Special Project 09/04 HW2

What can we do with Self-Supervised Learning?

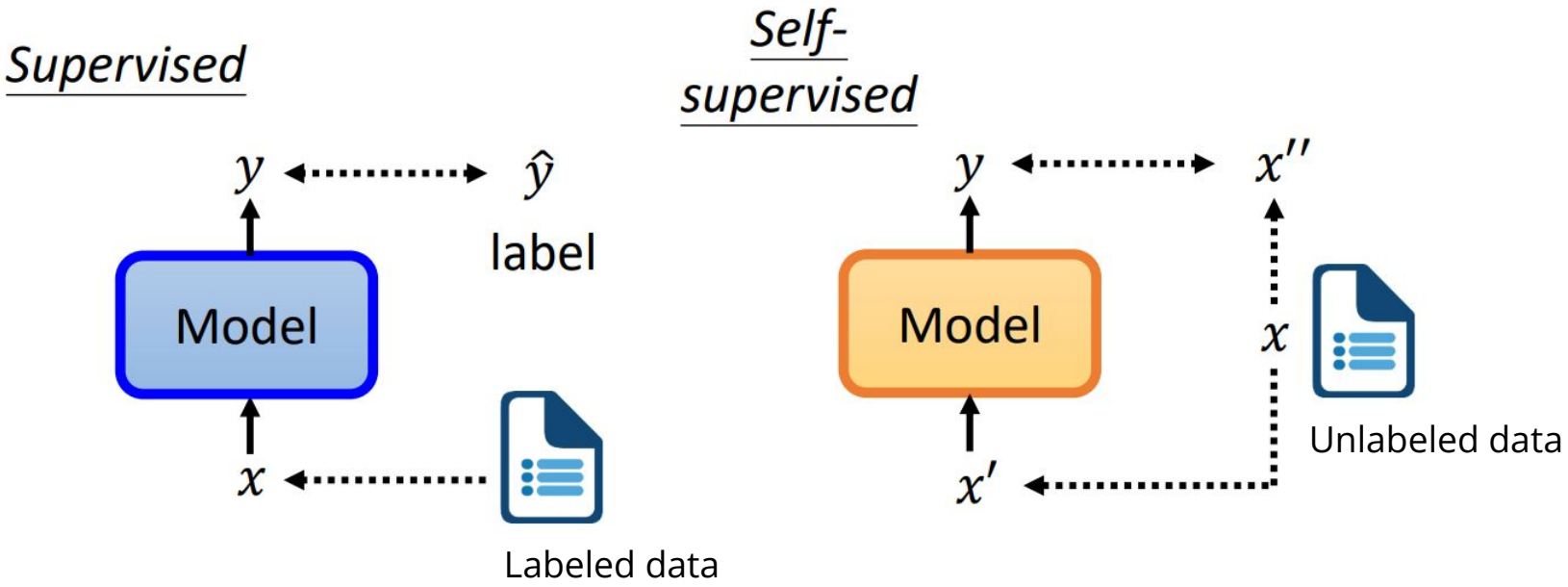
Evan Yi-Cheng Lin
r12942075@ntu.edu.tw

For **S**elf-**S**upervised **S**peech
Pre-training and **R**epresentation **L**earning

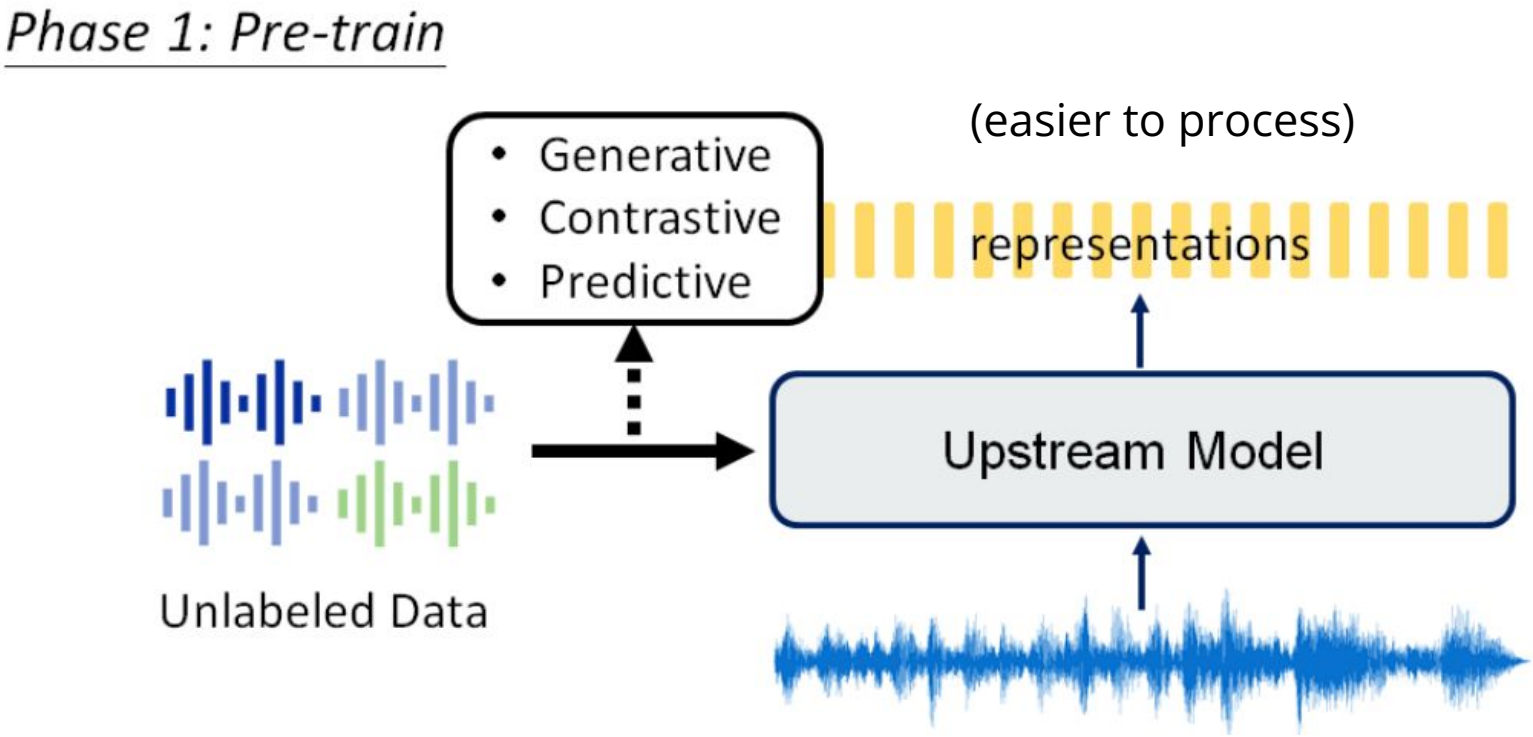
S3PRL
SPEECH TOOLKIT

S3PRL Repo
S3PRL Doc
Superb Doc
Tutorial link (optional)

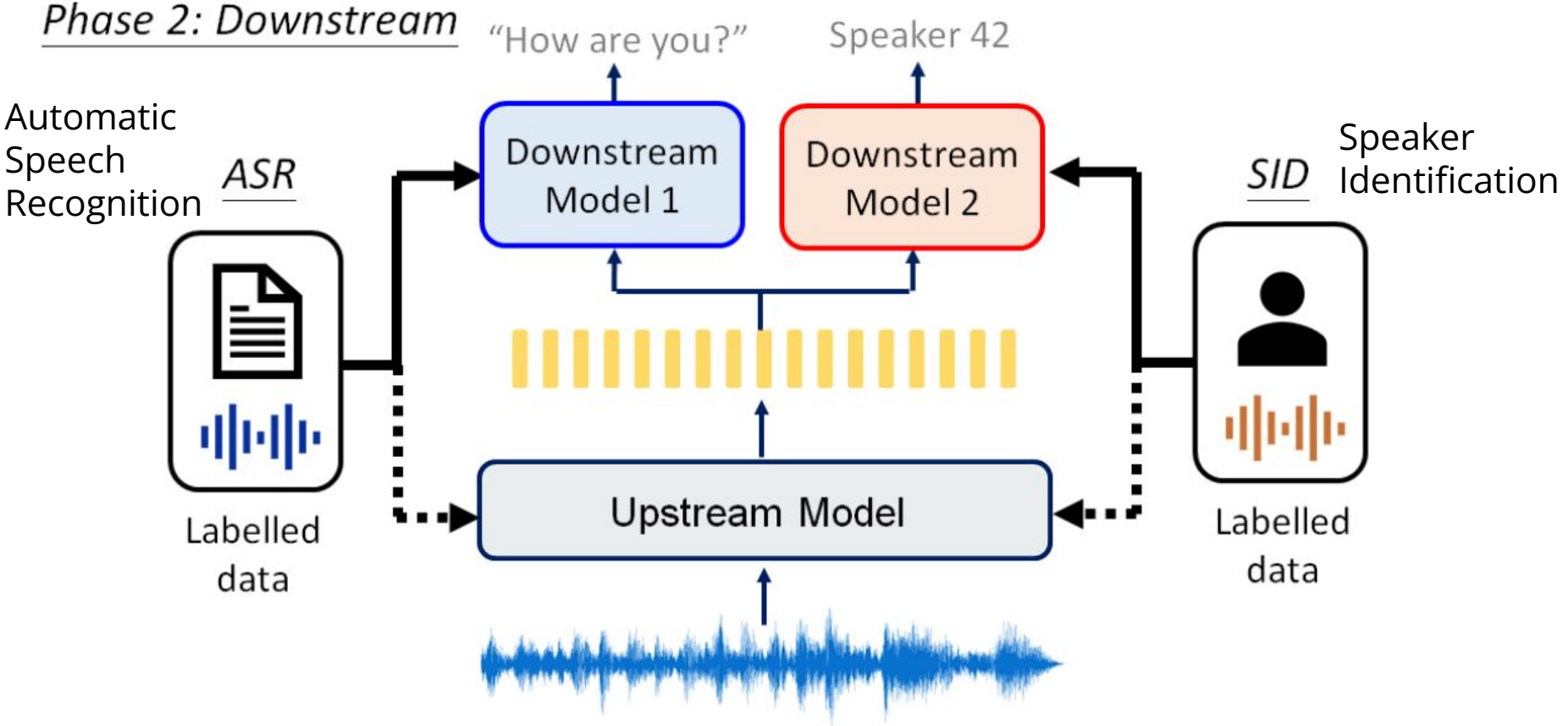
Supervised v.s. Self-supervised Learning



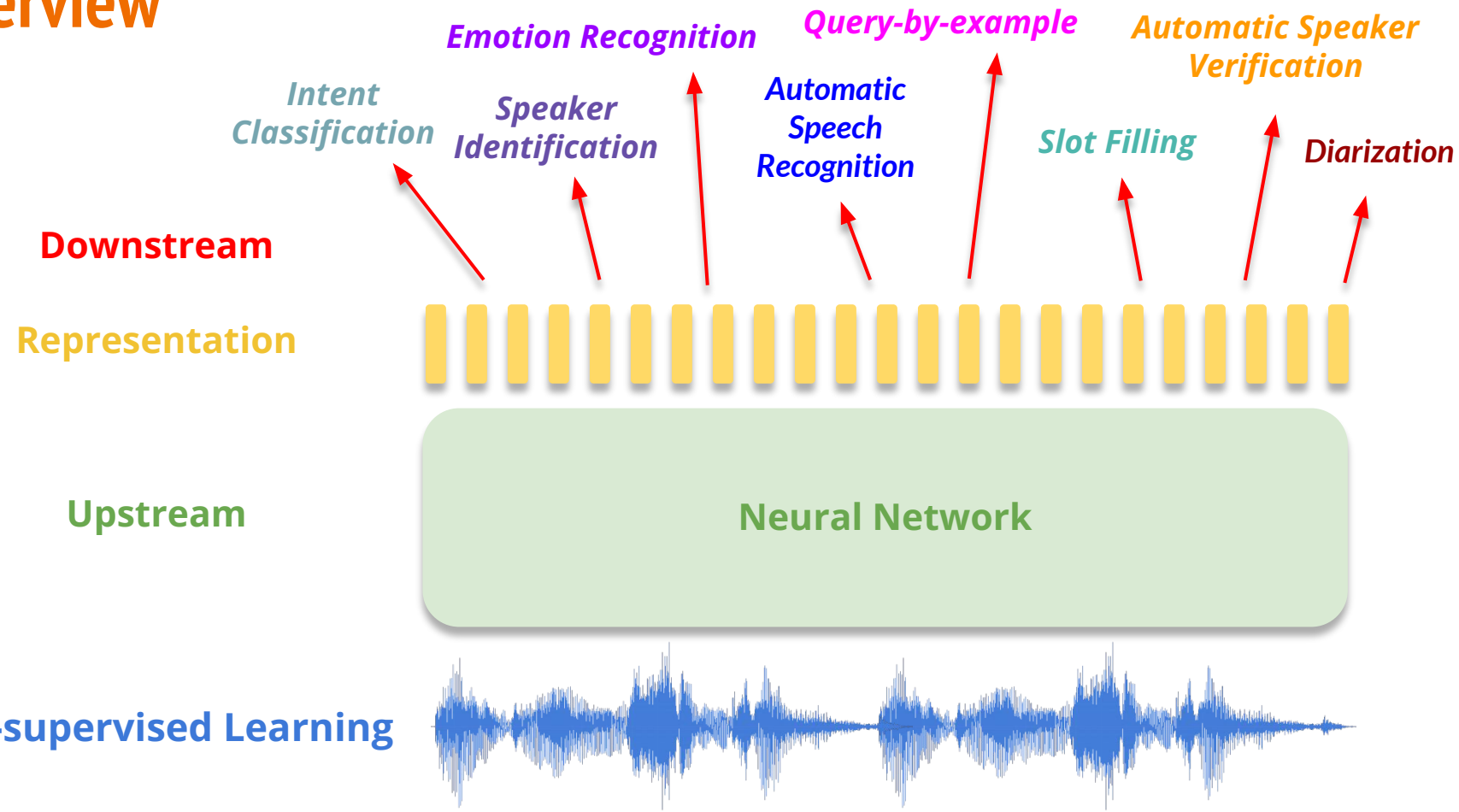
Overview


Phase 1: Pre-train
(easier to process)
• Generative
• Contrastive
• Predictive
representations
Upstream Model
Unlabeled Data

Overview


Phase 2: Downstream
"How are you?"
Speaker 42
Automatic Speech Recognition
ASR
Downstream Model 1
Downstream Model 2
SID
Speaker Identification
Labelled data
Upstream Model
Labelled data

Overview

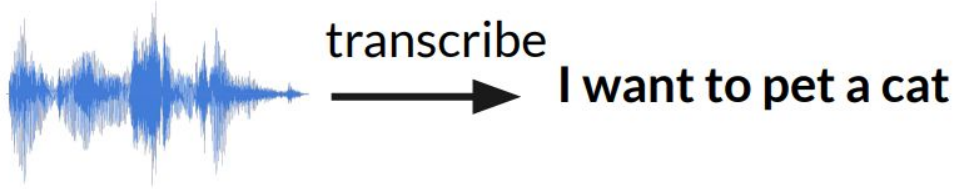

Intent Classification
Speaker Identification
Emotion Recognition
Automatic Speech Recognition
Query-by-example
Slot Filling
Automatic Speaker Verification
Diarization
Downstream
Representation
Upstream
Neural Network
Self-supervised Learning

Overview of downstream tasks (Content related)

Phoneme Recognition (PR)

Keyword spotting (KS)

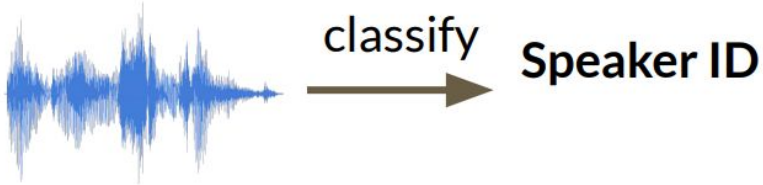
Speech Recognition (SR)



Query by Example (QbE)

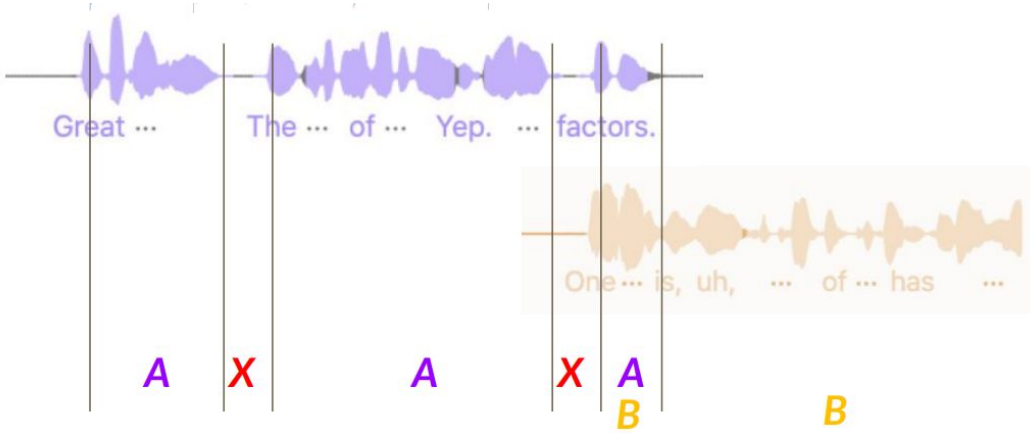
Overview of downstream tasks (Speaker related)

Speaker Identification (SID)



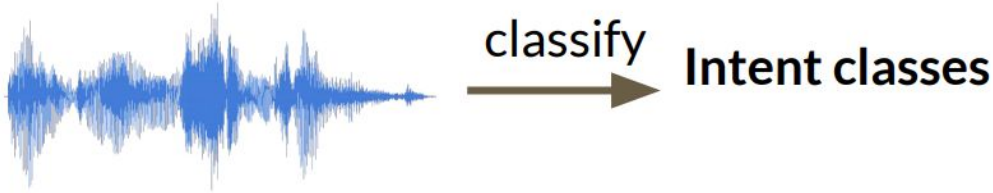
Speaker Verification (SV)

Speaker Diarization (SD)

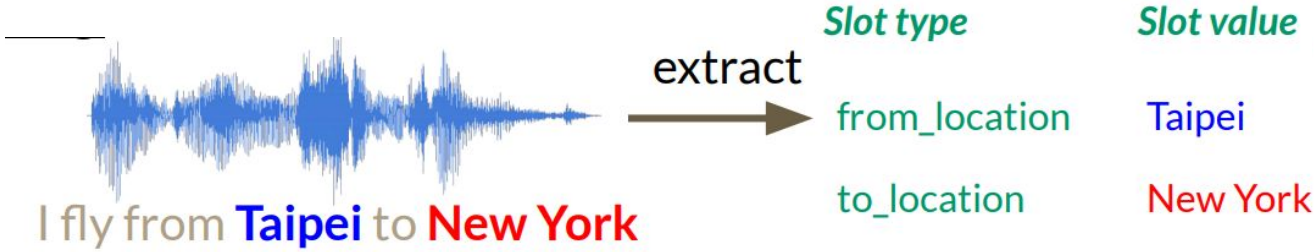


Overview of downstream tasks (Semantic)

Intent Classification (IC)



Slot Filling (SF)



Overview of downstream tasks (Emotion)

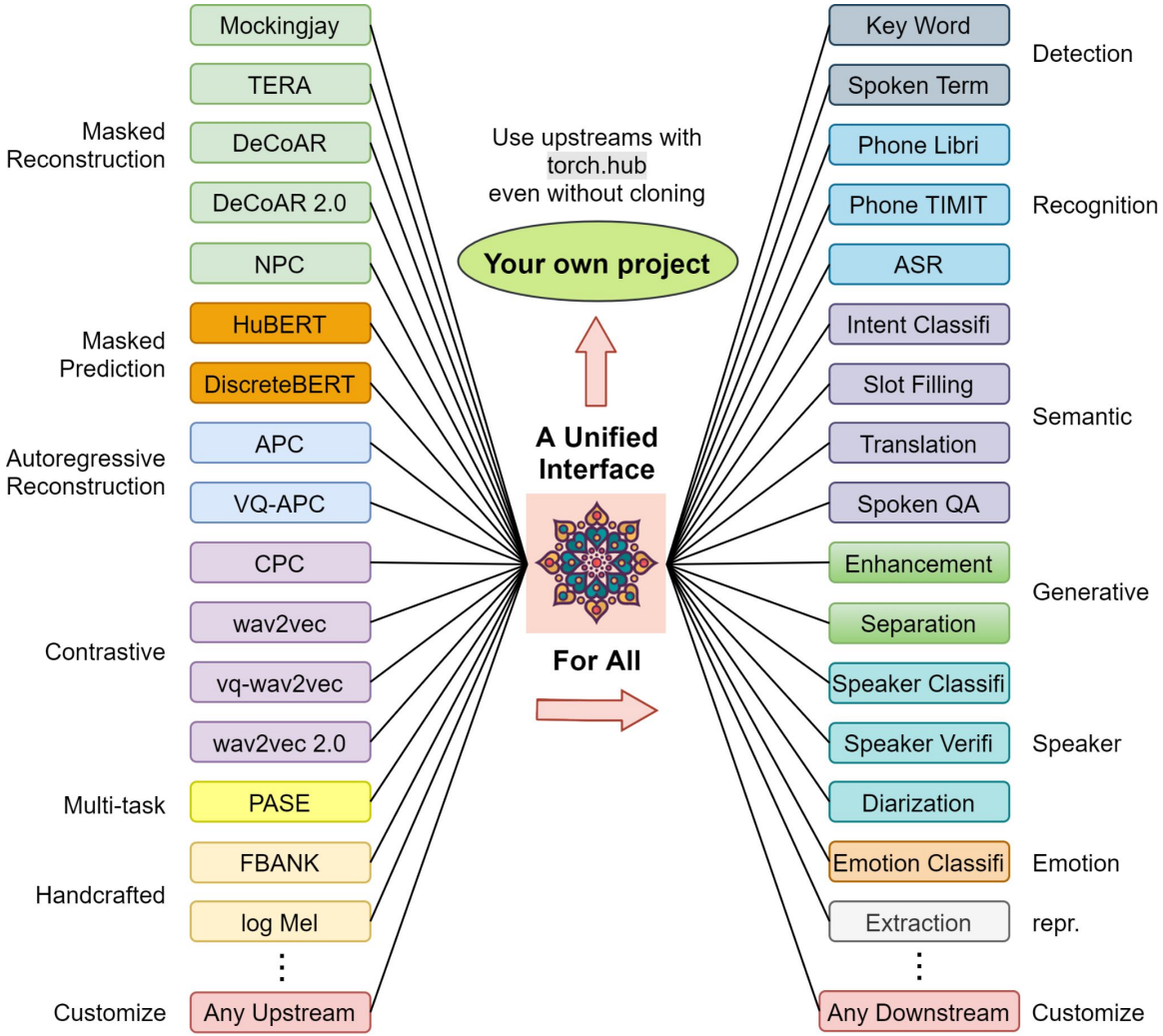
Emotion Recognition (ER)

classify

Happy / Angry / Sad .....

S3PRL Features

- Upstream Pre-training
- Upstream Hub
- Downstream fine-tuning
- SUPERB Challenge



How to pass this exercise?

Goal: Each team must earn >= 3 points in total

- **Track 1: 1 point**
- **Track 2-1: 2 points**
- **Track 2-2: 3 points**
- **Track 3-1: 2 points**
- **Track 3-2 : 3 points**
- **Track 4: 5 points (special)**

Easy

Hard

Examples:

- “2 points” + “1 points”
- “2 points” x2
- “1 point” x3

Friendly warning: Some are easy some are hard, first come first serve!
Also keep in mind that **you’ll have to present what you did when its due**.

Track 1

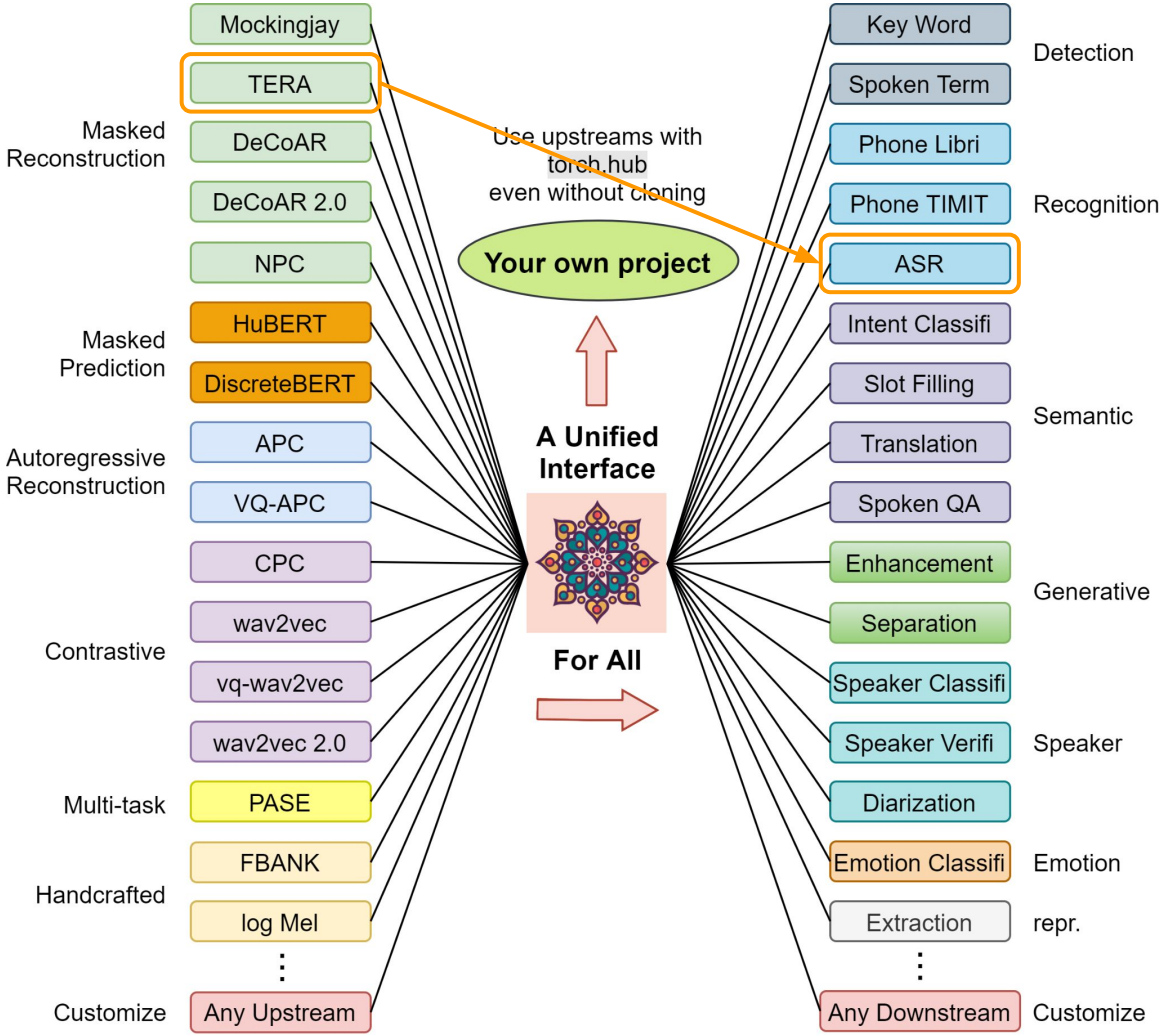
1. Run an **“upstream and downstream” pair** of your choice **(1 point)**
2. Read the description of the downstream task and prepare data (link below).
3. Run the experiemnt and report your results / observations.
4. An unique *“upstream and downstream” pair* counts for 1 point,
   you can do different pairs to get more points.
5. Setting different hyperparameters for an existing pair also counts for 1 point.

Documentation: link

SUPERB Challenge: link

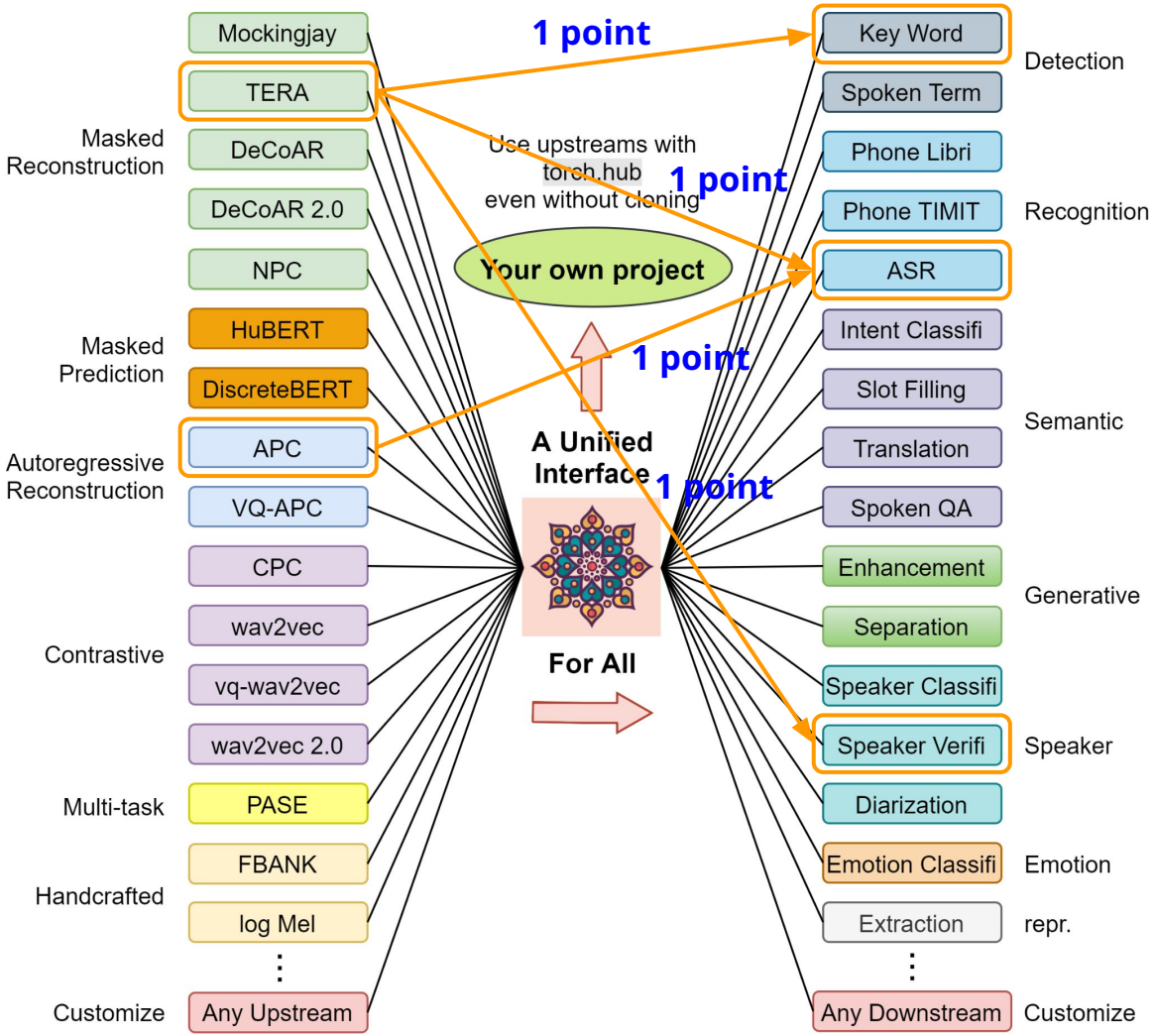
Track 1 (example)

- Choose one from the left.
- Choose one from the right.



Track 1 (example)

- Choose one from the left.
- Choose one from the right.
- Each unique pair counts **1 point**.

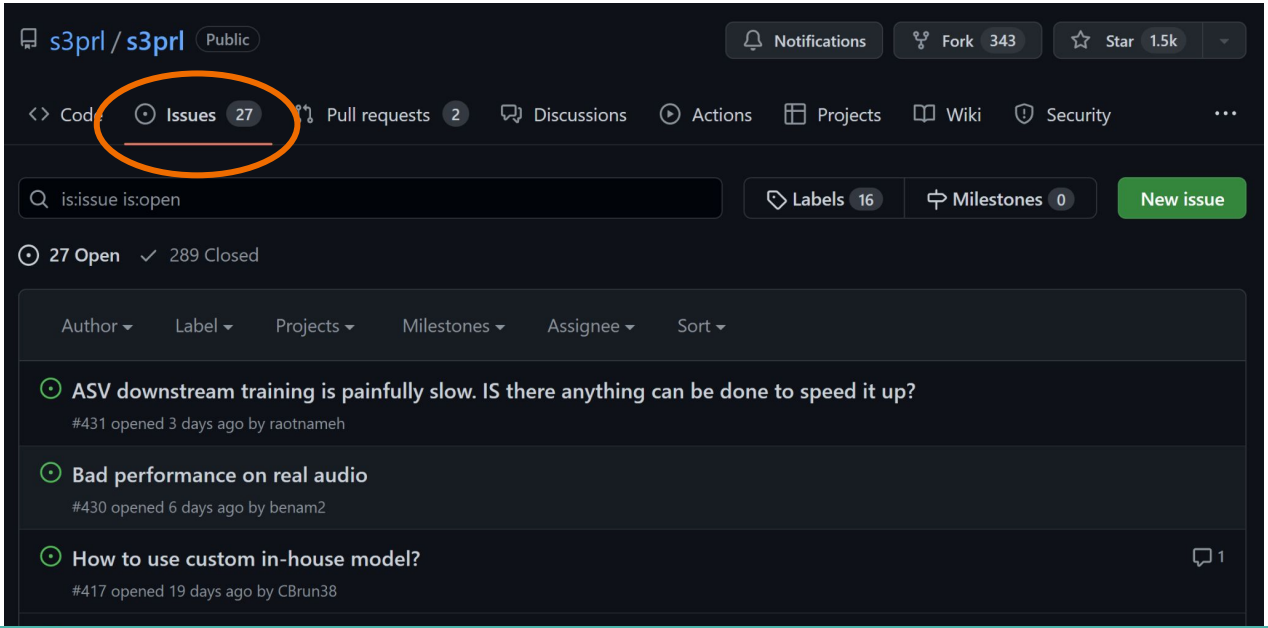


Friendly warning:
Some are easy some are hard, first come first serve!

Track 2-1

Resolve an open issue without making a pull request (2 points)

Comment in the issue and **tag me @dlion168**.



Friendly warning:
Some are easy some are hard, first come first serve!

Track 2-2

Resolve an open issue with a pull request (3 points)

Start a pull request.
See here for detailed steps:

Development pattern for contributors

1. Create a personal fork of the main S3PRL repository in GitHub.
2. Make your changes in a named branch different from `master`, e.g. you create a branch `new-awesome-feature`.
3. Contact us if you have any questions during development.
4. Generate a pull request through the Web interface of GitHub.
5. Please verify that your code is free of basic mistakes, we appreciate any contribution!

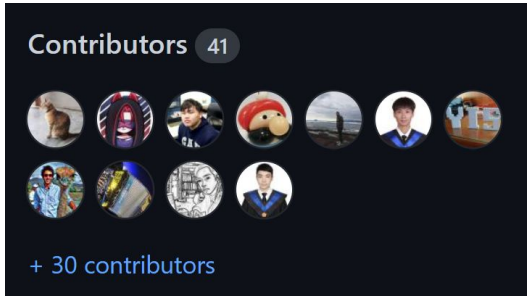
Friendly warning:
Some are easy some are hard, first come first serve!

Track 2-2

Resolve an open issue with a pull request (3 points)

Start a pull request,.
See here for detailed steps.
Comment in the issue and **tag me @dlion168**.

Note: if the issue is about adding a new feature rather than fixing bug, please discuss with me first! The reason is the S3PRL dev team might not want to include that feature!



Bonus: This will make you an official contributor to the project!

Track 3-1

Open a new issue without making a pull request (2 points)

Comment in the issue and **tag me @dlion168**.

Has to be a valid issue / problem / bug that applies to all users,
asking random questions or your own enviroment issues do NOT count.

Track 3-2

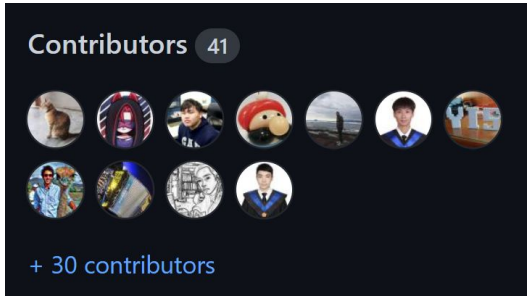
Open a new issue and resolve with pull request (3 points)

Comment in the issue and **tag me @dlion168**.

Has to be a valid issue / problem / bug that applies to all users,
asking random questions or your own enviroment issues do NOT count.

PS. s3prl currently do not accept new features.

Bonus: This will make you a official contributor to the project!



Track 4

Use this toolkit on your own application! (5 points)

- New dataset
- New downstream task
- Compare and analyze different upstreams (in any interesting and not obvious aspect)
- Design an experiment with an upstream model
- Anything reasonable or relatable to S3PRL or SUPERB
- Anything related to SSL (don't have to use this toolkit)

See this doc for customization usage: link

Bonus for choosing this track:
You **DO NOT need to file any PR, and DO NOT have to be fully finished by the deadline**, you can present a reasonable amount of progress that you've made.

Track 4 - Example: Anything related to SSL

https://youtu.be/sWz4e-DM4JU

Submission

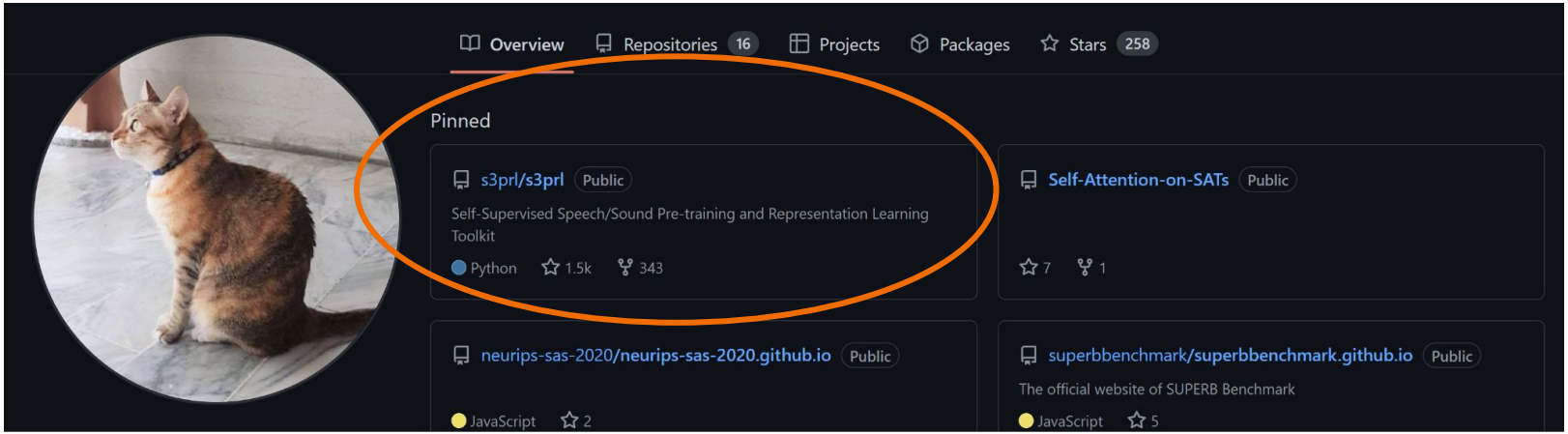
1. **For tracks 2~4, comment on FB post**
   a. a single sentence is sufficient (e.g. paste the open issue link)
   b. briefly describing what you plan to do (especially track 5)
   (This is to avoid overlap between groups)

2. **Presentation: on 9/11, and 9/18**
   Things you might want to cover, **not limited to or mandatory of**:
   a. The track(s) you choose
   b. The setting of your track
   c. Screen shots (proofs) of your training / inference results
   d. Screen shots (proofs) of your contribution / pull request / issue / comments
      i. What was it about & how did you solve the issue
   e. Your observations
   f. Your user feedback of the toolkit

3. **Do not have to submit any code (unless your make PRs)**
4. **No restriction on presentation format (slides in general)**

Recommandations and Goals

1. Team up on the harder tracks
2. Personally recommend **Track 4**, you can continue on this research direction in the next phase
3. If you became a contributor, then you can happily **pin this project on your Github homepage**:



What to do if you encounter any problem

1. **First** check if there are any related **existing** (open / closed) issues first.

2. Ask your question:
   a. Project related: Comment on the *FB post*
   b. Technical related:
      i. Open a new issue on the *S3PRL Github Repo*
      ii. Comment the issue link on the *FB post*
      iii. NO points.

3. Email / DM me
   a. If it is something personal, or you don't want to discuss in public.

Questions?

mailto:r12942075@ntu.edu.tw