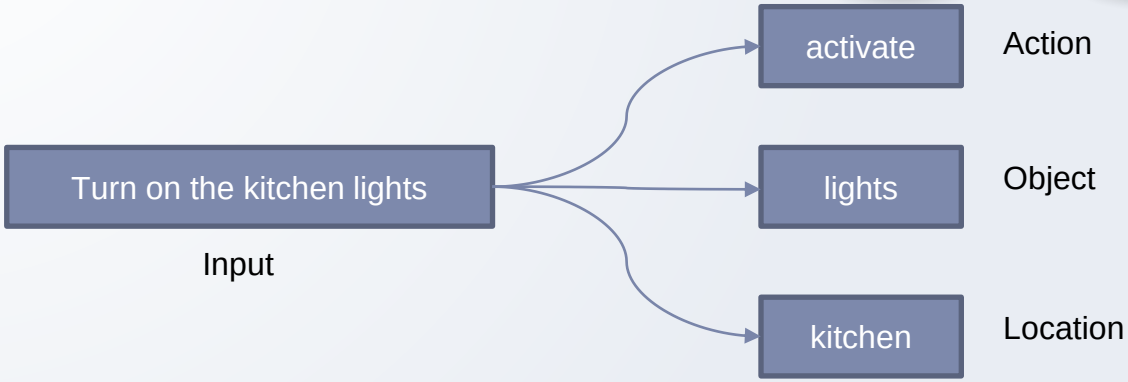
ACTION, ENTITY EXTRACTION WITH ROBERTA MODEL

Jose Moti
CE17B118

DATASET



11566 train examples

3118 validation examples

- Task → Identify action, object and location from each sentence given (none if not applicable).
- Maximum sentence length → 11
- 6 possible actions, 14 possible objects, 4 possible locations (all including none)
- So basically a classification task. Classify the given sentence to one of the given action, object and location

APPROACH

- Relatively simple and clean dataset. Basic NLP models will give good score.
- Implemented Roberta so that this can be scaled up to more complicated sentences and input datasets

- **Achieved 100% accuracy on given validation dataset.**

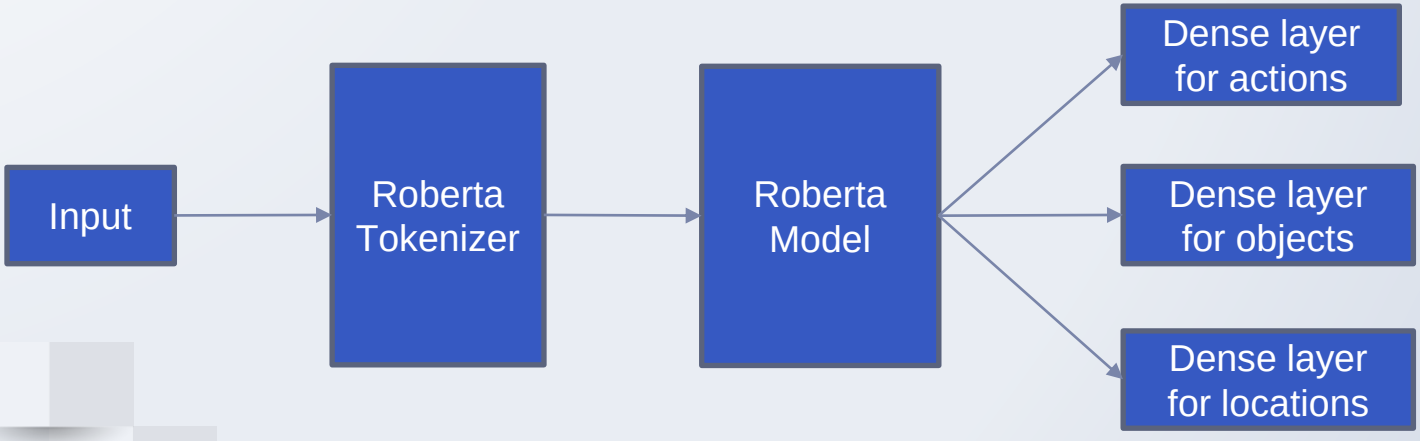
- Trained for 5 epochs – took 5 minutes on Colab TPUs and about 10 minutes on Colab GPUs.
- Same code can be used for training on TPUs, single/multiple GPUs or CPUs. Tensorflow
- Using tensorflow API, you can distribute your existing models and training code with minimal code changes.
- tf.distribute.experimental.TPUStrategy – strategy for training with TPUs
- tf.distribute.MirrorredStrategy – strategy for training with one or more GPUs

MODEL

- Roberta - Implemented by Facebook. Transformer based model. Improved version of Bidirectional Encoder Representation from Transformers (BERT) by Google.

Why Roberta?

- It is one of the state-of-the-art NLP model giving the best accuracies on many NLP tasks.



MODEL (continued)

- Trained with sparse categorical cross entropy loss
- Used Adam optimizer with learning rate 1e-5 and clipnorm of 1.
- Used Sparse Categorical Accuracy and Sparse Top3 Categorical Accuracy as metric.
- Only used dense layers other than the pretrained Roberta Model.

- While Inference – loads model weights, and encoder mappings. Calculates the F1 score for all actions, objects and locations.

See report for more info