Step 0: python environment

* The model is trained under “v3-8” type, “TPU VM” architecture, and with TPU software version “tpu-vm-tf 2.8.0”.
* Additional libraries needed are as follows:

# install dependencies  
pip install pandas scikit-learn tensorflow\_addons==0.16.1 tensorflow\_probability==0.16.0

Step 1: copy the entire “final\_submssion” folder to a designated folder.

Step 2: make sure train/test data are available and accessible from the system, e.g., in a directory named “ROOT\_PATH/Dream-Challenges-Data”. This folder should contain “train\_sequences.txt” and “test\_sequences.txt” files.

Step 3: modify “final\_submission/utils/train\_utils.py” file accordingly, we’ll most likely need to change in the “ROOT\_PATH” of folders “ROOT\_PATH/final\_submission/architecture” and “ROOT\_PATH/final\_submission/utils”. Also change data\_path line using directory path from step 2.

Step 4: train and predict:

* Change working directory to “ROOT\_PATH/final\_submission”
* Start training process with the following command:

# training  
python3 run\_config\_v1.py --model new\_BioNML\_CrossAtt\_v3 --config config39v3 --reset --train\_only

* Start predicting process with the following command (after training process completes and model weight files presenting under local folder as “model/ new\_BioNML\_CrossAtt\_v3/config39v3/\*.h5”):

# predicting  
python3 run\_config\_v1.py --model new\_BioNML\_CrossAtt\_v3 --config config39v3 --reset --predict\_only

Step 5: prediction output will be available as a text file like “ROOT\_PATH/final\_submission/prediction/new\_BioNML\_CrossAtt\_v3/config39v3/ test\_sequences\_fine\_tune\_v7\_config39v3\_\*.txt”.