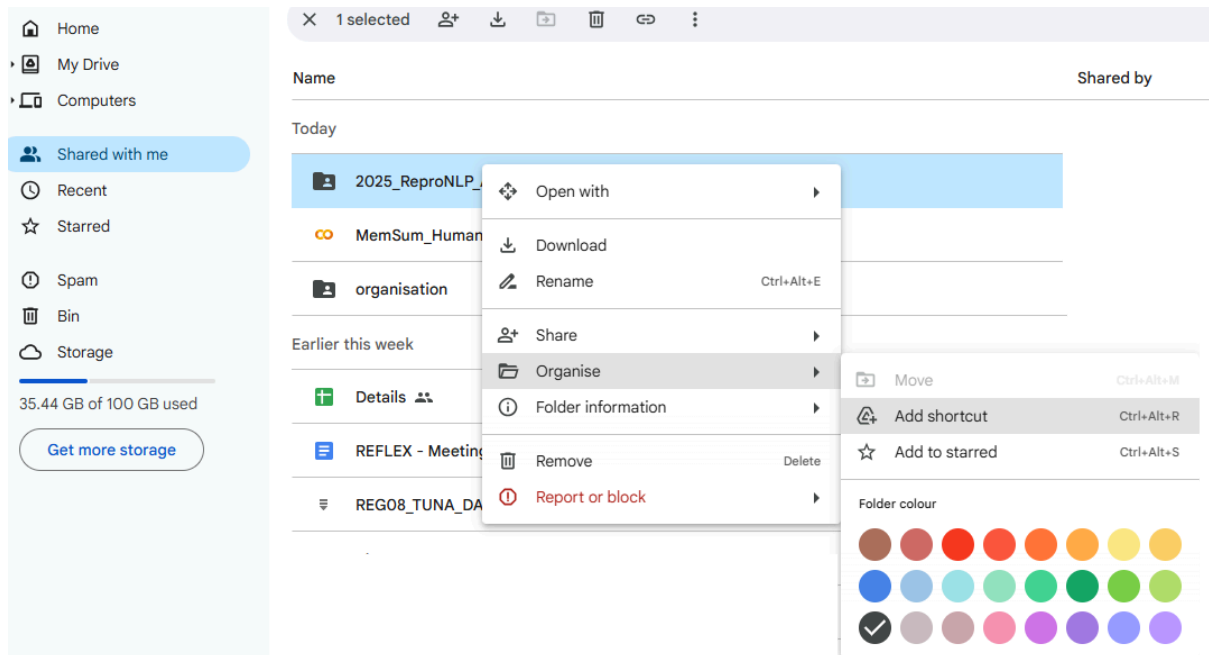
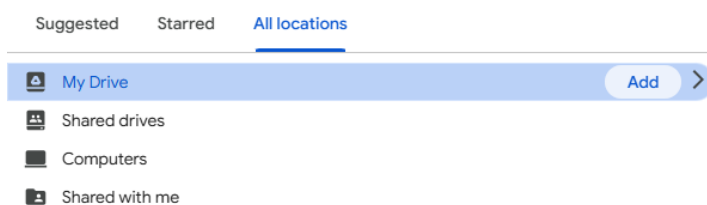


1. This is an important step as it will allow (among other things) to save automatically your annotations as they are done. Once opened, **create a shortcut to your main drive**, by right-clicking on the shared folder and then going to “Organise” > “Add shortcut”:



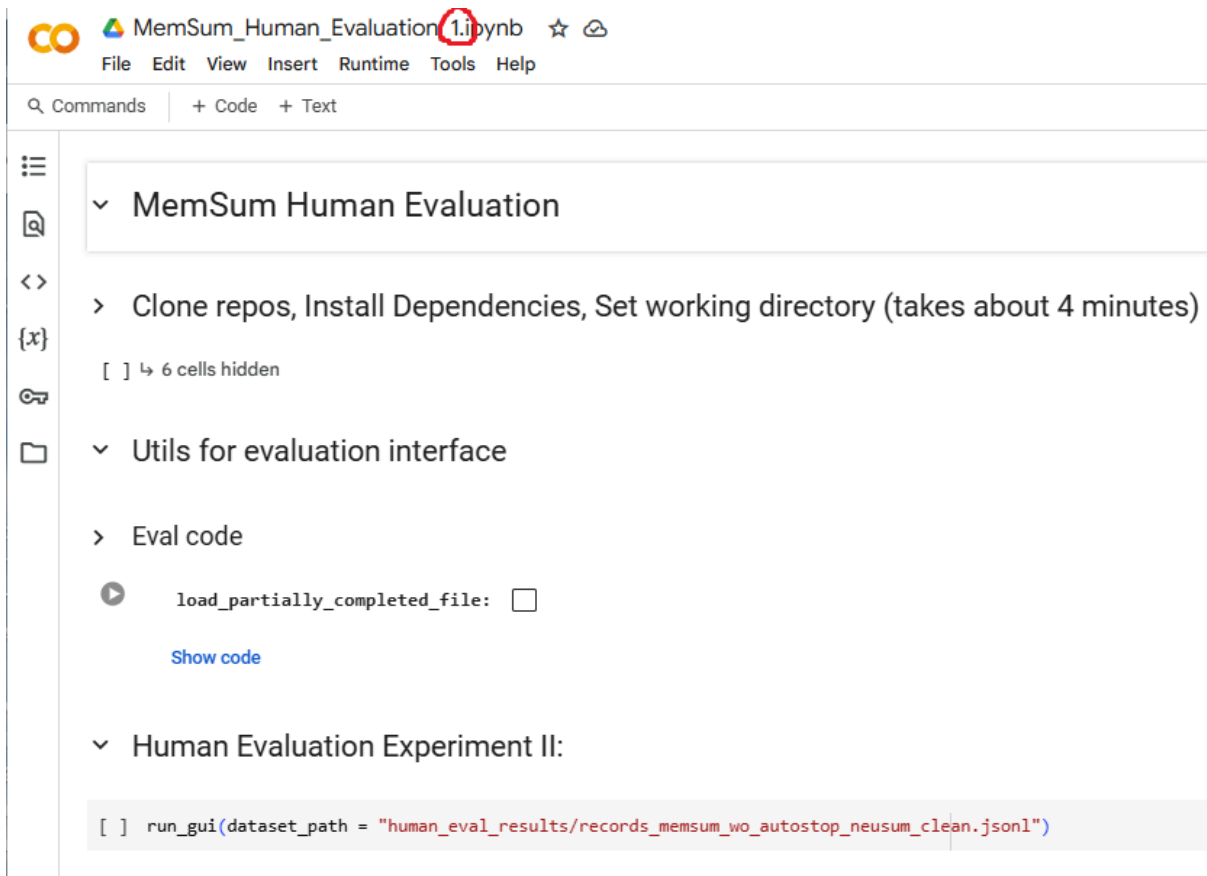
Then click on All location, and select “My Drive”, then click on “Add” on the right:

Add shortcut to 2025\_ReproNLP\_Annotators\_DCU

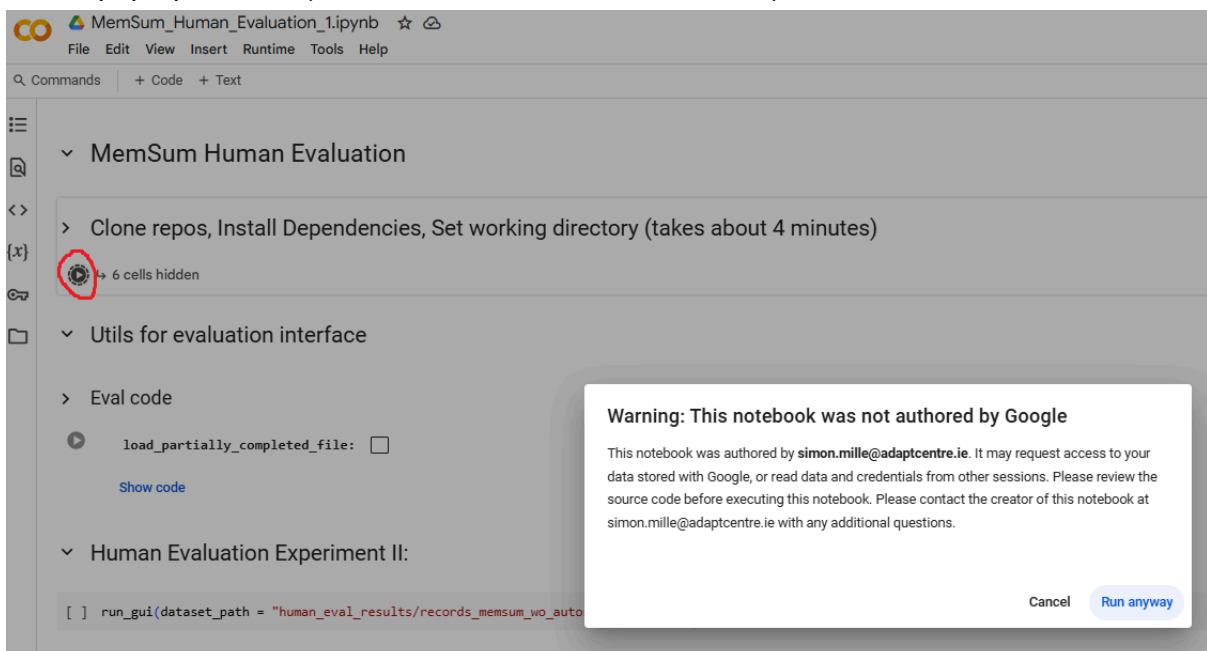


There should now be a shortcut in your Drive! You can of course remove this shortcut as soon as you are done with the evaluation.

2. Now go and click the link to the Notebook that is assigned to your email address in the [Annotators notebook assignments](#) spreadsheet of the shared drive. Make sure the ID at the end of your notebook name (**red circle** below) matches the ID in the spreadsheet:

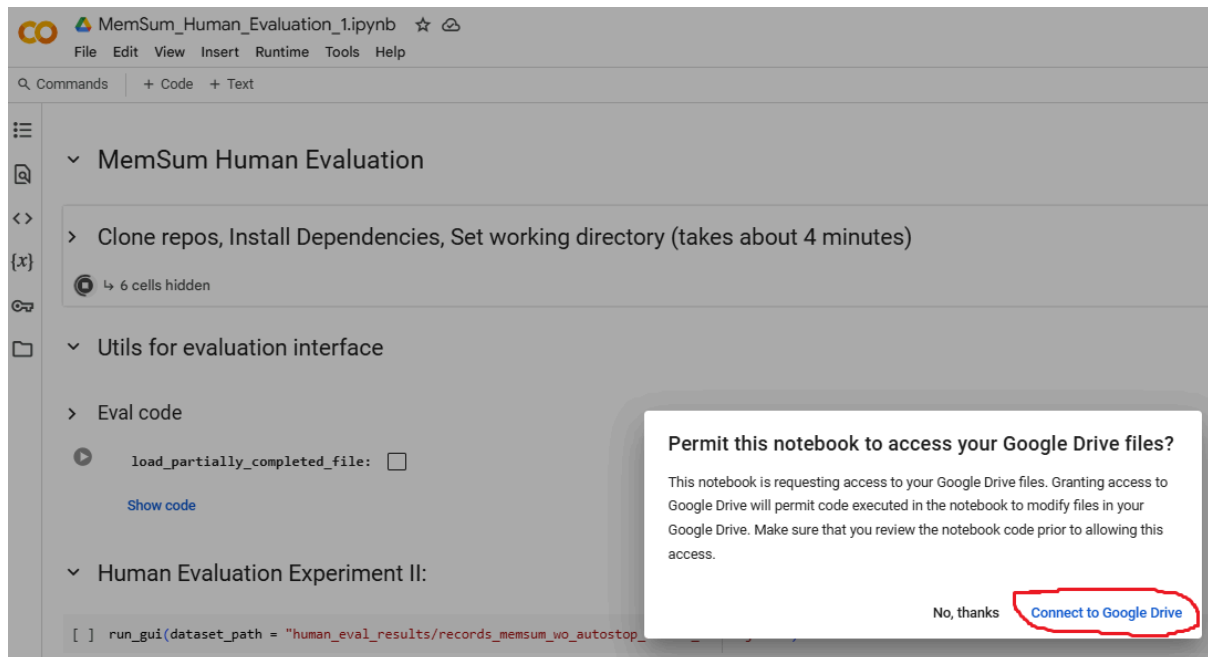


3. Click on the first “Play” button (**red circle** below), and then on “Run anyway” in the pop-up window (this runs the 6 hidden cells in a row):



4. You will then be prompted to allow access to Google Drive, click on “Connect to Google Drive” in the pop-up window (**red circle** below), and then choose the email account you provided to the team that carries out the experiment. At least allow Drive to modify files/folders, and proceed with logging in, until the pop-up window disappears. Once you are done, it will take about 4 minutes for the resources to be

loaded, during which the “Play” button becomes a “Stop” button. When the button is back to being “Play”, you can move to the next step.



5. Now click on the second “Play” button, in the “Eval code” cell (**red circle** below). This takes only a few seconds. The first time you use this interface, leave the “load\_partially\_completed\_file” unticked. But if for any reason you stop before finishing all 63 evaluations, and pick the task at a later time after restarting the Notebook runtime, ticking this box (**orange circle** below) and running the cell again will load your existing file. If you don’t tick that box, your existing file will be erased and you will have to start over.

MemSum\_Human\_Evaluation\_1.ipynb ☆

File Edit View Insert Runtime Tools Help

Commands + Code + Text

- MemSum Human Evaluation
  - Clone repos, Install Dependencies, Set working directory (takes about 4 minutes)
    - [ ] ↳ 6 cells hidden
  - Utils for evaluation interface
    - Eval code
      - load\_partially\_completed\_file: ☐
      - Show code
  - Human Evaluation Experiment II:
    - [ ] run\_gui(dataset\_path = "human\_eval\_results/records\_memsum\_wo\_autostop\_neusum\_clean.jsonl")

6. Finally, click on the third and last “Play” button (red circle below):

MemSum\_Human\_Evaluation\_1.ipynb ☆

File Edit View Insert Runtime Tools Help

Commands + Code + Text

- MemSum Human Evaluation
  - Clone repos, Install Dependencies, Set working directory (takes about 4 minutes)
    - [ ] ↳ 6 cells hidden
  - Utils for evaluation interface
    - Eval code
      - load\_partially\_completed\_file: ☐
      - Show code
    - Human Evaluation Experiment II:
      - run\_gui(dataset\_path = "human\_eval\_results/records\_memsum\_wo\_autostop\_neusum\_clean.jsonl")

7. This will immediately launch the annotation interface, which looks like the figure below. You then have to follow the annotation instructions provided in the Notebook. You will need to read the reference summary on the left, and select which of Summary A or Summary B has an overall quality closer to that of the reference summary. To help you in your task, you can highlight text spans with similar meanings across all 3 summaries by copy/pasting a span in the top right text field and then clicking the “Highlight relevant sentences...” button to the left of it.

The screenshot shows a Jupyter Notebook environment with a file named 'MemSum\_Human\_Evaluation\_Tipynb'. The notebook contains a code cell that runs a GUI application for human evaluation. The application window is titled 'Human Evaluation Experiment II:' and features a 'Read' section with three columns: 'Reference Summary', 'Summary A', and 'Summary B'. Each column contains a paragraph of text. Below the text columns is an 'Evaluation' section with radio buttons for 'Summary A' and 'Summary B', and a 'Submit & Eval Next' button. The interface also includes a search bar at the top right and a 'Show Source Document >>>' link.

run\_gui(dataset\_path = "human\_eval\_results/records\_memsum\_wo\_autostop\_neusum\_clean.jsonl")

**Read**

Highlight relevant sentences given a query

Reference Summary	Summary A	Summary B
We experienced a case of malignant hyperthermia ( mh ) in 6-year - old boy during anesthesia induction for strabismus surgery . It has been generally reported that sevoflurane can induce the delayed onset of mh in the absence of succinylcholine . Our case of mh was elicited after about 2 - 3 min of sevoflurane administration with n2o , o2 and rocuronium . However , we successfully treated the patient by early recognition of his condition and administering symptomatic treatment and dantrolene .	A 6-year - old , 25 kg boy was scheduled for strabismus surgery in both eyes , neither the patient nor his family had any history of neuromuscular disease or a special family history . Approximately 90 min after the onset of mh , the patient recovered consciousness and the tracheal tube was extubated . The patient was discharged from the hospital four days after the onset of mh without any problem . reported that mh emerged about 150 minutes after general anesthesia with sevoflurane in a 24-year - old man who was undergoing bilateral sagittal split ramus osteotomy . We know that the time of onset of a fulminant episode of mh is unpredictable , and it can vary from within minutes to within several hours of induction , and the fulminant mh episodes are apparently the result of a rapid , sustained rise in myoplasmic ca . Our attention was drawn to the patient ' strabismus , and strabismus has been observed in mh susceptible patients . This case report demonstrates that sevoflurane can trigger mh within few minutes after exposure and so the patients who are susceptible to mh must be carefully monitored even during anesthetic induction .	A 6-year - old , 25 kg boy was scheduled for strabismus surgery in both eyes , neither the patient nor his family had any history of neuromuscular disease or a special family history . The patient was discharged from the hospital four days after the onset of mh without any problem . Generally , sevoflurane and desflurane have been reported to be less potent triggers , they produce a more gradual onset of mh and the onset of mh with sevoflurane in humans has been reported to occur both at an early period and after prolonged anesthesia . reported that mh emerged about 150 minutes after general anesthesia with sevoflurane in a 24-year - old man who was undergoing bilateral sagittal split ramus osteotomy . We know that the time of onset of a fulminant episode of mh is unpredictable , and it can vary from within minutes to within several hours of induction , and the fulminant mh episodes are apparently the result of a rapid , sustained rise in myoplasmic ca . The halothane and caffeine contracture tests are bioassays and they currently remain the most reliable indicators , but they are not generally used in our country , and so a diagnostic contracture test was not performed . This case report demonstrates that sevoflurane can trigger mh within few minutes after exposure and so the patients who are susceptible to mh must be carefully monitored even during anesthetic induction .

Show Source Document >>>

**Evaluation (choose one that is closer to the reference summary)**

Overall:

☐ Summary A

☐ Summary B

**Submit & Eval Next**

You have evaluated 0/63 examples.