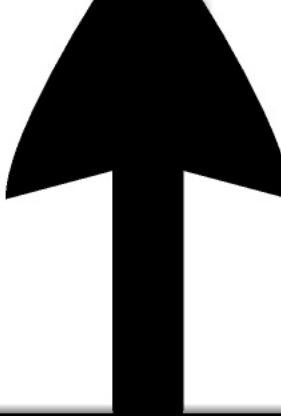




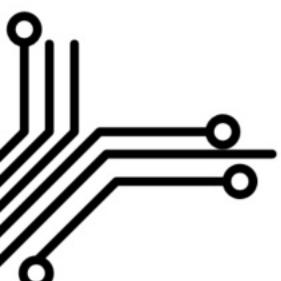
Adrian Gheorghiу



Cristian Bassotto



NLI Project Report



Giuseppe Cianci



Nikolay Kormushev





Task 1 Domain Classification/Dialogue act prediction

1

2





Task 1 Domain Classification/Dialogue act prediction



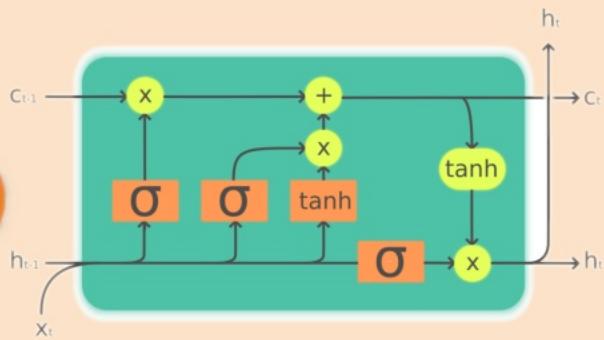
BERT



Task 1 Domain Classification/Dialogue act prediction



BERT



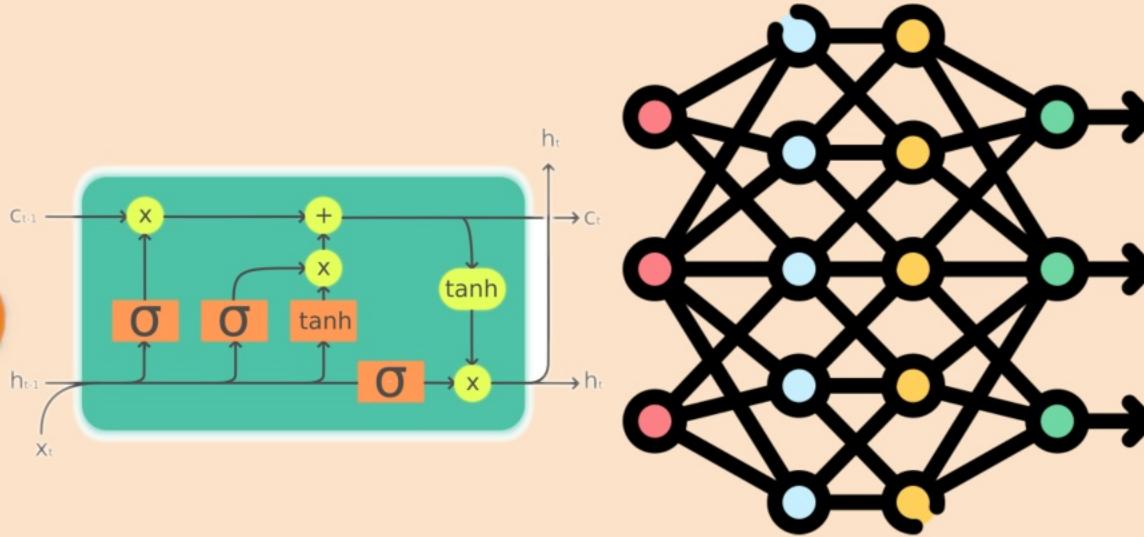
LSTM



Task 1 Domain Classification/Dialogue act prediction



BERT



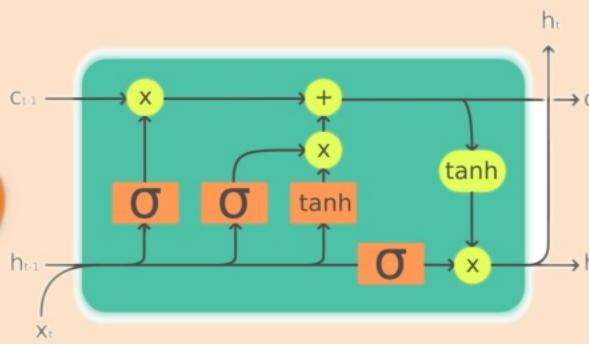
Multilayer Perceptron
(MLP) with BatchNorm
and Dropout Layers



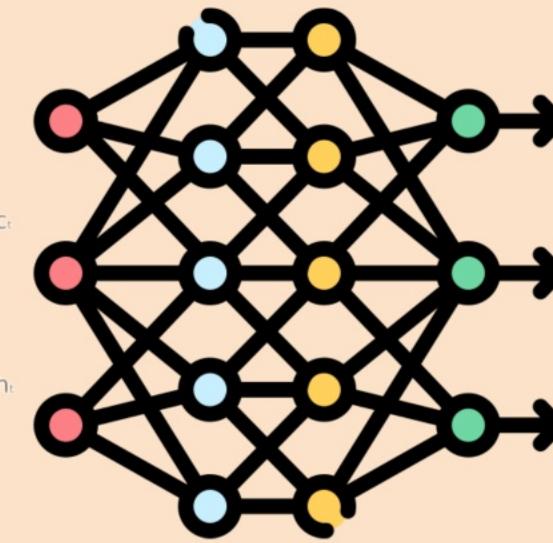
Task 1 Domain Classification/Dialogue act prediction



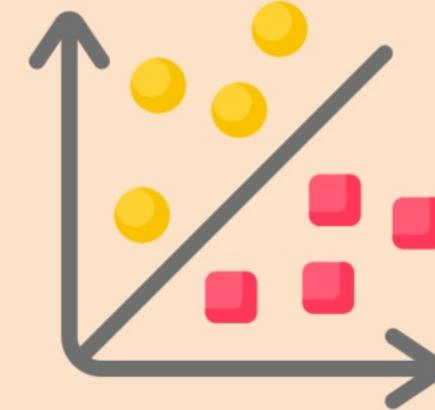
BERT



LSTM



Multilayer Perceptron
(MLP) with BatchNorm
and Dropout Layers



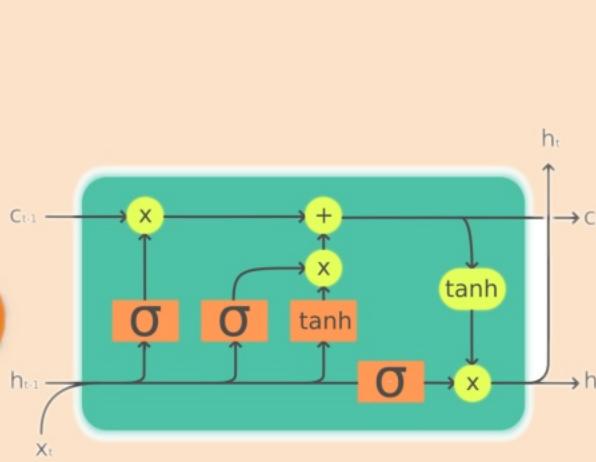
Support
Vector
Machines



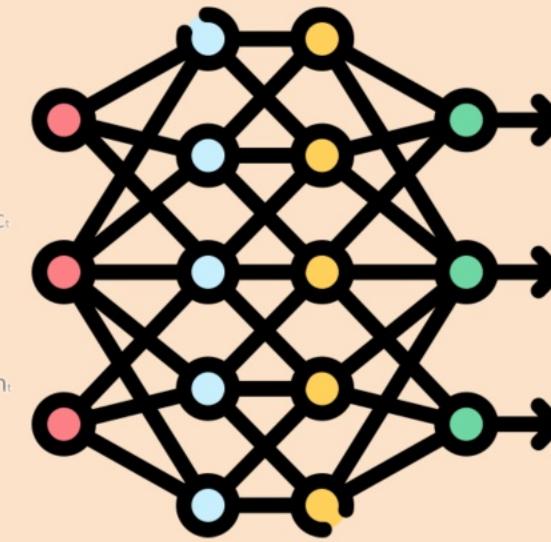
Task 1 Domain Classification/Dialogue act prediction



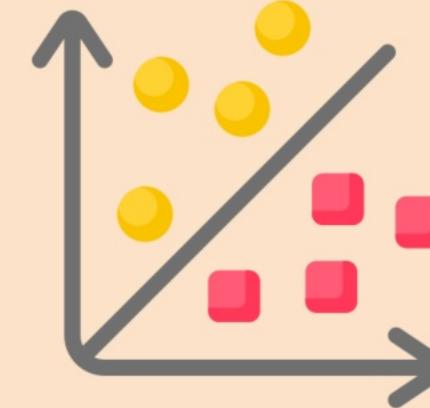
BERT



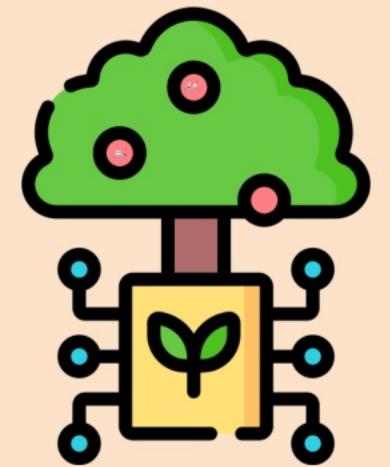
LSTM



Multilayer Perceptron
(MLP) with BatchNorm
and Dropout Layers



Support
Vector
Machines



Extreme Gradient
Boosted Classifier
(XGBClassifier)

Methods tried XGBClassifier



Word unigram frequencies + Part of Speech (POS) unigram frequencies: 59.4%

Term Frequency-Inverse Document Frequency (TF-IDF): 60.7%

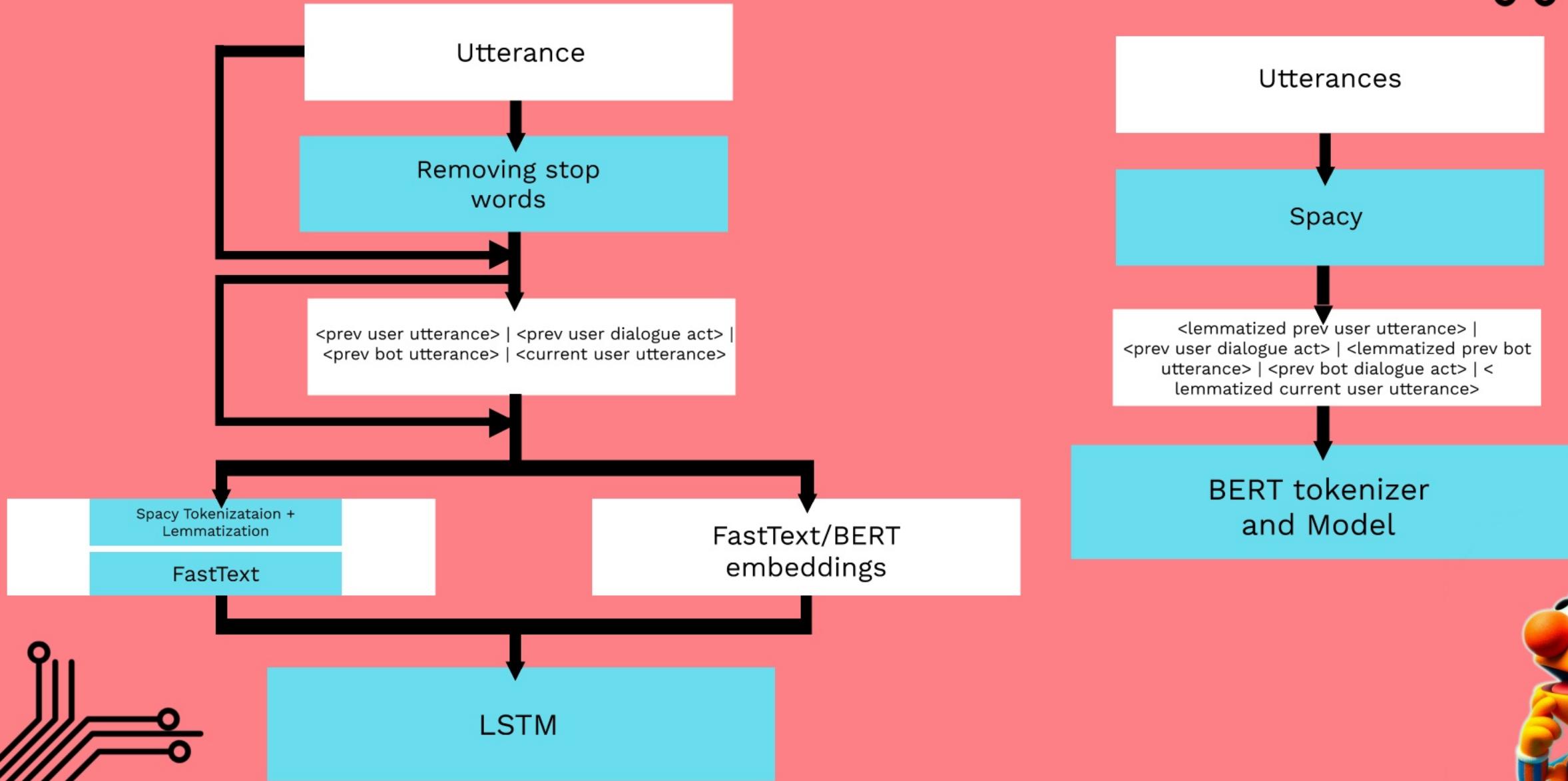
TF-IDF + POS unigram frequencies: 61.6%

TF-IDF + Spacy embeddings: 59.2%

lemmatization + TF-IDF + fastText embeddings + top-k feature selection based on Analysis of Variance (ANOVA) score: 72.7%



Preprocessing



Methods tried LSTM



fastText embeddings: 77.7%

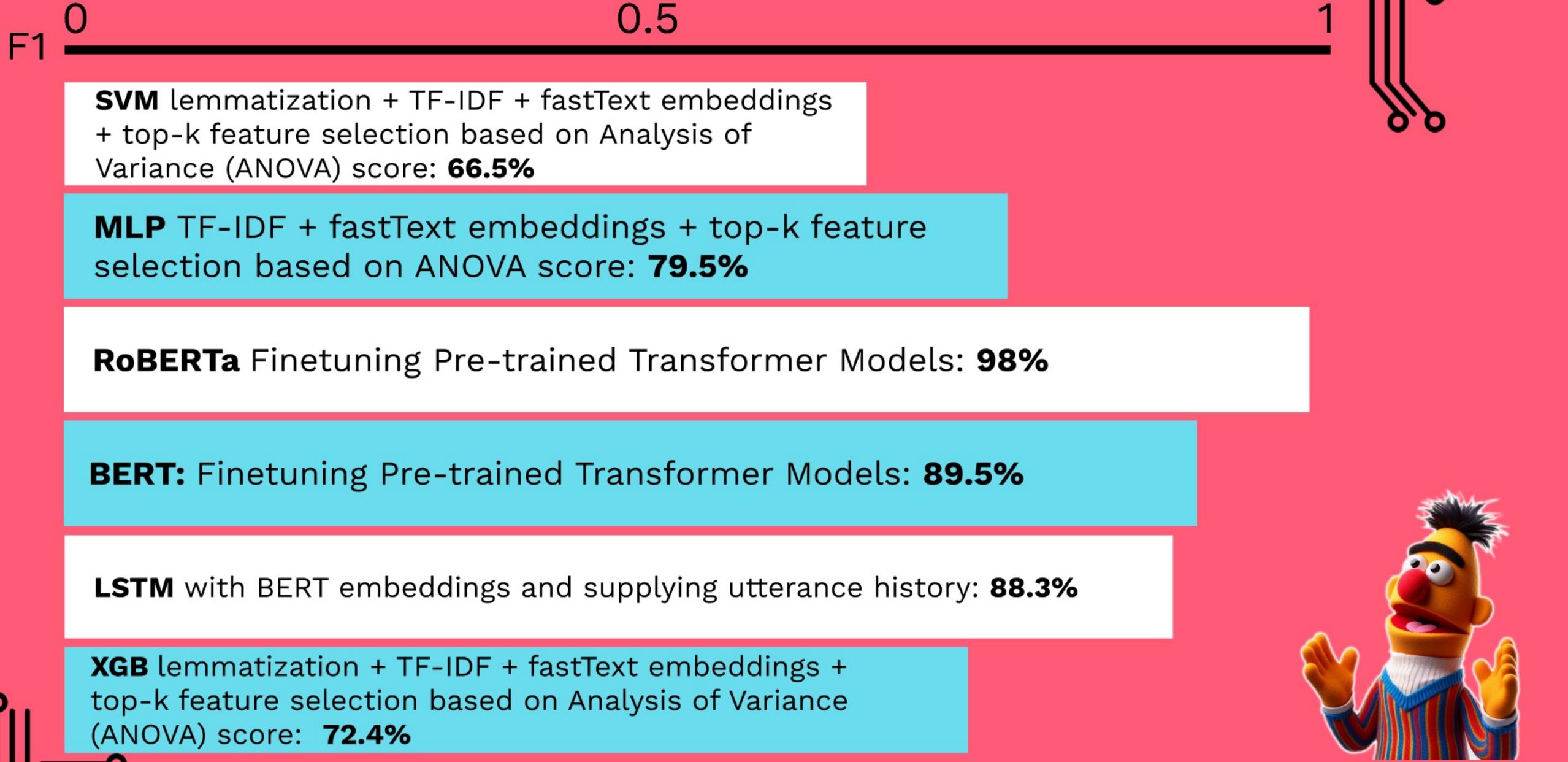
BERT embeddings: 75.1%

fastText embeddings and supplying utterance history: 87.3%

BERT embeddings and supplying utterance history: 88.3%



All Methods comparison



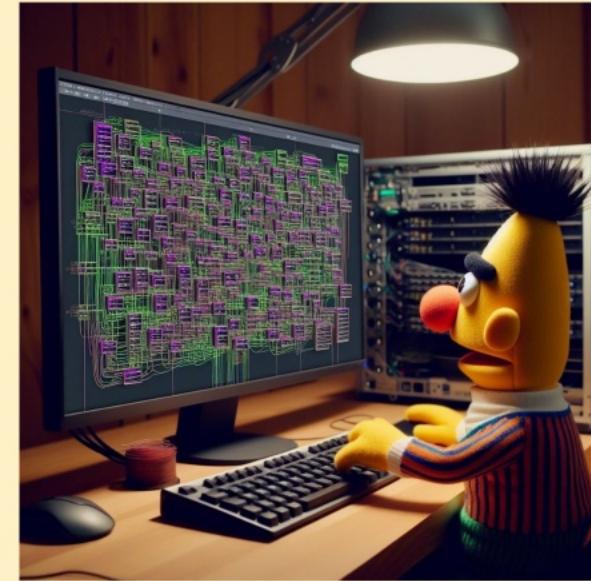
Task 2 Slot Filling



BERT and RoBERTa

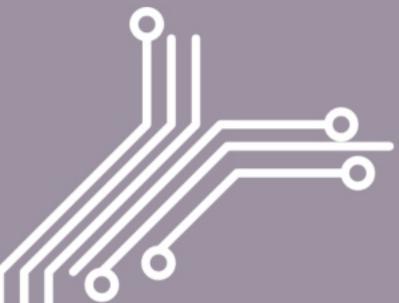
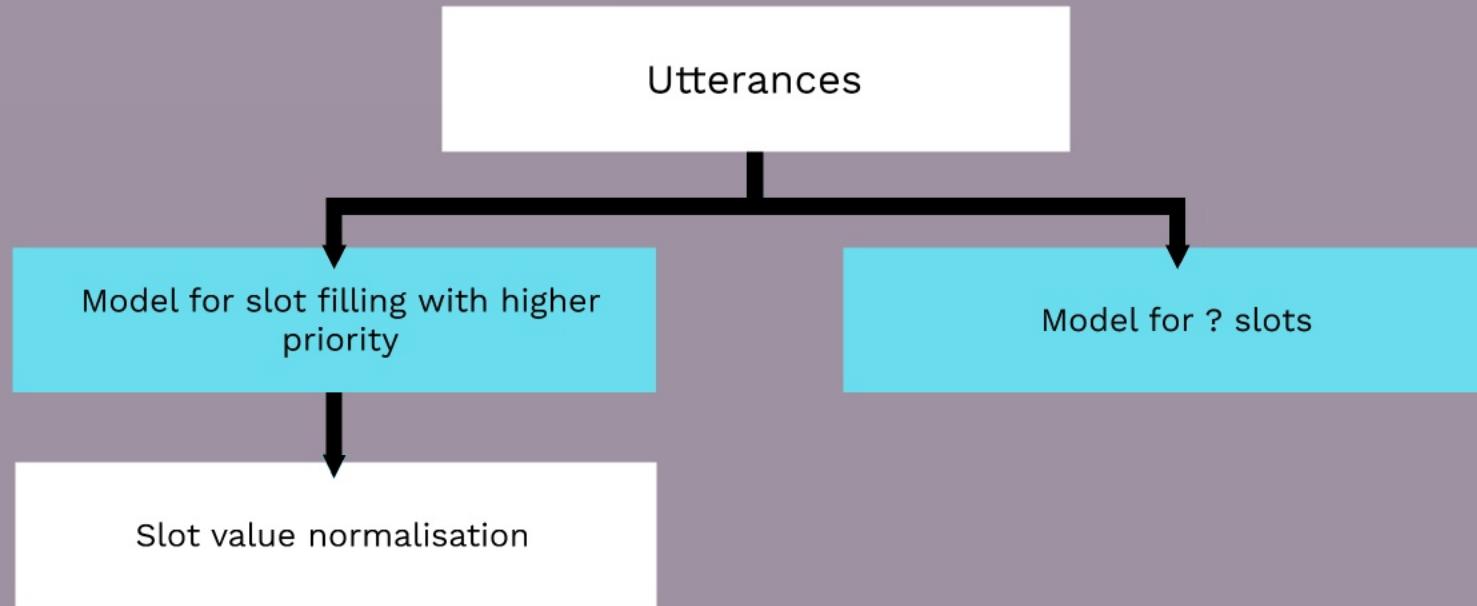


GPT-2

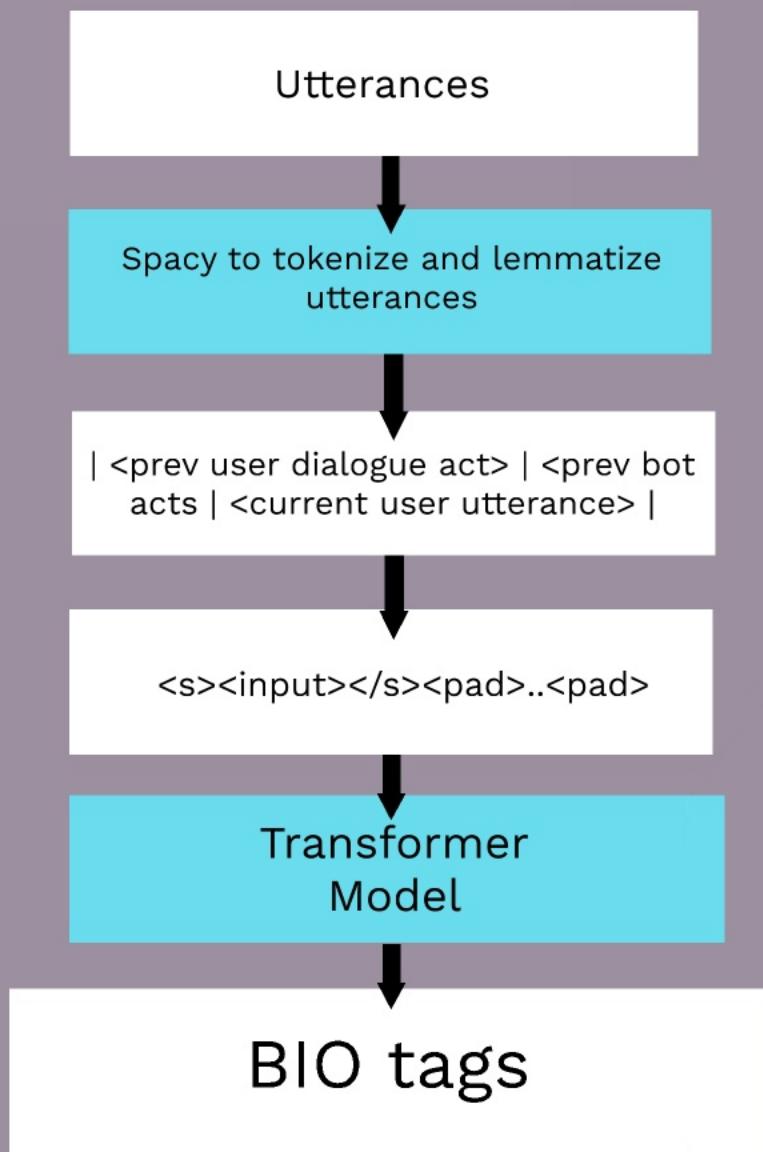
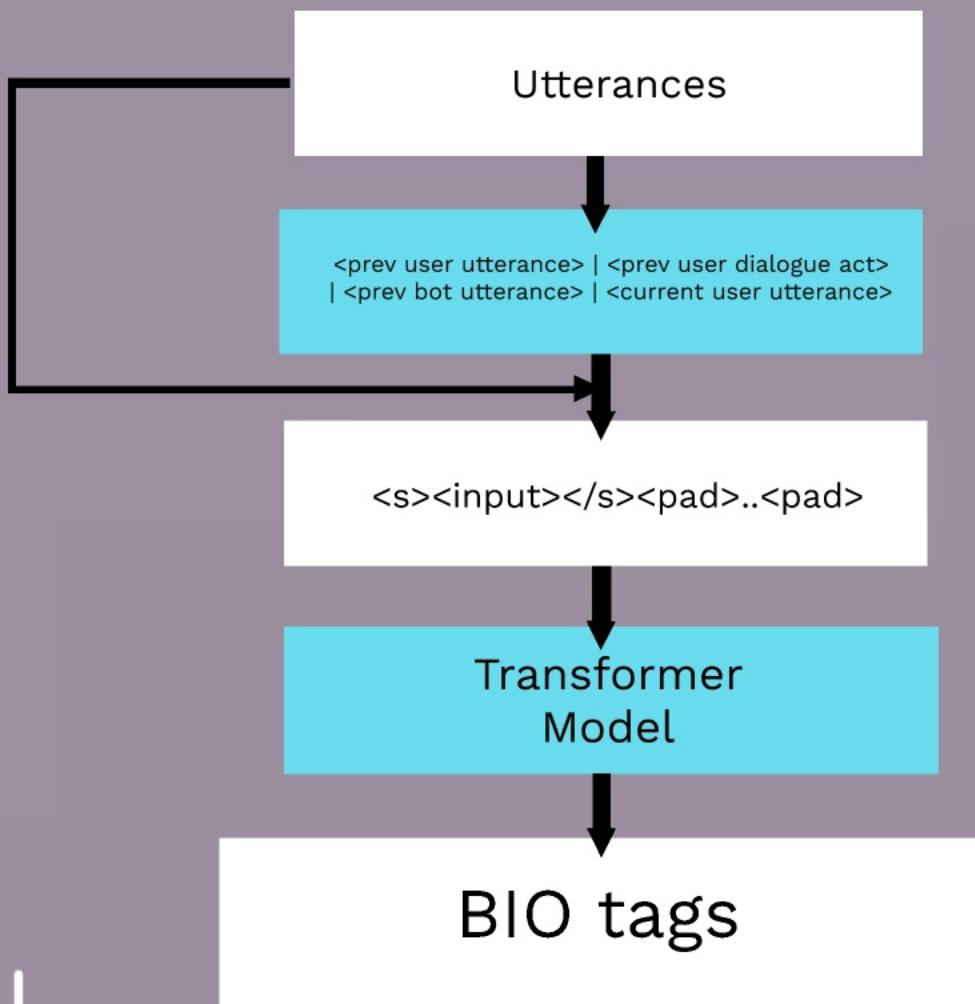


ConvBERT

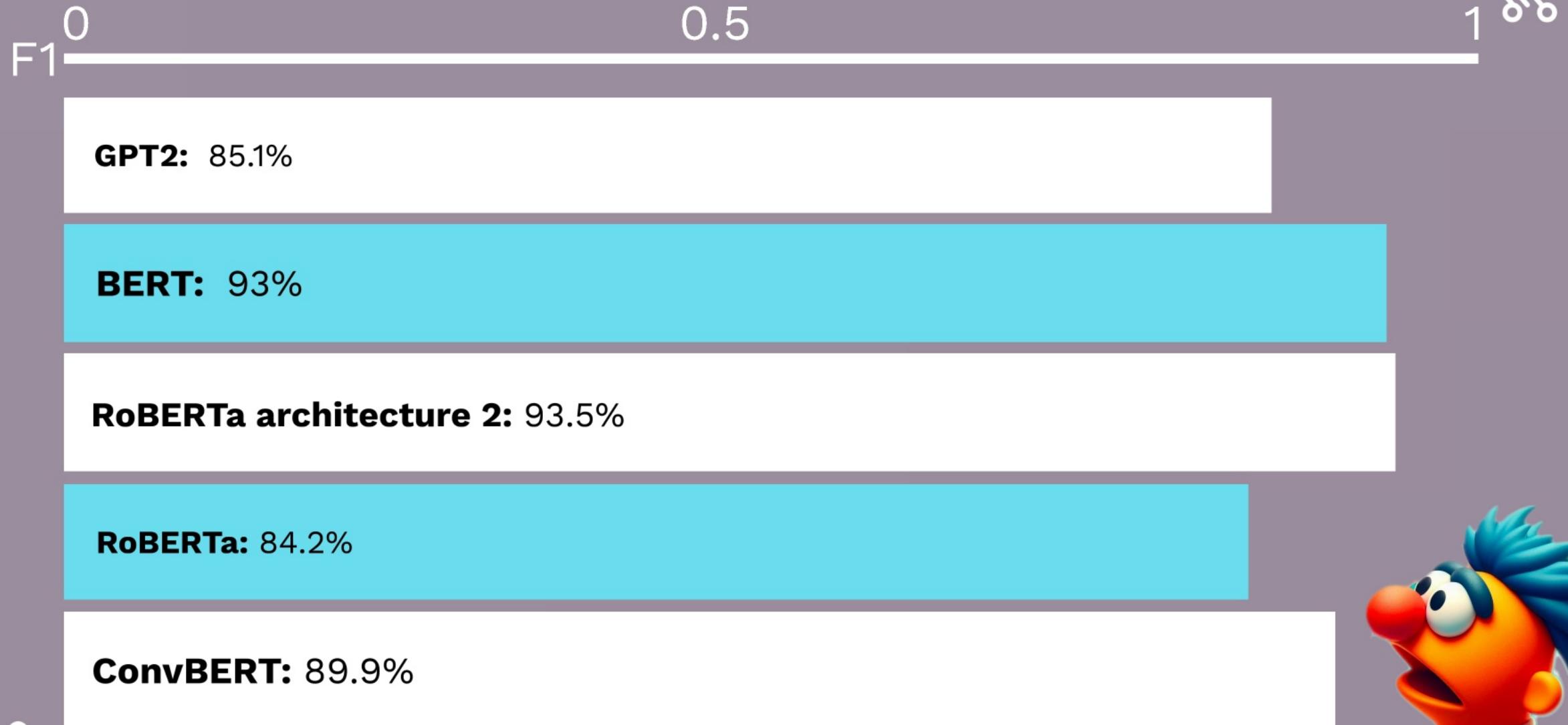
General Pipeline



Pipelines for slot filling



Methods comparison slot filling



Pipelines for ? slots



Methods comparison ? slots



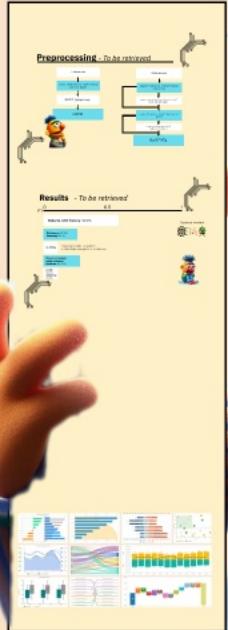
RoBERTa: 90.4%

LSTM: 82%

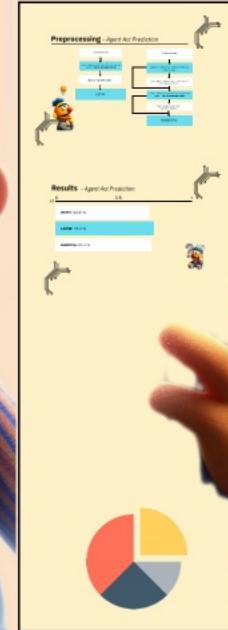


Task 3: Agent Move Prediction

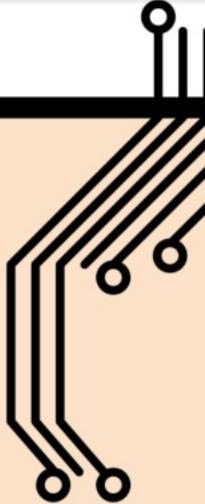
To be retrieved



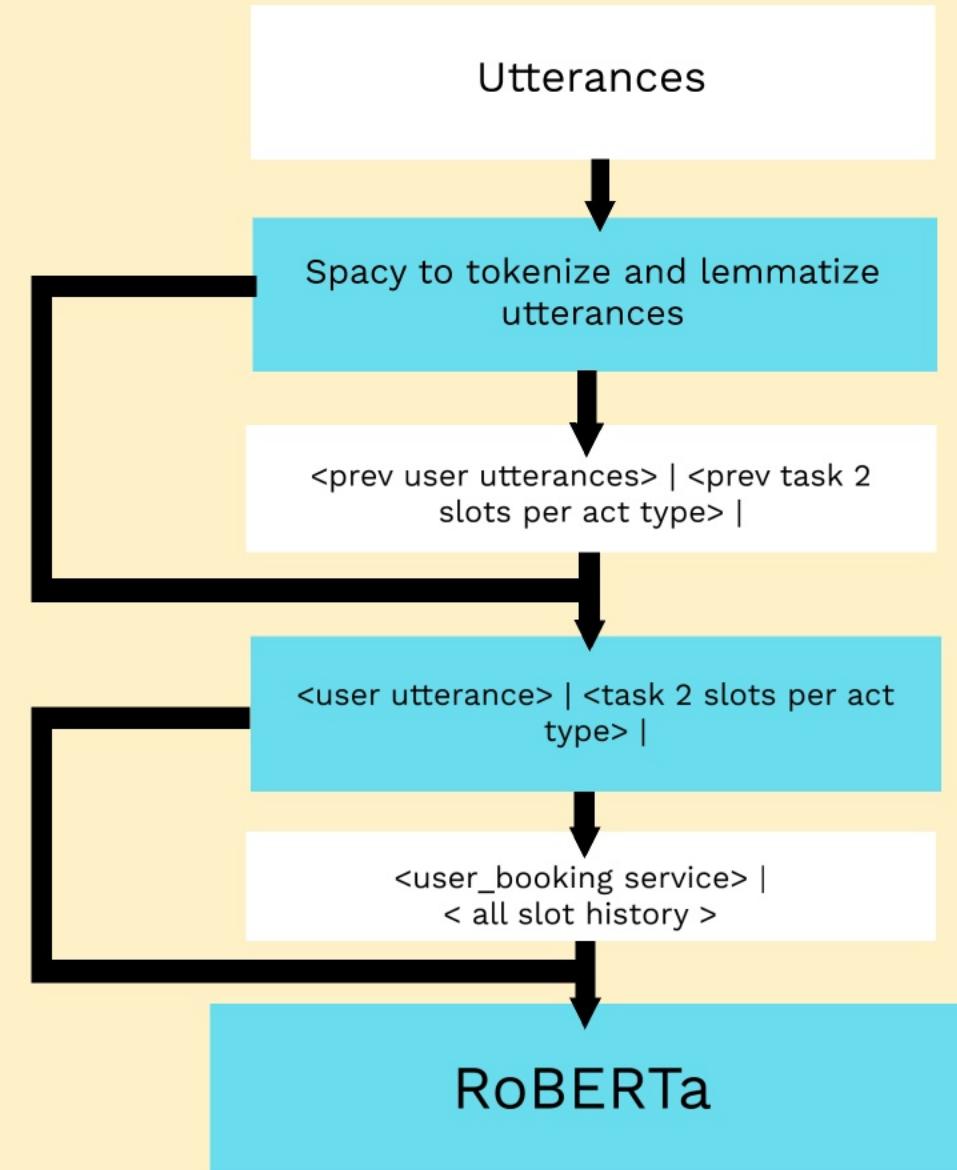
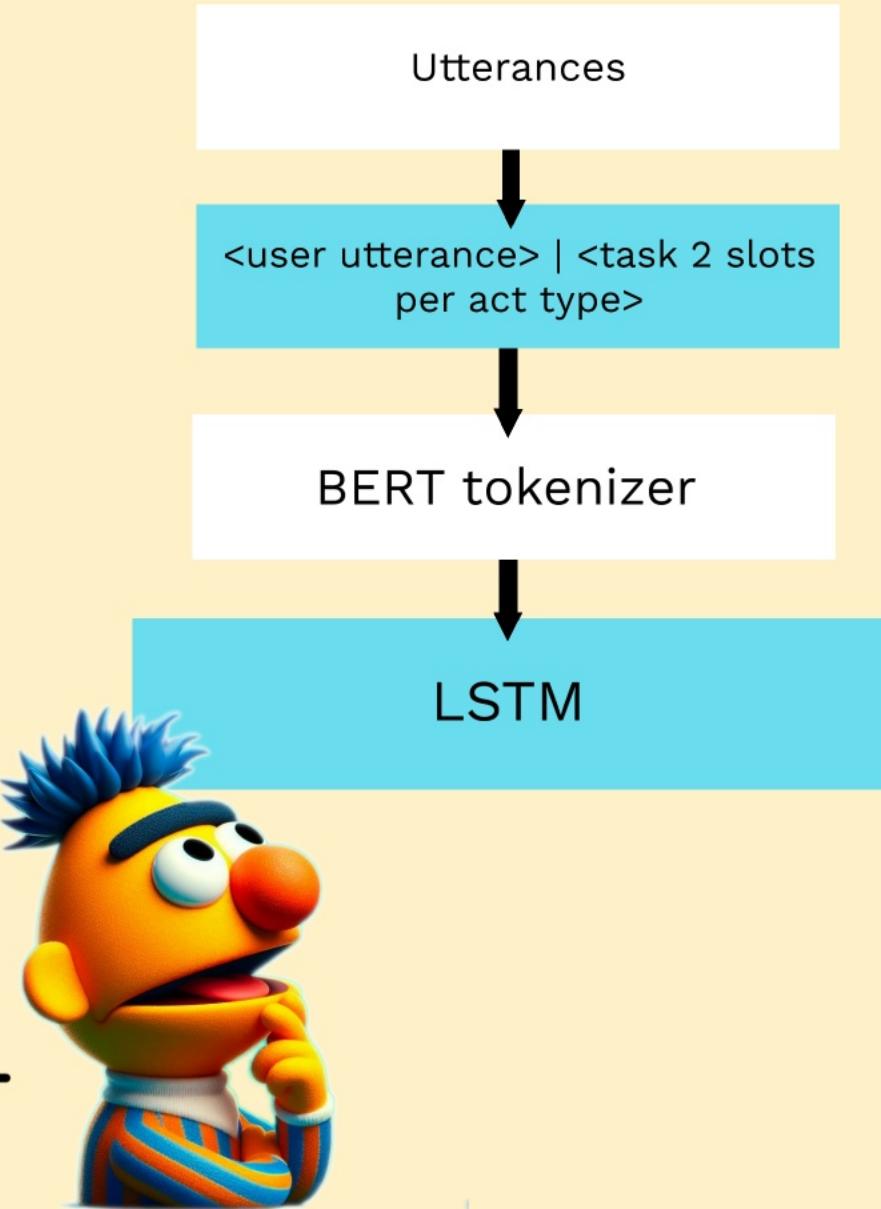
Agent dialogue act



To be requested



Preprocessing - To be retrieved

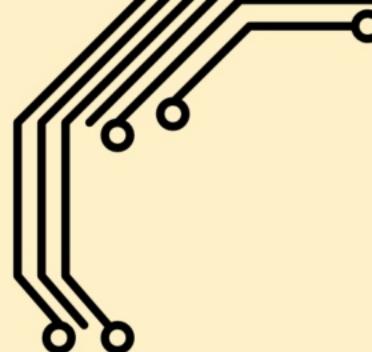


Results - To be retrieved

F1 0

0.5

1



Roberta with history: 59.6%

Dictionary: 29.4%
Accuracy: 52 %

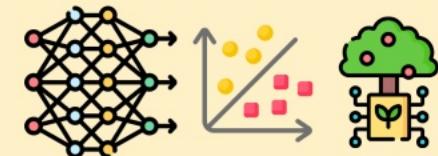
0-15%

Classical models using BERT embeddings averaged out as features

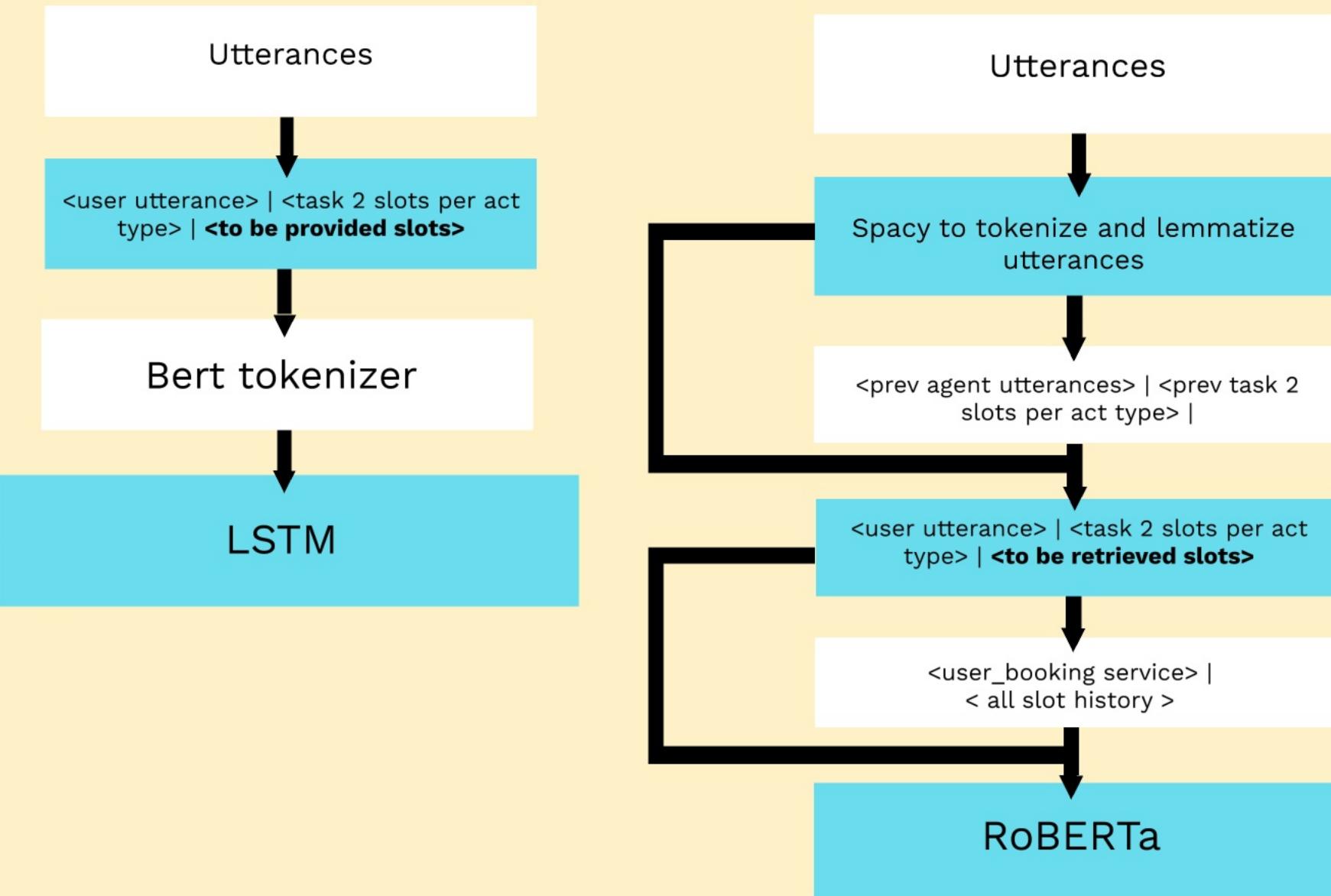
Classical models using indicator features: 30-31%

LSTM:
10.8 %
Accuracy:
29.9%

Classical models



Preprocessing - Agent Act Prediction



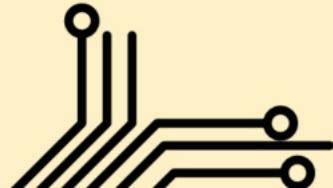
Results - Agent Act Prediction



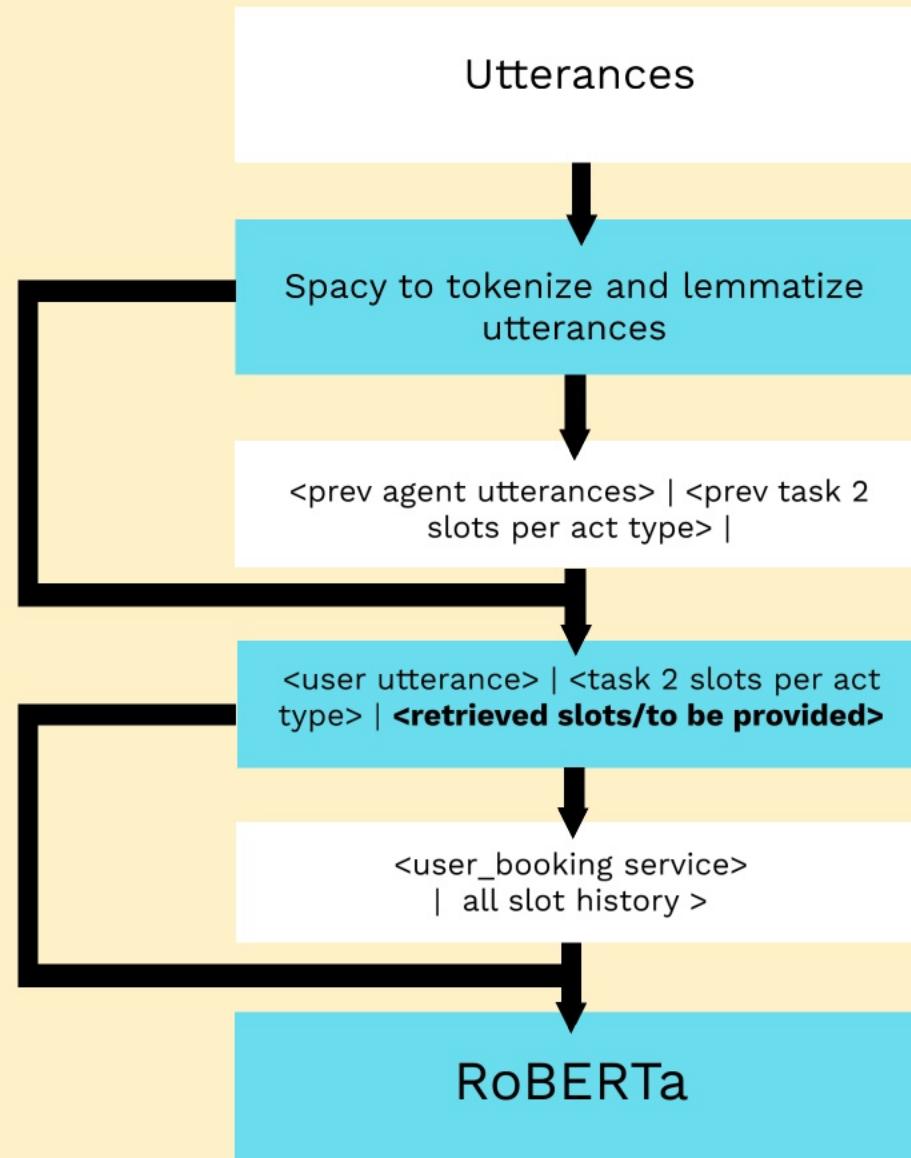
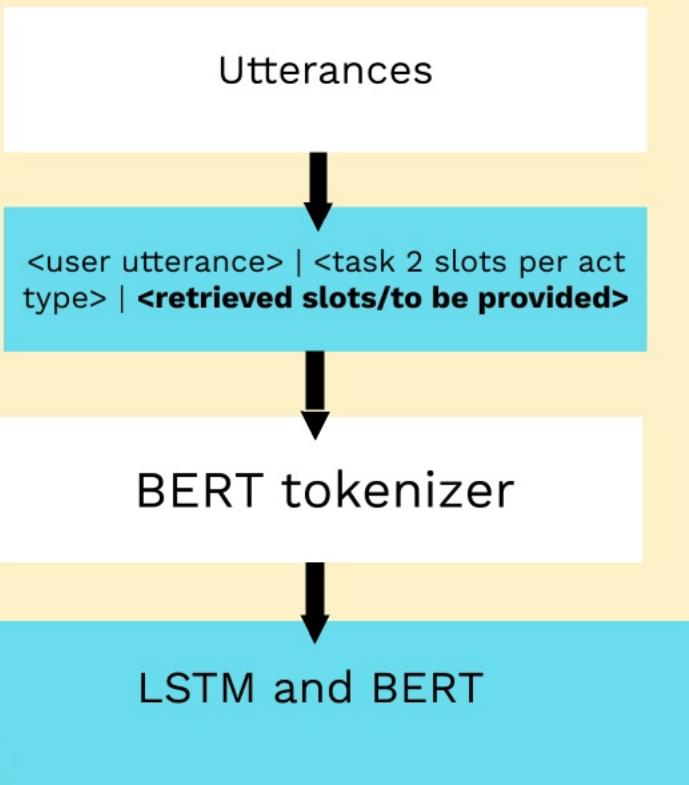
BERT: 63.8 %

LSTM: 73.2 %

RoBERTa: 65.4 %



Preprocessing - to be requested



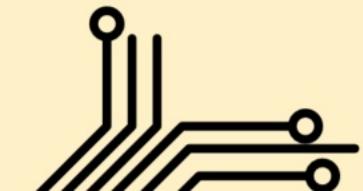
Results - to be requested

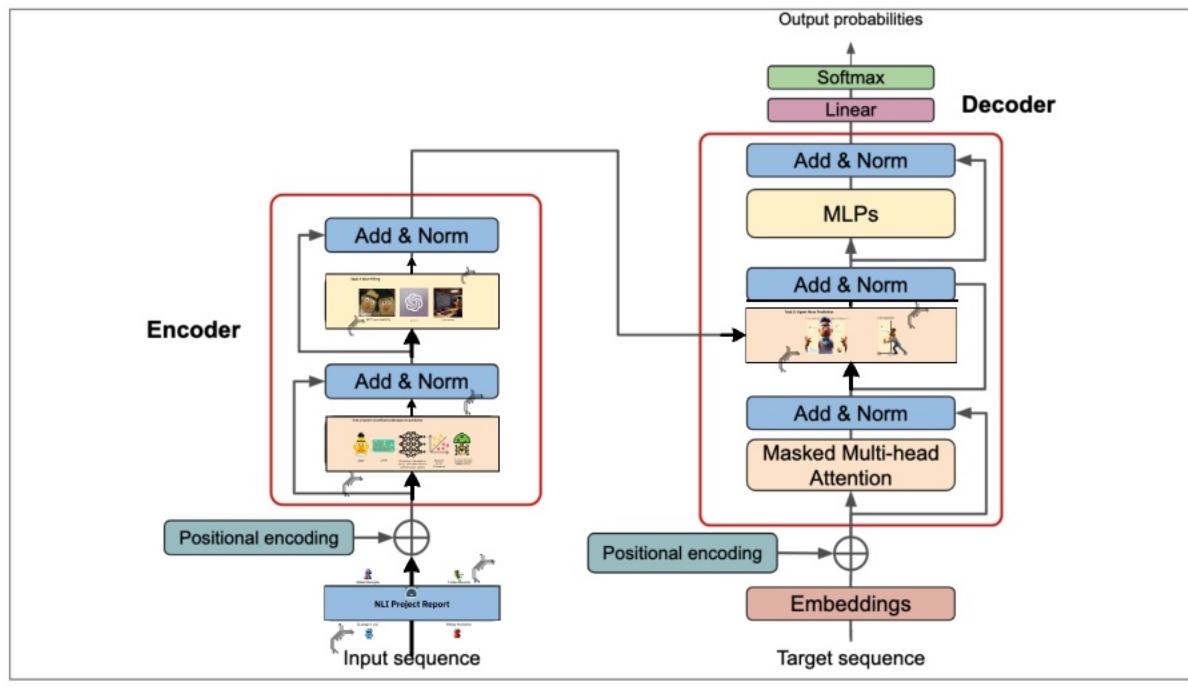


LSTM: 65.0 %

BERT: 76.8%

RoBERTa: 64.3%





Questions?/ Demo

