009.01797
- A wholistic view of continual learning with deep neural networks: Forgotten lessons and the bridge to active and open world learning, https://www.sciencedirect.com/science/article/pii/S089360802300014X

Distillation.

- Knowledge Distillation: A Survey, https://arxiv.org/abs/2006.05525

E.2 Graph Representation Learning

- Architectures of Topological Deep Learning: A Survey of Message-Passing Topological Neural Networks, https://arxiv.org/abs/2304.10031
- Foundations and Frontiers of Graph Learning Theory, https://arxiv.org/abs/2407.03125
- Deep Learning on Graphs: A Survey, https://arxiv.org/abs/1812.04202
- Comprehensible Artificial Intelligence on Knowledge Graphs: A survey, https://www.sciencedirect.com/science/article/pii/S1570826823000355
- A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future directions, https://journalofbigdata.springeropen.com/articles/10.1186/s40537-023-00876-4
- A survey of graph neural networks in various learning paradigms: methods, applications, and challenges, https://link.springer.com/article/10.1007/s10462-022-10321-2
- A Comprehensive Survey on Automatic Knowledge Graph Construction, https://dl.acm.org/doi/10.1145/3618295
- A Comprehensive Survey on Deep Graph Representation Learning Methods, https://dl.acm.org/doi/pdf/10.1613/jair.1.14768
- Graph Self-Supervised Learning: A Survey, https://arxiv.org/abs/2103.00111
- Position Paper: Challenges and Opportunities in Topological Deep Learning, https://arxiv.org/abs/2402.08871
- Uncertainty in Graph Neural Networks: A Survey, https://arxiv.org/abs/2403.07185
- Attending to Graph Transformers, https://arxiv.org/abs/2302.04181

Temporal Graphs.

- Signal Processing over Time-Varying Graphs: A Systematic Review, https://arxiv.org/abs/2412.00462
- A Survey on Temporal Knowledge Graph Completion: Taxonomy, Progress, and Prospects, https://arxiv.org/abs/2308.02457
- A Comprehensive Survey of Dynamic Graph Neural Networks: Models, Frameworks, Benchmarks, Experiments and Challenges, https://arxiv.org/abs/2405.00476

Graphs meet LLMs.

- A Survey of Large Language Models for Graphs, https://arxiv.org/abs/2405.08011

Continual Graph Learning.

- Continual Graph Learning: A Survey, https://arxiv.org/abs/2301.12230

Graphs and Heterophily.

- The Heterophilic Graph Learning Handbook: Benchmarks, Models, Theoretical Analysis, Applications and Challenges, https://arxiv.org/abs/2407.09618

Summarization of Graphs.

- A Survey on Graph Condensation, https://arxiv.org/abs/2402.02000
- A Comprehensive Survey on Graph Summarization with Graph Neural Networks, https://arxiv.org/abs/2302.06114
- A Survey on Extractive Knowledge Graph Summarization: Applications, Approaches, Evaluation, and Future Directions, https://arxiv.org/abs/2402.12001
- A Comprehensive Survey on Graph Reduction: Sparsification, Coarsening, and Condensation, https://arxiv.org/abs/2402.03358

- A Survey on Structure-Preserving Graph Transformers, https://arxiv.org/abs/2401.16176

Structural Graph Summarization.

- Structural Summarization of Semantic Graphs Using Quotients, https://doi.org/10.4230/TGDK.1.1.12

Link Prediction.

- Beyond Transduction: A Survey on Inductive, Few Shot, and Zero Shot Link Prediction in Knowledge Graphs, https://arxiv.org/abs/2312.04997
- A Survey on Graph Classification and Link Prediction based on GNN, https://arxiv.org/abs/2307.00865

Recommender.

- A survey of graph neural network based recommendation in social networks, https://www.sciencedirect.com/science/article/abs/p
- Graph Neural Networks in Recommender Systems: A Survey, https://arxiv.org/abs/2011.02260
- A Survey of Graph Neural Networks for Recommender Systems: Challenges, Methods, and Directions, https://dl.acm.org/doi/full/10.1145/3568022

Distillation.

- Graph-based Knowledge Distillation: A survey and experimental evaluation, https://arxiv.org/abs/2302.14643
- Knowledge Distillation on Graphs: A Survey, https://arxiv.org/abs/2302.00219

Continual Learning.

- Continual Learning on Graphs: Challenges, Solutions, and Opportunities, https://arxiv.org/abs/2402.11565
- Graph Learning under Distribution Shifts: A Comprehensive Survey on Domain Adaptation, Out-of-distribution, and Continual Learning, https://arxiv.org/abs/2402.16374

Knowledge Graphs meet LLMs.

- Combining Knowledge Graphs and Large Language Models, https://arxiv.org/abs/2407.06564
  TL;DR: An analysis of 28 papers outlining methods for KG-powered LLMs, LLM-based KGs, and LLM-KG hybrid approaches

E.3 Natural Language Processing / LLMs

- On the Opportunities and Risks of Foundation Models, https://arxiv.org/abs/2108.07258 [argue among others that due to the huge resources required by language models, research on them is pushed into the hands of a few global industrial players only]

LLM. * A Comprehensive Overview of Large Language Models, https://arxiv.org/abs/2307.06435
* Large Language Models: A Survey, https://arxiv.org/abs/2402.06196
* The Life Cycle of Knowledge in Big Language Models: A Survey, https://arxiv.org/abs/2303.07616
* A Survey of Large Language Models, https://arxiv.org/abs/2303.18223
* A Survey of Knowledge Enhanced Pre-Trained Language Models, https://ieeexplore.ieee.org/document/10234662
* A Survey on Deep Semi-Supervised Learning, https://ieeexplore.ieee.org/document/9941371
* The Prompt Report: A Systematic Survey of Prompting Techniques, https://arxiv.org/abs/2406.06608

Prompting LLM. * The Prompt Report: A Systematic Survey of Prompting Techniques, https://arxiv.org/abs/2406.06608

Efficient LLM. * Efficient Large Language Models: A Survey, https://arxiv.org/abs/2312.03863 and https://github.com/AIoT-MLSys-Lab/Efficient-LLMs-Survey

Personalized LLMs. * Personalization of Large Language Models: A Survey, https://arxiv.org/abs/2411.00027

Continual LLM. * Continual Learning for Large Language Models: A Survey, https://arxiv.org/abs/2402.01364

Classification. * A Survey of Text Classification With Transformers: How Wide? How Large? How Long? How Accurate? How Expensive? How Safe?, https://ieeexplore.ieee.org/document/10380590

* Graph Neural Networks for Text Classification: A Survey, https://arxiv.org/abs/2304.11534

* Recent Advances in Hierarchical Multi-label Text Classification: A Survey, https://arxiv.org/abs/2307.16265

* Text classification using embeddings: a survey, https://link.springer.com/article/10.1007/s10115-023-01856-z

* Deep learning, graph-based text representation and classification: a survey, perspectives and challenges, https://link.springer.com/article/10.1007/s10462-022-10265-7

* A Survey of Cross-Lingual Text Classification and Its Applications on Fake News Detection, https://www.worldscientific.com/doi/10.1142/S2811

Augmentation. * A Survey on Data Augmentation for Text Classification, https://dl.acm.org/doi/10.1145/3544558

* Augmented Language Models: a Survey, https://arxiv.org/abs/2302.07842

* Retrieval-Augmented Generation for Large Language Models: A Survey, https://arxiv.org/abs/2312.10997

KG Editing with LLM. * Knowledge Editing for Large Language Models: A Survey, https://arxiv.org/abs/2310.16218

Graphs for NLP. * Graph Neural Networks for Natural Language Processing, https://arxiv.org/abs/2106.06090

Interpretability. * Post-hoc Interpretability for Neural NLP: A Survey, https://arxiv.org/abs/2108.04840

Specific Domains. * Artificial Intelligence for Literature Reviews: Opportunities and Challenges, https://arxiv.org/abs/2402.08565

Instance Selection. * A Comparative Survey of Instance Selection Methods applied to Non-Neural and Transformer-Based Text Classification, https://dl.acm.org/doi/10.1145/3582000

Question Answering. * Deep learning-based question answering: a survey, https://link.springer.com/article/10.1007/s10115-022-01783-5

* Modern Question Answering Datasets and Benchmarks: A Survey, https://arxiv.org/abs/2206.15030

LLMs and code

* Large Language Models for Code Completion: A Systematic Literature Review, https://www.sciencedirect.com/science/article/pii/S0920548924000862

E.4 Information Retrieval

* A Survey on RAG Meets LLMs: Towards Retrieval-Augmented Large Language Models, https://arxiv.org/abs/2405.06211

* A Survey of Generative Search and Recommendation in the Era of Large Language Models, https://arxiv.org/abs/2404.16924

* Large Language Models for Information Retrieval: A Survey, https://arxiv.org/abs/2308.07107

* Neural Approaches to Conversational Information Retrieval, https://arxiv.org/pdf/2201.05176.pdf

Ranking. * Pretrained Transformers for Text Ranking: BERT and Beyond, https://arxiv.org/abs/2010.06467

* Utilizing BERT for Information Retrieval: Survey, Applications, Resources, and Challenges, https://dl.acm.org/doi/10.1145/3648471

* Large Language Models are Effective Text Rankers with Pairwise Ranking Prompting, https://arxiv.org/abs/2306.17563

Code Search
* Survey of Code Search Based on Deep Learning, https://arxiv.org/abs/2305.05959

E.5 Time Series Analysis

- Foundation Models for Time Series Analysis: A Tutorial and Survey, https://arxiv.org/abs/2403.14735

F Conference Planner

Table 1. Conferences on Graphs, Semantic Web, and Web. Entries are grouped based on CORE and then submission deadline (DL). Rank *u* means currently unranked.

Name	Acronym	CORE	DL
ACM International Conference on Web Search and Data Mining	WSDM	A*	
International World Wide Web Conference	WWW	A*	
IEEE International Conference on Web Services	ICWS	A	February
Extended Semantic Web Conference	ESWC	A	
International Semantic Web Conference	ISWC	A	
International Conference on Web Engineering	ICWE	B	January
International Conference on Graph Transformation	ICGT	B	February
International Conference on Web Information Systems Engineering	WISE	B	June
IEEE/WIC/ACM International Conference on Web Intelligence	WI/WI-IAT	B	May
Web and Big Data	APWEB	C	March
Latin-American Algorithms, Graphs and Optimization Symposium	LAGOS	C	March
International Conference on Internet and Web Applications and Services	ICIW	C	March
Information Integration and Web-based Applications and Services	iiWAS	C	June
International Conference on Web Information Systems and Technologies	WEBIST	C	June
Cologne-Twente Workshop on Graphs and Combinatorial Optimization	CTW	C	
EuroConference on Combinatorics, Graph Theory and Applications	EUroComb	C	(bi-anual; odd years);
Workshop on Algorithms And Models For The Web Graph	WAW	C	
Learning on Graphs Conference	LoG	u	September

Graphs, Semantic Web, Web.

G Fun

See also paper templates, but in other disciplines.

tinyurl.com/paper-template → https://drive.google.com/file/d/1IaQpS5blxHNIKEBoXh0kQPRGjAtXr6XZ/view

and

tinyurl.com/papertemplate → https://www.kidzone.ws/magic/walkthrough-t.htm

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009