

**Input sentence**

**Pretrained (Glove)  
Embedding Layer**

**Bert Embedder**

**Linear Block**  
- Linear Layer  
- ReLU

**Linear Block**  
- Linear Layer  
- ReLU

**Attention Layer**  
- 12 heads

**Attention Layer**  
- 24 heads

**Concatenation**

**BiLSTM (x2)**

**Prediction Block**  
- Linear Layer  
- Softmax

