

```
!python3 example-doc.py
```

```
Seperating Preamble took 0.023373126983642578
Creating doc for jdgement took 8.093775987625122
78
Creating doc for preamble took 0.3533329963684082
/content/drive/MyDrive/legal-roberta-base/legal_ner.py:39: UserWarning: [W101] Skipping Doc custom extension 'trf_data' while merging docs.
  combined_doc = spacy.tokens.Doc.from_docs([doc_preamble, doc_judgment])
Combining took 0.00970315933227539
```

!python3 example-sent.py

```
Seperating Preamble took 0.023441553115844727
/content/drive/MyDrive/legal-roberta-base/data_preparation.py:261: UserWarning: [W101] Skipping Doc custom extension 'trf_data' while merging docs.
  combined_docs=spacy.tokens.Doc.from_docs(docs)
Creating doc for jdgement took 9.762348175048828
99
Creating doc for preamble took 0.36974358558654785
/content/drive/MyDrive/legal-roberta-base/legal_ner.py:39: UserWarning: [W101] Skipping Doc custom extension 'trf_data' while merging docs.
  combined_doc = spacy.tokens.Doc.from_docs([doc_preamble, doc_judgment])
Combining took 0.010306119918823242
```

!python -m spacy train training/config.cfg --output ./output --paths.train training/Combined_Data_Judgement_Preamble_spacy_format/train.spacy --paths.dev training/Combined_Data_Judgement_Preamble_spacy_format/dev.spacy

```
i Saving to output directory: output
i Using GPU: 0
```

===== Initializing pipeline =====

```
[2022-12-12 14:54:14,568] [INFO] Set up nlp object from config
INFO:spacy:Set up nlp object from config
[2022-12-12 14:54:14,577] [INFO] Pipeline: ['transformer', 'ner']
INFO:spacy:Pipeline: ['transformer', 'ner']
[2022-12-12 14:54:14,580] [INFO] Created vocabulary
INFO:spacy:Created vocabulary
[2022-12-12 14:54:14,581] [INFO] Finished initializing nlp object
INFO:spacy:Finished initializing nlp object
Some weights of the model checkpoint at saibo/legal-roberta-base were not used when initializing RobertaModel: ['lm_head.decoder.bias', 'lm_head.decoder.weight', 'lm_head.layer_norm.weight', '
- This IS expected if you are initializing RobertaModel from the checkpoint of a model trained on another task or with another architecture (e.g. initializing a BertForSequenceClassification m
- This IS NOT expected if you are initializing RobertaModel from the checkpoint of a model that you expect to be exactly identical (initializing a BertForSequenceClassification model from a Be
Some weights of RobertaModel were not initialized from the model checkpoint at saibo/legal-roberta-base and are newly initialized: ['roberta.pooler.dense.bias', 'roberta.pooler.dense.weight']
You should probably TRAIN this model on a down-stream task to be able to use it for predictions and inference.
[2022-12-12 14:54:50,915] [INFO] Initialized pipeline components: ['transformer', 'ner']
INFO:spacy:Initialized pipeline components: ['transformer', 'ner']
✓ Initialized pipeline
```

===== Training pipeline =====

```
i Pipeline: ['transformer', 'ner']
i Initial learn rate: 0.0
```

E	#	LOSS	TRANS...	LOSS NER	ENTS_F	ENTS_P	ENTS_R	SCORE
---	---	------	----------	----------	--------	--------	--------	-------

Token indices sequence length is longer than the specified maximum sequence length for this model (514 > 512). Running this sequence through the model will result in indexing errors

0	0	13868.19	704.23	0.10	0.07	0.20	0.00
2	1000	1338703.62	274560.47	74.73	74.65	74.80	0.75
4	2000	47594.88	149375.47	85.28	86.30	84.28	0.85

⚠ Aborting and saving the final best model. Encountered exception: OutOfMemoryError('CUDA out of memory. Tried to allocate 216.00 MiB (GPU 0; 14.76 GiB total capacity; 12.89 GiB already allocated; 177.75 MiB free; 13.61 GiB reserved in total by PyTorch) If reserved memory is >> allocated memory try setting max_split_size_mb to avoid fragmentation. See documentation for Memory Management and PYTORCH_CUDA_ALLOC_CONF')

Traceback (most recent call last):
File "/usr/lib/python3.8/runpy.py", line 194, in _run_module_as_main
return _run_code(code, main_globals, None,
File "/usr/lib/python3.8/runpy.py", line 87, in _run_code
exec(code, run_globals)
File "/usr/local/lib/python3.8/dist-packages/spacy/__main__.py", line 4, in <module>
setup_cli()
File "/usr/local/lib/python3.8/dist-packages/spacy/cli/_util.py", line 71, in setup_cli
command(prog_name=COMMAND)
File "/usr/local/lib/python3.8/dist-packages/click/core.py", line 829, in __call__
return self.main(*args, **kwargs)
File "/usr/local/lib/python3.8/dist-packages/click/core.py", line 782, in main
rv = self.invoke(ctx)
File "/usr/local/lib/python3.8/dist-packages/click/core.py", line 1259, in invoke
return _process_result(sub_ctx.command.invoke(sub_ctx))
File "/usr/local/lib/python3.8/dist-packages/click/core.py", line 1066, in invoke
return ctx.invoke(self.callback, **ctx.params)
File "/usr/local/lib/python3.8/dist-packages/click/core.py", line 610, in invoke
return callback(*args, **kwargs)
File "/usr/local/lib/python3.8/dist-packages/typer/main.py", line 532, in wrapper
return callback(**use_params) # type: ignore
File "/usr/local/lib/python3.8/dist-packages/spacy/cli/train.py", line 45, in train_cli