ACTION, ENTITY EXTRACTION WITH ROBERTA MODEL

Jose Moti
CE17B118

DATASET

Input		Label	Type

```
Turn on the kitchen lights  →  activate      Action
Input                       →  lights        Object
                            →  kitchen       Location
```

200 train examples

After cleaning

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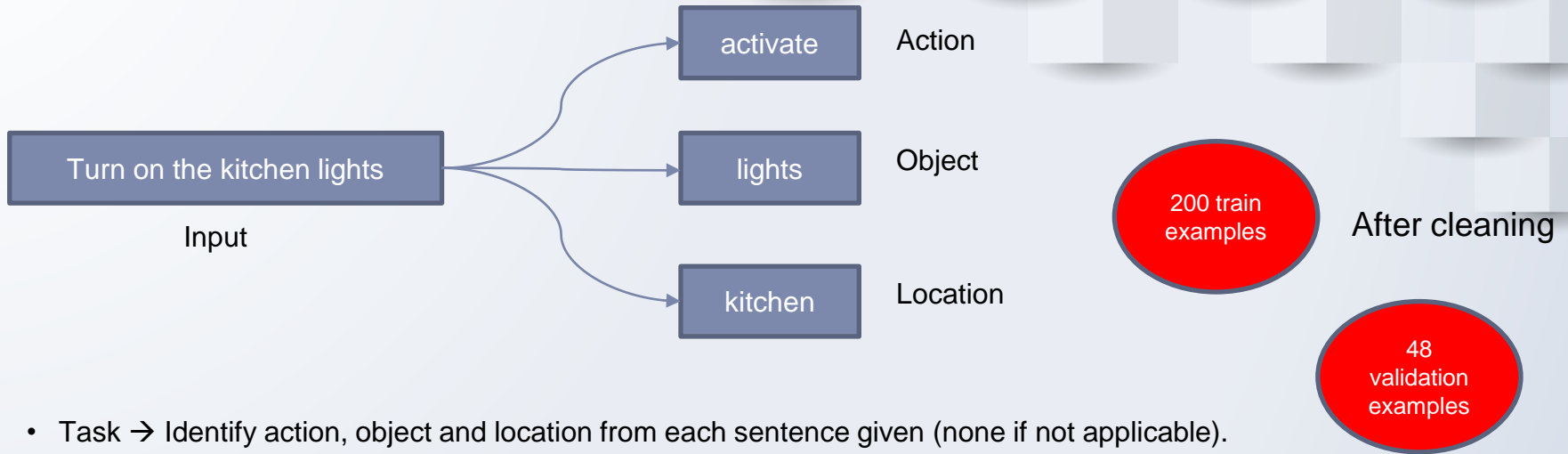
- Task → Identify action, object and location from each sentence given (none if not applicable).
- Maximum sentence length → 11
- 6 possible actions, 14 possible objects, 4 possible locations (all including none)
- So basically a classification task. Classify the given sentence to one of the given action,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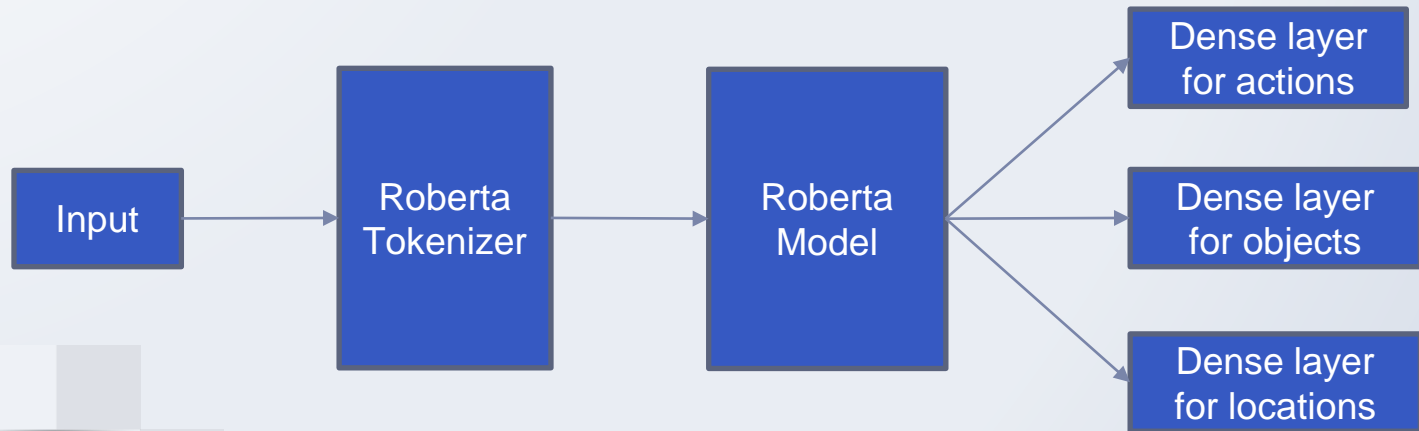
APPROACH

- 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 validation dataset..**

- Made sure that validation dataset and train dataset are completely disjoint

- Trained for 20 epochs – took <5 minutes on Colab TPUs (cluster of 8) and 10+ minutes on Colab single GPU.

- 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.

```
Input → Roberta Tokenizer → Roberta Model → Dense layer for actions
                                           → Dense layer for objects
                                           → Dense layer for locations
```

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