

# Swahili transformer model bake off

pretrained swahili transformer, lr=0.0001, pos tagging

```
2022-06-11 17:40:51,753 - INFO - allennlp.common.util - Metrics: {
  "best_epoch": 4,
  "peak_worker_0_memory_MB": 5448.7890625,
  "peak_gpu_0_memory_MB": 9332.94873046875,
  "training_duration": "1:26:00.349669",
  "epoch": 9,
  "training_accuracy": 0.996923352964769,
  "training_accuracy3": 0.9999548500435081,
  "training_precision": 0.9857814311981201,
  "training_recall": 0.9809743165969849,
  "training_fscore": 0.9833330512046814,
  "training_loss": 0.008818848443899869,
  "training_worker_0_memory_MB": 5448.7890625, "training_gpu_0_memory_MB":
9332.94873046875,
  "validation_accuracy": 0.9915935820231985,
  "validation_accuracy3": 0.9994341834054076,
  "validation_precision": 0.9857875108718872,
  "validation_recall": 0.9809861779212952,
  "validation_fscore": 0.9833420515060425,
  "validation_loss": 0.06279067079275392,
  "best_validation_accuracy": 0.989734470355252, "best_validation_accuracy3":
0.9997575071737461,
  "best_validation_precision": 0.9777423143386841, "best_validation_recall":
0.9695621728897095,
  "best_validation_fscore": 0.9735202789306641, "best_validation_loss":
0.05510239019787424,
  "test_accuracy": 0.9954022988505747,
  "test_accuracy3": 0.9996934865900383,
  "test_precision": 0.9858008623123169,
```

```
"test_recall": 0.9810028076171875,  
"test_fscore": 0.9833571910858154,  
"test_loss": 0.0319001576157281  
}
```

pretrained swahili transformer, lr=0.00001, pos tagging

```
"best_epoch": 7,  
"peak_worker_0_memory_MB": 5449.3203125,  
"peak_gpu_0_memory_MB": 9332.94873046875,  
"training_duration": "1:25:54.586565",  
"epoch": 9,  
"training_accuracy": 0.9972294344879995,  
"training_accuracy3": 0.9999730272987192,  
"training_precision": 0.9764115214347839,  
"training_recall": 0.9566909074783325,  
"training_fscore": 0.9659566879272461,  
"training_loss": 0.007170079944642032,  
"training_worker_0_memory_MB": 5449.3203125, "training_gpu_0_memory_MB":  
9332.94873046875,  
"validation_accuracy": 0.9921189831467486,  
"validation_accuracy3": 0.9996766762316615,  
"validation_precision": 0.9764279127120972,  
"validation_recall": 0.9567330479621887,  
"validation_fscore": 0.9659876227378845,  
"validation_loss": 0.0693623513392432,  
"best_validation_accuracy": 0.991916905791537, "best_validation_accuracy3":  
0.9996362607606192,  
"best_validation_precision": 0.9715831279754639, "best_validation_recall":  
0.9473865032196045,  
"best_validation_fscore": 0.9585748910903931, "best_validation_loss":  
0.05559036874520399,  
"test_accuracy": 0.9962835249042146,  
"test_accuracy3": 0.9997318007662835,  
"test_precision": 0.9764541387557983,  
"test_recall": 0.9567817449569702,  
"test_fscore": 0.9660264253616333,  
"test_loss": 0.032172908316944124  
}
```

pretrained swahili transformer, lr=0.0001, crf, pos tagging

```
2022-06-11 22:44:50,984 - INFO - allennlp.common.util - Metrics: {
  "best_epoch": 4,
  "peak_worker_0_memory_MB": 5698.4609375,
  "peak_gpu_0_memory_MB": 9332.60205078125,
  "training_duration": "2:12:03.720062",
  "epoch": 9,
  "training_accuracy": 0.9975718705216521,
  "training_accuracy3": 0.9978339748145334,
  "training_loss": 1.5775271221179867,
  "training_worker_0_memory_MB": 5698.4609375, "training_gpu_0_memory_MB":
9332.60205078125,
  "validation_accuracy": 0.9919573212625793,
  "validation_accuracy3": 0.9926039687992564,
  "validation_loss": 8.17987885513926,
  "best_validation_accuracy": 0.9915127510811138, "best_validation_accuracy3":
0.9925231378571717,
  "best_validation_loss": 7.377661480167048,
  "test_accuracy": 0.9957471264367816,
  "test_accuracy3": 0.9962835249042146,
  "test_loss": 3.963569987903942
}
```

pretrained swahili transformer, lr=0.0001, helsinki pos tagging

```
{
  "best_epoch": 5,
  "peak_worker_0_memory_MB": 5583.07421875,
  "peak_gpu_0_memory_MB": 9335.52783203125,
  "training_duration": "1:24:44.563542",
  "epoch": 9,
  "training_accuracy": 0.9965898688553921,
  "training_accuracy3": 0.9999364222373633,
  "training_precision": 0.9467342495918274,
  "training_recall": 0.8755066990852356,
  "training_fscore": 0.8997867107391357,
  "training_loss": 0.01112793169544195,
  "training_worker_0_memory_MB": 5583.07421875, "training_gpu_0_memory_MB":
9335.52783203125,
```

```
"validation_accuracy": 0.9915556244189649,  
"validation_accuracy3": 0.999535171986365,  
"validation_precision": 0.9467248916625977,  
"validation_recall": 0.8757268190383911,  
"validation_fscore": 0.8999462723731995,  
"validation_loss": 0.06682281847298145,  
"best_validation_accuracy": 0.9905484970560893, "best_validation_accuracy3":  
0.9994964363185621,  
"best_validation_precision": 0.9082106351852417, "best_validation_recall":  
0.8251690864562988,  
"best_validation_fscore": 0.8477039337158203, "best_validation_loss":  
0.05897250342555344,  
"test_accuracy": 0.9952177563611142,  
"test_accuracy3": 0.9998174716168364,  
"test_precision": 0.9467272162437439,  
"test_recall": 0.8757625222206116,  
"test_fscore": 0.8999730348587036,  
"test_loss": 0.02968891077139415  
}
```

---

Revision #3

Created Sat, Jun 11, 2022 9:51 PM by [kenneth](#)

Updated Mon, Jun 13, 2022 2:36 AM by [kenneth](#)