xml-roberta-base

Layers	12
Hidden size	768
Parameters	110 M
Attention heads	12

xml-roberta-large

Layers	24
Hidden size	1.024
Parameters	355 M
Attention heads	16

bert-base-multilingual-cased

Layers	12
Hidden size	768
Parameters	110 M
Attention heads	12