

Nama : Agis Satria Mandala
NIM : 2000018075

Set Hyperparameter menjadi jumlah head=16, dan jumlah layer-10

The screenshot shows a Jupyter Notebook titled "2000018075_Agis Satria Mandala_Postest 8.ipynb". The left sidebar contains a "Table of contents" with sections 5.4 through 11. The main area displays the following content:

10.1 Set hyperparameters

Mengeset Hyperparameter yang sesuai

Agar modelnya tetap kecil dan cukup cepat trainingnya, pada praktikum ini *num_layers*, *d_model*, and *dff* sudah diperkecil (dibandingkan aslinya Transformer).

Aslinya; *num_layers*=6, *d_model* = 512, *dff* = 2048. Lihat [paper](#) .

```
[43] num_layers = 10
      d_model = 128
      dff = 512
      num_heads = 16

      input_vocab_size = tokenizer_pt.vocab_size + 2
      target_vocab_size = tokenizer_en.vocab_size + 2
      dropout_rate = 0.1
```

10.2 Optimizer

Gunakan Adam optimizer untuk melakukan optimasi

$$lr_{rate} = d_{model}^{-0.5} * \min(step_num^{-0.5}, step_num * warmup_steps^{-1.5})$$

Automatic saving failed. This file was updated remotely or in another tab. [Show diff](#)

5m 28s completed at 06:24

1. Epoch 5

The screenshot shows the same Jupyter Notebook interface, now displaying the training progress for Epoch 1 and Epoch 2. The left sidebar is the same. The main area displays the following content:

```
Epoch 1 Batch 0 Loss 4.4968 Accuracy 0.1631
Epoch 1 Batch 50 Loss 4.4973 Accuracy 0.1474
Epoch 1 Batch 100 Loss 4.5017 Accuracy 0.1478
Epoch 1 Batch 150 Loss 4.5046 Accuracy 0.1486
Epoch 1 Batch 200 Loss 4.5044 Accuracy 0.1476
Epoch 1 Batch 250 Loss 4.5054 Accuracy 0.1477
Epoch 1 Batch 300 Loss 4.5048 Accuracy 0.1475
Epoch 1 Batch 350 Loss 4.5105 Accuracy 0.1476
Epoch 1 Batch 400 Loss 4.5098 Accuracy 0.1475
Epoch 1 Batch 450 Loss 4.5087 Accuracy 0.1478
Epoch 1 Batch 500 Loss 4.5077 Accuracy 0.1479
Epoch 1 Batch 550 Loss 4.5069 Accuracy 0.1478
Epoch 1 Batch 600 Loss 4.5067 Accuracy 0.1479
Epoch 1 Batch 650 Loss 4.5074 Accuracy 0.1480
Epoch 1 Batch 700 Loss 4.5077 Accuracy 0.1480
Epoch 1 Loss 4.5079 Accuracy 0.1481
Time taken for 1 epoch: 313.19347858428955 secs

Epoch 2 Batch 0 Loss 4.4380 Accuracy 0.1515
Epoch 2 Batch 50 Loss 4.5113 Accuracy 0.1471
Epoch 2 Batch 100 Loss 4.5083 Accuracy 0.1464
Epoch 2 Batch 150 Loss 4.4992 Accuracy 0.1466
Epoch 2 Batch 200 Loss 4.4921 Accuracy 0.1470
Epoch 2 Batch 250 Loss 4.4924 Accuracy 0.1475
Epoch 2 Batch 300 Loss 4.4928 Accuracy 0.1479
Epoch 2 Batch 350 Loss 4.4952 Accuracy 0.1478
Epoch 2 Batch 400 Loss 4.4924 Accuracy 0.1479
Epoch 2 Batch 450 Loss 4.4906 Accuracy 0.1482
Epoch 2 Batch 500 Loss 4.4908 Accuracy 0.1483
Epoch 2 Batch 550 Loss 4.4920 Accuracy 0.1485
Epoch 2 Batch 600 Loss 4.4933 Accuracy 0.1485
Epoch 2 Batch 650 Loss 4.4927 Accuracy 0.1487
Epoch 2 Batch 700 Loss 4.4915 Accuracy 0.1489
Epoch 2 Loss 4.4916 Accuracy 0.1489
Time taken for 1 epoch: 112.62078523635864 secs
```

7s completed at 07:41

2000018075_Agis Satria Mandala_Postest 8.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Table of contents

5.4. Melakukan Filtering

5.5. Positional Encoding

5.6. Masking

6. Scaled Dot Product Attention

7. Multi Head Attention

8. Point wise feed forward network

9. Membangun Encoder dan Decoder

9.1 Encoder Layer

9.2 Decoder Layer

9.3 Encoder

9.4 Decoder

10. Membangun Model Transformer

10.1 Set hyperparameters

10.2 Optimizer

10.3 Loss and metrics

10.4 Training Model Transformer

11. Evaluasi Model

Summary

POST TEST:

Section

+ Code + Text

Epoch 3 Batch 0 Loss 4.6866 Accuracy 0.1426

Epoch 3 Batch 50 Loss 4.4662 Accuracy 0.1495

Epoch 3 Batch 100 Loss 4.4785 Accuracy 0.1486

Epoch 3 Batch 150 Loss 4.4788 Accuracy 0.1482

Epoch 3 Batch 200 Loss 4.4755 Accuracy 0.1486

Epoch 3 Batch 250 Loss 4.4784 Accuracy 0.1489

Epoch 3 Batch 300 Loss 4.4813 Accuracy 0.1490

Epoch 3 Batch 350 Loss 4.4832 Accuracy 0.1488

Epoch 3 Batch 400 Loss 4.4809 Accuracy 0.1488

Epoch 3 Batch 450 Loss 4.4833 Accuracy 0.1489

Epoch 3 Batch 500 Loss 4.4828 Accuracy 0.1491

Epoch 3 Batch 550 Loss 4.4794 Accuracy 0.1491

Epoch 3 Batch 600 Loss 4.4776 Accuracy 0.1492

Epoch 3 Batch 650 Loss 4.4785 Accuracy 0.1492

Epoch 3 Batch 700 Loss 4.4767 Accuracy 0.1492

Epoch 3 Loss 4.4768 Accuracy 0.1492

Time taken for 1 epoch: 108.44540190696716 secs

Epoch 4 Batch 0 Loss 4.4528 Accuracy 0.1545

Epoch 4 Batch 50 Loss 4.4457 Accuracy 0.1500

Epoch 4 Batch 100 Loss 4.4517 Accuracy 0.1497

Epoch 4 Batch 150 Loss 4.4578 Accuracy 0.1496

Epoch 4 Batch 200 Loss 4.4604 Accuracy 0.1493

Epoch 4 Batch 250 Loss 4.4577 Accuracy 0.1496

Epoch 4 Batch 300 Loss 4.4579 Accuracy 0.1497

Epoch 4 Batch 350 Loss 4.4585 Accuracy 0.1494

Epoch 4 Batch 400 Loss 4.4587 Accuracy 0.1495

Epoch 4 Batch 450 Loss 4.4574 Accuracy 0.1500

Epoch 4 Batch 500 Loss 4.4596 Accuracy 0.1500

Epoch 4 Batch 550 Loss 4.4594 Accuracy 0.1501

Epoch 4 Batch 600 Loss 4.4594 Accuracy 0.1500

Epoch 4 Batch 650 Loss 4.4604 Accuracy 0.1500

Epoch 4 Batch 700 Loss 4.4612 Accuracy 0.1500

Epoch 4 Loss 4.4612 Accuracy 0.1501

Time taken for 1 epoch: 107.17480564117432 secs

7s completed at 07:41

2000018075_Agis Satria Mandala_Postest 8.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Table of contents

5.4. Melakukan Filtering

5.5. Positional Encoding

5.6. Masking

6. Scaled Dot Product Attention

7. Multi Head Attention

8. Point wise feed forward network

9. Membangun Encoder dan Decoder

9.1 Encoder Layer

9.2 Decoder Layer

9.3 Encoder

9.4 Decoder

10. Membangun Model Transformer

10.1 Set hyperparameters

10.2 Optimizer

10.3 Loss and metrics

10.4 Training Model Transformer

11. Evaluasi Model

Summary

POST TEST:

Section

+ Code + Text

Epoch 4 Batch 350 Loss 4.4585 Accuracy 0.1494

Epoch 4 Batch 400 Loss 4.4587 Accuracy 0.1495

Epoch 4 Batch 450 Loss 4.4574 Accuracy 0.1500

Epoch 4 Batch 500 Loss 4.4596 Accuracy 0.1500

Epoch 4 Batch 550 Loss 4.4594 Accuracy 0.1501

Epoch 4 Batch 600 Loss 4.4594 Accuracy 0.1500

Epoch 4 Batch 650 Loss 4.4604 Accuracy 0.1500

Epoch 4 Batch 700 Loss 4.4612 Accuracy 0.1500

Epoch 4 Loss 4.4612 Accuracy 0.1501

Time taken for 1 epoch: 107.17480564117432 secs

Epoch 5 Batch 0 Loss 4.4365 Accuracy 0.1330

Epoch 5 Batch 50 Loss 4.4184 Accuracy 0.1527

Epoch 5 Batch 100 Loss 4.4301 Accuracy 0.1509

Epoch 5 Batch 150 Loss 4.4301 Accuracy 0.1514

Epoch 5 Batch 200 Loss 4.4389 Accuracy 0.1512

Epoch 5 Batch 250 Loss 4.4413 Accuracy 0.1508

Epoch 5 Batch 300 Loss 4.4427 Accuracy 0.1510

Epoch 5 Batch 350 Loss 4.4426 Accuracy 0.1509

Epoch 5 Batch 400 Loss 4.4429 Accuracy 0.1508

Epoch 5 Batch 450 Loss 4.4438 Accuracy 0.1508

Epoch 5 Batch 500 Loss 4.4444 Accuracy 0.1507

Epoch 5 Batch 550 Loss 4.4445 Accuracy 0.1506

Epoch 5 Batch 600 Loss 4.4472 Accuracy 0.1506

Epoch 5 Batch 650 Loss 4.4469 Accuracy 0.1506

Epoch 5 Batch 700 Loss 4.4468 Accuracy 0.1506

Saving checkpoint for epoch 5 at gdrive/My Drive/Colab Notebooks/Deep Learning Labs/LabDL-08-Data/ckpt-8

Epoch 5 Loss 4.4465 Accuracy 0.1506

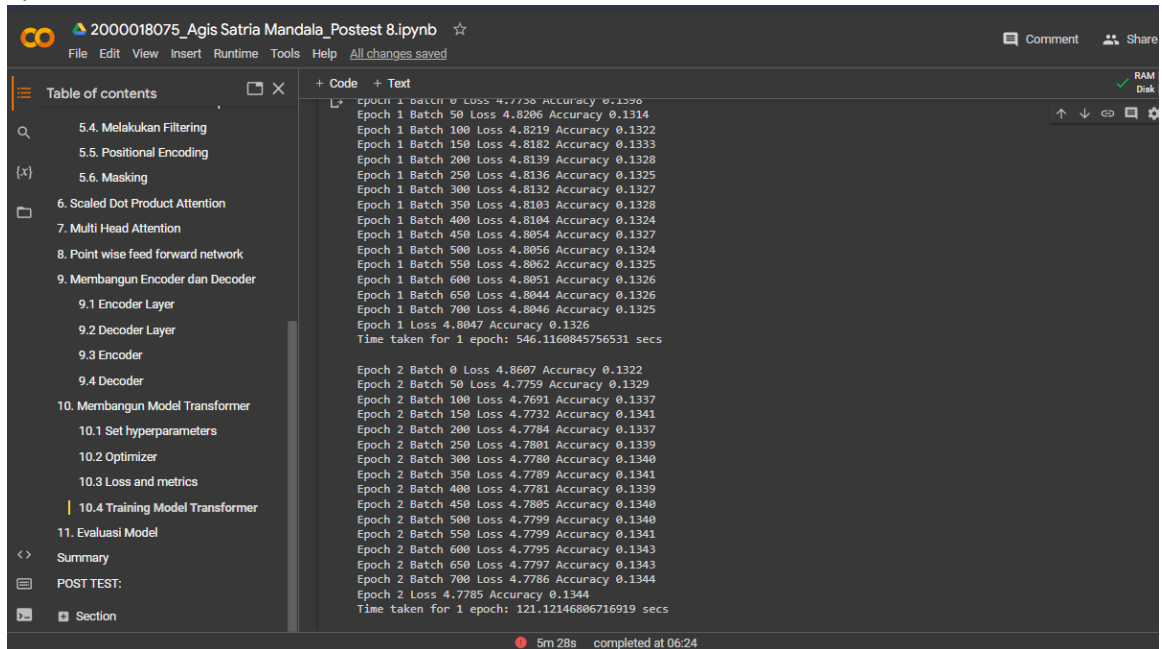
Time taken for 1 epoch: 108.49511742591858 secs

7s completed at 07:41

11. Evaluasi Model

Dapat terlihat pada saat menggunakan epoch 5, akurasi yang mampu dihasilkan pada setiap epochnya adalah 0.1481 untuk epoch pertama, 0.1489 untuk epoch kedua, 0.1492 untuk epoch ketiga, 0.1501 untuk epoch keempat dan 0.1506 untuk epoch kelima, berdasarkan data tersebut, dapat terlihat bahwa pada setiap training yang dilakukan nilai akurasinya akan semakin meningkat.

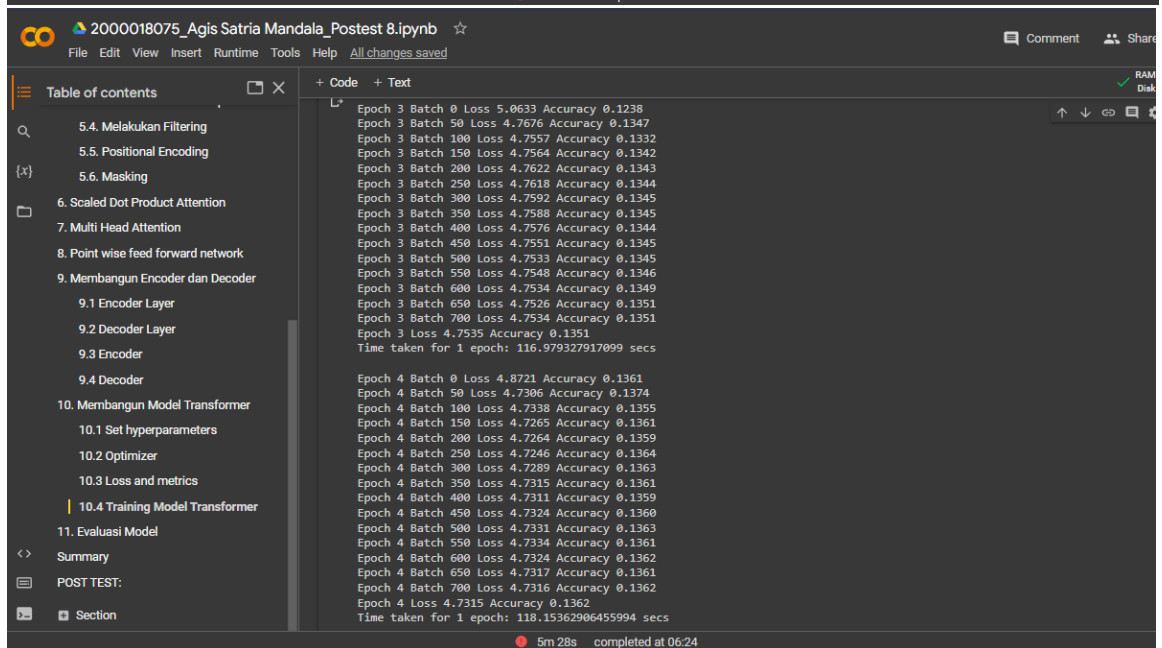
2. Epoch 8



The screenshot shows a Jupyter Notebook interface with a dark theme. The top bar includes the Colab logo, the file name '2000018075_Agis Satria Mandala_Postest 8.ipynb', and icons for comment and share. Below the top bar is a menu bar with 'File', 'Edit', 'View', 'Insert', 'Runtime', 'Tools', and 'Help'. The left sidebar contains a 'Table of contents' panel with a search icon and a list of sections: 5.4. Melakukan Filtering, 5.5. Positional Encoding, 5.6. Masking, 6. Scaled Dot Product Attention, 7. Multi Head Attention, 8. Point wise feed forward network, 9. Membangun Encoder dan Decoder (with sub-sections 9.1 Encoder Layer, 9.2 Decoder Layer, 9.3 Encoder, 9.4 Decoder), 10. Membangun Model Transformer (with sub-sections 10.1 Set hyperparameters, 10.2 Optimizer, 10.3 Loss and metrics, 10.4 Training Model Transformer), 11. Evaluasi Model, Summary, POST TEST:, and Section. The main area is split into two tabs: '+ Code' and '+ Text'. The '+ Code' tab is active, displaying the output of the notebook. The output shows the results for Epoch 1 and Epoch 2. For Epoch 1, the loss decreases from 4.7738 to 4.8047, and accuracy increases from 0.1326 to 0.1326. For Epoch 2, the loss decreases from 4.8607 to 4.7785, and accuracy increases from 0.1322 to 0.1344. The time taken for 1 epoch is 546.1160845756531 secs for Epoch 1 and 121.12146806716919 secs for Epoch 2. The bottom status bar shows '5m 28s completed at 06:24'.

```
Epoch 1 Batch 0 Loss 4.7738 Accuracy 0.1326
Epoch 1 Batch 50 Loss 4.8206 Accuracy 0.1314
Epoch 1 Batch 100 Loss 4.8219 Accuracy 0.1322
Epoch 1 Batch 150 Loss 4.8182 Accuracy 0.1333
Epoch 1 Batch 200 Loss 4.8139 Accuracy 0.1328
Epoch 1 Batch 250 Loss 4.8136 Accuracy 0.1325
Epoch 1 Batch 300 Loss 4.8132 Accuracy 0.1327
Epoch 1 Batch 350 Loss 4.8103 Accuracy 0.1328
Epoch 1 Batch 400 Loss 4.8104 Accuracy 0.1324
Epoch 1 Batch 450 Loss 4.8054 Accuracy 0.1327
Epoch 1 Batch 500 Loss 4.8056 Accuracy 0.1324
Epoch 1 Batch 550 Loss 4.8062 Accuracy 0.1325
Epoch 1 Batch 600 Loss 4.8051 Accuracy 0.1326
Epoch 1 Batch 650 Loss 4.8044 Accuracy 0.1326
Epoch 1 Batch 700 Loss 4.8046 Accuracy 0.1325
Epoch 1 Loss 4.8047 Accuracy 0.1326
Time taken for 1 epoch: 546.1160845756531 secs

Epoch 2 Batch 0 Loss 4.8607 Accuracy 0.1322
Epoch 2 Batch 50 Loss 4.7759 Accuracy 0.1329
Epoch 2 Batch 100 Loss 4.7601 Accuracy 0.1337
Epoch 2 Batch 150 Loss 4.7732 Accuracy 0.1341
Epoch 2 Batch 200 Loss 4.7784 Accuracy 0.1337
Epoch 2 Batch 250 Loss 4.7801 Accuracy 0.1339
Epoch 2 Batch 300 Loss 4.7780 Accuracy 0.1340
Epoch 2 Batch 350 Loss 4.7789 Accuracy 0.1341
Epoch 2 Batch 400 Loss 4.7781 Accuracy 0.1339
Epoch 2 Batch 450 Loss 4.7805 Accuracy 0.1340
Epoch 2 Batch 500 Loss 4.7799 Accuracy 0.1340
Epoch 2 Batch 550 Loss 4.7799 Accuracy 0.1341
Epoch 2 Batch 600 Loss 4.7795 Accuracy 0.1343
Epoch 2 Batch 650 Loss 4.7797 Accuracy 0.1343
Epoch 2 Batch 700 Loss 4.7786 Accuracy 0.1344
Epoch 2 Loss 4.7785 Accuracy 0.1344
Time taken for 1 epoch: 121.12146806716919 secs
```



The screenshot shows a Jupyter Notebook interface with a dark theme. The top bar includes the Colab logo, the file name '2000018075_Agis Satria Mandala_Postest 8.ipynb', and icons for comment and share. Below the top bar is a menu bar with 'File', 'Edit', 'View', 'Insert', 'Runtime', 'Tools', and 'Help'. The left sidebar contains a 'Table of contents' panel with a search icon and a list of sections: 5.4. Melakukan Filtering, 5.5. Positional Encoding, 5.6. Masking, 6. Scaled Dot Product Attention, 7. Multi Head Attention, 8. Point wise feed forward network, 9. Membangun Encoder dan Decoder (with sub-sections 9.1 Encoder Layer, 9.2 Decoder Layer, 9.3 Encoder, 9.4 Decoder), 10. Membangun Model Transformer (with sub-sections 10.1 Set hyperparameters, 10.2 Optimizer, 10.3 Loss and metrics, 10.4 Training Model Transformer), 11. Evaluasi Model, Summary, POST TEST:, and Section. The main area is split into two tabs: '+ Code' and '+ Text'. The '+ Code' tab is active, displaying the output of the notebook. The output shows the results for Epoch 3 and Epoch 4. For Epoch 3, the loss decreases from 5.0633 to 4.7535, and accuracy increases from 0.1238 to 0.1351. For Epoch 4, the loss decreases from 4.8721 to 4.7315, and accuracy increases from 0.1361 to 0.1362. The time taken for 1 epoch is 116.979327917099 secs for Epoch 3 and 118.15362906455994 secs for Epoch 4. The bottom status bar shows '5m 28s completed at 06:24'.

```
Epoch 3 Batch 0 Loss 5.0633 Accuracy 0.1238
Epoch 3 Batch 50 Loss 4.7676 Accuracy 0.1347
Epoch 3 Batch 100 Loss 4.7557 Accuracy 0.1332
Epoch 3 Batch 150 Loss 4.7564 Accuracy 0.1342
Epoch 3 Batch 200 Loss 4.7622 Accuracy 0.1343
Epoch 3 Batch 250 Loss 4.7618 Accuracy 0.1344
Epoch 3 Batch 300 Loss 4.7592 Accuracy 0.1345
Epoch 3 Batch 350 Loss 4.7588 Accuracy 0.1345
Epoch 3 Batch 400 Loss 4.7576 Accuracy 0.1344
Epoch 3 Batch 450 Loss 4.7551 Accuracy 0.1345
Epoch 3 Batch 500 Loss 4.7533 Accuracy 0.1345
Epoch 3 Batch 550 Loss 4.7548 Accuracy 0.1346
Epoch 3 Batch 600 Loss 4.7534 Accuracy 0.1349
Epoch 3 Batch 650 Loss 4.7526 Accuracy 0.1351
Epoch 3 Batch 700 Loss 4.7534 Accuracy 0.1351
Epoch 3 Loss 4.7535 Accuracy 0.1351
Time taken for 1 epoch: 116.979327917099 secs

Epoch 4 Batch 0 Loss 4.8721 Accuracy 0.1361
Epoch 4 Batch 50 Loss 4.7386 Accuracy 0.1374
Epoch 4 Batch 100 Loss 4.7338 Accuracy 0.1355
Epoch 4 Batch 150 Loss 4.7265 Accuracy 0.1361
Epoch 4 Batch 200 Loss 4.7264 Accuracy 0.1359
Epoch 4 Batch 250 Loss 4.7246 Accuracy 0.1364
Epoch 4 Batch 300 Loss 4.7289 Accuracy 0.1363
Epoch 4 Batch 350 Loss 4.7315 Accuracy 0.1361
Epoch 4 Batch 400 Loss 4.7311 Accuracy 0.1359
Epoch 4 Batch 450 Loss 4.7324 Accuracy 0.1360
Epoch 4 Batch 500 Loss 4.7331 Accuracy 0.1363
Epoch 4 Batch 550 Loss 4.7334 Accuracy 0.1361
Epoch 4 Batch 600 Loss 4.7324 Accuracy 0.1362
Epoch 4 Batch 650 Loss 4.7317 Accuracy 0.1361
Epoch 4 Batch 700 Loss 4.7316 Accuracy 0.1362
Epoch 4 Loss 4.7315 Accuracy 0.1362
Time taken for 1 epoch: 118.15362906455994 secs
```

```
Epoch 5 Batch 0 Loss 4.7352 Accuracy 0.1370
Epoch 5 Batch 100 Loss 4.7185 Accuracy 0.1363
Epoch 5 Batch 200 Loss 4.7171 Accuracy 0.1366
Epoch 5 Batch 300 Loss 4.7191 Accuracy 0.1362
Epoch 5 Batch 400 Loss 4.7188 Accuracy 0.1367
Epoch 5 Batch 500 Loss 4.7158 Accuracy 0.1367
Epoch 5 Batch 600 Loss 4.7123 Accuracy 0.1367
Epoch 5 Batch 700 Loss 4.7065 Accuracy 0.1370
Epoch 5 Batch 800 Loss 4.7064 Accuracy 0.1371
Epoch 5 Batch 900 Loss 4.7078 Accuracy 0.1372
Epoch 5 Batch 1000 Loss 4.7049 Accuracy 0.1372
Epoch 5 Batch 1100 Loss 4.7069 Accuracy 0.1373
Epoch 5 Batch 1200 Loss 4.7083 Accuracy 0.1374
Epoch 5 Batch 1300 Loss 4.7094 Accuracy 0.1375
Saving checkpoint for epoch 5 at gdrive/My Drive/Colab Notebooks/Deep Learning Labs/LabDL-08-Data/ckpt-5
Epoch 5 Loss 4.7090 Accuracy 0.1375
Time taken for 1 epoch: 118.6752287368774 secs

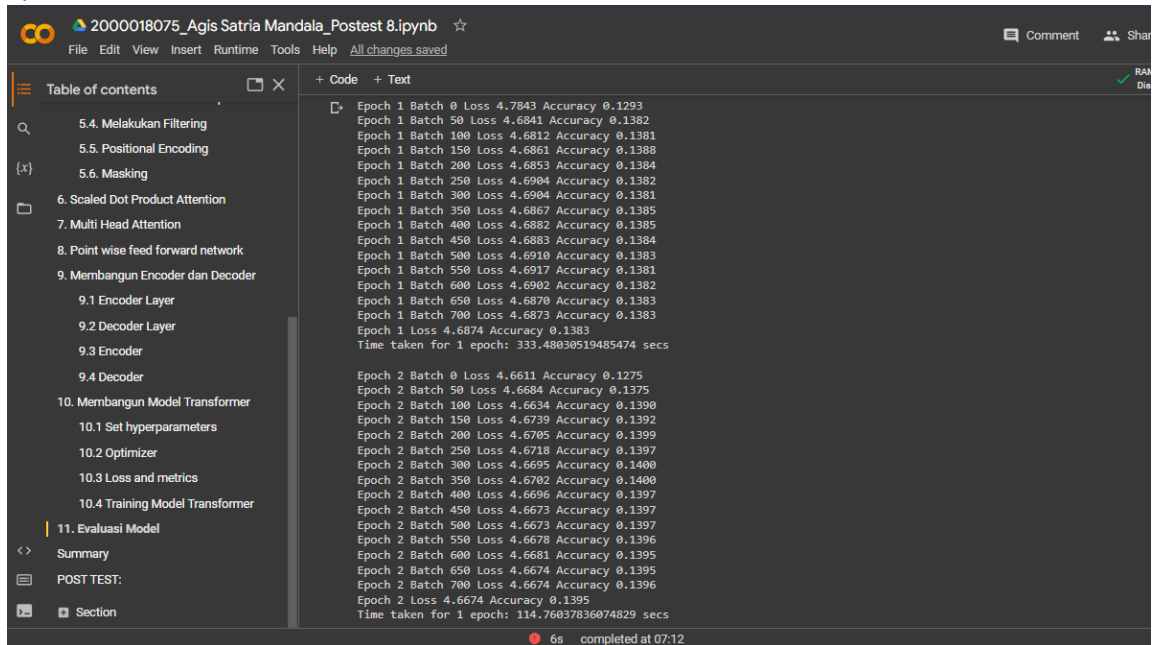
Epoch 6 Batch 0 Loss 4.6234 Accuracy 0.1364
Epoch 6 Batch 100 Loss 4.7019 Accuracy 0.1365
Epoch 6 Batch 200 Loss 4.6919 Accuracy 0.1375
Epoch 6 Batch 300 Loss 4.6919 Accuracy 0.1381
Epoch 6 Batch 400 Loss 4.6933 Accuracy 0.1381
Epoch 6 Batch 500 Loss 4.6952 Accuracy 0.1383
Epoch 6 Batch 600 Loss 4.6907 Accuracy 0.1384
Epoch 6 Batch 700 Loss 4.6919 Accuracy 0.1383
Epoch 6 Batch 800 Loss 4.6951 Accuracy 0.1382
Epoch 6 Batch 900 Loss 4.6936 Accuracy 0.1383
Epoch 6 Batch 1000 Loss 4.6942 Accuracy 0.1385
Epoch 6 Batch 1100 Loss 4.6913 Accuracy 0.1385
Epoch 6 Batch 1200 Loss 4.6906 Accuracy 0.1385
Epoch 6 Batch 1300 Loss 4.6890 Accuracy 0.1385
Epoch 6 Batch 1400 Loss 4.6878 Accuracy 0.1385
Epoch 6 Loss 4.6881 Accuracy 0.1385
Time taken for 1 epoch: 120.0010022354126 secs
```

```
Epoch 7 Batch 0 Loss 4.8324 Accuracy 0.1271
Epoch 7 Batch 100 Loss 4.6709 Accuracy 0.1394
Epoch 7 Batch 200 Loss 4.6555 Accuracy 0.1390
Epoch 7 Batch 300 Loss 4.6634 Accuracy 0.1387
Epoch 7 Batch 400 Loss 4.6698 Accuracy 0.1391
Epoch 7 Batch 500 Loss 4.6678 Accuracy 0.1388
Epoch 7 Batch 600 Loss 4.6704 Accuracy 0.1389
Epoch 7 Batch 700 Loss 4.6691 Accuracy 0.1393
Epoch 7 Batch 800 Loss 4.6683 Accuracy 0.1393
Epoch 7 Batch 900 Loss 4.6669 Accuracy 0.1394
Epoch 7 Batch 1000 Loss 4.6685 Accuracy 0.1392
Epoch 7 Batch 1100 Loss 4.6692 Accuracy 0.1391
Epoch 7 Batch 1200 Loss 4.6684 Accuracy 0.1392
Epoch 7 Batch 1300 Loss 4.6678 Accuracy 0.1393
Epoch 7 Loss 4.6674 Accuracy 0.1393
Time taken for 1 epoch: 115.54340481758118 secs

Epoch 8 Batch 0 Loss 4.7783 Accuracy 0.1377
Epoch 8 Batch 100 Loss 4.6457 Accuracy 0.1410
Epoch 8 Batch 200 Loss 4.6448 Accuracy 0.1406
Epoch 8 Batch 300 Loss 4.6458 Accuracy 0.1408
Epoch 8 Batch 400 Loss 4.6460 Accuracy 0.1403
Epoch 8 Batch 500 Loss 4.6429 Accuracy 0.1399
Epoch 8 Batch 600 Loss 4.6476 Accuracy 0.1398
Epoch 8 Batch 700 Loss 4.6457 Accuracy 0.1401
Epoch 8 Batch 800 Loss 4.6482 Accuracy 0.1402
Epoch 8 Batch 900 Loss 4.6496 Accuracy 0.1403
Epoch 8 Batch 1000 Loss 4.6488 Accuracy 0.1402
Epoch 8 Batch 1100 Loss 4.6460 Accuracy 0.1402
Epoch 8 Batch 1200 Loss 4.6470 Accuracy 0.1403
Epoch 8 Batch 1300 Loss 4.6467 Accuracy 0.1404
Epoch 8 Batch 1400 Loss 4.6473 Accuracy 0.1404
Epoch 8 Loss 4.6475 Accuracy 0.1404
Time taken for 1 epoch: 112.0259566497803 secs
```

Pada saat menggunakan epoch 8, nilai akurasi yang didapatkan adalah 0.1326 untuk epoch pertama, 0.1344 untuk epoch kedua, 0.1351 untuk epoch ketiga, 0.1362 untuk epoch keempat, 0.1375 untuk epoch kelima, 0.1385 untuk epoch keenam, 0.1393 untuk epoch ketujuh, dan 0.1404 untuk epoch kedelapan. Walaupun setingan pada 2 data training tersebut sama yang berbeda hanyalah epochnya, namun disaat menggunakan epoch 5 nilai akurasi yang dihasilkan lebih baik daripada saat menggunakan epoch 8.

3. Epoch 10



2000018075_Agis Satria Mandala_Postest 8.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Table of contents

- 5.4. Melakukan Filtering
- 5.5. Positional Encoding
- 5.6. Masking
- 6. Scaled Dot Product Attention
- 7. Multi Head Attention
- 8. Point wise feed forward network
- 9. Membangun Encoder dan Decoder
 - 9.1 Encoder Layer
 - 9.2 Decoder Layer
 - 9.3 Encoder
 - 9.4 Decoder
- 10. Membangun Model Transformer
 - 10.1 Set hyperparameters
 - 10.2 Optimizer
 - 10.3 Loss and metrics
 - 10.4 Training Model Transformer
- 11. Evaluasi Model

Summary

POST TEST:

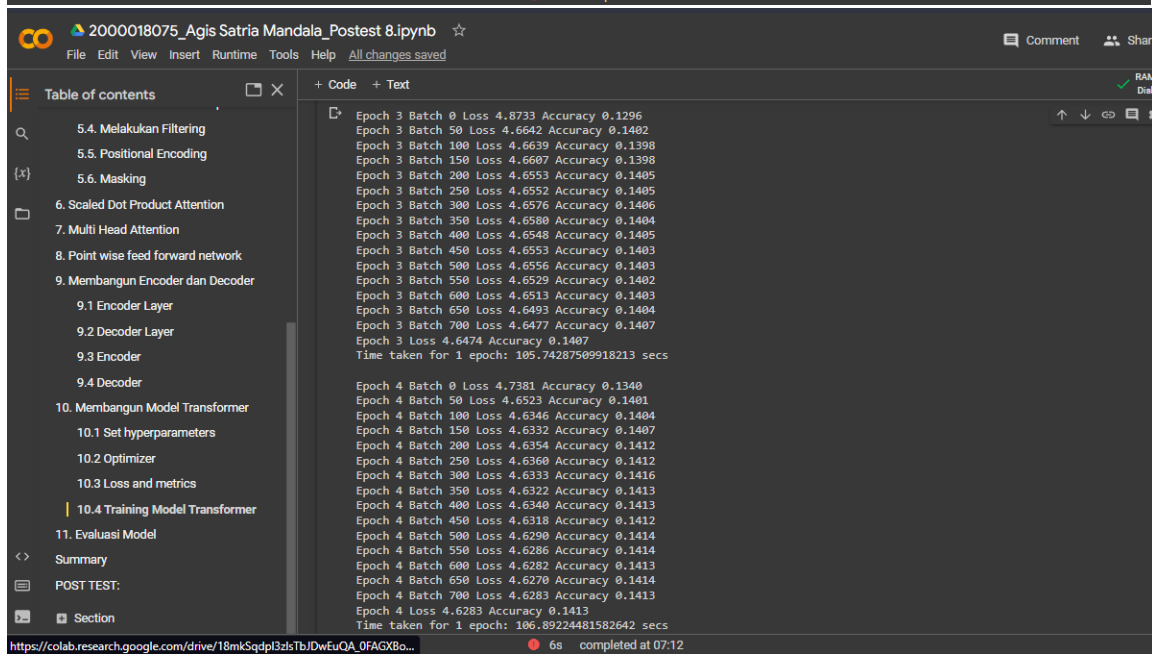
Section

+ Code + Text

```
Epoch 1 Batch 0 Loss 4.7843 Accuracy 0.1293
Epoch 1 Batch 50 Loss 4.6841 Accuracy 0.1382
Epoch 1 Batch 100 Loss 4.6812 Accuracy 0.1381
Epoch 1 Batch 150 Loss 4.6861 Accuracy 0.1388
Epoch 1 Batch 200 Loss 4.6853 Accuracy 0.1384
Epoch 1 Batch 250 Loss 4.6904 Accuracy 0.1382
Epoch 1 Batch 300 Loss 4.6904 Accuracy 0.1381
Epoch 1 Batch 350 Loss 4.6867 Accuracy 0.1385
Epoch 1 Batch 400 Loss 4.6882 Accuracy 0.1385
Epoch 1 Batch 450 Loss 4.6883 Accuracy 0.1384
Epoch 1 Batch 500 Loss 4.6910 Accuracy 0.1383
Epoch 1 Batch 550 Loss 4.6917 Accuracy 0.1381
Epoch 1 Batch 600 Loss 4.6902 Accuracy 0.1382
Epoch 1 Batch 650 Loss 4.6870 Accuracy 0.1383
Epoch 1 Batch 700 Loss 4.6873 Accuracy 0.1383
Epoch 1 Loss 4.6874 Accuracy 0.1383
Time taken for 1 epoch: 333.48030519485474 secs

Epoch 2 Batch 0 Loss 4.6611 Accuracy 0.1275
Epoch 2 Batch 50 Loss 4.6684 Accuracy 0.1375
Epoch 2 Batch 100 Loss 4.6634 Accuracy 0.1390
Epoch 2 Batch 150 Loss 4.6739 Accuracy 0.1392
Epoch 2 Batch 200 Loss 4.6705 Accuracy 0.1399
Epoch 2 Batch 250 Loss 4.6718 Accuracy 0.1397
Epoch 2 Batch 300 Loss 4.6695 Accuracy 0.1400
Epoch 2 Batch 350 Loss 4.6702 Accuracy 0.1400
Epoch 2 Batch 400 Loss 4.6696 Accuracy 0.1397
Epoch 2 Batch 450 Loss 4.6673 Accuracy 0.1397
Epoch 2 Batch 500 Loss 4.6673 Accuracy 0.1397
Epoch 2 Batch 550 Loss 4.6678 Accuracy 0.1396
Epoch 2 Batch 600 Loss 4.6681 Accuracy 0.1395
Epoch 2 Batch 650 Loss 4.6674 Accuracy 0.1395
Epoch 2 Batch 700 Loss 4.6674 Accuracy 0.1396
Epoch 2 Loss 4.6674 Accuracy 0.1395
Time taken for 1 epoch: 114.76037836074829 secs
```

6s completed at 07:12



2000018075_Agis Satria Mandala_Postest 8.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Table of contents

- 5.4. Melakukan Filtering
- 5.5. Positional Encoding
- 5.6. Masking
- 6. Scaled Dot Product Attention
- 7. Multi Head Attention
- 8. Point wise feed forward network
- 9. Membangun Encoder dan Decoder
 - 9.1 Encoder Layer
 - 9.2 Decoder Layer
 - 9.3 Encoder
 - 9.4 Decoder
- 10. Membangun Model Transformer
 - 10.1 Set hyperparameters
 - 10.2 Optimizer
 - 10.3 Loss and metrics
 - 10.4 Training Model Transformer
- 11. Evaluasi Model

Summary

POST TEST:

Section

+ Code + Text

```
Epoch 3 Batch 0 Loss 4.8733 Accuracy 0.1296
Epoch 3 Batch 50 Loss 4.6642 Accuracy 0.1402
Epoch 3 Batch 100 Loss 4.6639 Accuracy 0.1398
Epoch 3 Batch 150 Loss 4.6607 Accuracy 0.1398
Epoch 3 Batch 200 Loss 4.6553 Accuracy 0.1405
Epoch 3 Batch 250 Loss 4.6552 Accuracy 0.1405
Epoch 3 Batch 300 Loss 4.6576 Accuracy 0.1406
Epoch 3 Batch 350 Loss 4.6580 Accuracy 0.1404
Epoch 3 Batch 400 Loss 4.6548 Accuracy 0.1405
Epoch 3 Batch 450 Loss 4.6553 Accuracy 0.1403
Epoch 3 Batch 500 Loss 4.6556 Accuracy 0.1403
Epoch 3 Batch 550 Loss 4.6529 Accuracy 0.1402
Epoch 3 Batch 600 Loss 4.6513 Accuracy 0.1403
Epoch 3 Batch 650 Loss 4.6493 Accuracy 0.1404
Epoch 3 Batch 700 Loss 4.6477 Accuracy 0.1407
Epoch 3 Loss 4.6474 Accuracy 0.1407
Time taken for 1 epoch: 105.74287509918213 secs

Epoch 4 Batch 0 Loss 4.7381 Accuracy 0.1340
Epoch 4 Batch 50 Loss 4.6523 Accuracy 0.1401
Epoch 4 Batch 100 Loss 4.6346 Accuracy 0.1404
Epoch 4 Batch 150 Loss 4.6332 Accuracy 0.1407
Epoch 4 Batch 200 Loss 4.6354 Accuracy 0.1412
Epoch 4 Batch 250 Loss 4.6360 Accuracy 0.1412
Epoch 4 Batch 300 Loss 4.6333 Accuracy 0.1416
Epoch 4 Batch 350 Loss 4.6322 Accuracy 0.1413
Epoch 4 Batch 400 Loss 4.6340 Accuracy 0.1413
Epoch 4 Batch 450 Loss 4.6318 Accuracy 0.1412
Epoch 4 Batch 500 Loss 4.6290 Accuracy 0.1414
Epoch 4 Batch 550 Loss 4.6286 Accuracy 0.1414
Epoch 4 Batch 600 Loss 4.6282 Accuracy 0.1413
Epoch 4 Batch 650 Loss 4.6270 Accuracy 0.1414
Epoch 4 Batch 700 Loss 4.6283 Accuracy 0.1413
Epoch 4 Loss 4.6283 Accuracy 0.1413
Time taken for 1 epoch: 106.89224481582642 secs
```

6s completed at 07:12

https://colab.research.google.com/drive/18mkSqdp13zIsTbJDwEuQA_0FAGXBo...

2000018075_Agis Satria Mandala_Postest 8.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Table of contents

5.4. Melakukan Filtering

5.5. Positional Encoding

5.6. Masking

6. Scaled Dot Product Attention

7. Multi Head Attention

8. Point wise feed forward network

9. Membangun Encoder dan Decoder

9.1 Encoder Layer

9.2 Decoder Layer

9.3 Encoder

9.4 Decoder

10. Membangun Model Transformer

10.1 Set hyperparameters

10.2 Optimizer

10.3 Loss and metrics

10.4 Training Model Transformer

11. Evaluasi Model

Summary

POST TEST:

Section

+ Code + Text

Epoch 5 Batch 0 Loss 4.7234 Accuracy 0.1320
Epoch 5 Batch 50 Loss 4.6343 Accuracy 0.1412
Epoch 5 Batch 100 Loss 4.6150 Accuracy 0.1409
Epoch 5 Batch 150 Loss 4.6123 Accuracy 0.1419
Epoch 5 Batch 200 Loss 4.6136 Accuracy 0.1410
Epoch 5 Batch 250 Loss 4.6075 Accuracy 0.1412
Epoch 5 Batch 300 Loss 4.6113 Accuracy 0.1418
Epoch 5 Batch 350 Loss 4.6111 Accuracy 0.1418
Epoch 5 Batch 400 Loss 4.6097 Accuracy 0.1421
Epoch 5 Batch 450 Loss 4.6116 Accuracy 0.1421
Epoch 5 Batch 500 Loss 4.6108 Accuracy 0.1420
Epoch 5 Batch 550 Loss 4.6126 Accuracy 0.1420
Epoch 5 Batch 600 Loss 4.6123 Accuracy 0.1423
Epoch 5 Batch 650 Loss 4.6104 Accuracy 0.1425
Epoch 5 Batch 700 Loss 4.6095 Accuracy 0.1425
Saving checkpoint for epoch 5 at gdrive/My Drive/Colab Notebooks/Deep Learning Labs/LabDL-08-Data/ckpt-6
Epoch 5 Loss 4.6093 Accuracy 0.1425
Time taken for 1 epoch: 109.38544797897339 secs

Epoch 6 Batch 0 Loss 4.4883 Accuracy 0.1306
Epoch 6 Batch 50 Loss 4.6193 Accuracy 0.1416
Epoch 6 Batch 100 Loss 4.6114 Accuracy 0.1426
Epoch 6 Batch 150 Loss 4.5982 Accuracy 0.1433
Epoch 6 Batch 200 Loss 4.5980 Accuracy 0.1436
Epoch 6 Batch 250 Loss 4.5940 Accuracy 0.1434
Epoch 6 Batch 300 Loss 4.5957 Accuracy 0.1437
Epoch 6 Batch 350 Loss 4.5990 Accuracy 0.1436
Epoch 6 Batch 400 Loss 4.5961 Accuracy 0.1437
Epoch 6 Batch 450 Loss 4.5950 Accuracy 0.1438
Epoch 6 Batch 500 Loss 4.5923 Accuracy 0.1438
Epoch 6 Batch 550 Loss 4.5910 Accuracy 0.1437
Epoch 6 Batch 600 Loss 4.5920 Accuracy 0.1437
Epoch 6 Batch 650 Loss 4.5912 Accuracy 0.1436
Epoch 6 Batch 700 Loss 4.5918 Accuracy 0.1435
Epoch 6 Loss 4.5920 Accuracy 0.1435
Time taken for 1 epoch: 106.3685781955719 secs

6s completed at 07:12

2000018075_Agis Satria Mandala_Postest 8.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Table of contents

5.4. Melakukan Filtering

5.5. Positional Encoding

5.6. Masking

6. Scaled Dot Product Attention

7. Multi Head Attention

8. Point wise feed forward network

9. Membangun Encoder dan Decoder

9.1 Encoder Layer

9.2 Decoder Layer

9.3 Encoder

9.4 Decoder

10. Membangun Model Transformer

10.1 Set hyperparameters

10.2 Optimizer

10.3 Loss and metrics

10.4 Training Model Transformer

11. Evaluasi Model

Summary

POST TEST:

Section

+ Code + Text

Epoch 7 Batch 0 Loss 4.6959 Accuracy 0.1500
Epoch 7 Batch 50 Loss 4.5717 Accuracy 0.1451
Epoch 7 Batch 100 Loss 4.5903 Accuracy 0.1435
Epoch 7 Batch 150 Loss 4.5899 Accuracy 0.1435
Epoch 7 Batch 200 Loss 4.5891 Accuracy 0.1438
Epoch 7 Batch 250 Loss 4.5901 Accuracy 0.1435
Epoch 7 Batch 300 Loss 4.5854 Accuracy 0.1438
Epoch 7 Batch 350 Loss 4.5846 Accuracy 0.1439
Epoch 7 Batch 400 Loss 4.5839 Accuracy 0.1439
Epoch 7 Batch 450 Loss 4.5811 Accuracy 0.1442
Epoch 7 Batch 500 Loss 4.5783 Accuracy 0.1445
Epoch 7 Batch 550 Loss 4.5758 Accuracy 0.1446
Epoch 7 Batch 600 Loss 4.5741 Accuracy 0.1447
Epoch 7 Batch 650 Loss 4.5733 Accuracy 0.1447
Epoch 7 Batch 700 Loss 4.5745 Accuracy 0.1446
Epoch 7 Loss 4.5744 Accuracy 0.1446
Time taken for 1 epoch: 104.668616771698 secs

Epoch 8 Batch 0 Loss 4.5227 Accuracy 0.1495
Epoch 8 Batch 50 Loss 4.5723 Accuracy 0.1455
Epoch 8 Batch 100 Loss 4.5689 Accuracy 0.1452
Epoch 8 Batch 150 Loss 4.5630 Accuracy 0.1461
Epoch 8 Batch 200 Loss 4.5612 Accuracy 0.1456
Epoch 8 Batch 250 Loss 4.5604 Accuracy 0.1458
Epoch 8 Batch 300 Loss 4.5638 Accuracy 0.1458
Epoch 8 Batch 350 Loss 4.5596 Accuracy 0.1458
Epoch 8 Batch 400 Loss 4.5583 Accuracy 0.1457
Epoch 8 Batch 450 Loss 4.5574 Accuracy 0.1456
Epoch 8 Batch 500 Loss 4.5557 Accuracy 0.1456
Epoch 8 Batch 550 Loss 4.5579 Accuracy 0.1456
Epoch 8 Batch 600 Loss 4.5580 Accuracy 0.1455
Epoch 8 Batch 650 Loss 4.5577 Accuracy 0.1455
Epoch 8 Batch 700 Loss 4.5578 Accuracy 0.1454
Epoch 8 Loss 4.5575 Accuracy 0.1454
Time taken for 1 epoch: 105.33035802841187 secs

6s completed at 07:12

```
Epoch 9 Batch 0 Loss 4.6812 Accuracy 0.1410
Epoch 9 Batch 50 Loss 4.5378 Accuracy 0.1448
Epoch 9 Batch 100 Loss 4.5350 Accuracy 0.1459
Epoch 9 Batch 150 Loss 4.5326 Accuracy 0.1465
Epoch 9 Batch 200 Loss 4.5368 Accuracy 0.1459
Epoch 9 Batch 250 Loss 4.5386 Accuracy 0.1463
Epoch 9 Batch 300 Loss 4.5412 Accuracy 0.1459
Epoch 9 Batch 350 Loss 4.5406 Accuracy 0.1458
Epoch 9 Batch 400 Loss 4.5430 Accuracy 0.1459
Epoch 9 Batch 450 Loss 4.5434 Accuracy 0.1459
Epoch 9 Batch 500 Loss 4.5416 Accuracy 0.1459
Epoch 9 Batch 550 Loss 4.5420 Accuracy 0.1461
Epoch 9 Batch 600 Loss 4.5433 Accuracy 0.1460
Epoch 9 Batch 650 Loss 4.5435 Accuracy 0.1461
Epoch 9 Batch 700 Loss 4.5404 Accuracy 0.1462
Epoch 9 Loss 4.5481 Accuracy 0.1462
Time taken for 1 epoch: 104.85443639755249 secs

Epoch 10 Batch 0 Loss 4.6175 Accuracy 0.1634
Epoch 10 Batch 50 Loss 4.5135 Accuracy 0.1464
Epoch 10 Batch 100 Loss 4.5281 Accuracy 0.1464
Epoch 10 Batch 150 Loss 4.5232 Accuracy 0.1463
Epoch 10 Batch 200 Loss 4.5277 Accuracy 0.1462
Epoch 10 Batch 250 Loss 4.5345 Accuracy 0.1464
Epoch 10 Batch 300 Loss 4.5306 Accuracy 0.1464
Epoch 10 Batch 350 Loss 4.5279 Accuracy 0.1465
Epoch 10 Batch 400 Loss 4.5273 Accuracy 0.1466
Epoch 10 Batch 450 Loss 4.5286 Accuracy 0.1465
Epoch 10 Batch 500 Loss 4.5278 Accuracy 0.1465
Epoch 10 Batch 550 Loss 4.5252 Accuracy 0.1467
Epoch 10 Batch 600 Loss 4.5246 Accuracy 0.1467
Epoch 10 Batch 650 Loss 4.5252 Accuracy 0.1467
Epoch 10 Batch 700 Loss 4.5241 Accuracy 0.1469
Saving checkpoint for epoch 10 at gdrive/My Drive/Colab Notebooks/Deep Learning Labs/Lab01-08-Data/ckpt-7
Epoch 10 Loss 4.5243 Accuracy 0.1469
Time taken for 1 epoch: 119.64482760429382 secs
```

Untuk pelatihan yang menggunakan epoch berjumlah 10, akurasi yang didapatkan adalah 0.1383 untuk epoch pertama, 0.1395 untuk epoch kedua, 0.1407 untuk epoch ketiga, 0.1413 untuk epoch keempat, 0.1425 untuk epoch kelima, 0.1435 untuk epoch keenam, 0.1446 untuk epoch ketujuh, 0.1454 untuk epoch kedelapan, 0.1462 untuk epoch kesembilan, dan 0.1469 untuk epoch kesepuluh. Walaupun akurasi yang didapatkan lebih baik dari saat menggunakan epoch berjumlah 8, namun akurasi yang dilatih dengan jumlah 5 epoch masih lebih baik daripada 10 epoch.

```
Epoch 1 Batch 50 Loss 5.1342 Accuracy 0.1117
Epoch 1 Batch 100 Loss 5.0380 Accuracy 0.1171
Epoch 1 Batch 150 Loss 4.9830 Accuracy 0.1197
Epoch 1 Batch 200 Loss 4.9474 Accuracy 0.1219
Epoch 1 Batch 250 Loss 4.9272 Accuracy 0.1233
Epoch 1 Batch 300 Loss 4.9055 Accuracy 0.1247
Epoch 1 Batch 350 Loss 4.8884 Accuracy 0.1259
Epoch 1 Batch 400 Loss 4.8755 Accuracy 0.1270
Epoch 1 Batch 450 Loss 4.8617 Accuracy 0.1280
Epoch 1 Batch 500 Loss 4.8514 Accuracy 0.1284
Epoch 1 Batch 550 Loss 4.8431 Accuracy 0.1290
Epoch 1 Batch 600 Loss 4.8358 Accuracy 0.1295
Epoch 1 Batch 650 Loss 4.8286 Accuracy 0.1301
Epoch 1 Batch 700 Loss 4.8204 Accuracy 0.1306
Epoch 1 Loss 4.8196 Accuracy 0.1306
Time taken for 1 epoch: 511.1683325767517 secs

Epoch 2 Batch 0 Loss 4.7305 Accuracy 0.1306
Epoch 2 Batch 50 Loss 4.6942 Accuracy 0.1372
Epoch 2 Batch 100 Loss 4.6967 Accuracy 0.1381
Epoch 2 Batch 150 Loss 4.7023 Accuracy 0.1372
Epoch 2 Batch 200 Loss 4.6952 Accuracy 0.1369
Epoch 2 Batch 250 Loss 4.6947 Accuracy 0.1373
Epoch 2 Batch 300 Loss 4.6934 Accuracy 0.1375
Epoch 2 Batch 350 Loss 4.6875 Accuracy 0.1379
Epoch 2 Batch 400 Loss 4.6872 Accuracy 0.1379
Epoch 2 Batch 450 Loss 4.6822 Accuracy 0.1381
Epoch 2 Batch 500 Loss 4.6802 Accuracy 0.1382
Epoch 2 Batch 550 Loss 4.6794 Accuracy 0.1386
Epoch 2 Batch 600 Loss 4.6750 Accuracy 0.1386
Epoch 2 Batch 650 Loss 4.6748 Accuracy 0.1388
Epoch 2 Batch 700 Loss 4.6723 Accuracy 0.1390
Epoch 2 Loss 4.6722 Accuracy 0.1390
Time taken for 1 epoch: 54.11385226249695 secs

Epoch 3 Batch 0 Loss 4.6081 Accuracy 0.1540
```


The image displays two screenshots of a Jupyter Notebook interface, showing the execution of a Transformer model training and evaluation script.

Top Screenshot: The notebook is titled "2000018075_Agis Satria Mandala_Postest 8.ipynb". The left sidebar shows the "Table of contents" with sections 5.4 through 11. The main code cell shows the output of the training process, including loss and accuracy metrics for Epochs 3 and 4. The output indicates that the training process is completed at 12:37.

Bottom Screenshot: The notebook is titled "2000018075_Agis Satria Mandala_Postest 8.ipynb". The left sidebar shows the "Table of contents" with sections 5.4 through 11. The main code cell shows the output of the evaluation process, including loss and accuracy metrics for Epochs 4 and 5. The output indicates that the evaluation process is completed at 12:37.

The evaluation results show that the model's performance is significantly lower when using 4 layers and 8 heads compared to using 10 layers and 16 heads, and even lower when using 10 epochs.

Itu merupakan hasil yang didapatkan dimana menggunakan jumlah layer sebanyak 4 layer, jumlah heads sebanyak 8 dan epoch 5, jika dibandingkan dengan data akurasi yang didapatkan dimana menggunakan 10 layer dan 16 head, akurasi yang didapatkan ini jauh lebih kecil bahkan jika dibandingkan dengan yang menggunakan 10 epoch.