

eval_plot

April 13, 2025

1 Creates plots from the eval runs over the checkpoints.

2024-12-31_21-52-49_rlvr_gsm8k_only_smollm2_instruct is the only real GSM8K only run.
https://wandb.ai/julesgm/open_instruct_rl/runs/jtobyw4t/overview is the associated Wandb run.

2024-12-31_22-56-50_rlvr_gsm8k_only_smollm2_instruct is the only real MATH only run.

```
[3]: # Standard library imports
import collections
import json
import os
import pathlib
import re

# Third party imports
import matplotlib.pyplot as plt
import more_itertools as mit
import numpy as np
import pandas as pd
import rich
import sklearn.linear_model as skl_lin

def extract_step(path):
    """ Look at the path name and extract the step number.

    Big assumption: Just takes the last number in the path name.
    """
    return int(re.findall(r"\d+", path.parent.name)[-1])

def extract_info_file(*, file, task_key, is_baseline=False):
    """ Reads the lighteval JSON file and extract the performance.
    """

    file = pathlib.Path(file)
```

```

if is_baseline:
    step = 0
else:
    step = extract_step(file)

d = json.loads(file.read_text())
task_results = d["results"][task_key]
qem = task_results["qem"]
qem_stderr = task_results["qem_stderr"]

return dict(step=step, qem=qem, qem_stderr=qem_stderr)

def plot_task(*, baseline_file, files, task_key, legend, marker="o",
↳linestyle="--"):
    """ Extracts the performance from the files
    """

    files = sorted(files, key=extract_step)

    shared = files[0].parent.parts
    for file_ in files:
        parts = file_.parent.parts
        for i, shared_part in enumerate(shared):
            if shared_part != parts[i]:
                shared = shared[:i]
                break
        if len(shared) == 0:
            break
    merged_shared = pathlib.Path(*shared)
    # rich.print(f"[bold]Shared path: {merged_shared}")
    # for file_ in files:
    #     print(file_.relative_to(merged_shared))

    if baseline_file:
        results = [
            extract_info_file(
                file=baseline_file,
                task_key=task_key,
                is_baseline=True,
            )]
    else:
        results = []

    for file in files:
        results.append(
            extract_info_file(

```

```

        file=file,
        task_key=task_key,
    )
)

results = pd.DataFrame(results)
x = results["step"]
y = results["qem"]
plt.plot(x, y, label=legend, marker=marker, linestyle=linestyle)
# Get the color of the last plot
color = plt.gca().lines[-1].get_color()

# # Smooth with a linear regression
# model = skl_lin.LinearRegression(fit_intercept=False)
# model.fit(x.values.reshape(-1, 1), y - y[0])
# x_pred = pd.Series(range(0, x.max() + 1))
# y_pred = model.predict(x_pred.values.reshape(-1, 1)) + y[0]
# plt.plot(x_pred, y_pred, linestyle="--", color=color)

```

[4]: *# Code to make plots for specific tasks.*

```

MATH_BASELINE = "./all_eval_outputs_important/previous_outputs/outputs_math/
↳results/HuggingFaceTB/SmolLM2-1.7B-Instruct/results_2025-01-02T20-51-13.
↳105551.json"
GSM8K_BASELINE = "./all_eval_outputs_important/previous_outputs/outputs_gsm8k/
↳results/HuggingFaceTB/SmolLM2-1.7B-Instruct/results_2025-01-02T15-48-12.
↳453113.json"

def gsm8k():
    task_key = "custom|gsm8k|8"
    eval_output_root_path = pathlib.Path("./all_eval_outputs_important/
↳previous_outputs/outputs_gsm8k/results/")
    plt.figure()
    plt.title("GSM8K")
    gsm8k_baseline_file = eval_output_root_path / "HuggingFaceTB/SmolLM2-1.
↳7B-Instruct/results_2025-01-02T15-48-12.453113.json"

    # GSM8K only
    glob_pattern =
↳"_network_scratch_g_gagnonju_open_instruct_output_2024-12-31_21-52-49_rlvr_gsm8k_only_smol
↳*.json"
    files = eval_output_root_path.glob(glob_pattern)
    plot_task(baseline_file=gsm8k_baseline_file, files=files,
↳task_key=task_key, legend="Trained on GSM8K only")
    print()

```

```

    # GSM8K and MATH
    glob_pattern = _
    ↪ "_network_scratch_g_gagnonju_open_instruct_output_2024-12-31_21-22-51_rlv_r_gsm8k_only_small
    ↪ *.json"
    files = eval_output_root_path.glob(glob_pattern)
    plot_task(baseline_file=gsm8k_baseline_file, files=files, _
    ↪ task_key=task_key, legend="Trained on GSM8K and MATH")

    plt.legend()

def math():
    task_key = "custom|math|5"
    eval_output_root_path = pathlib.Path("./all_eval_outputs_important/
    ↪ previous_outputs/outputs_math/results/")

    plt.figure()
    plt.title("Hendrycks MATH")

    # Baseline
    math_baseline_file = eval_output_root_path / "HuggingFaceTB/SmolLM2-1.
    ↪ 7B-Instruct/results_2025-01-02T20-51-13.105551.json"

    # GSM8K and MATH
    glob_pattern = _
    ↪ "_network_scratch_g_gagnonju_open_instruct_output_2024-12-31_21-22-51_rlv_r_gsm8k_only_small
    ↪ *.json"
    files = eval_output_root_path.glob(glob_pattern)
    plot_task(
        baseline_file=math_baseline_file,
        files=files,
        task_key=task_key,
        legend="Trained on GSM8K and MATH"
    )

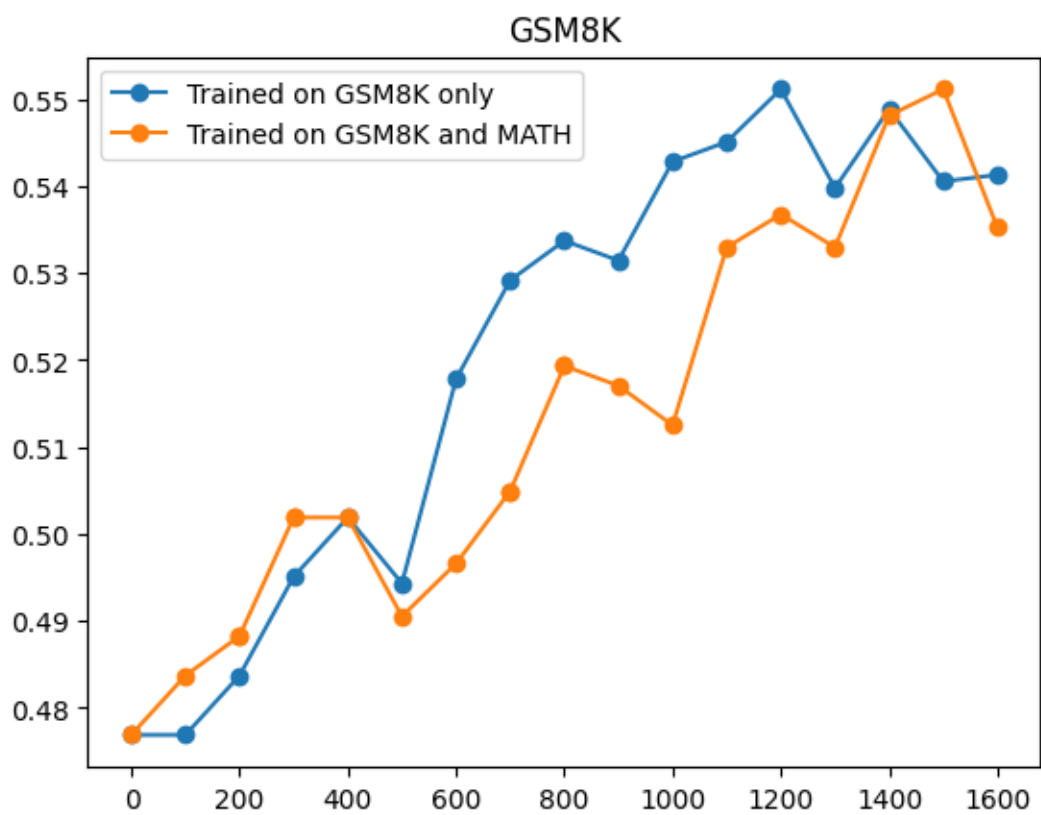
    glob_pattern = _
    ↪ "_network_scratch_g_gagnonju_open_instruct_output_2024-12-31_22-56-50_rlv_r_gsm8k_only_small
    ↪ *.json"
    files = eval_output_root_path.glob(glob_pattern)
    plot_task(
        baseline_file=math_baseline_file,
        files=files,
        task_key=task_key,
        legend="Trained MATH",
    )

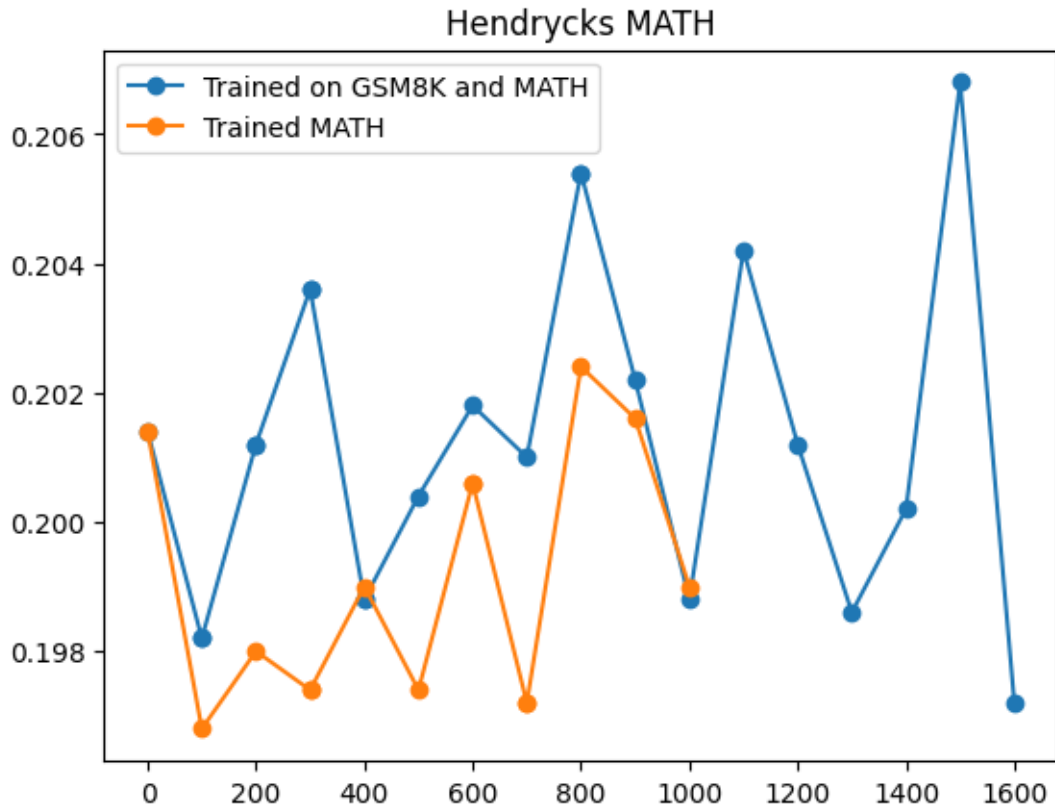
```

```
plt.legend()
```

```
gsm8k()
```

```
math()
```





```
[5]: def gsm8k_lambdal():
    task_key = "custom|gsm8k|8"
    eval_output_root_path = pathlib.Path("./all_eval_outputs_important/
↳ lambdal_outputs_gsm8k/results/")

    plt.figure()
    plt.title("GSM8K")

    gsm8k_baseline_file = GSM8K_BASELINE

    # GSM8K only
    glob_pattern = _
↳ "_network_scratch_g_gagnonju_lambdal_marglicot_openinstruct_open_instruct_output_2025-02-10
↳ *.json"
    files = eval_output_root_path.glob(glob_pattern)
    assert files, "GSM8K Only: No files found"
    plot_task(baseline_file=gsm8k_baseline_file, files=files, _
↳ task_key=task_key, legend="Trained on GSM8K only")
    print()
```

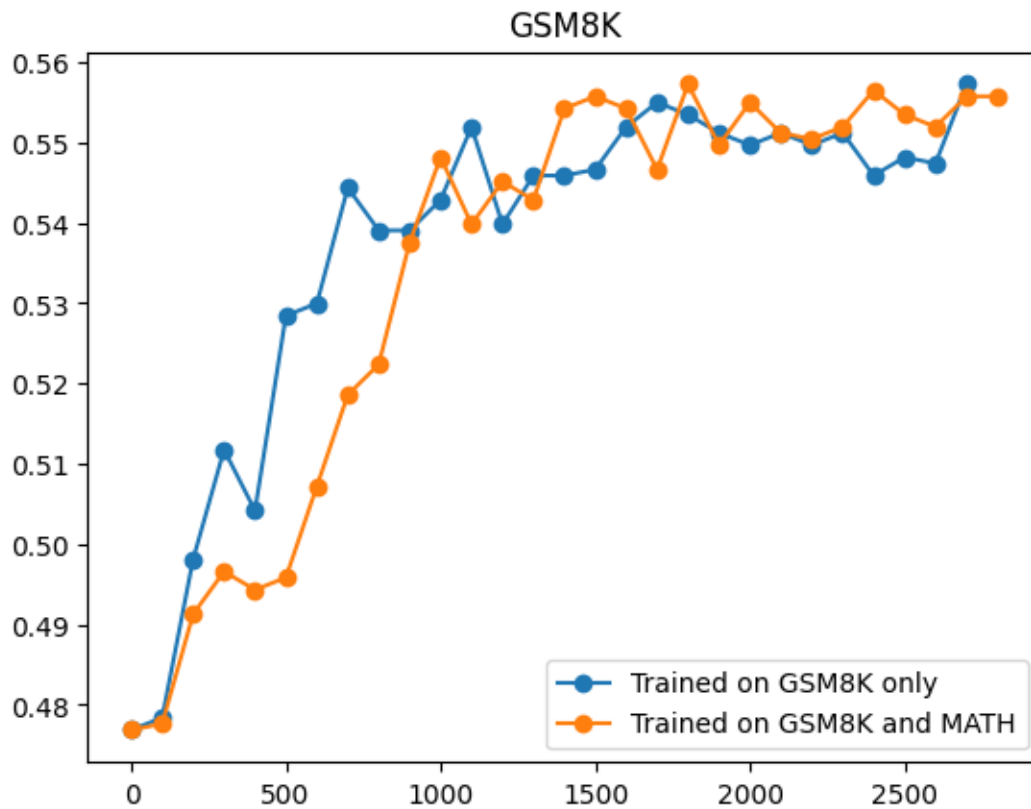
```

eval_output_root_path = pathlib.Path("./all_eval_outputs_important/
↳ lambdal_outputs_task_gsm8k_train_gsm8k_math/results")

# GSM8K and MATH
glob_pattern = _
↳ "_network_scratch_g_gagnonju_lambdal_marglicot_openinstruct_open_instruct_output_2025-02-10
↳ *.json"
files = eval_output_root_path.glob(glob_pattern)
assert files, "GSM8K & MATH: No files found"
plot_task(baseline_file=gsm8k_baseline_file, files=files, _
↳ task_key=task_key, legend="Trained on GSM8K and MATH")

plt.legend()
gsm8k_lambdal()

```



```

[6]: def from_path_and_glob(*, path, glob_pattern, baseline_file, task_key, _
↳ setting_name):
    assert path.exists(), path

```

```

path_glob_pattern = pathlib.Path(glob_pattern)
assert len(pathlib.Path(glob_pattern).parts) == 2

files = sorted(path.glob(glob_pattern))
assert files, f"{setting_name}: No files found"

plot_task(
    baseline_file=baseline_file,
    files=files,
    task_key=task_key,
    legend=f"Trained on {setting_name}"
)

def math_lambdal():
    task_key = "custom|math|5"

    plt.figure()
    plt.title("Hendrycks MATH")

    # Baseline
    math_baseline_file = pathlib.Path(MATH_BASELINE)
    assert math_baseline_file.exists(), "Baseline file not found"

    # GSM8K and MATH
    task_math_train_gsm8k_math_path = pathlib.Path("./
↳all_eval_outputs_important/lambdal_outputs_task_math_train_gsm8k_math/
↳results/")
    glob_pattern = _
↳"_network_scratch_g_gagnonju_lambdal_marglicot_openinstruct_open_instruct_output_2025-02-10
↳*.json"
    from_path_and_glob(
        path=task_math_train_gsm8k_math_path,
        glob_pattern=glob_pattern,
        baseline_file=math_baseline_file,
        task_key=task_key,
        setting_name="GSM8K and MATH"
    )

    # MATH
    task_math_train_math = pathlib.Path("./all_eval_outputs_important/
↳lambdal_outputs_math/results/")
    glob_pattern = _
↳"_network_scratch_g_gagnonju_lambdal_marglicot_openinstruct_open_instruct_output_2025-02-10
↳*.json"
    from_path_and_glob(
        path=task_math_train_math,
        glob_pattern=glob_pattern,

```



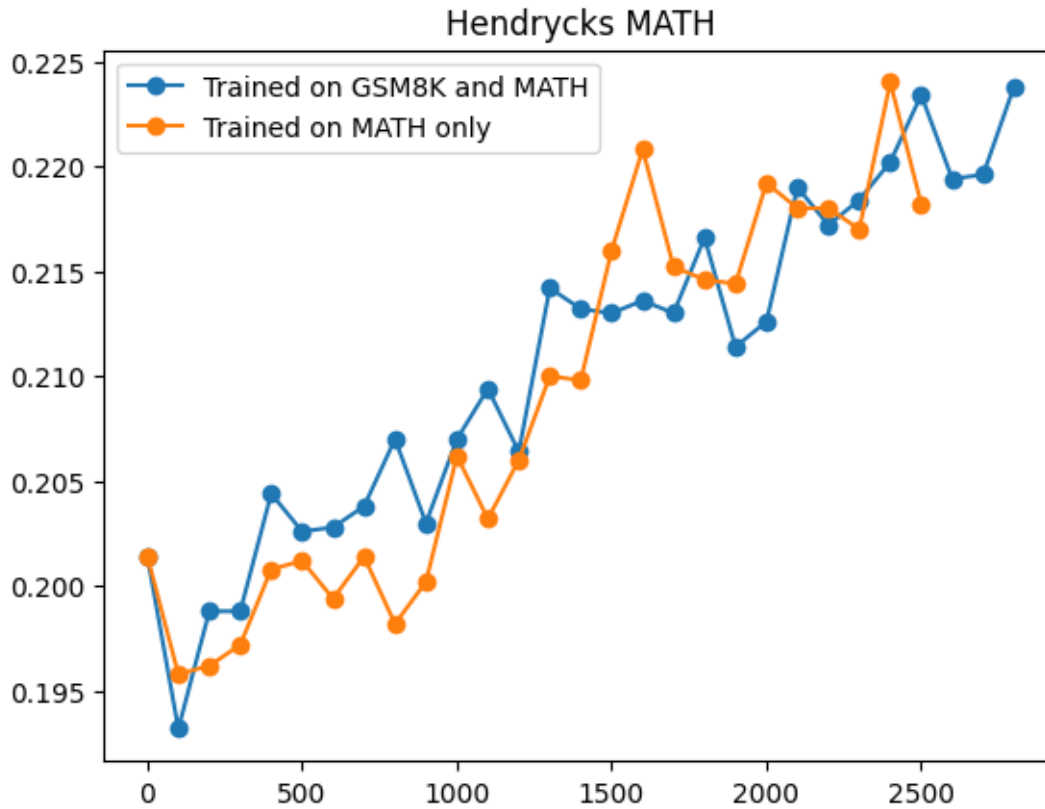
```

        baseline_file=math_baseline_file,
        task_key=task_key,
        setting_name="MATH only"
    )

    plt.legend()

    math_lambdal()

```



```

[53]: def sft_outputs(n_shots: int, task):
    assert isinstance(n_shots, int), type(n_shots).mro()
    task_key = f"custom|{task}|{n_shots}"
    eval_output_root_path = pathlib.Path(f"./all_eval_outputs_important/
    ↪sft_outputs_{task}/")

    if task == "gsm8k":
        other_task = "math"
    elif task == "math":
        other_task = "gsm8k"
    else:
        raise ValueError(f"Invalid task: {task}")

```

```

runs = []
for path in list(eval_output_root_path.iterdir()):

    shared = collections.defaultdict(list)
    for file in (path / "results").iterdir():
        if task in file.name.lower() and not other_task in file.name.
↳lower():
            shared_file_name = file.name.rsplit("_", 2)[0]
            shared[shared_file_name].append(file)

    for shared_file_name, files in shared.items():
        runs.append((path / "results" / shared_file_name, files))

initialized_figure = False

for path, run in runs:
    learning_rates = set()
    max_epochs = set()
    is_cot = set()

    for indiv in run:
        meta_infos = list(indiv.glob("meta_info.json"))
        if not meta_infos:
            final = indiv.name.rsplit("sft_saves_", 1)[-1].rsplit("_model",
↳1)[0].rsplit("_", 1)[0]
            epoch = indiv.name.rsplit("_", 2)[-2]

            path = pathlib.Path(os.environ["SCRATCH"]) / "marglicot_saves" /
↳ "sft_saves" / indiv.parts[2] / "checkpoints" / final / str(epoch) /
↳ "meta_info.json"
            assert path.exists(), path
            meta_infos.append(path)

        meta_infos = mit.one(meta_infos)
        meta_info = json.loads(meta_infos.read_text().strip())
        learning_rates.add(meta_info["cfg"]["learning_rate"])
        max_epochs.add(meta_info["cfg"]["max_num_epochs"])
        is_cot.add(meta_info["cfg"]["output_type"]["enum"] ==
↳ "chain_of_thought_then_answer")

    if len(is_cot) == 1:
        is_cot = mit.one(is_cot)
    else:
        is_cot = "unknown"

```

```

if len(learning_rates) == 1:
    learning_rate = mit.one(learning_rates)
else:
    learning_rate = "unknown"

if len(max_epochs) == 1:
    max_epoch = mit.one(max_epochs)
else:
    max_epoch = "unknown"

n_shots_set = set()
files = []
for path in run:
    result = mit.first(path.glob("results*.json"), None)
    if result is not None:
        result_json = json.loads(result.read_text())
        task_keys = set(result_json["summary_tasks"].keys())
        task_key = mit.one(task_keys)
        config_n_shots = int(task_key.rsplit("|", 1)[-1])
        n_shots_set.add(config_n_shots)
        if config_n_shots == n_shots:
            files.append(result)

if not n_shots_set:
    n_shots_config = "unknown"
else:
    n_shots_config = mit.one(n_shots_set, ValueError("Multiple n_shots_
↳found"))

    # print(f"n_shots: {n_shots_config}, lr: {learning_rate}, max_epoch:
↳{max_epoch}")

if files:
    if not initialized_figure:
        plt.figure(figsize=(15, 11))
        plt.title(f"{task.upper()} {n_shots} shots")
        plt.yticks(np.arange(0, 0.51, 0.01))
        initialized_figure = True

    plot_task(
        baseline_file=None,
        files=files,
        task_key=task_key,
        legend=f"n_shots: {n_shots}, lr: {learning_rate}, max_epoch:
↳{max_epoch} is_cot: {is_cot}",
        marker="o" if is_cot else "x",

```

```

        linestyle="--" if n_shots == "0" else "--",
    )
    else:
        pass
        # rich.print(f"[bold]Skipped:[/bold] lr: {learning_rate} max_epoch: {max_epoch} is_cot: {is_cot}")

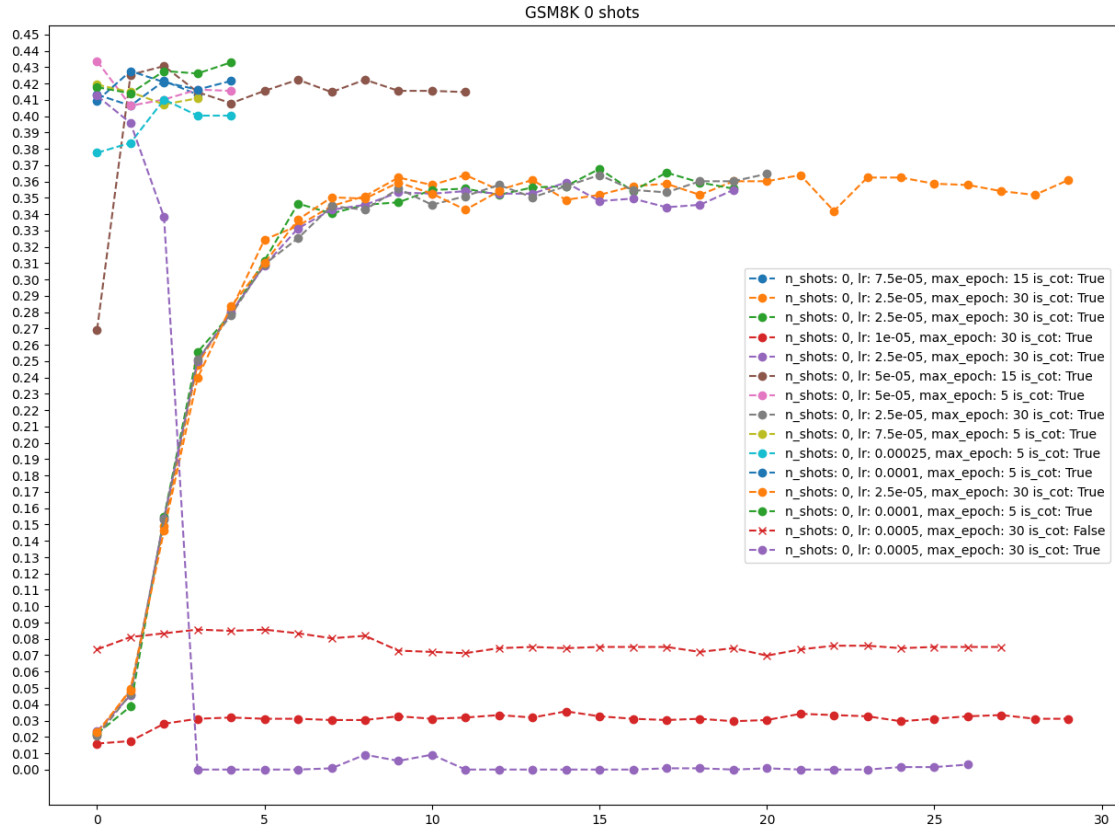
if initialized_figure:
    plt.legend()
    plt.show()

for n_shots in range(0, 10):
    print(f"n_shots: {n_shots}")
    sft_outputs(n_shots, "gsm8k")

for n_shots in range(0, 10):
    print(f"n_shots: {n_shots}")
    sft_outputs(n_shots, "math")

```

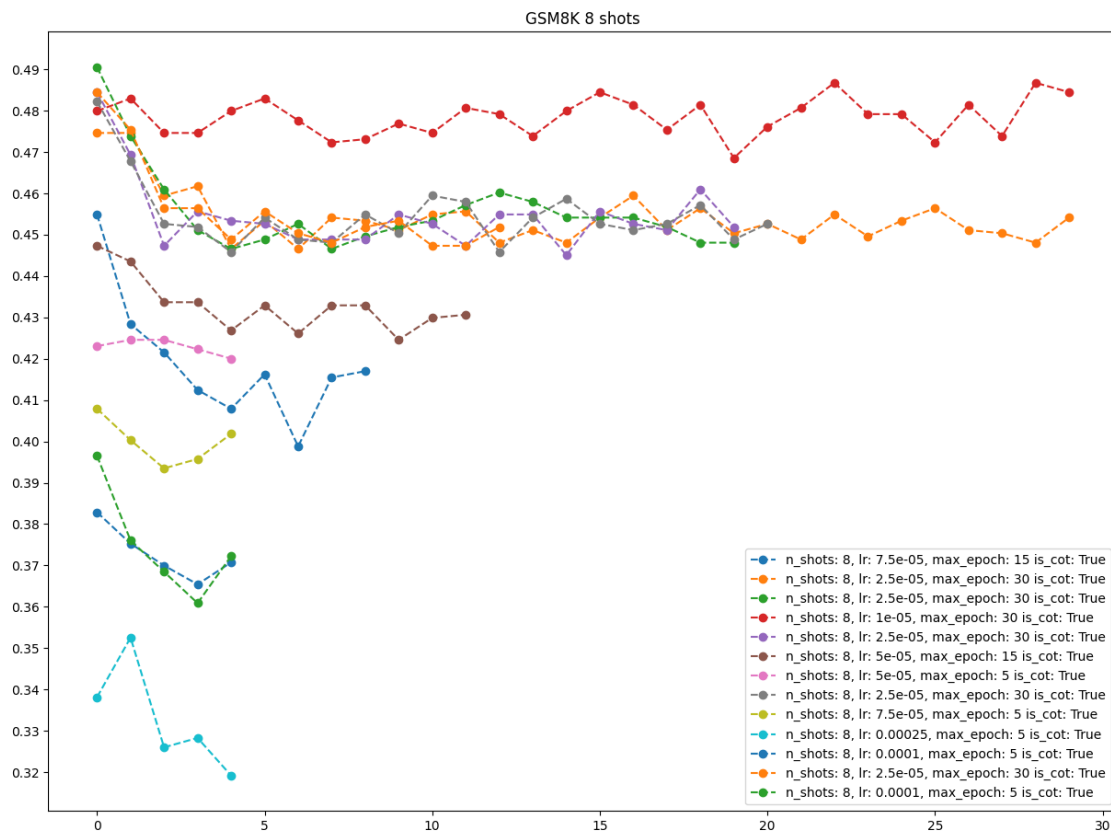
n_shots: 0



```

n_shots: 1
n_shots: 2
n_shots: 3
n_shots: 4
n_shots: 5
n_shots: 6
n_shots: 7
n_shots: 8

```



```

n_shots: 9
n_shots: 0

```

```

-----
FileNotFoundError                                Traceback (most recent call last)
Cell In[53], line 114
    112 for n_shots in range(0, 10):
    113     print(f"n_shots: {n_shots}")
--> 114     sft_outputs(n_shots, "math")

Cell In[53], line 17, in sft_outputs(n_shots, task)
    14 for path in list(eval_output_root_path.iterdir()):

```

```

    16     shared = collections.defaultdict(list)
---> 17     for file in (path / "results").iterdir():
    18         if task in file.name.lower() and not other_task in file.name.
↳lower():
    19             shared_file_name = file.name.rsplit("_", 2)[0]

File ~/mambaforge/lib/python3.10/pathlib.py:1017, in Path.iterdir(self)
    1013 def iterdir(self):
    1014     """Iterate over the files in this directory. Does not yield any
    1015     result for the special paths '.' and '..'.
    1016     """
-> 1017     for name in self._accessor.listdir(self):
    1018         if name in {'.', '..'}:
    1019             # Yielding a path object for these makes little sense
    1020             continue

FileNotFoundError: [Errno 2] No such file or directory:
↳'all_eval_outputs_important/sft_outputs_math/0_shot/results'

```

```

[52]: import dataclasses
import itertools

@dataclasses.dataclass(order=True, frozen=True, slots=True)
class EpochFile:
    epoch_number: int
    path: pathlib.Path = dataclasses.field(compare=False)

def supervised_sampling_outputs(n_shots_trained: int, n_shots_eval: int, task:
↳str):
    assert isinstance(n_shots_trained, int), type(n_shots_trained).mro()
    assert isinstance(n_shots_eval, int), type(n_shots_eval).mro()

    task_key = f"custom|{task}|{n_shots_eval}"
    eval_output_root_path = pathlib.Path(f"./all_eval_outputs_important/
↳rejection_sampling_outputs_{task}_trained_{n_shots_trained}_shot/
↳{n_shots_eval}_shot")
    if not eval_output_root_path.exists():
        return

    epoch_folders = eval_output_root_path / "results"

    shared_start_files = collections.defaultdict(list)
    shared_start_learning_rates = collections.defaultdict(set)
    shared_start_epochs = collections.defaultdict(list)
    for epoch_folder in epoch_folders.iterdir():
        json_results = mit.first(epoch_folder.glob("results*.json"), None)

```

```

        if json_results is None:
            continue
        results_json = json.loads(json_results.read_text())
        hydra_config = json.loads((epoch_folder / "hydra_config.json").
↪read_text())
        shared_first_part, epoch_number = epoch_folder.name.rsplit("epoch_", 1)
        epoch_number = int(epoch_number)

        print(f"Epoch {epoch_number}")
        print(results_json)
        print(hydra_config)

        shared_start_files[shared_first_part].
↪append(EpochFile(epoch_number=epoch_number, path=json_results))
        shared_start_learning_rates[shared_first_part].
↪add(hydra_config["training"]["learning_rate"])

    for k in shared_start_files:
        shared_start_files[k].sort()

    if not any(itertools.chain.from_iterable(shared_start_files.values())):
        return

    for shared_first_part, files in shared_start_files.items():
        learning_rate = mit.one(shared_start_learning_rates[shared_first_part])
        assert isinstance(files, list), type(files)
        plot_task(
            baseline_file=None,
            files=[x.path for x in files],
            task_key=task_key,
            legend=f"n_shots_trained: {n_shots_trained}, n_shots_eval: ↪
↪{n_shots_eval}, lr: {learning_rate}",
            marker="o",
            linestyle="-",
        )

for task in ["gsm8k", "math"]:
    plt.figure()
    plt.title(f"{task.upper()} {n_shots_trained} shots")
    for n_shots_trained in range(0, 10):
        for n_shots_eval in range(0, 10):
            # rich.print(f"[bold]{task}:[/bold] {task}, [bold]n_shots_trained:[/
↪bold] {n_shots_trained}, [bold]n_shots_eval:[/bold] {n_shots_eval}")
            supervised_sampling_outputs(n_shots_trained=n_shots_trained, ↪
↪n_shots_eval=n_shots_eval, task=task)

```

```
plt.legend()
plt.show()
```

Epoch 2

```
{'config_general': {'lighteval_sha': '?', 'num_fewshot_seeds': 1,
'override_batch_size': -1, 'max_samples': None, 'job_id': '', 'start_time':
14296438.947322655, 'end_time': 14296512.580050832,
'total_evaluation_time_seconds': '73.63272817619145', 'model_name': '_network_s
cratch_g_gagnonju_rejection_sampling_saves_gsm8k_2025-03-16_02-32-57_epoch_2',
'model_sha': '', 'model_dtype': None, 'model_size': None}, 'results':
{'custom|gsm8k|0': {'qem': 0.006065200909780136, 'qem_stderr':
0.002138670301460471}, 'all': {'qem': 0.006065200909780136, 'qem_stderr':
0.002138670301460471}}, 'versions': {'custom|gsm8k|0': 0}, 'config_tasks':
{'custom|gsm8k': {'name': 'gsm8k', 'prompt_function': 'jules_gsm8k', 'hf_repo':
'openai/gsm8k', 'hf_subset': 'main', 'metric': [{'metric_name': 'qem',
'higher_is_better': True, 'category': '3', 'use_case': '5', 'sample_level_fn':
'compute', 'corpus_level_fn': 'mean'}]}, 'hf_revision':
'e53f048856ff4f594e959d75785d2c2d37b678ee', 'hf_filter': None,
'hf_avail_splits': ['train', 'test'], 'trust_dataset': False,
'evaluation_splits': ['test'], 'few_shots_split': None, 'few_shots_select':
'random_sampling_from_train', 'generation_size': 256, 'generation_grammar':
None, 'stop_sequence': ['Question:', 'Question'], 'output_regex': None,
'num_samples': None, 'suite': ['custom'], 'original_num_docs': 1319,
'effective_num_docs': 1319, 'must_remove_duplicate_docs': False, 'version': 0,
'frozen': False}}, 'summary_tasks': {'custom|gsm8k|0': {'hashes':
{'hash_examples': '4c0843a5d99bcfdc', 'hash_full_prompts': 'c53d40c3dafc71d9',
'hash_input_tokens': 'df9416030feff162', 'hash_cont_tokens':
'037178dce68193cd'}, 'truncated': 0, 'non_truncated': 1319, 'padded': 0,
'non_padded': 1319, 'effective_few_shots': 0.0, 'num_truncated_few_shots': 0}},
'summary_general': {'hashes': {'hash_examples': '18b756b7813d1bdf',
'hash_full_prompts': 'bbae0e5faff2dc56', 'hash_input_tokens':
'6d816743a921fc1e', 'hash_cont_tokens': '7fb9c3c0def9ce4e'}, 'truncated': 0,
'non_truncated': 1319, 'padded': 0, 'non_padded': 1319,
'num_truncated_few_shots': 0}}
{'accelerate': {'gpu_ids': [0], 'seed': 42}, 'dataset': {'answer_field':
'answer', 'load_dataset_args': ['main'], 'name': 'gsm8k', 'question_field':
'question', 'split': 'train', 'valid_split': 'test'}, 'evaluation':
{'eval_batch_size': 1500, 'eval_percentage': 0.05, 'eval_subset': None},
'experiment_name': 'gsm8k_experiment', 'internal_master_port': 25124,
'master_port': 25123, 'model': {'name': 'HuggingFaceTB/SmolLM2-1.7B-Instruct'},
'output_dir': '/network/scratch/g/gagnonju/rejection_sampling_saves/gsm8k/2025-
03-16_02-32-57/', 'training': {'batch_size': 12, 'learning_rate': 0.0001,
'max_length': 768, 'num_epochs': 30}, 'vllm': {'gpu_id': 1}, 'vllm_sampling':
{'num_candidates': 48, 'temperature': 1, 'top_p': None}, 'wandb': {'entity':
'julesgm', 'log_interval': 10, 'project': 'expert_iteration'}}
```

Epoch 3


```

{'config_general': {'lighteval_sha': '?', 'num_fewshot_seeds': 1,
'override_batch_size': -1, 'max_samples': None, 'job_id': '', 'start_time':
14296438.952243285, 'end_time': 14296512.26563578,
'total_evaluation_time_seconds': '73.31339249387383', 'model_name': '_network_s
cratch_g_gagnonju_rejection_sampling_saves_gsm8k_2025-03-16_02-32-57_epoch_3',
'model_sha': '', 'model_dtype': None, 'model_size': None}, 'results':
{'custom|gsm8k|0': {'qem': 0.0075815011372251705, 'qem_stderr':
0.0023892815120772357}, 'all': {'qem': 0.0075815011372251705, 'qem_stderr':
0.0023892815120772357}}, 'versions': {'custom|gsm8k|0': 0}, 'config_tasks':
{'custom|gsm8k': {'name': 'gsm8k', 'prompt_function': 'jules_gsm8k', 'hf_repo':
'openai/gsm8k', 'hf_subset': 'main', 'metric': [{'metric_name': 'qem',
'higher_is_better': True, 'category': '3', 'use_case': '5', 'sample_level_fn':
'compute', 'corpus_level_fn': 'mean'}]}, 'hf_revision':
'e53f048856ff4f594e959d75785d2c2d37b678ee', 'hf_filter': None,
'hf_avail_splits': ['train', 'test'], 'trust_dataset': False,
'evaluation_splits': ['test'], 'few_shots_split': None, 'few_shots_select':
'random_sampling_from_train', 'generation_size': 256, 'generation_grammar':
None, 'stop_sequence': ['Question:', 'Question'], 'output_regex': None,
'num_samples': None, 'suite': ['custom'], 'original_num_docs': 1319,
'effective_num_docs': 1319, 'must_remove_duplicate_docs': False, 'version': 0,
'frozen': False}}, 'summary_tasks': {'custom|gsm8k|0': {'hashes':
{'hash_examples': '4c0843a5d99bcfdc', 'hash_full_prompts': 'c53d40c3dafc71d9',
'hash_input_tokens': 'df9416030feff162', 'hash_cont_tokens':
'd484d6543472ecc9'}, 'truncated': 0, 'non_truncated': 1319, 'padded': 0,
'non_padded': 1319, 'effective_few_shots': 0.0, 'num_truncated_few_shots': 0}},
'summary_general': {'hashes': {'hash_examples': '18b756b7813d1bdf',
'hash_full_prompts': 'bbae0e5faff2dc56', 'hash_input_tokens':
'6d816743a921fc1e', 'hash_cont_tokens': '4a3d10a6f5ba8286'}, 'truncated': 0,
'non_truncated': 1319, 'padded': 0, 'non_padded': 1319,
'num_truncated_few_shots': 0}}
{'accelerate': {'gpu_ids': [0], 'seed': 42}, 'dataset': {'answer_field':
'answer', 'load_dataset_args': ['main'], 'name': 'gsm8k', 'question_field':
'question', 'split': 'train', 'valid_split': 'test'}, 'evaluation':
{'eval_batch_size': 1500, 'eval_percentage': 0.05, 'eval_subset': None},
'experiment_name': 'gsm8k_experiment', 'internal_master_port': 25124,
'master_port': 25123, 'model': {'name': 'HuggingFaceTB/SmolLM2-1.7B-Instruct'},
'output_dir': '/network/scratch/g/gagnonju/rejection_sampling_saves/gsm8k/2025-
03-16_02-32-57/', 'training': {'batch_size': 12, 'learning_rate': 0.0001,
'max_length': 768, 'num_epochs': 30}, 'vllm': {'gpu_id': 1}, 'vllm_sampling':
{'num_candidates': 48, 'temperature': 1, 'top_p': None}, 'wandb': {'entity':
'julesgm', 'log_interval': 10, 'project': 'expert_iteration'}}
Epoch 0
{'config_general': {'lighteval_sha': '?', 'num_fewshot_seeds': 1,
'override_batch_size': -1, 'max_samples': None, 'job_id': '', 'start_time':
14296449.905404476, 'end_time': 14296523.594348352,
'total_evaluation_time_seconds': '73.68894387595356', 'model_name': '_network_s
cratch_g_gagnonju_rejection_sampling_saves_gsm8k_2025-03-16_02-32-57_epoch_0',
'model_sha': '', 'model_dtype': None, 'model_size': None}, 'results':

```

```
{'custom|gsm8k|0': {'qem': 0.013646702047005308, 'qem_stderr':
0.0031957470754808257}, 'all': {'qem': 0.013646702047005308, 'qem_stderr':
0.0031957470754808257}}, 'versions': {'custom|gsm8k|0': 0}, 'config_tasks':
{'custom|gsm8k': {'name': 'gsm8k', 'prompt_function': 'jules_gsm8k', 'hf_repo':
'openai/gsm8k', 'hf_subset': 'main', 'metric': [{'metric_name': 'qem',
'higher_is_better': True, 'category': '3', 'use_case': '5', 'sample_level_fn':
'compute', 'corpus_level_fn': 'mean'}]}, 'hf_revision':
'e53f048856ff4f594e959d75785d2c2d37b678ee', 'hf_filter': None,
'hf_avail_splits': ['train', 'test'], 'trust_dataset': False,
'evaluation_splits': ['test'], 'few_shots_split': None, 'few_shots_select':
'random_sampling_from_train', 'generation_size': 256, 'generation_grammar':
None, 'stop_sequence': ['Question:', 'Question'], 'output_regex': None,
'num_samples': None, 'suite': ['custom'], 'original_num_docs': 1319,
'effective_num_docs': 1319, 'must_remove_duplicate_docs': False, 'version': 0,
'frozen': False}}, 'summary_tasks': {'custom|gsm8k|0': {'hashes':
{'hash_examples': '4c0843a5d99bcfdc', 'hash_full_prompts': 'c53d40c3dafc71d9',
'hash_input_tokens': 'df9416030feff162', 'hash_cont_tokens':
'a981ac14d8acff78'}, 'truncated': 0, 'non_truncated': 1319, 'padded': 0,
'non_padded': 1319, 'effective_few_shots': 0.0, 'num_truncated_few_shots': 0}},
'summary_general': {'hashes': {'hash_examples': '18b756b7813d1bdf',
'hash_full_prompts': 'bbae0e5faff2dc56', 'hash_input_tokens':
'6d816743a921fc1e', 'hash_cont_tokens': '22da611f78f1b333'}, 'truncated': 0,
'non_truncated': 1319, 'padded': 0, 'non_padded': 1319,
'num_truncated_few_shots': 0}}
{'accelerate': {'gpu_ids': [0], 'seed': 42}, 'dataset': {'answer_field':
'answer', 'load_dataset_args': ['main'], 'name': 'gsm8k', 'question_field':
'question', 'split': 'train', 'valid_split': 'test'}, 'evaluation':
{'eval_batch_size': 1500, 'eval_percentage': 0.05, 'eval_subset': None},
'experiment_name': 'gsm8k_experiment', 'internal_master_port': 25124,
'master_port': 25123, 'model': {'name': 'HuggingFaceTB/SmolLM2-1.7B-Instruct'},
'output_dir': '/network/scratch/g/gagnonju/rejection_sampling_saves/gsm8k/2025-
03-16_02-32-57/', 'training': {'batch_size': 12, 'learning_rate': 0.0001,
'max_length': 768, 'num_epochs': 30}, 'vllm': {'gpu_id': 1}, 'vllm_sampling':
{'num_candidates': 48, 'temperature': 1, 'top_p': None}, 'wandb': {'entity':
'julesgm', 'log_interval': 10, 'project': 'expert_iteration'}}
Epoch 0
{'config_general': {'lighteval_sha': '?', 'num_fewshot_seeds': 1,
'override_batch_size': -1, 'max_samples': None, 'job_id': '', 'start_time':
15917358.636957237, 'end_time': 15917434.03633844,
'total_evaluation_time_secondes': '75.39938120171428', 'model_name': '_network_s
cratch_g_gagnonju_rejection_sampling_saves_gsm8k_2025-03-16_19-48-52_epoch_0',
'model_sha': '', 'model_dtype': None, 'model_size': None}, 'results':
{'custom|gsm8k|0': {'qem': 0.025018953752843062, 'qem_stderr':
0.004302045046564291}, 'all': {'qem': 0.025018953752843062, 'qem_stderr':
0.004302045046564291}}, 'versions': {'custom|gsm8k|0': 0}, 'config_tasks':
{'custom|gsm8k': {'name': 'gsm8k', 'prompt_function': 'jules_gsm8k', 'hf_repo':
'openai/gsm8k', 'hf_subset': 'main', 'metric': [{'metric_name': 'qem',
'higher_is_better': True, 'category': '3', 'use_case': '5', 'sample_level_fn':
```

```

'compute', 'corpus_level_fn': 'mean']], 'hf_revision':
'e53f048856ff4f594e959d75785d2c2d37b678ee', 'hf_filter': None,
'hf_avail_splits': ['train', 'test'], 'trust_dataset': False,
'evaluation_splits': ['test'], 'few_shots_split': None, 'few_shots_select':
'random_sampling_from_train', 'generation_size': 256, 'generation_grammar':
None, 'stop_sequence': ['Question:', 'Question'], 'output_regex': None,
'num_samples': None, 'suite': ['custom'], 'original_num_docs': 1319,
'effective_num_docs': 1319, 'must_remove_duplicate_docs': False, 'version': 0,
'frozen': False}}, 'summary_tasks': {'custom|gsm8k|0': {'hashes':
{'hash_examples': '4c0843a5d99bcfdc', 'hash_full_prompts': 'c53d40c3dafc71d9',
'hash_input_tokens': 'df9416030feff162', 'hash_cont_tokens':
'3ecc2708162cdcf4'}}, 'truncated': 0, 'non_truncated': 1319, 'padded': 0,
'non_padded': 1319, 'effective_few_shots': 0.0, 'num_truncated_few_shots': 0}},
'summary_general': {'hashes': {'hash_examples': '18b756b7813d1bdf',
'hash_full_prompts': 'bbae0e5faff2dc56', 'hash_input_tokens':
'6d816743a921fc1e', 'hash_cont_tokens': '697b632bfedabff0'}, 'truncated': 0,
'non_truncated': 1319, 'padded': 0, 'non_padded': 1319,
'num_truncated_few_shots': 0}}
{'accelerate': {'gpu_ids': [0], 'seed': 42}, 'dataset': {'answer_field':
'answer', 'load_dataset_args': ['main'], 'name': 'gsm8k', 'question_field':
'question', 'split': 'train', 'valid_split': 'test'}, 'evaluation':
{'eval_batch_size': 1500, 'eval_percentage': 0.05, 'eval_subset': None},
'experiment_name': 'gsm8k_experiment', 'internal_master_port': 25124,
'master_port': 25123, 'model': {'name': 'HuggingFaceTB/SmolLM2-1.7B-Instruct'},
'output_dir': '/network/scratch/g/gagnonju/rejection_sampling_saves/gsm8k/2025-
03-16_19-48-52/', 'training': {'batch_size': 12, 'learning_rate': 1e-05,
'max_length': 768, 'num_epochs': 30}, 'vllm': {'gpu_id': 1}, 'vllm_sampling':
{'num_candidates': 48, 'temperature': 1, 'top_p': None}, 'wandb': {'entity':
'julesgm', 'log_interval': 10, 'project': 'expert_iteration'}}
Epoch 4
{'config_general': {'lighteval_sha': '?', 'num_fewshot_seeds': 1,
'override_batch_size': -1, 'max_samples': None, 'job_id': '', 'start_time':
14296438.956267824, 'end_time': 14296513.201798344,
'total_evaluation_time_seconds': '74.24553051963449', 'model_name': '_network_s
cratch_g_gagnonju_rejection_sampling_saves_gsm8k_2025-03-16_02-32-57_epoch_4',
'model_sha': '', 'model_dtype': None, 'model_size': None}, 'results':
{'custom|gsm8k|0': {'qem': 0.006823351023502654, 'qem_stderr':
0.0022675371022544805}, 'all': {'qem': 0.006823351023502654, 'qem_stderr':
0.0022675371022544805}}, 'versions': {'custom|gsm8k|0': 0}, 'config_tasks':
{'custom|gsm8k': {'name': 'gsm8k', 'prompt_function': 'jules_gsm8k', 'hf_repo':
'openai/gsm8k', 'hf_subset': 'main', 'metric': [{'metric_name': 'qem',
'higher_is_better': True, 'category': '3', 'use_case': '5', 'sample_level_fn':
'compute', 'corpus_level_fn': 'mean'}], 'hf_revision':
'e53f048856ff4f594e959d75785d2c2d37b678ee', 'hf_filter': None,
'hf_avail_splits': ['train', 'test'], 'trust_dataset': False,
'evaluation_splits': ['test'], 'few_shots_split': None, 'few_shots_select':
'random_sampling_from_train', 'generation_size': 256, 'generation_grammar':
None, 'stop_sequence': ['Question:', 'Question'], 'output_regex': None,

```

```

'num_samples': None, 'suite': ['custom'], 'original_num_docs': 1319,
'effective_num_docs': 1319, 'must_remove_duplicate_docs': False, 'version': 0,
'frozen': False}}, 'summary_tasks': {'custom|gsm8k|0': {'hashes':
{'hash_examples': '4c0843a5d99bcfdc', 'hash_full_prompts': 'c53d40c3dafc71d9',
'hash_input_tokens': 'df9416030feff162', 'hash_cont_tokens':
'1671f3e23c632463'}}, 'truncated': 0, 'non_truncated': 1319, 'padded': 0,
'non_padded': 1319, 'effective_few_shots': 0.0, 'num_truncated_few_shots': 0}},
'summary_general': {'hashes': {'hash_examples': '18b756b7813d1bdf',
'hash_full_prompts': 'bbae0e5faff2dc56', 'hash_input_tokens':
'6d816743a921fc1e', 'hash_cont_tokens': '5a87a4476695929a'}, 'truncated': 0,
'non_truncated': 1319, 'padded': 0, 'non_padded': 1319,
'num_truncated_few_shots': 0}}
{'accelerate': {'gpu_ids': [0], 'seed': 42}, 'dataset': {'answer_field':
'answer', 'load_dataset_args': ['main'], 'name': 'gsm8k', 'question_field':
'question', 'split': 'train', 'valid_split': 'test'}, 'evaluation':
{'eval_batch_size': 1500, 'eval_percentage': 0.05, 'eval_subset': None},
'experiment_name': 'gsm8k_experiment', 'internal_master_port': 25124,
'master_port': 25123, 'model': {'name': 'HuggingFaceTB/SmolLM2-1.7B-Instruct'},
'output_dir': '/network/scratch/g/gagnonju/rejection_sampling_saves/gsm8k/2025-
03-16_02-32-57/', 'training': {'batch_size': 12, 'learning_rate': 0.0001,
'max_length': 768, 'num_epochs': 30}, 'vllm': {'gpu_id': 1}, 'vllm_sampling':
{'num_candidates': 48, 'temperature': 1, 'top_p': None}, 'wandb': {'entity':
'julesgm', 'log_interval': 10, 'project': 'expert_iteration'}}
Epoch 1
{'config_general': {'lighteval_sha': '?', 'num_fewshot_seeds': 1,
'override_batch_size': -1, 'max_samples': None, 'job_id': '', 'start_time':
10988430.484439556, 'end_time': 10988503.649234096,
'total_evaluation_time_seconds': '73.16479454003274', 'model_name': '_network_s
cratch_g_gagnonju_rejection_sampling_saves_gsm8k_2025-03-16_02-32-57_epoch_1',
'model_sha': '', 'model_dtype': None, 'model_size': None}, 'results':
{'custom|gsm8k|0': {'qem': 0.0075815011372251705, 'qem_stderr':
0.0023892815120772357}, 'all': {'qem': 0.0075815011372251705, 'qem_stderr':
0.0023892815120772357}}, 'versions': {'custom|gsm8k|0': 0}, 'config_tasks':
{'custom|gsm8k': {'name': 'gsm8k', 'prompt_function': 'jules_gsm8k', 'hf_repo':
'openai/gsm8k', 'hf_subset': 'main', 'metric': [{'metric_name': 'qem',
'higher_is_better': True, 'category': '3', 'use_case': '5', 'sample_level_fn':
'compute', 'corpus_level_fn': 'mean'}]}, 'hf_revision':
'e53f048856ff4f594e959d75785d2c2d37b678ee', 'hf_filter': None,
'hf_avail_splits': ['train', 'test'], 'trust_dataset': False,
'evaluation_splits': ['test'], 'few_shots_split': None, 'few_shots_select':
'random_sampling_from_train', 'generation_size': 256, 'generation_grammar':
None, 'stop_sequence': ['Question:', 'Question'], 'output_regex': None,
'num_samples': None, 'suite': ['custom'], 'original_num_docs': 1319,
'effective_num_docs': 1319, 'must_remove_duplicate_docs': False, 'version': 0,
'frozen': False}}, 'summary_tasks': {'custom|gsm8k|0': {'hashes':
{'hash_examples': '4c0843a5d99bcfdc', 'hash_full_prompts': 'c53d40c3dafc71d9',
'hash_input_tokens': 'df9416030feff162', 'hash_cont_tokens':
'676dc72c254a076f'}, 'truncated': 0, 'non_truncated': 1319, 'padded': 0,

```

```

'non_padded': 1319, 'effective_few_shots': 0.0, 'num_truncated_few_shots': 0}},
'summary_general': {'hashes': {'hash_examples': '18b756b7813d1bdf',
'hash_full_prompts': 'bbae0e5faff2dc56', 'hash_input_tokens':
'6d816743a921fc1e', 'hash_cont_tokens': '4b3f29864d6bea1b'}, 'truncated': 0,
'non_truncated': 1319, 'padded': 0, 'non_padded': 1319,
'num_truncated_few_shots': 0}}
{'accelerate': {'gpu_ids': [0], 'seed': 42}, 'dataset': {'answer_field':
'answer', 'load_dataset_args': ['main'], 'name': 'gsm8k', 'question_field':
'question', 'split': 'train', 'valid_split': 'test'}, 'evaluation':
{'eval_batch_size': 1500, 'eval_percentage': 0.05, 'eval_subset': None},
'experiment_name': 'gsm8k_experiment', 'internal_master_port': 25124,
'master_port': 25123, 'model': {'name': 'HuggingFaceTB/SmolLM2-1.7B-Instruct'},
'output_dir': '/network/scratch/g/gagnonju/rejection_sampling_saves/gsm8k/2025-
03-16_02-32-57/', 'training': {'batch_size': 12, 'learning_rate': 0.0001,
'max_length': 768, 'num_epochs': 30}, 'vllm': {'gpu_id': 1}, 'vllm_sampling':
{'num_candidates': 48, 'temperature': 1, 'top_p': None}, 'wandb': {'entity':
'julesgm', 'log_interval': 10, 'project': 'expert_iteration'}}
Epoch 5
{'config_general': {'lighteval_sha': '?', 'num_fewshot_seeds': 1,
'override_batch_size': -1, 'max_samples': None, 'job_id': '', 'start_time':
3305591.985749133, 'end_time': 3305668.312336878,
'total_evaluation_time_seconds': '76.3265877449885', 'model_name': '_network_sc
cratch_g_gagnonju_rejection_sampling_saves_gsm8k_2025-03-16_02-32-57_epoch_5',
'model_sha': '', 'model_dtype': None, 'model_size': None}, 'results':
{'custom|gsm8k|0': {'qem': 0.006823351023502654, 'qem_stderr':
0.0022675371022544805}, 'all': {'qem': 0.006823351023502654, 'qem_stderr':
0.0022675371022544805}}, 'versions': {'custom|gsm8k|0': 0}, 'config_tasks':
{'custom|gsm8k': {'name': 'gsm8k', 'prompt_function': 'jules_gsm8k', 'hf_repo':
'openai/gsm8k', 'hf_subset': 'main', 'metric': [{'metric_name': 'qem',
'higher_is_better': True, 'category': '3', 'use_case': '5', 'sample_level_fn':
'compute', 'corpus_level_fn': 'mean'}]}, 'hf_revision':
'e53f048856ff4f594e959d75785d2c2d37b678ee', 'hf_filter': None,
'hf_avail_splits': ['train', 'test'], 'trust_dataset': False,
'evaluation_splits': ['test'], 'few_shots_split': None, 'few_shots_select':
'random_sampling_from_train', 'generation_size': 256, 'generation_grammar':
None, 'stop_sequence': ['Question:', 'Question'], 'output_regex': None,
'num_samples': None, 'suite': ['custom'], 'original_num_docs': 1319,
'effective_num_docs': 1319, 'must_remove_duplicate_docs': False, 'version': 0,
'frozen': False}}, 'summary_tasks': {'custom|gsm8k|0': {'hashes':
{'hash_examples': '4c0843a5d99bcfdc', 'hash_full_prompts': 'c53d40c3dafc71d9',
'hash_input_tokens': 'df9416030feff162', 'hash_cont_tokens':
'1671f3e23c632463'}, 'truncated': 0, 'non_truncated': 1319, 'padded': 0,
'non_padded': 1319, 'effective_few_shots': 0.0, 'num_truncated_few_shots': 0}},
'summary_general': {'hashes': {'hash_examples': '18b756b7813d1bdf',
'hash_full_prompts': 'bbae0e5faff2dc56', 'hash_input_tokens':
'6d816743a921fc1e', 'hash_cont_tokens': '5a87a4476695929a'}, 'truncated': 0,
'non_truncated': 1319, 'padded': 0, 'non_padded': 1319,
'num_truncated_few_shots': 0}}

```

```

{'accelerate': {'gpu_ids': [0], 'seed': 42}, 'dataset': {'answer_field':
'answer', 'load_dataset_args': ['main'], 'name': 'gsm8k', 'question_field':
'question', 'split': 'train', 'valid_split': 'test'}, 'evaluation':
{'eval_batch_size': 1500, 'eval_percentage': 0.05, 'eval_subset': None},
'experiment_name': 'gsm8k_experiment', 'internal_master_port': 25124,
'master_port': 25123, 'model': {'name': 'HuggingFaceTB/SmolLM2-1.7B-Instruct'},
'output_dir': '/network/scratch/g/gagnonju/rejection_sampling_saves/gsm8k/2025-
03-16_02-32-57/', 'training': {'batch_size': 12, 'learning_rate': 0.0001,
'max_length': 768, 'num_epochs': 30}, 'vllm': {'gpu_id': 1}, 'vllm_sampling':
{'num_candidates': 48, 'temperature': 1, 'top_p': None}, 'wandb': {'entity':
'julesgm', 'log_interval': 10, 'project': 'expert_iteration'}}
Epoch 2
{'config_general': {'lighteval_sha': '?', 'num_fewshot_seeds': 1,
'override_batch_size': -1, 'max_samples': None, 'job_id': '', 'start_time':
15919396.897363015, 'end_time': 15919601.671870537,
'total_evaluation_time_seconds': '204.774507522583', 'model_name': '_network_sc
cratch_g_gagnonju_rejection_sampling_saves_gsm8k_2025-03-16_02-32-57_epoch_2',
'model_sha': '', 'model_dtype': None, 'model_size': None}, 'results':
{'custom|gsm8k|8': {'qem': 0.47536012130401817, 'qem_stderr':
0.013755751352764918}, 'all': {'qem': 0.47536012130401817, 'qem_stderr':
0.013755751352764918}}, 'versions': {'custom|gsm8k|8': 0}, 'config_tasks':
{'custom|gsm8k': {'name': 'gsm8k', 'prompt_function': 'jules_gsm8k', 'hf_repo':
'openai/gsm8k', 'hf_subset': 'main', 'metric': [{'metric_name': 'qem',
'higher_is_better': True, 'category': '3', 'use_case': '5', 'sample_level_fn':
'compute', 'corpus_level_fn': 'mean'}]}, 'hf_revision':
'e53f048856ff4f594e959d75785d2c2d37b678ee', 'hf_filter': None,
'hf_avail_splits': ['train', 'test'], 'trust_dataset': False,
'evaluation_splits': ['test'], 'few_shots_split': None, 'few_shots_select':
'random_sampling_from_train', 'generation_size': 256, 'generation_grammar':
None, 'stop_sequence': ['Question:', 'Question'], 'output_regex': None,
'num_samples': None, 'suite': ['custom'], 'original_num_docs': 1319,
'effective_num_docs': 1319, 'must_remove_duplicate_docs': False, 'version': 0,
'frozen': False}}, 'summary_tasks': {'custom|gsm8k|8': {'hashes':
{'hash_examples': '4c0843a5d99bcfdc', 'hash_full_prompts': 'bc3e523598c718cb',
'hash_input_tokens': '9b0240e60aaa0b21', 'hash_cont_tokens':
'597d9aaafb898ff6'}, 'truncated': 0, 'non_truncated': 1319, 'padded': 0,
'non_padded': 1319, 'effective_few_shots': 8.0, 'num_truncated_few_shots': 0}},
'summary_general': {'hashes': {'hash_examples': '18b756b7813d1bdf',
'hash_full_prompts': 'ceeb6a330a123191', 'hash_input_tokens':
'6949d8c1d2d40dcd', 'hash_cont_tokens': '7c32fbaa116238f4'}, 'truncated': 0,
'non_truncated': 1319, 'padded': 0, 'non_padded': 1319,
'num_truncated_few_shots': 0}}
{'accelerate': {'gpu_ids': [0], 'seed': 42}, 'dataset': {'answer_field':
'answer', 'load_dataset_args': ['main'], 'name': 'gsm8k', 'question_field':
'question', 'split': 'train', 'valid_split': 'test'}, 'evaluation':
{'eval_batch_size': 1500, 'eval_percentage': 0.05, 'eval_subset': None},
'experiment_name': 'gsm8k_experiment', 'internal_master_port': 25124,
'master_port': 25123, 'model': {'name': 'HuggingFaceTB/SmolLM2-1.7B-Instruct'},

```

```
'output_dir': '/network/scratch/g/gagnonju/rejection_sampling_saves/gsm8k/2025-03-16_02-32-57/', 'training': {'batch_size': 12, 'learning_rate': 0.0001, 'max_length': 768, 'num_epochs': 30}, 'vllm': {'gpu_id': 1}, 'vllm_sampling': {'num_candidates': 48, 'temperature': 1, 'top_p': None}, 'wandb': {'entity': 'julesgm', 'log_interval': 10, 'project': 'expert_iteration'}}
```

Epoch 3

```
{'config_general': {'lighteval_sha': '?', 'num_fewshot_seeds': 1, 'override_batch_size': -1, 'max_samples': None, 'job_id': '', 'start_time': 14298497.041787483, 'end_time': 14298701.981319644, 'total_evaluation_time_seconds': '204.93953216075897', 'model_name': '_network_scratch_g_gagnonju_rejection_sampling_saves_gsm8k_2025-03-16_02-32-57_epoch_3', 'model_sha': '', 'model_dtype': None, 'model_size': None}, 'results': {'custom|gsm8k|8': {'qem': 0.46853677028051555, 'qem_stderr': 0.013745189948450412}, 'all': {'qem': 0.46853677028051555, 'qem_stderr': 0.013745189948450412}}, 'versions': {'custom|gsm8k|8': 0}, 'config_tasks': {'custom|gsm8k': {'name': 'gsm8k', 'prompt_function': 'jules_gsm8k', 'hf_repo': 'openai/gsm8k', 'hf_subset': 'main', 'metric': [{'metric_name': 'qem', 'higher_is_better': True, 'category': '3', 'use_case': '5', 'sample_level_fn': 'compute', 'corpus_level_fn': 'mean'}]}, 'hf_revision': 'e53f048856ff4f594e959d75785d2c2d37b678ee', 'hf_filter': None, 'hf_avail_splits': ['train', 'test'], 'trust_dataset': False, 'evaluation_splits': ['test'], 'few_shots_split': None, 'few_shots_select': 'random_sampling_from_train', 'generation_size': 256, 'generation_grammar': None, 'stop_sequence': ['Question:', 'Question'], 'output_regex': None, 'num_samples': None, 'suite': ['custom'], 'original_num_docs': 1319, 'effective_num_docs': 1319, 'must_remove_duplicate_docs': False, 'version': 0, 'frozen': False}}, 'summary_tasks': {'custom|gsm8k|8': {'hashes': {'hash_examples': '4c0843a5d99bcfdc', 'hash_full_prompts': 'bc3e523598c718cb', 'hash_input_tokens': '9b0240e60aaa0b21', 'hash_cont_tokens': 'e80f6dd144486402'}, 'truncated': 0, 'non_truncated': 1319, 'padded': 0, 'non_padded': 1319, 'effective_few_shots': 8.0, 'num_truncated_few_shots': 0}}, 'summary_general': {'hashes': {'hash_examples': '18b756b7813d1bdf', 'hash_full_prompts': 'ceeb6a330a123191', 'hash_input_tokens': '6949d8c1d2d40dcd', 'hash_cont_tokens': 'b205fd5e9d827de7'}, 'truncated': 0, 'non_truncated': 1319, 'padded': 0, 'non_padded': 1319, 'num_truncated_few_shots': 0}} {'accelerate': {'gpu_ids': [0], 'seed': 42}, 'dataset': {'answer_field': 'answer', 'load_dataset_args': ['main'], 'name': 'gsm8k', 'question_field': 'question', 'split': 'train', 'valid_split': 'test'}, 'evaluation': {'eval_batch_size': 1500, 'eval_percentage': 0.05, 'eval_subset': None}, 'experiment_name': 'gsm8k_experiment', 'internal_master_port': 25124, 'master_port': 25123, 'model': {'name': 'HuggingFaceTB/SmolLM2-1.7B-Instruct'}, 'output_dir': '/network/scratch/g/gagnonju/rejection_sampling_saves/gsm8k/2025-03-16_02-32-57/', 'training': {'batch_size': 12, 'learning_rate': 0.0001, 'max_length': 768, 'num_epochs': 30}, 'vllm': {'gpu_id': 1}, 'vllm_sampling': {'num_candidates': 48, 'temperature': 1, 'top_p': None}, 'wandb': {'entity': 'julesgm', 'log_interval': 10, 'project': 'expert_iteration'}}
```

Epoch 4

```

{'config_general': {'lighteval_sha': '?', 'num_fewshot_seeds': 1,
'override_batch_size': -1, 'max_samples': None, 'job_id': '', 'start_time':
14298475.568214037, 'end_time': 14298676.508912284,
'total_evaluation_time_seconds': '200.9406982474029', 'model_name': '_network_s
cratch_g_gagnonju_rejection_sampling_saves_gsm8k_2025-03-16_02-32-57_epoch_4',
'model_sha': '', 'model_dtype': None, 'model_size': None}, 'results':
{'custom|gsm8k|8': {'qem': 0.46626231993934797, 'qem_stderr':
0.01374109641222676}, 'all': {'qem': 0.46626231993934797, 'qem_stderr':
0.01374109641222676}}, 'versions': {'custom|gsm8k|8': 0}, 'config_tasks':
{'custom|gsm8k': {'name': 'gsm8k', 'prompt_function': 'jules_gsm8k', 'hf_repo':
'openai/gsm8k', 'hf_subset': 'main', 'metric': [{'metric_name': 'qem',
'higher_is_better': True, 'category': '3', 'use_case': '5', 'sample_level_fn':
'compute', 'corpus_level_fn': 'mean'}]}, 'hf_revision':
'e53f048856ff4f594e959d75785d2c2d37b678ee', 'hf_filter': None,
'hf_avail_splits': ['train', 'test'], 'trust_dataset': False,
'evaluation_splits': ['test'], 'few_shots_split': None, 'few_shots_select':
'random_sampling_from_train', 'generation_size': 256, 'generation_grammar':
None, 'stop_sequence': ['Question:', 'Question'], 'output_regex': None,
'num_samples': None, 'suite': ['custom'], 'original_num_docs': 1319,
'effective_num_docs': 1319, 'must_remove_duplicate_docs': False, 'version': 0,
'frozen': False}}, 'summary_tasks': {'custom|gsm8k|8': {'hashes':
{'hash_examples': '4c0843a5d99bcfdc', 'hash_full_prompts': 'bc3e523598c718cb',
'hash_input_tokens': '9b0240e60aaa0b21', 'hash_cont_tokens':
'8008ce1fa1f61391'}, 'truncated': 0, 'non_truncated': 1319, 'padded': 0,
'non_padded': 1319, 'effective_few_shots': 8.0, 'num_truncated_few_shots': 0}},
'summary_general': {'hashes': {'hash_examples': '18b756b7813d1bdf',
'hash_full_prompts': 'ceeb6a330a123191', 'hash_input_tokens':
'6949d8c1d2d40dcd', 'hash_cont_tokens': '68114122ae405869'}, 'truncated': 0,
'non_truncated': 1319, 'padded': 0, 'non_padded': 1319,
'num_truncated_few_shots': 0}}
{'accelerate': {'gpu_ids': [0], 'seed': 42}, 'dataset': {'answer_field':
'answer', 'load_dataset_args': ['main'], 'name': 'gsm8k', 'question_field':
'question', 'split': 'train', 'valid_split': 'test'}, 'evaluation':
{'eval_batch_size': 1500, 'eval_percentage': 0.05, 'eval_subset': None},
'experiment_name': 'gsm8k_experiment', 'internal_master_port': 25124,
'master_port': 25123, 'model': {'name': 'HuggingFaceTB/SmolLM2-1.7B-Instruct'},
'output_dir': '/network/scratch/g/gagnonju/rejection_sampling_saves/gsm8k/2025-
03-16_02-32-57/', 'training': {'batch_size': 12, 'learning_rate': 0.0001,
'max_length': 768, 'num_epochs': 30}, 'vllm': {'gpu_id': 1}, 'vllm_sampling':
{'num_candidates': 48, 'temperature': 1, 'top_p': None}, 'wandb': {'entity':
'julesgm', 'log_interval': 10, 'project': 'expert_iteration'}}
Epoch 1
{'config_general': {'lighteval_sha': '?', 'num_fewshot_seeds': 1,
'override_batch_size': -1, 'max_samples': None, 'job_id': '', 'start_time':
14298478.647958167, 'end_time': 14298680.889067454,
'total_evaluation_time_seconds': '202.24110928736627', 'model_name': '_network_
scratch_g_gagnonju_rejection_sampling_saves_gsm8k_2025-03-16_02-32-57_epoch_1',
'model_sha': '', 'model_dtype': None, 'model_size': None}, 'results':

```



```

{'custom|gsm8k|8': {'qem': 0.4746019711902957, 'qem_stderr':
0.013754705089112314}, 'all': {'qem': 0.4746019711902957, 'qem_stderr':
0.013754705089112314}}, 'versions': {'custom|gsm8k|8': 0}, 'config_tasks':
{'custom|gsm8k': {'name': 'gsm8k', 'prompt_function': 'jules_gsm8k', 'hf_repo':
'openai/gsm8k', 'hf_subset': 'main', 'metric': [{'metric_name': 'qem',
'higher_is_better': True, 'category': '3', 'use_case': '5', 'sample_level_fn':
'compute', 'corpus_level_fn': 'mean'}]}, 'hf_revision':
'e53f048856ff4f594e959d75785d2c2d37b678ee', 'hf_filter': None,
'hf_avail_splits': ['train', 'test'], 'trust_dataset': False,
'evaluation_splits': ['test'], 'few_shots_split': None, 'few_shots_select':
'random_sampling_from_train', 'generation_size': 256, 'generation_grammar':
None, 'stop_sequence': ['Question:', 'Question'], 'output_regex': None,
'num_samples': None, 'suite': ['custom'], 'original_num_docs': 1319,
'effective_num_docs': 1319, 'must_remove_duplicate_docs': False, 'version': 0,
'frozen': False}}, 'summary_tasks': {'custom|gsm8k|8': {'hashes':
{'hash_examples': '4c0843a5d99bcfdc', 'hash_full_prompts': 'bc3e523598c718cb',
'hash_input_tokens': '9b0240e60aaa0b21', 'hash_cont_tokens':
'9402dfbd35793ccb'}, 'truncated': 0, 'non_truncated': 1319, 'padded': 0,
'non_padded': 1319, 'effective_few_shots': 8.0, 'num_truncated_few_shots': 0}},
'summary_general': {'hashes': {'hash_examples': '18b756b7813d1bdf',
'hash_full_prompts': 'ceeb6a330a123191', 'hash_input_tokens':
'6949d8c1d2d40dcd', 'hash_cont_tokens': '5855d3d0ab6f4d81'}, 'truncated': 0,
'non_truncated': 1319, 'padded': 0, 'non_padded': 1319,
'num_truncated_few_shots': 0}}
{'accelerate': {'gpu_ids': [0], 'seed': 42}, 'dataset': {'answer_field':
'answer', 'load_dataset_args': ['main'], 'name': 'gsm8k', 'question_field':
'question', 'split': 'train', 'valid_split': 'test'}, 'evaluation':
{'eval_batch_size': 1500, 'eval_percentage': 0.05, 'eval_subset': None},
'experiment_name': 'gsm8k_experiment', 'internal_master_port': 25124,
'master_port': 25123, 'model': {'name': 'HuggingFaceTB/SmolLM2-1.7B-Instruct'},
'output_dir': '/network/scratch/g/gagnonju/rejection_sampling_saves/gsm8k/2025-
03-16_02-32-57/', 'training': {'batch_size': 12, 'learning_rate': 0.0001,
'max_length': 768, 'num_epochs': 30}, 'vllm': {'gpu_id': 1}, 'vllm_sampling':
{'num_candidates': 48, 'temperature': 1, 'top_p': None}, 'wandb': {'entity':
'julesgm', 'log_interval': 10, 'project': 'expert_iteration'}}
Epoch 0
{'config_general': {'lighteval_sha': '?', 'num_fewshot_seeds': 1,
'override_batch_size': -1, 'max_samples': None, 'job_id': '', 'start_time':
3307557.476750867, 'end_time': 3307760.964996366,
'total_evaluation_time_secondes': '203.48824549885467', 'model_name': '_network_
scratch_g_gagnonju_rejection_sampling_saves_gsm8k_2025-03-16_02-32-57_epoch_0',
'model_sha': '', 'model_dtype': None, 'model_size': None}, 'results':
{'custom|gsm8k|8': {'qem': 0.4746019711902957, 'qem_stderr':
0.013754705089112314}, 'all': {'qem': 0.4746019711902957, 'qem_stderr':
0.013754705089112314}}, 'versions': {'custom|gsm8k|8': 0}, 'config_tasks':
{'custom|gsm8k': {'name': 'gsm8k', 'prompt_function': 'jules_gsm8k', 'hf_repo':
'openai/gsm8k', 'hf_subset': 'main', 'metric': [{'metric_name': 'qem',
'higher_is_better': True, 'category': '3', 'use_case': '5', 'sample_level_fn':

```

```

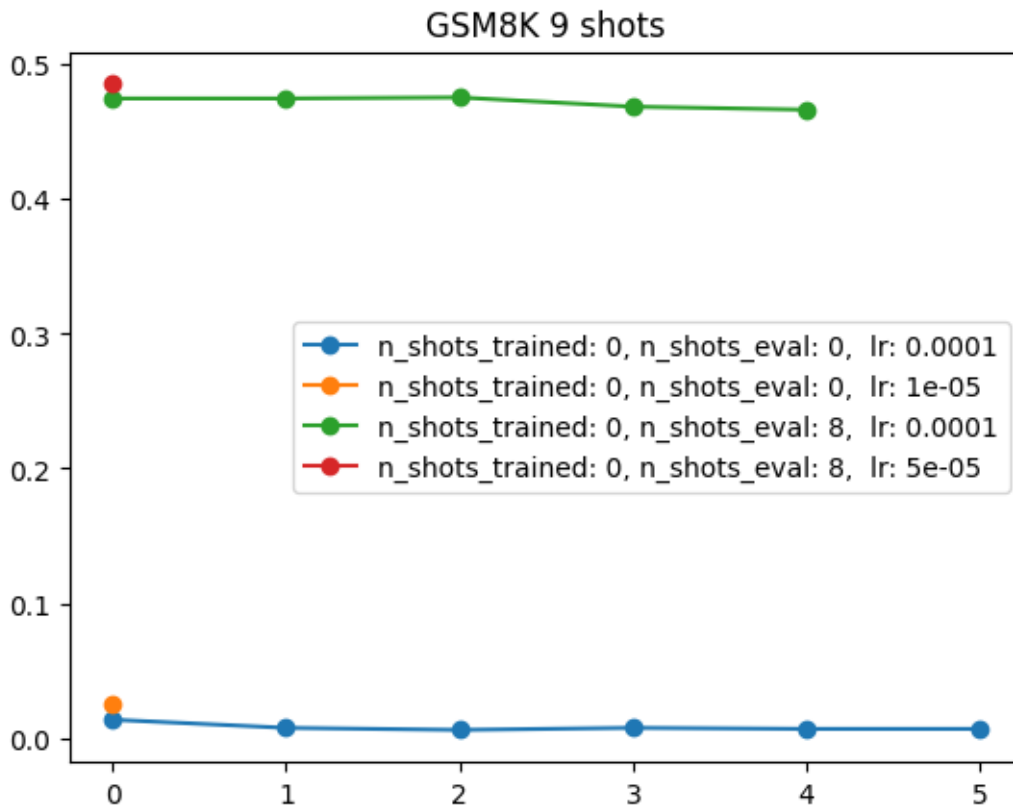
'compute', 'corpus_level_fn': 'mean']], 'hf_revision':
'e53f048856ff4f594e959d75785d2c2d37b678ee', 'hf_filter': None,
'hf_avail_splits': ['train', 'test'], 'trust_dataset': False,
'evaluation_splits': ['test'], 'few_shots_split': None, 'few_shots_select':
'random_sampling_from_train', 'generation_size': 256, 'generation_grammar':
None, 'stop_sequence': ['Question:', 'Question'], 'output_regex': None,
'num_samples': None, 'suite': ['custom'], 'original_num_docs': 1319,
'effective_num_docs': 1319, 'must_remove_duplicate_docs': False, 'version': 0,
'frozen': False}}, 'summary_tasks': {'custom|gsm8k|8': {'hashes':
{'hash_examples': '4c0843a5d99bcfdc', 'hash_full_prompts': 'bc3e523598c718cb',
'hash_input_tokens': '9b0240e60aaa0b21', 'hash_cont_tokens':
'b953241efa7ee8b0'}}, 'truncated': 0, 'non_truncated': 1319, 'padded': 0,
'non_padded': 1319, 'effective_few_shots': 8.0, 'num_truncated_few_shots': 0}},
'summary_general': {'hashes': {'hash_examples': '18b756b7813d1bdf',
'hash_full_prompts': 'ceeb6a330a123191', 'hash_input_tokens':
'6949d8c1d2d40dcd', 'hash_cont_tokens': '66467b97838c0153'}}, 'truncated': 0,
'non_truncated': 1319, 'padded': 0, 'non_padded': 1319,
'num_truncated_few_shots': 0}}
{'accelerate': {'gpu_ids': [0], 'seed': 42}, 'dataset': {'answer_field':
'answer', 'load_dataset_args': ['main'], 'name': 'gsm8k', 'question_field':
'question', 'split': 'train', 'valid_split': 'test'}, 'evaluation':
{'eval_batch_size': 1500, 'eval_percentage': 0.05, 'eval_subset': None},
'experiment_name': 'gsm8k_experiment', 'internal_master_port': 25124,
'master_port': 25123, 'model': {'name': 'HuggingFaceTB/SmolLM2-1.7B-Instruct'},
'output_dir': '/network/scratch/g/gagnonju/rejection_sampling_saves/gsm8k/2025-
03-16_02-32-57/', 'training': {'batch_size': 12, 'learning_rate': 0.0001,
'max_length': 768, 'num_epochs': 30}, 'vllm': {'gpu_id': 1}, 'vllm_sampling':
{'num_candidates': 48, 'temperature': 1, 'top_p': None}, 'wandb': {'entity':
'julesgm', 'log_interval': 10, 'project': 'expert_iteration'}}
Epoch 0
{'config_general': {'lighteval_sha': '?', 'num_fewshot_seeds': 1,
'override_batch_size': -1, 'max_samples': None, 'job_id': '', 'start_time':
14298496.037849374, 'end_time': 14298703.892703576,
'total_evaluation_time_seconds': '207.85485420189798', 'model_name': '_network_
scratch_g_gagnonju_rejection_sampling_saves_gsm8k_2025-03-16_19-48-52_epoch_0',
'model_sha': '', 'model_dtype': None, 'model_size': None}, 'results':
{'custom|gsm8k|8': {'qem': 0.4852160727824109, 'qem_stderr':
0.013766463050787598}, 'all': {'qem': 0.4852160727824109, 'qem_stderr':
0.013766463050787598}}, 'versions': {'custom|gsm8k|8': 0}, 'config_tasks':
{'custom|gsm8k': {'name': 'gsm8k', 'prompt_function': 'jules_gsm8k', 'hf_repo':
'openai/gsm8k', 'hf_subset': 'main', 'metric': [{'metric_name': 'qem',
'higher_is_better': True, 'category': '3', 'use_case': '5', 'sample_level_fn':
'compute', 'corpus_level_fn': 'mean'}]}, 'hf_revision':
'e53f048856ff4f594e959d75785d2c2d37b678ee', 'hf_filter': None,
'hf_avail_splits': ['train', 'test'], 'trust_dataset': False,
'evaluation_splits': ['test'], 'few_shots_split': None, 'few_shots_select':
'random_sampling_from_train', 'generation_size': 256, 'generation_grammar':
None, 'stop_sequence': ['Question:', 'Question'], 'output_regex': None,

```

```

'num_samples': None, 'suite': ['custom'], 'original_num_docs': 1319,
'effective_num_docs': 1319, 'must_remove_duplicate_docs': False, 'version': 0,
'frozen': False}}, 'summary_tasks': {'custom|gsm8k|8': {'hashes':
{'hash_examples': '4c0843a5d99bcfdc', 'hash_full_prompts': 'bc3e523598c718cb',
'hash_input_tokens': '9b0240e60aaa0b21', 'hash_cont_tokens':
'3c889bdd790b7e9f'}, 'truncated': 0, 'non_truncated': 1319, 'padded': 0,
'non_padded': 1319, 'effective_few_shots': 8.0, 'num_truncated_few_shots': 0}},
'summary_general': {'hashes': {'hash_examples': '18b756b7813d1bdf',
'hash_full_prompts': 'ceeb6a330a123191', 'hash_input_tokens':
'6949d8c1d2d40dcd', 'hash_cont_tokens': '639b431b83ea9278'}, 'truncated': 0,
'non_truncated': 1319, 'padded': 0, 'non_padded': 1319,
'num_truncated_few_shots': 0}}
{'accelerate': {'gpu_ids': [0], 'seed': 42}, 'dataset': {'answer_field':
'answer', 'load_dataset_args': ['main'], 'name': 'gsm8k', 'question_field':
'question', 'split': 'train', 'valid_split': 'test'}, 'evaluation':
{'eval_batch_size': 1500, 'eval_percentage': 0.05, 'eval_subset': None},
'experiment_name': 'gsm8k_experiment', 'internal_master_port': 25124,
'master_port': 25123, 'model': {'name': 'HuggingFaceTB/SmolLM2-1.7B-Instruct'},
'output_dir': '/network/scratch/g/gagnonju/rejection_sampling_saves/gsm8k/2025-
03-16_19-48-52/', 'training': {'batch_size': 12, 'learning_rate': 5e-05,
'max_length': 768, 'num_epochs': 30}, 'vllm': {'gpu_id': 1}, 'vllm_sampling':
{'num_candidates': 48, 'temperature': 1, 'top_p': None}, 'wandb': {'entity':
'julesgm', 'log_interval': 10, 'project': 'expert_iteration'}}

```



Epoch 0

```
{'config_general': {'lighteval_sha': '?', 'num_fewshot_seeds': 1,
'override_batch_size': -1, 'max_samples': None, 'job_id': '', 'start_time':
14296459.009183617, 'end_time': 14296572.86289152,
'total_evaluation_time_seconds': '113.85370790213346', 'model_name': '_network_
scratch_g_gagnonju_rejection_sampling_saves_hendrycks_math_2025-03-16_19-48-
52_epoch_0', 'model_sha': '', 'model_dtype': None, 'model_size': None},
'results': {'custom|math|0': {'qem': 0.0292, 'qem_stderr':
0.0023813034500509713}, 'all': {'qem': 0.0292, 'qem_stderr':
0.0023813034500509713}}, 'versions': {'custom|math|0': 0}, 'config_tasks':
{'custom|math': {'name': 'math', 'prompt_function': 'prompt_math', 'hf_repo':
'HuggingFaceTB/math_tasks', 'hf_subset': 'math', 'metric': [{'metric_name':
'qem', 'higher_is_better': True, 'category': '3', 'use_case': '5',
'sample_level_fn': 'compute', 'corpus_level_fn': 'mean'}]}, 'hf_revision':
'3d34f1076f279000b9315583dcdacfd288898283', 'hf_filter': None,
'hf_avail_splits': ['train', 'test', 'demo'], 'trust_dataset': True,
'evaluation_splits': ['test'], 'few_shots_split': 'demo', 'few_shots_select':
'sequential', 'generation_size': 1024, 'generation_grammar': None,
'stop_sequence': ['\n\n'], 'output_regex': None, 'num_samples': None, 'suite':
['custom'], 'original_num_docs': 5000, 'effective_num_docs': 5000,
'must_remove_duplicate_docs': False, 'version': 0, 'frozen': False}},
'summary_tasks': {'custom|math|0': {'hashes': {'hash_examples':
'8db5c745df18b44d', 'hash_full_prompts': '9aa85ba2b28f0792',
'hash_input_tokens': '32b26e8e3b25e439', 'hash_cont_tokens':
'8647b1b1b6b1303d'}, 'truncated': 0, 'non_truncated': 5000, 'padded': 0,
'non_padded': 5000, 'effective_few_shots': 0.0, 'num_truncated_few_shots': 0}},
'summary_general': {'hashes': {'hash_examples': '687658c8f867347e',
'hash_full_prompts': '6c307d3e23a42b74', 'hash_input_tokens':
'83d1ca64363bd827', 'hash_cont_tokens': '3ed527cffb9deb5d'}, 'truncated': 0,
'non_truncated': 5000, 'padded': 0, 'non_padded': 5000,
'num_truncated_few_shots': 0}}
{'accelerate': {'gpu_ids': [0], 'seed': 42}, 'dataset': {'answer_field':
'solution', 'load_dataset_args': [], 'name': 'hendrycks/competition_math',
'question_field': 'problem', 'split': 'train', 'valid_split': 'test'},
'evaluation': {'eval_batch_size': 1500, 'eval_percentage': 0.05, 'eval_subset':
None}, 'experiment_name': 'hendrycks_math_experiment', 'internal_master_port':
25124, 'master_port': 25123, 'model': {'name':
'HuggingFaceTB/SmolLM2-1.7B-Instruct'}, 'output_dir': '/network/scratch/g/gagnon
ju/rejection_sampling_saves/hendrycks_math/2025-03-16_19-48-52/', 'training':
{'batch_size': 12, 'learning_rate': 1e-05, 'max_length': 768, 'num_epochs': 30},
'vllm': {'gpu_id': 1}, 'vllm_sampling': {'num_candidates': 48, 'temperature': 1,
'top_p': None}, 'wandb': {'entity': 'julesgm', 'log_interval': 10, 'project':
'expert_iteration'}}
```

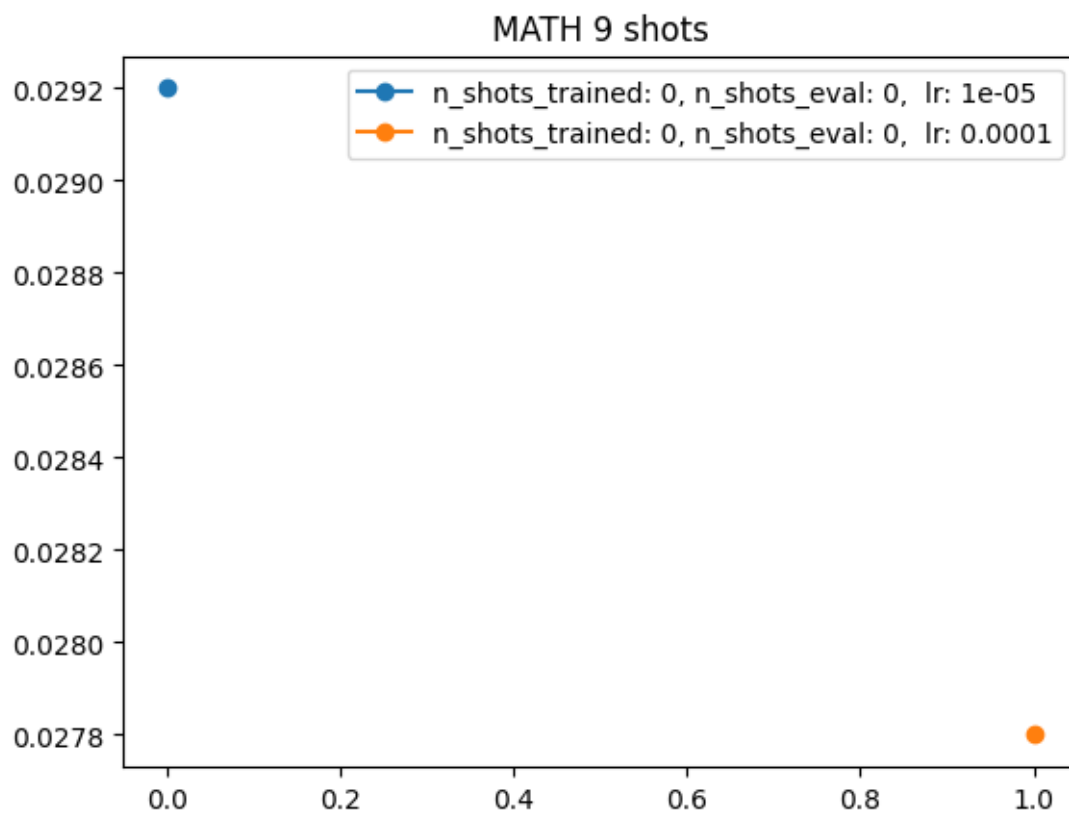
Epoch 1

```
{'config_general': {'lighteval_sha': '?', 'num_fewshot_seeds': 1,
```

```

'override_batch_size': -1, 'max_samples': None, 'job_id': '', 'start_time':
15917084.092109412, 'end_time': 15917194.611858476,
'total_evaluation_time_seconds': '110.51974906399846', 'model_name': '_network_
scratch_g_gagnonju_rejection_sampling_saves_hendrycks_math_2025-03-16_11-36-
38_epoch_1', 'model_sha': '', 'model_dtype': None, 'model_size': None},
'results': {'custom|math|0': {'qem': 0.0278, 'qem_stderr': 0.00232519102498275},
'all': {'qem': 0.0278, 'qem_stderr': 0.00232519102498275}}, 'versions':
{'custom|math|0': 0}, 'config_tasks': {'custom|math': {'name': 'math',
'prompt_function': 'prompt_math', 'hf_repo': 'HuggingFaceTB/math_tasks',
'hf_subset': 'math', 'metric': [{'metric_name': 'qem', 'higher_is_better': True,
'category': '3', 'use_case': '5', 'sample_level_fn': 'compute',
'corpus_level_fn': 'mean'}]}, 'hf_revision':
'3d34f1076f279000b9315583dcdacfd288898283', 'hf_filter': None,
'hf_avail_splits': ['train', 'test', 'demo'], 'trust_dataset': True,
'evaluation_splits': ['test'], 'few_shots_split': 'demo', 'few_shots_select':
'sequential', 'generation_size': 1024, 'generation_grammar': None,
'stop_sequence': ['\\n\\n'], 'output_regex': None, 'num_samples': None, 'suite':
['custom'], 'original_num_docs': 5000, 'effective_num_docs': 5000,
'must_remove_duplicate_docs': False, 'version': 0, 'frozen': False}},
'summary_tasks': {'custom|math|0': {'hashes': {'hash_examples':
'8db5c745df18b44d', 'hash_full_prompts': '9aa85ba2b28f0792',
'hash_input_tokens': '32b26e8e3b25e439', 'hash_cont_tokens':
'9ad273b14e47cfb9'}}, 'truncated': 0, 'non_truncated': 5000, 'padded': 0,
'non_padded': 5000, 'effective_few_shots': 0.0, 'num_truncated_few_shots': 0}},
'summary_general': {'hashes': {'hash_examples': '687658c8f867347e',
'hash_full_prompts': '6c307d3e23a42b74', 'hash_input_tokens':
'83d1ca64363bd827', 'hash_cont_tokens': '7bd667c5c734f5fe'}}, 'truncated': 0,
'non_truncated': 5000, 'padded': 0, 'non_padded': 5000,
'num_truncated_few_shots': 0}}
{'accelerate': {'gpu_ids': [0], 'seed': 42}, 'dataset': {'answer_field':
'solution', 'load_dataset_args': [], 'name': 'hendrycks/competition_math',
'question_field': 'problem', 'split': 'train', 'valid_split': 'test'},
'evaluation': {'eval_batch_size': 1500, 'eval_percentage': 0.05, 'eval_subset':
None}, 'experiment_name': 'hendrycks_math_experiment', 'internal_master_port':
25124, 'master_port': 25123, 'model': {'name':
'HuggingFaceTB/SmolLM2-1.7B-Instruct'}, 'output_dir': '/network/scratch/g/gagnon
ju/rejection_sampling_saves/hendrycks_math/2025-03-16_11-36-38/', 'training':
{'batch_size': 12, 'learning_rate': 0.0001, 'max_length': 768, 'num_epochs':
30}, 'vllm': {'gpu_id': 1}, 'vllm_sampling': {'num_candidates': 48,
'temperature': 1, 'top_p': None}, 'wandb': {'entity': 'julesgm', 'log_interval':
10, 'project': 'expert_iteration'}}

```



[]: