

CS 197 Notes

Course link - <https://www.cs197.seas.harvard.edu/>

A Framework for Generating Research Ideas

Lecture 10+11

How to generate research ideas?

Apply framework to:

Identify gaps in a research paper like research questions, experiment setup and findings.

Generate ideas to build on a research paper, task of interest, evaluation strategy, proposed method

Iterate on your ideas to improve quality

Identifying Gaps In A Research Paper

Gaps can be in *what questions were asked, what experiments were setup* and the way the *paper is in with previous works*. Gaps point to future directions for research.

Identify gaps in research question

First, write the research question of the paper. Then write research hypothesis supporting that central research question. Research hypothesis is a precise testable statement of what the authors predict will be the outcome of the study. May have to infer hypothesis from experiments performed. Then, look at gaps between the research question and hypothesis, what hypothesis have not been tested?

Example Answer:

Research Question:

How well can an algorithm detect diseases without explicit annotation?

Research Hypothesis:

1. A self-supervised model trained on chest X-ray reports (CheXzero) can perform pathology-classification tasks with accuracies comparable to those of radiologists.
2. CheXzero can outperform fully supervised models on pathology detection.
3. CheXzero can outperform previous self-supervised approaches (MoCo-CXR, MedAug, and ConVIRT) on disease classification.

Gaps:

1. Can CheXZero detect diseases that have never been implicitly seen in reports?
2. Can CheXZero maintain high-level of performance even when using a small corpus of image-text reports.

Identify gaps in experiment setup

Pay attention to gaps. Any lack of experiments the way the method was evaluated? In the comparisons, implementation? Does the experimental setup test the research hypothesis convincingly? We are interested in the setup of experiments, not results.

Example Answer:

Research Hypothesis (with Experimental Setups):

1. *A self-supervised model trained on chest X-ray reports (CheXzero) can perform pathology-classification tasks with accuracies comparable to those of radiologists.*
 - a. Evaluated on a test set of 500 studies from a single institution with a reference standard set by a majority vote - similar to what was used by previous studies. Comparison is performed to the average of 3 board-certified radiologists on the F1 and MCC metrics on 5 diseases. Gaps:
2. *CheXzero can outperform fully supervised models on pathology detection.*
 - a. Evaluated on the AUC metric on the average of 5 pathologies on the CheXpert test set (500 studies). Methods evaluated include a baseline supervised DenseNet121 model along with the DAM method with the reasoning "The DAM supervised method is included as a comparison and currently is state-of-the-art on the CheXpert dataset. An additional supervised baseline, DenseNet121, trained on the CheXpert dataset is included as a comparison since DenseNet121 is commonly used in self-supervised approaches."
3. *CheXzero can outperform previous self-supervised approaches (MoCo-CXR, MedAug, and ConVIRT) on disease classification.*
 - a. Setup as above.

Gaps:

1. On hypothesis 1, The number of radiologists is maybe too small to decisively argue for being absolutely comparable to radiologists. Maybe the experience/training of the radiologists needs to be understood to qualify more precisely what constitutes radiologist level performance.
2. On hypotheses 2/3, The number of pathologies evaluated for were limited by the number of samples in the test set. A larger set of pathologies evaluated would support the hypotheses more.
3. On hypothesis 3, the number of self-supervised approaches compared to are limited – the choice of label-efficient approaches, ConVIRT, MedAug and MoCo-CXR. There are more self-supervised learning algorithms which can be compared to.
4. On hypothesis 3, unclear also whether the comparisons are single models or ensemble models, or whether they use the same training source.

Identify gaps in limitation, implicit and explicit

Now that we have identified gaps in the experimental setup, we make our way to the results and discussion. Here, we're on the lookout for expressed limitations of the work. Part of this work is easy: sometimes, there's an explicit limitation section that we can directly use, or we can infer it from statements of future work. However, sometimes the limitations of a method are expressed in the results themselves: where the methods fail

Example Answer:

Gaps:

Explicitly Listed:

1. "the self-supervised method still requires repeatedly querying performance on a labelled validation set for hyperparameter selection and to determine condition-specific probability thresholds when calculating MCC and F1 statistics.
2. "the self-supervised method is currently limited to classifying image data; however, medical datasets often combine different imaging modalities, can incorporate non-imaging data from electronic health records or other sources, or can be a time series. For instance,

magnetic resonance imaging and computed tomography produce three-dimensional data that have been used to train other machine-learning pipelines.

3. "On the same note, it would be of interest to apply the method to other tasks in which medical data are paired with some form of unstructured text. For instance, the self-supervised method could leverage the availability of pathology reports that describe diagnoses such as cancer present in histopathology scans"
4. "Lastly, future work should develop approaches to scale this method to larger image sizes to better classify smaller pathologies."

Implicit through results:

1. The model's MCC performance is lower than radiologists on atelectasis and pleural effusion.
2. The model's AUC performance on Padchest is < 0.700 on 19 findings out of 57 radiographic findings where $n > 50$.
3. The CheXzero method severely underperforms on detection of "No Finding" on Padchest, with an AUC of 0.755.

Generating Ideas For Building on a Research Paper

1. Change task of interest

Apply main idea to a different modality?

Apply idea to different data type?

Apply idea to a different task? (Object detection or semantic segmentation, VQA?)

Change outcome of interest?

- Robustness properties of the model
- Data efficiency compared to other learning paradigms (fully supervised)

2. Change the evaluation strategy?

Evaluate on a different dataset?

Evaluate on a different metric? Give more insights.

Explain why it works/breaks?

Make different comparisons?

3. Change the proposed method?

Change training dataset?

Change pre-training/training strategy? (New checkpoints other than ImageNet, like LION-5B)

Change deep learning architecture?

Change problem formulation? (Multiple views as input)

Iterate on your idea

1. Search whether your idea has been tried?

Construct titles for you paper and search.

Example: if I am interested in the application of a CheXzero-like approach to other kinds of data, I might search for:

- contrastive learning histopathology text (no relevant results)
- contrastive learning histopathology genomic alteration (returned a [match](#))

2. Read Important Related Works and Follow Up Works

See google scholar cited by, to find related works of the paper you are interested/building on. Arxiv, papers with code

3. Get feedback


Once draft is ready, send it to authors your building on, ask what they thing.

Tips to Manage Your Time and Efforts

Lecture 18

2. Making the most of supervisor meetings

- How:
 - For each meeting, prepare a couple of PowerPoint slides:
 - 1 slide with date of meeting
 - 1 slide with talking points for the meeting (meeting agenda)
 - 3-5 slides with main progress (e.g. result tables, findings, summary of literature, problems...)
 - At the end of the meeting:
 - 1 slide with action points for next week
- Why:
 - Basis for productive discussion
 - Keeps you on track
 - Reminds your supervisor of the current status of your project
 - Documentation of your progress & valuable future reference

Think of the project tracker as your journal or diary or a log. It can remind you of the progress that you're making, and the direction you're headed. Every entry in the project tracker is marked by the date. In addition, Bullet points are organized and in sufficient detail such that a search that a future version of you project team member will be able to understand. The project tracker often links to important reference documents including a related work document, a working manuscript document, and an experimental results report (such as a google sheet or a Weights and Biases table. 

Project Tracker in Google Docs Template:

Week of November 6-12

- Meeting Objective Description
- Progress Updates (per person)
 - Code
 - PRs
 - Text Descriptions
- Current Issues/Blockers
 - What has been done to address
 - What still needs to be attended to
- Next Steps (with assignments per person)

Making Progress and Impact in AI Research

Lecture 19

Mentor/Advisor Relationship

Your mentor/advisor is an input-output machine

Show your work

- what did you do

- why do it
- How you did it
- What did you find?
- Why do you think it does or does not make sense

Present failures

Provide context

Set expectations

Research skills

Imagine success

Work backward

Toy examples

Simple things first

One thing at a time "You don't learn anything by changing multiple things at the same time."

Create more impact

Put work out there

- Pick a good name
- Make results available
- Lower barrier for others to follow
- Make others' life easier
- Show your work

Tips for Creating High-Quality Slides

Lecture 19

Assertion-evidence approach

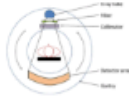
1. Build talk on messages, not topics

2. Support messages with visual evidence, not long text,
3. Explain evidence by fashioning words on the spot

Script: "A CT (computed tomography) scanner is a specialized medical imaging device that uses x-ray technology to create detailed 3D images of the inside of the body. The CT machine consists of a large, donut-shaped x-ray machine that rotates around the body while the patient lies on a table that slides into the center of the machine. As the machine rotates, it emits a beam of x-rays through the body, which is then detected by detectors on the other side of the body. The data collected by the detectors is then used to create a 3D image of the inside of the body."

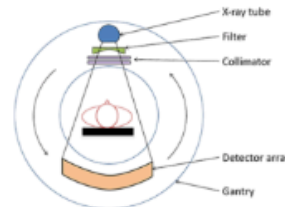
How does a CT Scanner Work

- CT scanner is a specialized medical imaging device
- Uses x-ray technology to create detailed 3D images of the inside of the body
- CT machine consists of a large, donut-shaped x-ray machine that rotates around the body
- Patient lies on a table that slides into the center of the machine
- X-ray beam is emitted through the body and detected by detectors on the other side
- Data collected by detectors is used to create a 3D image of the inside of the body



Bad ✗

A CT scanner is a medical device that uses x-ray technology to create 3D images of the inside of the body.



Good ✓

Script: "First, the CheXzero training pipeline involves using raw radiology reports as a natural source of supervision, allowing the model to learn features from them. Then, to predict pathologies in a chest X-ray image, we generated a positive and negative prompt for each pathology, such as 'consolidation' versus 'no consolidation'. By comparing the model output for the positive and negative prompts, the self-supervised method can compute a probability score for the pathology, which can then be used to classify its presence in the chest X-ray image."

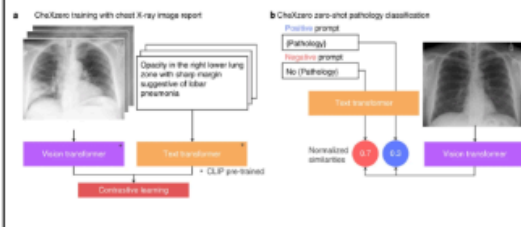
How does CheXzero training and inference work?

- Training pipeline uses raw radiology reports as a source of supervision to allow model to learn features
- To predict pathologies in chest X-ray images, positive and negative prompts are generated for each pathology
- Self-supervised method compares model output for positive and negative prompts to compute probability score for pathology
- Probability score can be used to classify presence of pathology in chest X-ray image



Bad ✗

CheXzero learns features from raw radiology reports to then predict pathologies in chest X-ray images using prompts.



Good ✓

Click to edit theme title style

- First level
 - Second level
 - Third level
 - Fourth level
 - Fifth level
 - Sixth level
 - Seventh level
 - Eighth level
 - Ninth level

Bad ❌

Write a sentence headline that states the main assertion of the slide

In the body of the slide, support the headline assertion with visual evidence, using text only where necessary.

Good ✔

Text telegrams

Strategies for Ensuring Effective Communication

- Use language that is easy to understand and avoids unnecessary jargon or technical terms
- Provide enough context so that the message can be understood without any additional information
- Consider the tone and level of formality that is appropriate for the audience and the situation
- Encourage two-way communication by actively listening to others and asking questions to clarify any misunderstandings
- Be aware of nonverbal cues and body language, as they can also convey important information and influence the way a message is received.

Bad ❌

Effective Communication strategies

Clear, simple language, no jargon

Adequate context

Appropriate tone, formality

Two-way comm., active listening, questions

Nonverbal cues, body language

Good ✔

Introduce Texts Sequentially

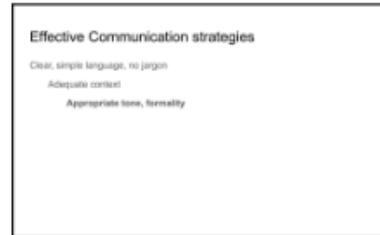
Second, introducing text sequentially rather than all at once. This makes reading or paraphrasing the slide to the audience more effective as the audience knows what point you are referring to. Comment fully on each line of text before moving on to the next one.



Build-up 1



Build-up 2



Build-up 3

Other tips

Use serif and sans-serif fonts together.

Serif fonts are typically used for body text because they are easier to read in large blocks of text due to the extra details on the letters. Sans-serif fonts are often used for headings and other large blocks of text because they are simpler and more modern-looking.

White background ,neutral and professional-looking

Using two colors for slide text can be effective because it can help to create a visual hierarchy and emphasize important points. For example, you could use a primary color for the main points and black for secondary points or supporting details. This can make it easier for the audience to follow the content and understand the structure of the presentation.

Use your slide footer for citations and page numbers

Including citations in the slide footer can be a helpful way to provide credit to the sources of your information and show that your presentation is based on credible research. To do this, you can use the slide footer feature in your presentation software to add a text box or field where you can include the citations for any sources you have used. This could include the author, title, and publication information for articles, books, or other resources.

Including the page number in the slide footer can also be useful for the audience, as it allows them to easily mark the page and ask questions about specific information during the presentation. This can be especially helpful if the audience wants to refer back to a specific slide or source later. To include the page number in the slide footer, you can use the built-in page numbering feature in your presentation software. This will automatically update the page number as you add or remove slides, so you don't have to manually update it yourself.

Overall, using the slide footer to include citations and the page number can help to make your presentation more professional and organized, and can also make it easier for the audience to follow along and ask questions.

Have an outline slide

The image shows two side-by-side slide layouts under the heading "Learning outcomes". The left slide, labeled "Bad" with a red X, contains a numbered list of four items. The right slide, labeled "Better" with a green checkmark, contains four separate boxes, each with a bolded key phrase. The boxes are: "Identify Essential Components", "Train Novel Models", "Focus on Experiment Reproducibility and Automation", and "Emphasize Modularity for Future Work".

Learning outcomes

1. Identify the essential components that go into building a novel model.

2. Train novel models, including:

- a. Manage experiment configurations using Hydra.
- b. Log model checkpoints.
- c. Manage and analyze experiment results using Google spreadsheet pivot tables.
- d. Log metric values and images using Weights & Biases.

3. Develop a good understanding of the value of code quality, experiment reproducibility, and experiment automation.

4. Develop a preliminary understanding of other best practices to explore further.

Learning outcomes

Identify Essential Components

Train Novel Models

Focus on Experiment Reproducibility and Automation

Emphasize Modularity for Future Work

Bad ❌

Better ✅

Reasoning:

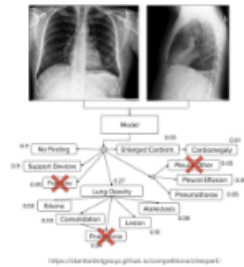
- Eliminated unnecessary points that can be iterated on by the speaker
- Created object representations of each concept with the key word bolded with the main idea

The dataset

CheXpert: 224,316 chest X-rays for 65,340 patients. Validation set consists of 234 chest X-rays from 200 patients. The test set consists of 666 chest X-rays from 500 patients.

Labeled for the presence of 14 observations (13 pathologies and an observation of "No Finding").

Our work concerns 10 of the 14 observations. "Fracture" and "Pneumothorax" not included because of low prevalence in test set, "Pneumonia" not included because it is a clinical (as opposed to a radiological) diagnosis; "No Finding" not included because not applicable to evaluating localization performance.



<https://stanfordmlgroup.github.io/competitions/chexpert/>

Bad ✗

CheXpert contains chest X-rays with pathology labels

CheXpert	# of chest X-Rays	# of patients
Training	~200k	~50k
Validation	~200	~200
Test	~500	~500



Our work concerns 10 of the 14 observations (13 pathologies + 1 observation "No Finding").

"Fracture" and "Pneumothorax" have low prevalence in test set
 "Pneumonia" is a clinical diagnosis
 "No Finding" is applicable to evaluating localization performance.



<https://stanfordmlgroup.github.io/competitions/chexpert/>

Better ✓

Reasoning:

- More descriptive title
- Less text
- Essential components extracted from figure with color
- Kept only orders of magnitude from the table

