

Weights and Biases Logger

This notebook logs the evaluation metrics of dev set and test set on every model.

```
In [ ]: import pandas as pd
import wandb
from configparser import ConfigParser
```

.cfg file is a config file that store your personal api key The format inside the file looks like this:

```
[<YOUR_API_KEY_NAME>]
auth_key: <HERE_IS_YOUR_API_KEY>
```

.gitignore is set on .cfg

```
In [7]: parser = ConfigParser()
_ = parser.read("../notebook.cfg")
wandb_api_auth_key = parser.get("wandb_api_key", "auth_key")
```

```
In [8]: wandb.login(key=wandb_api_auth_key)
```

```
wandb: Currently logged in as: haydenchiush. Use `wandb login --relogin` to
force relogin
wandb: WARNING If you're specifying your api key in code, ensure this code i
s not shared publicly.
wandb: WARNING Consider setting the WANDB_API_KEY environment variable, or r
unning `wandb login` from the command line.
wandb: Appending key for api.wandb.ai to your netrc file: /Users/haydenchi
u/.netrc
```

```
Out[8]: True
```

Validation set

```
In [18]: PROJECT_NAME = 'BioLaySumm2024'# wandb project name
ENTITY = 'bossy_beaver' #change to your wandb team name
PATH_TO_VAL_METRICS = './Proxy_val_Results'
VAL_METRICS_FILES = ['elife_scores.txt', 'plos_scores.txt', 'scores.txt']
```

```
In [26]: # For example we can use these parameters to define our metrics
# file = 'elife_scores.txt'
# data_src = 'eLife_dev'
# sample_portion = 0.1
# name = 'dummy_baseline'
# tags = ['baseline']
```

```
In [20]: # Initialize WandB
def wandb_log_eval_metrics(file, data_src, sample_portion, name, tags, job_type):
    run_config = {'data_src':data_src, 'sample_portion':sample_portion}
    run = wandb.init(project=PROJECT_NAME, entity=ENTITY,
                     job_type=job_type, tags=tags, name=name)

    # Open the file in read mode
    with open(PATH_TO_VAL_METRICS + "/" + file, 'r') as file:

        lines = file.readlines()

    # Create an empty dictionary to store the metrics
    metrics = {}

    # Iterate over each line
    for line in lines:
        # Split the line into key and value using ':' as the delimiter
        key, value = line.strip().split(': ')
        # Store the key-value pair in the dictionary
        metrics[key] = float(value)
    print(metrics)

    # Log metrics to WandB
    run.log(metrics)

    # Finish WandB run
    run.finish()
```

```
In [22]: # eLife dummy baseline on 10% dev set
wandb_log_eval_metrics(file='elife_scores.txt', data_src='eLife_dev', sample_portion=0.1,
                       name='elife_dummy_baseline', tags=['baseline'])
```

wandb: Currently logged in as: **haydenchiush** (**bossy_beaver**). Use ``wandb login --relogin`` to force relogin

wandb version 0.16.4 is available! To upgrade, please run: `$ pip install wandb --upgrade`

Tracking run with wandb version 0.16.3

Run data is saved locally in `/Users/haydenchiu/MDS-`

`CL/BioLaySumm2024/COLX_531_bossy_beaver/milestone2/wandb/run-`

`20240316_180321-kqrw4l9l`

Syncing run **elife_dummy_baseline** to [Weights & Biases](https://wandb.ai/bossy_beaver/BioLaySumm2024) ([docs](https://wandb.ai/bossy_beaver/BioLaySumm2024))

View project at https://wandb.ai/bossy_beaver/BioLaySumm2024

View run at https://wandb.ai/bossy_beaver/BioLaySumm2024/runs/kqrw4l9l

```
{'ROUGE1': 0.32312883506775875, 'ROUGE2': 0.065595999095099, 'ROUGEL': 0.300
8124655909614, 'BERTScore': 0.8375861893097559, 'FKGL': 15.133333333333335,
'DCRS': 11.590416666666664, 'CLI': 16.845416666666665, 'LENS': 39.6381475972
8811, 'AlignScore': 0.9943027173479398, 'SummaC': 0.9525209615627924}
VBox(children=(Label(value='0.001 MB of 0.001 MB uploaded\r'), FloatProgress
(value=1.0, max=1.0)))
```

Run history:

AlignScore	—
BERTScore	—
CLI	—
DCRS	—
FKGL	—
LENS	—
ROUGE1	—
ROUGE2	—
ROUGEL	—
SummaC	—

Run summary:

AlignScore	0.9943
BERTScore	0.83759
CLI	16.84542
DCRS	11.59042
FKGL	15.13333
LENS	39.63815
ROUGE1	0.32313
ROUGE2	0.0656
ROUGEL	0.30081
SummaC	0.95252

View run [elife_dummy_baseline](#) at:

https://wandb.ai/bossy_beaver/BioLaySumm2024/runs/kqrw4l9l

View job at

https://wandb.ai/bossy_beaver/BioLaySumm2024/jobs/QXJ0aWZhY3RDb2xsZWNOaW9uOjE0C

Synced 6 W&B file(s), 0 media file(s), 0 artifact file(s) and 0 other file(s)

Find logs at: `./wandb/run-20240316_180321-kqrw4l9l/logs`

```
In [23]: # PLOS dummy baseline on 10% dev set
wandb_log_eval_metrics(file='plos_scores.txt', data_src='PLOS_dev', sample_p
name='plos_dummy_baseline', tags=['baseline'])
```

wandb version 0.16.4 is available! To upgrade, please run: `$ pip install wandb --upgrade`

Tracking run with wandb version 0.16.3

Run data is saved locally in `/Users/haydenchiu/MDS-`

`CL/BioLaySumm2024/C0LX_531_bossy_beaver/milestone2/wandb/run-`
`20240316_180339-559am2mj`

Syncing run [plos_dummy_baseline](#) to [Weights & Biases \(docs\)](#)

View project at https://wandb.ai/bossy_beaver/BioLaySumm2024

View run at https://wandb.ai/bossy_beaver/BioLaySumm2024/runs/559am2mj

```
{'ROUGE1': 0.5125303136128874, 'ROUGE2': 0.21055582506879417, 'ROUGEL': 0.47
67545280909615, 'BERTScore': 0.8725961204888164, 'FKGL': 14.94927536231884,
'DCRS': 10.976449275362318, 'CLI': 16.293840579710146, 'LENS': 38.6190405789
40265, 'AlignScore': 0.9880770807680876, 'SummaC': 0.9578714357770007}
VBox(children=(Label(value='0.001 MB of 0.001 MB uploaded\r'), FloatProgress
(value=1.0, max=1.0)))
```

Run history:

AlignScore	—
BERTScore	—
CLI	—
DCRS	—
FKGL	—
LENS	—
ROUGE1	—
ROUGE2	—
ROUGEL	—
SummaC	—

Run summary:

AlignScore	0.98808
BERTScore	0.8726
CLI	16.29384
DCRS	10.97645
FKGL	14.94928
LENS	38.61904
ROUGE1	0.51253
ROUGE2	0.21056
ROUGEL	0.47675
SummaC	0.95787

View run **plos_dummy_baseline** at:

https://wandb.ai/bossy_beaver/BioLaySumm2024/runs/559am2mj

View job at

https://wandb.ai/bossy_beaver/BioLaySumm2024/jobs/QXJ0aWZhY3RDb2xsZWNOaW9uOjE0O

Synced 6 W&B file(s), 0 media file(s), 0 artifact file(s) and 0 other file(s)

Find logs at: `./wandb/run-20240316_180339-559am2mj/logs`

```
In [24]: # Combined dummy baseline on 10% dev set
wandb_log_eval_metrics(file='scores.txt', data_src='combined_dev', sample_pc
                    name='dummy_baseline', tags=['baseline', 'combined'])
```

wandb version 0.16.4 is available! To upgrade, please run: `$ pip install wandb --upgrade`

Tracking run with wandb version 0.16.3

Run data is saved locally in `/Users/haydenchiu/MDS-`

`CL/BioLaySumm2024/COLX_531_bossy_beaver/milestone2/wandb/run-`
`20240316_180533-wyn0doir`

Syncing run **dummy_baseline** to [Weights & Biases \(docs\)](#)

View project at https://wandb.ai/bossy_beaver/BioLaySumm2024

View run at https://wandb.ai/bossy_beaver/BioLaySumm2024/runs/wyn0doir

```
{'ROUGE1': 0.4178295743403231, 'ROUGE2': 0.13807591208194658, 'ROUGEL': 0.38
87834968409615, 'BERTScore': 0.8550911548992861, 'FKGL': 15.041304347826088,
'DCRS': 11.283432971014491, 'CLI': 16.569628623188407, 'LENS': 39.1285940881
14184, 'AlignScore': 0.9911898990580137, 'SummaC': 0.9551961986698966}
VBox(children=(Label(value='0.001 MB of 0.001 MB uploaded\r'), FloatProgress
(value=1.0, max=1.0)))
```

Run history:

AlignScore	—
BERTScore	—
CLI	—
DCRS	—
FKGL	—
LENS	—
ROUGE1	—
ROUGE2	—
ROUGEL	—
SummaC	—

Run summary:

AlignScore	0.99119
BERTScore	0.85509
CLI	16.56963
DCRS	11.28343
FKGL	15.0413
LENS	39.12859
ROUGE1	0.41783
ROUGE2	0.13808
ROUGEL	0.38878
SummaC	0.9552

View run **dummy_baseline** at:

https://wandb.ai/bossy_beaver/BioLaySumm2024/runs/wyn0doir

View job at

https://wandb.ai/bossy_beaver/BioLaySumm2024/jobs/QXJ0aWZhY3RDb2xsZWNOaW9uOjE0C

Synced 6 W&B file(s), 0 media file(s), 0 artifact file(s) and 0 other file(s)

Find logs at: `./wandb/run-20240316_180533-wyn0doir/logs`

Test Set

```
In [25]: PATH_TO_VAL_METRICS = './scoring_result'
TEST_METRICS_FILES = ['elife_scores.txt', 'plos_scores.txt', 'scores.txt']
```

```
In [27]: # elife dummy baseline on test set
wandb_log_eval_metrics(file='elife_scores.txt', data_src='eLife_test', sample_rate=1,
                        name='elife_dummy_baseline_test', tags=['baseline', 'elife'])
```

wandb version 0.16.4 is available! To upgrade, please run: `$ pip install wandb --upgrade`

Tracking run with wandb version 0.16.3

Run data is saved locally in `/Users/haydenchiu/MDS-`

`CL/BioLaySumm2024/COLX_531_bossy_beaver/milestone2/wandb/run-20240316_181600-a4kjesen`

Syncing run **elife_dummy_baseline_test** to [Weights & Biases \(docs\)](#)

View project at https://wandb.ai/bossy_beaver/BioLaySumm2024

View run at https://wandb.ai/bossy_beaver/BioLaySumm2024/runs/a4kjesen

```
{'ROUGE1': 0.3089138993381871, 'ROUGE2': 0.06957613212293816, 'ROUGEL': 0.2813751343330106, 'BERTScore': 0.8407322109585077, 'FKGL': 15.080985915492958, 'DCRS': 11.753802816901409, 'CLI': 17.258943661971827, 'LENS': 30.24145106148018, 'AlignScore': 0.9781640558175637, 'SummaC': 0.9402011343291108}
VBox(children=(Label(value='0.001 MB of 0.001 MB uploaded\r'), FloatProgress(value=1.0, max=1.0)))
```

Run history:

AlignScore	—
BERTScore	—
CLI	—
DCRS	—
FKGL	—
LENS	—
ROUGE1	—
ROUGE2	—
ROUGEL	—
SummaC	—

Run summary:

AlignScore	0.97816
BERTScore	0.84073
CLI	17.25894
DCRS	11.7538
FKGL	15.08099
LENS	30.24145
ROUGE1	0.30891
ROUGE2	0.06958
ROUGEL	0.28138
SummaC	0.9402

View run **elife_dummy_baseline_test** at:

https://wandb.ai/bossy_beaver/BioLaySumm2024/runs/a4kjesen

View job at

https://wandb.ai/bossy_beaver/BioLaySumm2024/jobs/QXJ0aWZhY3RDb2xsZWNOaW9uOjE0O

Synced 6 W&B file(s), 0 media file(s), 0 artifact file(s) and 0 other file(s)

Find logs at: `./wandb/run-20240316_181600-a4kjesen/logs`

```
In [28]: # PLOS dummy baseline on test set
wandb_log_eval_metrics(file='plos_scores.txt', data_src='PLOS_test', sample_
                    name='plos_dummy_baseline_test', tags=['baseline', 't
```

wandb version 0.16.4 is available! To upgrade, please run: `$ pip install wandb --upgrade`

Tracking run with wandb version 0.16.3

Run data is saved locally in `/Users/haydenchiu/MDS-`

`CL/BioLaySumm2024/COLX_531_bossy_beaver/milestone2/wandb/run-`
`20240316_181649-bdfffc27`

Syncing run **plos_dummy_baseline_test** to [Weights & Biases \(docs\)](#)

View project at https://wandb.ai/bossy_beaver/BioLaySumm2024

View run at https://wandb.ai/bossy_beaver/BioLaySumm2024/runs/bdfffc27

```
{'ROUGE1': 0.461164102179743, 'ROUGE2': 0.16057957580294918, 'ROUGEL': 0.4178593622680423, 'BERTScore': 0.8693121696861696, 'FKGL': 14.751408450704227, 'DCRS': 11.793098591549297, 'CLI': 16.704154929577463, 'LENS': 35.30267832931986, 'AlignScore': 0.9772088964220503, 'SummaC': 0.9488216630170043}
VBox(children=(Label(value='0.001 MB of 0.001 MB uploaded\r'), FloatProgress(value=1.0, max=1.0)))
```

Run history:

AlignScore	—
BERTScore	—
CLI	—
DCRS	—
FKGL	—
LENS	—
ROUGE1	—
ROUGE2	—
ROUGEL	—
SummaC	—

Run summary:

AlignScore	0.97721
BERTScore	0.86931
CLI	16.70415
DCRS	11.7931
FKGL	14.75141
LENS	35.30268
ROUGE1	0.46116
ROUGE2	0.16058
ROUGEL	0.41786
SummaC	0.94882

View run **plos_dummy_baseline_test** at:

https://wandb.ai/bossy_beaver/BioLaySumm2024/runs/bdffic27

View job at

https://wandb.ai/bossy_beaver/BioLaySumm2024/jobs/QXJ0aWZhY3RDb2xsZWNOaW9uOjE0O

Synced 6 W&B file(s), 0 media file(s), 0 artifact file(s) and 0 other file(s)

Find logs at: `./wandb/run-20240316_181649-bdffic27/logs`

```
In [29]: # Combined dummy baseline on test set
wandb_log_eval_metrics(file='scores.txt', data_src='combined_test', sample_p
name='dummy_baseline_test', tags=['baseline', 'combin
```

wandb version 0.16.4 is available! To upgrade, please run: `$ pip install wandb --upgrade`

Tracking run with wandb version 0.16.3

Run data is saved locally in `/Users/haydenchiu/MDS-`

`CL/BioLaySumm2024/COLX_531_bossy_beaver/milestone2/wandb/run-`
`20240316_181727-shde10wi`

Syncing run **dummy_baseline_test** to [Weights & Biases \(docs\)](#)

View project at https://wandb.ai/bossy_beaver/BioLaySumm2024

View run at https://wandb.ai/bossy_beaver/BioLaySumm2024/runs/shde10wi

```
{'ROUGE1': 0.385039000758965, 'ROUGE2': 0.11507785396294368, 'ROUGEL': 0.34961724830052643, 'BERTScore': 0.8550221903223387, 'FKGL': 14.916197183098593, 'DCRS': 11.773450704225354, 'CLI': 16.981549295774645, 'LENS': 32.77206469540002, 'AlignScore': 0.977686476119807, 'SummaC': 0.9445113986730576}
VBox(children=(Label(value='0.001 MB of 0.001 MB uploaded\r'), FloatProgress(value=1.0, max=1.0)))
```

Run history:

AlignScore	—
BERTScore	—
CLI	—
DCRS	—
FKGL	—
LENS	—
ROUGE1	—
ROUGE2	—
ROUGEL	—
SummaC	—

Run summary:

AlignScore	0.97769
BERTScore	0.85502
CLI	16.98155
DCRS	11.77345
FKGL	14.9162
LENS	32.77206
ROUGE1	0.38504
ROUGE2	0.11508
ROUGEL	0.34962
SummaC	0.94451

View run **dummy_baseline_test** at:
https://wandb.ai/bossy_beaver/BioLaySumm2024/runs/shde10wi
View job at
https://wandb.ai/bossy_beaver/BioLaySumm2024/jobs/QXJ0aWZhY3RDb2xsZWNOaW9uOjE0C
Synced 6 W&B file(s), 0 media file(s), 0 artifact file(s) and 0 other file(s)
Find logs at: `./wandb/run-20240316_181727-shde10wi/logs`

```
In [ ]:
```