In [1]:

```
import os
from pathlib import Path

import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
import matplotlib.pyplot as plt

import torch
from torch.utils.data import DataLoader
from torch.nn.utils.rnn import pad_sequence
from torch import nn

from tqdm.auto import tqdm,trange

from transformers import AdamW


BATCH_SIZE = 64
```

In [2]:

```
!rm -rf PoS-Tagging
!git clone https://github.com/Janluke0/PoS-Tagging/
os.chdir('PoS-Tagging')
out_dir = Path('/kaggle/working/')
out_dir.mkdir(exist_ok=True)
```

```
Cloning into 'PoS-Tagging'...
remote: Enumerating objects: 57, done.
remote: Counting objects: 100% (57/57), done.
remote: Compressing objects: 100% (35/35), done.
remote: Total 57 (delta 15), reused 57 (delta 15), pack-reused 0
Unpacking objects: 100% (57/57), 595.77 KiB | 1.84 MiB/s, done.
```

In [3]:

```
from model.transformers.italian import ItBERTCasedPos,ItBERTUncasedPos
from model import train_model
from dataset import TWITADS
```

# Common

this part should be adde to the repo, it's time to go lighting

In [4]:

```
def train_model(model, dl_train, dl_test, cuda=False, lr=0.001, epochs=10, show_plots=F
alse, save_dir=None):
    loss_function = nn.NLLLoss()
    optimizer = AdamW(model.parameters(), lr=lr,weight_decay=0.01)

    if cuda:
        model = model.cuda()
    if save_dir is not None:
        save_dir.mkdir(exist_ok=True)

    losses = []
    accuracies = []
    best_acc = 0
    best_loss = float('inf')
    pbar = trange(epochs)
    for epoch in pbar:
        model.train()
        for sample in tqdm(iter(dl_train), desc=f"Training {epoch}° epoch", leave=False
):
            x,m,y = sample['input_ids'],sample['attention_mask'], sample['labels']
            if cuda:
                x, m, y = x.cuda(), m.cuda(), y.cuda()
            optimizer.zero_grad()

            tag_scores =  model(input_ids=x,attention_mask=m)
            loss = loss_function(tag_scores.transpose(1, 2),y)

            loss.backward()
            optimizer.step()

        acc = []
        los = []
        ## evaluation
        model.eval()
        with torch.no_grad():
            for sample in tqdm(iter(dl_test), desc=f"Eval {epoch}° epoch", leave=False
):
                x,m,y = sample['input_ids'],sample['attention_mask'], sample['labels']
                if cuda:
                    x, m, y = x.cuda(), m.cuda(), y.cuda()

                tag_scores =  model(input_ids=x,attention_mask=m)
                if hasattr(tag_scores,'logits'):
                    tag_scores = tag_scores.logits

                loss = loss_function(tag_scores.transpose(1, 2),y)
                los.append(loss.cpu().item())

                acc.append(((tag_scores.argmax(2))==y)[m==1].float())

        acc = torch.cat(acc).mean().item()
        los = np.array(los).mean()

        losses.append(los)
        accuracies.append(acc)
        #show epoch results
        pbar.set_description(f"Loss:{los}\tAccurancy:{acc}")
        if show_plots:
            plt.subplot(121)
```

```
            plt.title("Test loss")
            plt.plot(losses)

            plt.subplot(122)
            plt.title("Test accuracy")
            plt.plot(accuracies)
        if save_dir is not None and acc >= best_acc:
            torch.save(model.state_dict(),save_dir/f"model_best_acc.pth")
        if save_dir is not None and loss <= best_loss:
            torch.save(model.state_dict(),save_dir/f"model_best_loss.pth")

        best_acc = max(acc,best_acc)
        best_loss = max(los,best_loss)

    return losses,accuracies

def show_pred(model, ds, i):
    REVTAG = {v:k for k,v in ds._TAGS.items()}
    model.cpu()
    sample = ds.collate([ds[i]])
    x,m,y = sample['input_ids'],sample['attention_mask'], sample['labels']
    with torch.no_grad():
        pred = model(input_ids=x,attention_mask=m)
    tkns = ds.tokenizer.convert_ids_to_tokens(x[0,1:-1])
    return list(zip(tkns,[REVTAG[v.item()] for v in pred[0].argmax(1)[1:-1]],[REVTAG[v.
item()] for v in y[0][1:-1]]))
```

In [5]:

```
def collate_fn(batch):
    input_ids, token_type_ids, attention_mask, labels = [[] for _ in range(4)]
    for sample in batch:
        input_ids.append(sample['input_ids'])
        token_type_ids.append(sample['token_type_ids'])
        attention_mask.append(sample['attention_mask'])
        labels.append(sample['labels'])
    return {
        'input_ids':pad_sequence(input_ids,batch_first=True),
        'token_type_ids': pad_sequence(token_type_ids,batch_first=True),
        'labels':pad_sequence(labels, padding_value=-100,batch_first=True),
        'attention_mask':pad_sequence(attention_mask,batch_first=True),
    }
```

In [6]:

```
def tokenize_and_align_labels(tokenizer, tokens, tags):
    tokens = list(tokens)
    tokenized_inputs = tokenizer(tokens, truncation=True, is_split_into_words=True)

    word_ids = tokenized_inputs.word_ids(batch_index=0)  # Map tokens to their respecti
ve word.
    previous_word_idx = None
    label_ids = []
    for word_idx in word_ids:                            # Set the special tokens to -1
00
        if word_idx is None:
            label_ids.append(-100)
        elif word_idx != previous_word_idx:              # Only label the first token o
f a given word.
            label_ids.append(tags[word_idx])

    tokenized_inputs["labels"] = label_ids

    return {k:torch.tensor(v) for k,v in tokenized_inputs.items()}
```

# Cased model

In [7]:

```
cased_model = ItBERTCasedPos(23)
```

Downloading: 433/433 [00:00<00:00, 16.0kB/s]

Downloading: 422M/422M [00:09<00:00, 44.4MB/s]

```
Some weights of the model checkpoint at dbmdz/bert-base-italian-cased were
not used when initializing BertForTokenClassification: ['cls.predictions.b
ias', 'cls.predictions.transform.dense.weight', 'cls.predictions.decoder.w
eight', 'cls.seq_relationship.weight', 'cls.predictions.transform.LayerNor
m.bias', 'cls.predictions.transform.LayerNorm.weight', 'cls.predictions.tr
ansform.dense.bias', 'cls.seq_relationship.bias']
- This IS expected if you are initializing BertForTokenClassification from
the checkpoint of a model trained on another task or with another architec
ture (e.g. initializing a BertForSequenceClassification model from a BertF
orPreTraining model).
- This IS NOT expected if you are initializing BertForTokenClassification
from the checkpoint of a model that you expect to be exactly identical (in
itializing a BertForSequenceClassification model from a BertForSequenceCla
ssification model).
Some weights of BertForTokenClassification were not initialized from the m
odel checkpoint at dbmdz/bert-base-italian-cased and are newly initialize
d: ['classifier.bias', 'classifier.weight']
You should probably TRAIN this model on a down-stream task to be able to u
se it for predictions and inference.
```

In [8]:

```
tokenizer = ItBERTCasedPos.tokenizer()
ds_train = TWITADS('resampled_train',
                   lambda w:[w],
                   transform=lambda a,b: tokenize_and_align_labels(tokenizer,a,b)
)
ds_val = TWITADS('resampled_validation',
                   lambda w:[w],
                   transform=lambda a,b: tokenize_and_align_labels(tokenizer,a,b)
)
```

Downloading: 59.0/59.0 [00:00<00:00, 2.22kB/s]

Downloading: 230k/230k [00:00<00:00, 822kB/s]

In [9]:

```
dl_train = DataLoader(ds_train, shuffle=True, batch_size=BATCH_SIZE, collate_fn=collate_fn)
dl_val  = DataLoader(ds_val, shuffle=True, batch_size=BATCH_SIZE, collate_fn=collate_fn)
```

In [10]:

```
?train_model
```

In [11]:

```
torch.manual_seed(42)
train_model(cased_model,dl_train,dl_val,cuda=torch.cuda.is_available(), lr=2e-5, epochs
=100, show_plots=True, save_dir=out_dir/"cased_model")
```

Loss:1.6926381587982178 Accurancy:0.9033723473548889: 100% 100/100 [57:51<00:00, 34.35s/it]

Training 0° epoch: 100% 91/91 [00:35<00:00, 3.19it/s]

Eval 0° epoch: 100% 10/10 [00:01<00:00, 6.44it/s]

Training 1° epoch: 100% 91/91 [00:31<00:00, 3.37it/s]

Eval 1° epoch: 100% 10/10 [00:01<00:00, 9.30it/s]

Training 2° epoch: 100% 91/91 [00:32<00:00, 3.20it/s]

Eval 2° epoch: 90% 9/10 [00:01<00:00, 8.41it/s]

Training 3° epoch: 100% 91/91 [00:32<00:00, 3.68it/s]

Eval 3° epoch: 90% 9/10 [00:01<00:00, 8.32it/s]

Training 4° epoch: 100% 91/91 [00:32<00:00, 3.42it/s]

Eval 4° epoch: 100% 10/10 [00:01<00:00, 9.02it/s]

Training 5° epoch: 100% 91/91 [00:32<00:00, 3.38it/s]

Eval 5° epoch: 100% 10/10 [00:01<00:00, 9.17it/s]

Training 6° epoch: 100% 91/91 [00:31<00:00, 3.45it/s]

Eval 6° epoch: 90% 9/10 [00:01<00:00, 9.07it/s]

Training 7° epoch: 100% 91/91 [00:32<00:00, 3.06it/s]

Eval 7° epoch: 100% 10/10 [00:01<00:00, 9.15it/s]

Training 8° epoch: 100% 91/91 [00:32<00:00, 3.61it/s]

Eval 8° epoch: 90% 9/10 [00:01<00:00, 8.55it/s]

Training 9° epoch: 100% 91/91 [00:32<00:00, 3.31it/s]

Eval 9° epoch: 90% 9/10 [00:01<00:00, 8.73it/s]

Training 10° epoch: 100% 91/91 [00:32<00:00, 3.13it/s]

Eval 10° epoch: 100% 10/10 [00:01<00:00, 8.77it/s]

Training 11° epoch: 100% 91/91 [00:32<00:00, 3.25it/s]

Eval 11° epoch: 100% 10/10 [00:01<00:00, 8.90it/s]

Training 12° epoch: 100% 91/91 [00:31<00:00, 3.28it/s]

Eval 12° epoch: 100% 10/10 [00:01<00:00, 8.54it/s]

Training 13° epoch: 100% 91/91 [00:32<00:00, 2.93it/s]

Eval 13° epoch: 100% 10/10 [00:01<00:00, 8.88it/s]

Training 14° epoch: 100% 91/91 [00:32<00:00, 3.13it/s]

Eval 14° epoch: 100% 10/10 [00:01<00:00, 8.85it/s]

Training 15° epoch: 100% 91/91 [00:31<00:00, 3.34it/s]

Eval 15° epoch: 100% 10/10 [00:01<00:00, 8.96it/s]

Training 16° epoch: 100% 91/91 [00:32<00:00, 3.43it/s]

Eval 16° epoch: 90% 9/10 [00:01<00:00, 8.84it/s]

Training 17° epoch: 100% 91/91 [00:32<00:00, 3.33it/s]

Eval 17° epoch: 90% 9/10 [00:01<00:00, 8.42it/s]

Training 18° epoch: 100% 91/91 [00:32<00:00, 3.31it/s]

Eval 18° epoch: 90% 9/10 [00:01<00:00, 8.72it/s]

Training 19° epoch: 100% 91/91 [00:32<00:00, 3.18it/s]

Eval 19° epoch: 100% 10/10 [00:01<00:00, 9.03it/s]

Training 20° epoch: 100% 91/91 [00:31<00:00, 3.21it/s]

Eval 20° epoch: 100% 10/10 [00:01<00:00, 8.93it/s]

Training 21° epoch: 100% 91/91 [00:32<00:00, 3.07it/s]

Eval 21° epoch: 100% 10/10 [00:01<00:00, 8.69it/s]

Training 22° epoch: 100% 91/91 [00:32<00:00, 3.25it/s]

Eval 22° epoch: 90% 9/10 [00:01<00:00, 8.63it/s]

Training 23° epoch: 100% 91/91 [00:31<00:00, 3.13it/s]

Eval 23° epoch: 100% 10/10 [00:01<00:00, 8.81it/s]

Training 24° epoch: 100% 91/91 [00:32<00:00, 3.26it/s]

Eval 24° epoch: 100% 10/10 [00:01<00:00, 9.13it/s]

Training 25° epoch: 100% 91/91 [00:32<00:00, 3.10it/s]

Eval 25° epoch: 100% 10/10 [00:01<00:00, 9.20it/s]

Training 26° epoch: 100% 91/91 [00:32<00:00, 3.42it/s]

Eval 26° epoch: 100% 10/10 [00:01<00:00, 9.41it/s]

Training 27° epoch: 100% 91/91 [00:32<00:00, 3.36it/s]

Eval 27° epoch: 100% 10/10 [00:01<00:00, 9.30it/s]

Training 28° epoch: 100% 91/91 [00:31<00:00, 3.54it/s]

Eval 28° epoch: 100% 10/10 [00:01<00:00, 8.79it/s]

Training 29° epoch: 100% 91/91 [00:32<00:00, 3.28it/s]

Eval 29° epoch: 90% 9/10 [00:01<00:00, 8.71it/s]

Training 30° epoch: 100% 91/91 [00:31<00:00, 3.30it/s]

Eval 30° epoch: 100% 10/10 [00:01<00:00, 9.14it/s]

Training 31° epoch: 100% 91/91 [00:31<00:00, 3.22it/s]

Eval 31° epoch: 90% 9/10 [00:01<00:00, 8.56it/s]

Training 32° epoch: 100% 91/91 [00:31<00:00, 3.57it/s]

Eval 32° epoch: 100% 10/10 [00:01<00:00, 8.93it/s]

Training 33° epoch: 100% 91/91 [00:32<00:00, 3.28it/s]

Eval 33° epoch: 90% 9/10 [00:01<00:00, 8.70it/s]

Training 34° epoch: 100% 91/91 [00:32<00:00, 3.32it/s]

Eval 34° epoch: 100% 10/10 [00:01<00:00, 9.06it/s]

Training 35° epoch: 100% 91/91 [00:31<00:00, 3.31it/s]

Eval 35° epoch: 100% 10/10 [00:01<00:00, 8.93it/s]

Training 36° epoch: 100% 91/91 [00:32<00:00, 3.26it/s]

Eval 36° epoch: 90% 9/10 [00:01<00:00, 8.64it/s]

Training 37° epoch: 100% 91/91 [00:32<00:00, 3.36it/s]

Eval 37° epoch: 100% 10/10 [00:01<00:00, 9.04it/s]

Training 38° epoch: 100% 91/91 [00:31<00:00, 3.14it/s]

Eval 38° epoch: 100% 10/10 [00:01<00:00, 9.27it/s]

Training 39° epoch: 100% 91/91 [00:32<00:00, 3.20it/s]

Eval 39° epoch: 90% 9/10 [00:01<00:00, 8.57it/s]

Training 40° epoch: 100% 91/91 [00:32<00:00, 3.23it/s]

Eval 40° epoch: 100% 10/10 [00:01<00:00, 8.76it/s]

Training 41° epoch: 100% 91/91 [00:31<00:00, 3.50it/s]

Eval 41° epoch: 90% 9/10 [00:01<00:00, 8.32it/s]

Training 42° epoch: 100% 91/91 [00:32<00:00, 3.32it/s]

Eval 42° epoch: 100% 10/10 [00:01<00:00, 8.88it/s]

Training 43° epoch: 100% 91/91 [00:31<00:00, 3.21it/s]

Eval 43° epoch: 100% 10/10 [00:01<00:00, 9.38it/s]

Training 44° epoch: 100% 91/91 [00:31<00:00, 3.33it/s]

Eval 44° epoch: 100% 10/10 [00:01<00:00, 8.99it/s]

Training 45° epoch: 100% 91/91 [00:32<00:00, 3.27it/s]

Eval 45° epoch: 100% 10/10 [00:01<00:00, 8.53it/s]

Training 46° epoch: 100% 91/91 [00:32<00:00, 3.37it/s]

Eval 46° epoch: 90% 9/10 [00:01<00:00, 8.55it/s]

Training 47° epoch: 100% 91/91 [00:31<00:00, 3.20it/s]

Eval 47° epoch: 90% 9/10 [00:01<00:00, 8.87it/s]

Training 48° epoch: 100% 91/91 [00:31<00:00, 3.44it/s]

Eval 48° epoch: 100% 10/10 [00:01<00:00, 8.82it/s]

Training 49° epoch: 100% 91/91 [00:31<00:00, 3.34it/s]

Eval 49° epoch: 90% 9/10 [00:01<00:00, 8.66it/s]

Training 50° epoch: 100% 91/91 [00:31<00:00, 3.20it/s]

Eval 50° epoch: 100% 10/10 [00:01<00:00, 8.54it/s]

Training 51° epoch: 100% 91/91 [00:31<00:00, 3.39it/s]

Eval 51° epoch: 100% 10/10 [00:01<00:00, 8.70it/s]

Training 52° epoch: 100% 91/91 [00:31<00:00, 3.34it/s]

Eval 52° epoch: 100% 10/10 [00:01<00:00, 8.59it/s]

Training 53° epoch: 100% 91/91 [00:32<00:00, 3.19it/s]

Eval 53° epoch: 90% 9/10 [00:01<00:00, 9.00it/s]

Training 54° epoch: 100% 91/91 [00:32<00:00, 3.12it/s]

Eval 54° epoch: 100% 10/10 [00:01<00:00, 9.11it/s]

Training 55° epoch: 100% 91/91 [00:32<00:00, 3.33it/s]

Eval 55° epoch: 90% 9/10 [00:01<00:00, 8.50it/s]

Training 56° epoch: 100% 91/91 [00:32<00:00, 3.47it/s]

Eval 56° epoch: 100% 10/10 [00:01<00:00, 9.08it/s]

Training 57° epoch: 100% 91/91 [00:32<00:00, 3.28it/s]

Eval 57° epoch: 100% 10/10 [00:01<00:00, 8.64it/s]

Training 58° epoch: 100% 91/91 [00:32<00:00, 3.23it/s]

Eval 58° epoch: 90% 9/10 [00:01<00:00, 8.46it/s]

Training 59° epoch: 100% 91/91 [00:32<00:00, 3.34it/s]

Eval 59° epoch: 90% 9/10 [00:01<00:00, 9.08it/s]

Training 60° epoch: 100% 91/91 [00:32<00:00, 3.15it/s]

Eval 60° epoch: 90% 9/10 [00:01<00:00, 8.96it/s]

Training 61° epoch: 100% 91/91 [00:32<00:00, 3.14it/s]

Eval 61° epoch: 90% 9/10 [00:01<00:00, 8.99it/s]

Training 62° epoch: 100% 91/91 [00:32<00:00, 3.07it/s]

Eval 62° epoch: 100% 10/10 [00:01<00:00, 8.56it/s]

Training 63° epoch: 100% 91/91 [00:31<00:00, 3.52it/s]

Eval 63° epoch: 90% 9/10 [00:01<00:00, 8.48it/s]

Training 64° epoch: 100% 91/91 [00:32<00:00, 3.26it/s]

Eval 64° epoch: 100% 10/10 [00:01<00:00, 8.62it/s]

Training 65° epoch: 100% 91/91 [00:32<00:00, 3.40it/s]

Eval 65° epoch: 100% 10/10 [00:01<00:00, 9.06it/s]

Training 66° epoch: 100% 91/91 [00:32<00:00, 3.34it/s]

Eval 66° epoch: 100% 10/10 [00:01<00:00, 9.12it/s]

Training 67° epoch: 100% 91/91 [00:32<00:00, 3.23it/s]

Eval 67° epoch: 90% 9/10 [00:01<00:00, 8.35it/s]

Training 68° epoch: 100% 91/91 [00:32<00:00, 3.17it/s]

Eval 68° epoch: 100% 10/10 [00:01<00:00, 8.93it/s]

Training 69° epoch: 100% 91/91 [00:32<00:00, 3.03it/s]

Eval 69° epoch: 100% 10/10 [00:01<00:00, 8.99it/s]

Training 70° epoch: 100% 91/91 [00:32<00:00, 3.21it/s]

Eval 70° epoch: 100% 10/10 [00:01<00:00, 8.85it/s]

Training 71° epoch: 100% 91/91 [00:32<00:00, 3.41it/s]

Eval 71° epoch: 100% 10/10 [00:01<00:00, 8.92it/s]

Training 72° epoch: 100% 91/91 [00:32<00:00, 3.26it/s]

Eval 72° epoch: 100% 10/10 [00:01<00:00, 8.87it/s]

Training 73° epoch: 100% 91/91 [00:32<00:00, 3.22it/s]

Eval 73° epoch: 90% 9/10 [00:01<00:00, 8.89it/s]

Training 74° epoch: 100% 91/91 [00:31<00:00, 3.37it/s]

Eval 74° epoch: 100% 10/10 [00:01<00:00, 8.55it/s]

Training 75° epoch: 100% 91/91 [00:32<00:00, 3.35it/s]

Eval 75° epoch: 100% 10/10 [00:01<00:00, 8.93it/s]

Training 76° epoch: 100% 91/91 [00:31<00:00, 3.01it/s]

Eval 76° epoch: 100% 10/10 [00:01<00:00, 8.94it/s]

Training 77° epoch: 100% 91/91 [00:32<00:00, 3.26it/s]

Eval 77° epoch: 100% 10/10 [00:01<00:00, 8.94it/s]

Training 78° epoch: 100% 91/91 [00:31<00:00, 3.44it/s]

Eval 78° epoch: 100% 10/10 [00:01<00:00, 8.80it/s]

Training 79° epoch: 100% 91/91 [00:32<00:00, 3.20it/s]

Eval 79° epoch: 100% 10/10 [00:01<00:00, 9.00it/s]

Training 80° epoch: 100% 91/91 [00:31<00:00, 3.20it/s]

Eval 80° epoch: 100% 10/10 [00:01<00:00, 9.24it/s]

Training 81° epoch: 100% 91/91 [00:31<00:00, 3.23it/s]

Eval 81° epoch: 90% 9/10 [00:01<00:00, 8.72it/s]

Training 82° epoch: 100% 91/91 [00:32<00:00, 3.33it/s]

Eval 82° epoch: 100% 10/10 [00:01<00:00, 8.59it/s]

Training 83° epoch: 100% 91/91 [00:32<00:00, 3.36it/s]

Eval 83° epoch: 100% 10/10 [00:01<00:00, 8.95it/s]

Training 84° epoch: 100% 91/91 [00:32<00:00, 3.08it/s]

Eval 84° epoch: 100% 10/10 [00:01<00:00, 8.68it/s]

Training 85° epoch: 100% 91/91 [00:32<00:00, 3.19it/s]

Eval 85° epoch: 90% 9/10 [00:01<00:00, 9.10it/s]

Training 86° epoch: 100% 91/91 [00:32<00:00, 3.31it/s]

Eval 86° epoch: 100% 10/10 [00:01<00:00, 8.75it/s]

Training 87° epoch: 100% 91/91 [00:32<00:00, 3.27it/s]

Eval 87° epoch: 100% 10/10 [00:01<00:00, 8.81it/s]

Training 88° epoch: 100% 91/91 [00:32<00:00, 3.23it/s]

Eval 88° epoch: 100% 10/10 [00:01<00:00, 9.10it/s]

Training 89° epoch: 100% 91/91 [00:31<00:00, 3.45it/s]

Eval 89° epoch: 100% 10/10 [00:01<00:00, 8.72it/s]

Training 90° epoch: 100% 91/91 [00:32<00:00, 3.36it/s]

Eval 90° epoch: 100% 10/10 [00:01<00:00, 8.93it/s]

Training 91° epoch: 100% 91/91 [00:31<00:00, 3.45it/s]

Eval 91° epoch: 100% 10/10 [00:01<00:00, 8.68it/s]

Training 92° epoch: 100% 91/91 [00:32<00:00, 3.42it/s]

Eval 92° epoch: 100% 10/10 [00:01<00:00, 8.89it/s]

Training 93° epoch: 100% 91/91 [00:32<00:00, 3.25it/s]

Eval 93° epoch: 100% 10/10 [00:01<00:00, 9.22it/s]

Training 94° epoch: 100% 91/91 [00:32<00:00, 3.33it/s]

Eval 94° epoch: 100% 10/10 [00:01<00:00, 9.49it/s]

Training 95° epoch: 100% 91/91 [00:32<00:00, 3.39it/s]

Eval 95° epoch: 100% 10/10 [00:01<00:00, 8.98it/s]

Training 96° epoch: 100% 91/91 [00:32<00:00, 3.29it/s]

Eval 96° epoch: 100% 10/10 [00:01<00:00, 8.61it/s]

Training 97° epoch: 100% 91/91 [00:32<00:00, 3.28it/s]

Eval 97° epoch: 90% 9/10 [00:01<00:00, 8.90it/s]

Training 98° epoch: 100% 91/91 [00:32<00:00, 3.15it/s]

Eval 98° epoch: 90% 9/10 [00:01<00:00, 9.02it/s]

Training 99° epoch: 100% 91/91 [00:31<00:00, 3.22it/s]

Eval 99° epoch: 90% 9/10 [00:01<00:00, 8.87it/s]

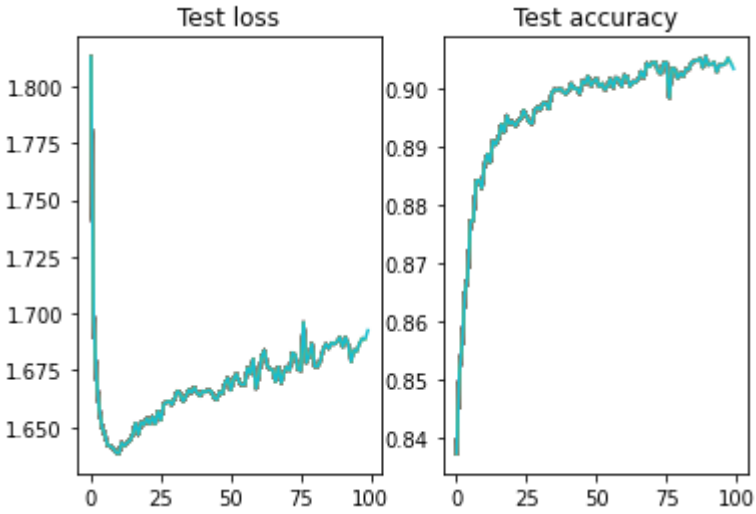
Out[11]:

```
([1.8130196571350097,
  1.6979915976524353,
  1.671607542037964,
  1.6579473495483399,
  1.6499605894088745,
  1.6468056559562683,
  1.6420766115188599,
  1.6416708827018738,
  1.6407026648521423,
  1.6388348460197448,
  1.6384066224098206,
  1.6431079983711243,
  1.641774797439575,
  1.6432533502578734,
  1.6440528988838197,
  1.646860945224762,
  1.651805865764618,
  1.646846842765808,
  1.6522601008415223,
  1.6513084769248962,
  1.653838562965393,
  1.6514398097991942,
  1.6545713543891907,
  1.6514313459396361,
  1.6569684028625489,
  1.6531487226486206,
  1.6598397016525268,
  1.6610154271125794,
  1.6614040374755858,
  1.6599397659301758,
  1.662954044342041,
  1.665926957130432,
  1.6644437432289123,
  1.661346936225891,
  1.6636237978935242,
  1.6665120601654053,
  1.665195059776306,
  1.6674719095230102,
  1.6650674700737,
  1.6638257265090943,
  1.6658761620521545,
  1.6654563546180725,
  1.6663202166557312,
  1.6658744812011719,
  1.6629128932952881,
  1.6621149063110352,
  1.6655591487884522,
  1.6648014068603516,
  1.6696534276008606,
  1.6715633034706117,
  1.6663417100906373,
  1.6710165381431579,
  1.673411738872528,
  1.6697793006896973,
  1.6686201214790344,
  1.6688056230545043,
  1.6761099815368652,
  1.6736240148544312,
  1.6800036311149598,
```

```
 1.6671542167663573,
 1.6723240494728089,
 1.6807734847068787,
 1.683759045600891,
 1.6771529912948608,
 1.675573468208313,
 1.675125205516815,
 1.6700300335884095,
 1.6764225959777832,
 1.6694926381111146,
 1.6740503191947937,
 1.6765450239181519,
 1.675078821182251,
 1.6827257752418519,
 1.6818488836288452,
 1.675560760498047,
 1.6748918056488038,
 1.6960821270942688,
 1.6784032821655273,
 1.6816184997558594,
 1.6867023587226868,
 1.6766635537147523,
 1.6766834139823914,
 1.6791154623031617,
 1.6840334296226502,
 1.6865848541259765,
 1.6846652030944824,
 1.6871513605117798,
 1.6866175293922425,
 1.6876469731330872,
 1.689544141292572,
 1.6853858470916747,
 1.6895296692848205,
 1.6865038871765137,
 1.6788011550903321,
 1.6839528560638428,
 1.6834118008613586,
 1.686885905265808,
 1.6890581607818604,
 1.6886680483818055,
 1.6926381587982178],
[0.8371599316596985,
 0.850985050201416,
 0.8560213446617126,
 0.8648101687431335,
 0.867723286151886,
 0.8771539926528931,
 0.8772527575492859,
 0.8840171694755554,
 0.8842146992683411,
 0.8828321695327759,
 0.8868315815925598,
 0.8887078762054443,
 0.8872759342193604,
 0.8909791111946106,
 0.8902878761291504,
 0.8914728760719299,
 0.893842875957489,
 0.8924604058265686,
 0.8952747583389282,
 0.8938922882080078,
```

0.8942872881889343,
0.8934972882270813,
0.8943366408348083,
0.8948304057121277,
0.8962129354476929,
0.8951760530471802,
0.8941391706466675,
0.8939909934997559,
0.8963610529899597,
0.8968054056167603,
0.8963610529899597,
0.8975460529327393,
0.8973485231399536,
0.8964598178863525,
0.8992248177528381,
0.8998172879219055,
0.8996691703796387,
0.8999654650688171,
0.8994717001914978,
0.8990272879600525,
0.8996198177337646,
0.9008048176765442,
0.8999160528182983,
0.9000148177146912,
0.8991260528564453,
0.9012492299079895,
0.9008541703224182,
0.9020885825157166,
0.9004591703414917,
0.9011504650115967,
0.90169358253479,
0.9003110527992249,
0.9008048176765442,
0.8997679352760315,
0.9004098176956177,
0.901644229888916,
0.900162935256958,
0.9018417000770569,
0.9015454649925232,
0.9003604650497437,
0.9022367000579834,
0.9019404649734497,
0.9003604650497437,
0.9011504650115967,
0.9010517001152039,
0.9015948176383972,
0.9028785824775696,
0.9018911123275757,
0.9043104648590088,
0.9037179946899414,
0.9044092297554016,
0.9045079946517944,
0.9030267000198364,
0.9025329351425171,
0.9043104648590088,
0.9044585824012756,
0.8983854055404663,
0.9034711122512817,
0.9033229947090149,
0.9017429351806641,
0.9026317000389099,

```
0.9020885825157166,
0.9033229947090149,
0.9032735824584961,
0.9045079946517944,
0.9047054648399353,
0.9048042297363281,
0.9051498770713806,
0.9036192297935486,
0.9054461121559143,
0.9040142297744751,
0.9041623473167419,
0.9045079946517944,
0.9028292298316956,
0.9040635824203491,
0.9040142297744751,
0.904211699962616,
0.9051992297172546,
0.9043598771095276,
0.9033723473548889])
```




## Uncased model

In [12]:

```
uncased_model = ItBERTUncasedPos(23)
```

Downloading: 433/433 [00:00<00:00, 17.5kB/s]

Downloading: 422M/422M [00:09<00:00, 50.8MB/s]

```
Some weights of the model checkpoint at dbmdz/bert-base-italian-uncased we
re not used when initializing BertForTokenClassification: ['cls.prediction
s.bias', 'cls.predictions.transform.dense.weight', 'cls.predictions.decode
r.weight', 'cls.seq_relationship.weight', 'cls.predictions.transform.Layer
Norm.bias', 'cls.predictions.transform.LayerNorm.weight', 'cls.prediction
s.transform.dense.bias', 'cls.seq_relationship.bias']
- This IS expected if you are initializing BertForTokenClassification from
the checkpoint of a model trained on another task or with another architec
ture (e.g. initializing a BertForSequenceClassification model from a BertF
orPreTraining model).
- This IS NOT expected if you are initializing BertForTokenClassification
from the checkpoint of a model that you expect to be exactly identical (in
itializing a BertForSequenceClassification model from a BertForSequenceCla
ssification model).
Some weights of BertForTokenClassification were not initialized from the m
odel checkpoint at dbmdz/bert-base-italian-uncased and are newly initializ
ed: ['classifier.bias', 'classifier.weight']
You should probably TRAIN this model on a down-stream task to be able to u
se it for predictions and inference.
```

In [13]:

```
tokenizer = ItBERTUncasedPos.tokenizer()
ds_train = TWITADS('resampled_train',
                   lambda w:[w],
                   transform=lambda a,b: tokenize_and_align_labels(tokenizer,a,b)
)
ds_val = TWITADS('resampled_validation',
                   lambda w:[w],
                   transform=lambda a,b: tokenize_and_align_labels(tokenizer,a,b)
)
```

Downloading: 59.0/59.0 [00:00<00:00, 2.39kB/s]

Downloading: 237k/237k [00:00<00:00, 799kB/s]

In [14]:

```
dl_train = DataLoader(ds_train, shuffle=True, batch_size=BATCH_SIZE, collate_fn=collate
_fn)
dl_val  = DataLoader(ds_val, shuffle=True, batch_size=BATCH_SIZE, collate_fn=collate_fn
)
```

In [15]:

```
torch.manual_seed(42)
train_model(uncased_model,dl_train,dl_val,cuda=torch.cuda.is_available(), lr=2e-5, epoc
hs=100, show_plots=True,save_dir=out_dir/"uncased_model")
```

Loss:1.6130800724029541 Accurancy:0.9009537696838379: 100% 100/100 [51:58<00:00, 31.16s/it]

Training 0° epoch: 100% 91/91 [00:31<00:00, 3.48it/s]

Eval 0° epoch: 100% 10/10 [00:01<00:00, 7.67it/s]

Training 1° epoch: 100% 91/91 [00:28<00:00, 3.73it/s]

Eval 1° epoch: 100% 10/10 [00:01<00:00, 10.08it/s]

Training 2° epoch: 100% 91/91 [00:28<00:00, 3.55it/s]

Eval 2° epoch: 90% 9/10 [00:00<00:00, 9.56it/s]

Training 3° epoch: 100% 91/91 [00:28<00:00, 3.87it/s]

Eval 3° epoch: 90% 9/10 [00:00<00:00, 9.60it/s]

Training 4° epoch: 100% 91/91 [00:28<00:00, 3.81it/s]

Eval 4° epoch: 90% 9/10 [00:00<00:00, 9.68it/s]

Training 5° epoch: 100% 91/91 [00:28<00:00, 3.85it/s]

Eval 5° epoch: 100% 10/10 [00:01<00:00, 10.02it/s]

Training 6° epoch: 100% 91/91 [00:28<00:00, 3.86it/s]

Eval 6° epoch: 100% 10/10 [00:01<00:00, 10.06it/s]

Training 7° epoch: 100% 91/91 [00:28<00:00, 3.42it/s]

Eval 7° epoch: 100% 10/10 [00:01<00:00, 9.93it/s]

Training 8° epoch: 100% 91/91 [00:28<00:00, 3.95it/s]

Eval 8° epoch: 90% 9/10 [00:00<00:00, 9.35it/s]

Training 9° epoch: 100% 91/91 [00:28<00:00, 3.82it/s]

Eval 9° epoch: 90% 9/10 [00:00<00:00, 9.81it/s]

Training 10° epoch: 100% 91/91 [00:28<00:00, 3.46it/s]

Eval 10° epoch: 100% 10/10 [00:01<00:00, 10.03it/s]

Training 11° epoch: 100% 91/91 [00:28<00:00, 3.60it/s]

Eval 11° epoch: 100% 10/10 [00:01<00:00, 9.88it/s]

Training 12° epoch: 100% 91/91 [00:28<00:00, 3.76it/s]

Eval 12° epoch: 90% 9/10 [00:00<00:00, 9.34it/s]

Training 13° epoch: 100% 91/91 [00:28<00:00, 3.61it/s]

Eval 13° epoch: 100% 10/10 [00:01<00:00, 10.00it/s]

Training 14° epoch: 100% 91/91 [00:28<00:00, 3.57it/s]

Eval 14° epoch: 90% 9/10 [00:00<00:00, 9.70it/s]

Training 15° epoch: 100% 91/91 [00:28<00:00, 3.74it/s]

Eval 15° epoch: 100% 10/10 [00:01<00:00, 9.99it/s]

Training 16° epoch: 100% 91/91 [00:28<00:00, 3.67it/s]

Eval 16° epoch: 90% 9/10 [00:00<00:00, 9.51it/s]

Training 17° epoch: 100% 91/91 [00:28<00:00, 3.84it/s]

Eval 17° epoch: 90% 9/10 [00:00<00:00, 9.64it/s]

Training 18° epoch: 100% 91/91 [00:28<00:00, 3.60it/s]

Eval 18° epoch: 90% 9/10 [00:00<00:00, 9.57it/s]

Training 19° epoch: 100% 91/91 [00:28<00:00, 3.73it/s]

Eval 19° epoch: 90% 9/10 [00:00<00:00, 9.83it/s]

Training 20° epoch: 100% 91/91 [00:28<00:00, 3.65it/s]

Eval 20° epoch: 100% 10/10 [00:01<00:00, 10.10it/s]

Training 21° epoch: 100% 91/91 [00:28<00:00, 3.41it/s]

Eval 21° epoch: 90% 9/10 [00:00<00:00, 9.72it/s]

Training 22° epoch: 100% 91/91 [00:28<00:00, 3.70it/s]

Eval 22° epoch: 100% 10/10 [00:01<00:00, 9.94it/s]

Training 23° epoch: 100% 91/91 [00:28<00:00, 3.71it/s]

Eval 23° epoch: 90% 9/10 [00:00<00:00, 9.58it/s]

Training 24° epoch: 100% 91/91 [00:28<00:00, 3.65it/s]

Eval 24° epoch: 90% 9/10 [00:00<00:00, 9.73it/s]

Training 25° epoch: 100% 91/91 [00:28<00:00, 3.39it/s]

Eval 25° epoch: 90% 9/10 [00:00<00:00, 9.89it/s]

Training 26° epoch: 100% 91/91 [00:28<00:00, 3.85it/s]

Eval 26° epoch: 90% 9/10 [00:00<00:00, 9.72it/s]

Training 27° epoch: 100% 91/91 [00:28<00:00, 3.84it/s]

Eval 27° epoch: 100% 10/10 [00:01<00:00, 10.08it/s]

Training 28° epoch: 100% 91/91 [00:28<00:00, 3.70it/s]

Eval 28° epoch: 90% 9/10 [00:00<00:00, 10.01it/s]

Training 29° epoch: 100% 91/91 [00:28<00:00, 3.70it/s]

Eval 29° epoch: 90% 9/10 [00:00<00:00, 9.89it/s]

Training 30° epoch: 100% 91/91 [00:28<00:00, 3.72it/s]

Eval 30° epoch: 90% 9/10 [00:00<00:00, 9.91it/s]

Training 31° epoch: 100% 91/91 [00:28<00:00, 3.52it/s]

Eval 31° epoch: 90% 9/10 [00:00<00:00, 8.99it/s]

Training 32° epoch: 100% 91/91 [00:28<00:00, 3.92it/s]

Eval 32° epoch: 90% 9/10 [00:00<00:00, 9.43it/s]

Training 33° epoch: 100% 91/91 [00:28<00:00, 3.56it/s]

Eval 33° epoch: 90% 9/10 [00:00<00:00, 9.88it/s]

Training 34° epoch: 100% 91/91 [00:28<00:00, 3.58it/s]

Eval 34° epoch: 90% 9/10 [00:00<00:00, 9.53it/s]

Training 35° epoch: 100% 91/91 [00:28<00:00, 3.65it/s]

Eval 35° epoch: 100% 10/10 [00:01<00:00, 10.08it/s]

Training 36° epoch: 100% 91/91 [00:28<00:00, 3.77it/s]

Eval 36° epoch: 90% 9/10 [00:00<00:00, 9.83it/s]

Training 37° epoch: 100% 91/91 [00:28<00:00, 3.82it/s]

Eval 37° epoch: 90% 9/10 [00:00<00:00, 9.90it/s]

Training 38° epoch: 100% 91/91 [00:28<00:00, 3.37it/s]

Eval 38° epoch: 100% 10/10 [00:01<00:00, 10.05it/s]

Training 39° epoch: 100% 91/91 [00:28<00:00, 3.58it/s]

Eval 39° epoch: 90% 9/10 [00:00<00:00, 9.89it/s]

Training 40° epoch: 100% 91/91 [00:28<00:00, 3.79it/s]

Eval 40° epoch: 90% 9/10 [00:00<00:00, 9.70it/s]

Training 41° epoch: 100% 91/91 [00:28<00:00, 3.82it/s]

Eval 41° epoch: 90% 9/10 [00:00<00:00, 9.83it/s]

Training 42° epoch: 100% 91/91 [00:28<00:00, 3.74it/s]

Eval 42° epoch: 100% 10/10 [00:01<00:00, 10.06it/s]

Training 43° epoch: 100% 91/91 [00:28<00:00, 3.79it/s]

Eval 43° epoch: 100% 10/10 [00:01<00:00, 10.02it/s]

Training 44° epoch: 100% 91/91 [00:28<00:00, 3.72it/s]

Eval 44° epoch: 100% 10/10 [00:01<00:00, 9.88it/s]

Training 45° epoch: 100% 91/91 [00:28<00:00, 3.65it/s]

Eval 45° epoch: 100% 10/10 [00:01<00:00, 9.89it/s]

Training 46° epoch: 100% 91/91 [00:28<00:00, 3.74it/s]

Eval 46° epoch: 90% 9/10 [00:00<00:00, 9.66it/s]

Training 47° epoch: 100% 91/91 [00:28<00:00, 3.32it/s]

Eval 47° epoch: 100% 10/10 [00:01<00:00, 9.99it/s]

Training 48° epoch: 100% 91/91 [00:28<00:00, 3.76it/s]

Eval 48° epoch: 90% 9/10 [00:00<00:00, 9.62it/s]

Training 49° epoch: 100% 91/91 [00:28<00:00, 3.85it/s]

Eval 49° epoch: 90% 9/10 [00:00<00:00, 9.76it/s]

Training 50° epoch: 100% 91/91 [00:28<00:00, 3.57it/s]

Eval 50° epoch: 100% 10/10 [00:01<00:00, 10.12it/s]

Training 51° epoch: 100% 91/91 [00:28<00:00, 3.81it/s]

Eval 51° epoch: 100% 10/10 [00:01<00:00, 9.83it/s]

Training 52° epoch: 100% 91/91 [00:28<00:00, 3.82it/s]

Eval 52° epoch: 90% 9/10 [00:00<00:00, 9.54it/s]

Training 53° epoch: 100% 91/91 [00:28<00:00, 3.69it/s]

Eval 53° epoch: 100% 10/10 [00:01<00:00, 10.00it/s]

Training 54° epoch: 100% 91/91 [00:28<00:00, 3.35it/s]

Eval 54° epoch: 100% 10/10 [00:01<00:00, 9.91it/s]

Training 55° epoch: 100% 91/91 [00:28<00:00, 3.83it/s]

Eval 55° epoch: 90% 9/10 [00:00<00:00, 9.90it/s]

Training 56° epoch: 100% 91/91 [00:28<00:00, 3.77it/s]

Eval 56° epoch: 90% 9/10 [00:00<00:00, 9.66it/s]

Training 57° epoch: 100% 91/91 [00:28<00:00, 3.73it/s]

Eval 57° epoch: 90% 9/10 [00:00<00:00, 9.89it/s]

Training 58° epoch: 100% 91/91 [00:28<00:00, 3.44it/s]

Eval 58° epoch: 90% 9/10 [00:00<00:00, 9.95it/s]

Training 59° epoch: 100% 91/91 [00:28<00:00, 3.54it/s]

Eval 59° epoch: 90% 9/10 [00:00<00:00, 10.16it/s]

Training 60° epoch: 100% 91/91 [00:28<00:00, 3.59it/s]

Eval 60° epoch: 100% 10/10 [00:01<00:00, 10.06it/s]

Training 61° epoch: 100% 91/91 [00:28<00:00, 3.40it/s]

Eval 61° epoch: 90% 9/10 [00:00<00:00, 9.88it/s]

Training 62° epoch: 100% 91/91 [00:28<00:00, 3.50it/s]

Eval 62° epoch: 90% 9/10 [00:00<00:00, 9.52it/s]

Training 63° epoch: 100% 91/91 [00:28<00:00, 3.75it/s]

Eval 63° epoch: 100% 10/10 [00:01<00:00, 9.99it/s]

Training 64° epoch: 100% 91/91 [00:28<00:00, 3.59it/s]

Eval 64° epoch: 90% 9/10 [00:00<00:00, 9.50it/s]

Training 65° epoch: 100% 91/91 [00:28<00:00, 3.79it/s]

Eval 65° epoch: 100% 10/10 [00:01<00:00, 9.99it/s]

Training 66° epoch: 100% 91/91 [00:28<00:00, 3.54it/s]

Eval 66° epoch: 100% 10/10 [00:01<00:00, 10.03it/s]

Training 67° epoch: 100% 91/91 [00:28<00:00, 3.67it/s]

Eval 67° epoch: 90% 9/10 [00:00<00:00, 9.69it/s]

Training 68° epoch: 100% 91/91 [00:28<00:00, 3.65it/s]

Eval 68° epoch: 90% 9/10 [00:00<00:00, 9.58it/s]

Training 69° epoch: 100% 91/91 [00:28<00:00, 3.72it/s]

Eval 69° epoch: 100% 10/10 [00:01<00:00, 10.02it/s]

Training 70° epoch: 100% 91/91 [00:28<00:00, 3.52it/s]

Eval 70° epoch: 90% 9/10 [00:00<00:00, 9.94it/s]

Training 71° epoch: 100% 91/91 [00:28<00:00, 3.77it/s]

Eval 71° epoch: 100% 10/10 [00:01<00:00, 9.78it/s]

Training 72° epoch: 100% 91/91 [00:28<00:00, 3.65it/s]

Eval 72° epoch: 100% 10/10 [00:01<00:00, 9.80it/s]

Training 73° epoch: 100% 91/91 [00:28<00:00, 3.73it/s]

Eval 73° epoch: 90% 9/10 [00:00<00:00, 9.79it/s]

Training 74° epoch: 100% 91/91 [00:28<00:00, 3.66it/s]

Eval 74° epoch: 90% 9/10 [00:00<00:00, 10.15it/s]

Training 75° epoch: 100% 91/91 [00:28<00:00, 3.77it/s]

Eval 75° epoch: 100% 10/10 [00:01<00:00, 9.99it/s]

Training 76° epoch: 100% 91/91 [00:28<00:00, 3.39it/s]

Eval 76° epoch: 100% 10/10 [00:01<00:00, 9.85it/s]

Training 77° epoch: 100% 91/91 [00:28<00:00, 3.52it/s]

Eval 77° epoch: 90% 9/10 [00:00<00:00, 10.06it/s]

Training 78° epoch: 100% 91/91 [00:28<00:00, 3.81it/s]

Eval 78° epoch: 100% 10/10 [00:01<00:00, 9.90it/s]

Training 79° epoch: 100% 91/91 [00:28<00:00, 3.51it/s]

Eval 79° epoch: 90% 9/10 [00:00<00:00, 9.88it/s]

Training 80° epoch: 100% 91/91 [00:28<00:00, 3.70it/s]

Eval 80° epoch: 100% 10/10 [00:01<00:00, 9.97it/s]

Training 81° epoch: 100% 91/91 [00:28<00:00, 3.45it/s]

Eval 81° epoch: 90% 9/10 [00:00<00:00, 9.83it/s]

Training 82° epoch: 100% 91/91 [00:28<00:00, 3.65it/s]

Eval 82° epoch: 90% 9/10 [00:00<00:00, 9.45it/s]

Training 83° epoch: 100% 91/91 [00:28<00:00, 3.74it/s]

Eval 83° epoch: 100% 10/10 [00:01<00:00, 9.86it/s]

Training 84° epoch: 100% 91/91 [00:28<00:00, 3.14it/s]

Eval 84° epoch: 100% 10/10 [00:01<00:00, 9.79it/s]

Training 85° epoch: 100% 91/91 [00:28<00:00, 3.68it/s]

Eval 85° epoch: 90% 9/10 [00:00<00:00, 9.75it/s]

Training 86° epoch: 100% 91/91 [00:28<00:00, 3.64it/s]

Eval 86° epoch: 100% 10/10 [00:01<00:00, 9.90it/s]

Training 87° epoch: 100% 91/91 [00:28<00:00, 3.76it/s]

Eval 87° epoch: 90% 9/10 [00:00<00:00, 9.93it/s]

Training 88° epoch: 100% 91/91 [00:28<00:00, 3.71it/s]

Eval 88° epoch: 90% 9/10 [00:00<00:00, 9.58it/s]

Training 89° epoch: 100% 91/91 [00:28<00:00, 3.75it/s]

Eval 89° epoch: 100% 10/10 [00:01<00:00, 10.04it/s]

Training 90° epoch: 100% 91/91 [00:28<00:00, 3.56it/s]

Eval 90° epoch: 100% 10/10 [00:01<00:00, 9.97it/s]

Training 91° epoch: 100% 91/91 [00:28<00:00, 3.85it/s]

Eval 91° epoch: 100% 10/10 [00:01<00:00, 9.82it/s]

Training 92° epoch: 100% 91/91 [00:28<00:00, 3.79it/s]

Eval 92° epoch: 100% 10/10 [00:01<00:00, 9.95it/s]

Training 93° epoch: 100% 91/91 [00:28<00:00, 3.56it/s]

Eval 93° epoch: 90% 9/10 [00:00<00:00, 10.07it/s]

Training 94° epoch: 100% 91/91 [00:28<00:00, 3.84it/s]

Eval 94° epoch: 100% 10/10 [00:01<00:00, 10.19it/s]

Training 95° epoch: 100% 91/91 [00:28<00:00, 3.76it/s]

Eval 95° epoch: 90% 9/10 [00:00<00:00, 9.75it/s]

Training 96° epoch: 100% 91/91 [00:28<00:00, 3.70it/s]

Eval 96° epoch: 100% 10/10 [00:01<00:00, 10.19it/s]

Training 97° epoch: 100% 91/91 [00:28<00:00, 3.44it/s]

Eval 97° epoch: 90% 9/10 [00:00<00:00, 9.89it/s]

Training 98° epoch: 100% 91/91 [00:28<00:00, 3.65it/s]

Eval 98° epoch: 100% 10/10 [00:01<00:00, 10.05it/s]

Training 99° epoch: 100% 91/91 [00:28<00:00, 3.55it/s]

Eval 99° epoch: 100% 10/10 [00:01<00:00, 10.15it/s]

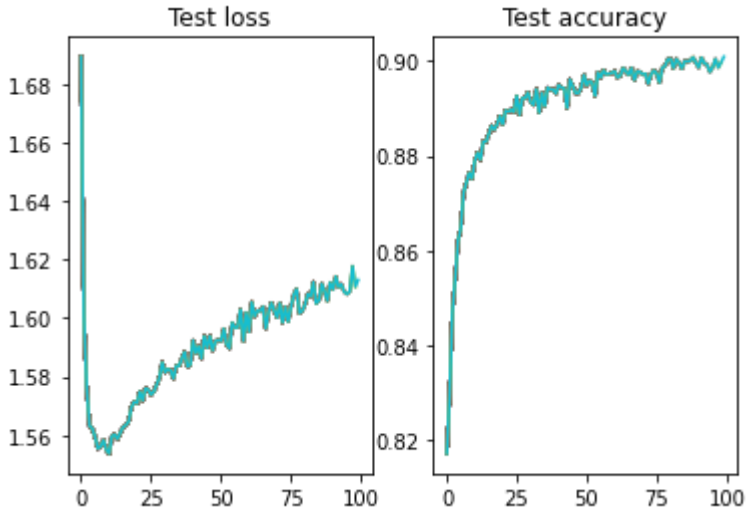
Out[15]:

```
([1.6894435286521912,
  1.600848376750946,
  1.5762154579162597,
  1.563267183303833,
  1.5622152447700501,
  1.5593457221984863,
  1.5553905487060546,
  1.5560644626617433,
  1.5582425951957704,
  1.5549959540367126,
  1.5536412239074706,
  1.5596275329589844,
  1.5601265549659729,
  1.558625626564026,
  1.560780692100525,
  1.5626933336257935,
  1.5633742451667785,
  1.5646281003952027,
  1.5694751501083375,
  1.5714045524597169,
  1.570802426338196,
  1.5748461723327636,
  1.5718544602394104,
  1.5759273290634155,
  1.5755036115646361,
  1.573784554004669,
  1.5748224973678588,
  1.5772189497947693,
  1.5796263694763184,
  1.5847679376602173,
  1.5814448356628419,
  1.5825689911842347,
  1.582095241546631,
  1.5793493390083313,
  1.583695363998413,
  1.583694589138031,
  1.5860363125801087,
  1.5883968234062196,
  1.5834315061569213,
  1.5839833855628966,
  1.5922879695892334,
  1.5880231022834779,
  1.590298855304718,
  1.5862347960472107,
  1.5938259482383728,
  1.5898569107055665,
  1.5938215374946594,
  1.589030110836029,
  1.5919887661933898,
  1.5925588607788086,
  1.5921477317810058,
  1.5963288068771362,
  1.591024684906006,
  1.5896793365478517,
  1.5984149336814881,
  1.5967358469963073,
  1.5968709111213684,
  1.6029447436332702,
  1.5922068238258362,
```

1.6018510937690735,
1.5953925371170044,
1.6053812146186828,
1.600671124458313,
1.6025359749794006,
1.6023690581321717,
1.603845226764679,
1.5965078711509704,
1.6020601034164428,
1.6051797270774841,
1.6026902079582215,
1.600213897228241,
1.6050945043563842,
1.5984504461288451,
1.6037792682647705,
1.5987818479537963,
1.603899931907654,
1.6079256772994994,
1.609902834892273,
1.6017167687416076,
1.6024322748184203,
1.604574191570282,
1.60822594165802,
1.6073840379714965,
1.6123082637786865,
1.6050231218338014,
1.606088638305664,
1.6117368340492249,
1.6091171264648438,
1.6059471249580384,
1.6121357083320618,
1.609833562374115,
1.6142178773880005,
1.6105773091316222,
1.611610770225525,
1.6090330600738525,
1.6082262516021728,
1.6090210795402526,
1.6177205324172974,
1.6106953978538514,
1.6130800724029541],
[0.8170003294944763,
0.829367995262146,
0.8468713760375977,
0.8538936972618103,
0.8620165586471558,
0.8646368384361267,
0.8721832036972046,
0.8742794394493103,
0.876375675201416,
0.8751703500747681,
0.8789435029029846,
0.8806204795837402,
0.8792055249214172,
0.8829787373542786,
0.8829787373542786,
0.8846033215522766,
0.8863326907157898,
0.8855990171432495,
0.88664710521698,
0.8883764743804932,

0.88664710521698,
0.8895817995071411,
0.8893197774887085,
0.8897390365600586,
0.8892673850059509,
0.8926737308502197,
0.8886908888816833,
0.8922020792961121,
0.8922020792961121,
0.8933550119400024,
0.8913111686706543,
0.8907347321510315,
0.8940362930297852,
0.8891101479530334,
0.8933026194572449,
0.8904727101325989,
0.8944031000137329,
0.8939838409423828,
0.8940362930297852,
0.893197774887085,
0.8949795365333557,
0.8944031000137329,
0.8942458629608154,
0.8901058435440063,
0.8961324691772461,
0.8947175145149231,
0.8932501673698425,
0.8932501673698425,
0.8944031000137329,
0.8970757722854614,
0.8949271440505981,
0.8958180546760559,
0.8953987956047058,
0.8926737308502197,
0.8975998163223267,
0.8965517282485962,
0.8980714678764343,
0.8972330093383789,
0.897337794303894,
0.8975998163223267,
0.8978618383407593,
0.8965517282485962,
0.8969709873199463,
0.898543119430542,
0.8978618383407593,
0.897128164768219,
0.898543119430542,
0.896027684211731,
0.8978094458580017,
0.8977570533752441,
0.8972330093383789,
0.897337794303894,
0.8952940106391907,
0.8972330093383789,
0.8980190753936768,
0.897442638874054,
0.8955036401748657,
0.898543119430542,
0.8990147709846497,
0.9004297256469727,
0.9004821181297302,

```
  0.898543119430542,
  0.9004297256469727,
  0.8996960520744324,
  0.8984907269477844,
  0.9000104665756226,
  0.899958074092865,
  0.9001153111457825,
  0.9007965922355652,
  0.8996960520744324,
  0.8987003564834595,
  0.9007441401481628,
  0.8995912671089172,
  0.8992767930030823,
  0.8977570533752441,
  0.8987003564834595,
  0.9004821181297302,
  0.8987003564834595,
  0.8996960520744324,
  0.9009537696838379])
```




In [16]:

```
#remove repo from saved output
!rm -rf /kagle/working/PoS-Tagging
```

```
huggingface/tokenizers: The current process just got forked, after paralle
lism has already been used. Disabling parallelism to avoid deadlocks...
To disable this warning, you can either:
        - Avoid using `tokenizers` before the fork if possible
        - Explicitly set the environment variable TOKENIZERS_PARALLELISM=
(true | false)
```