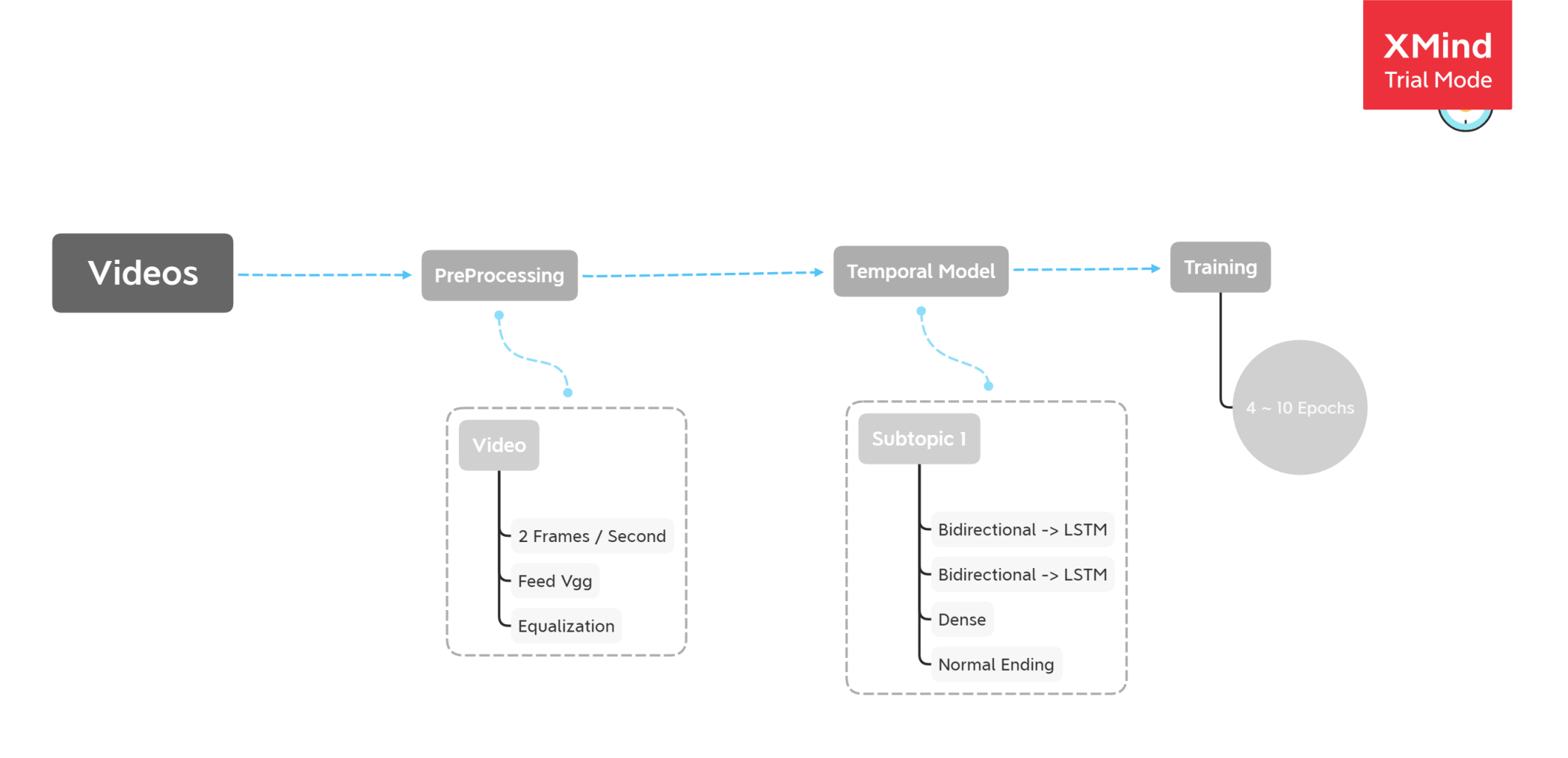
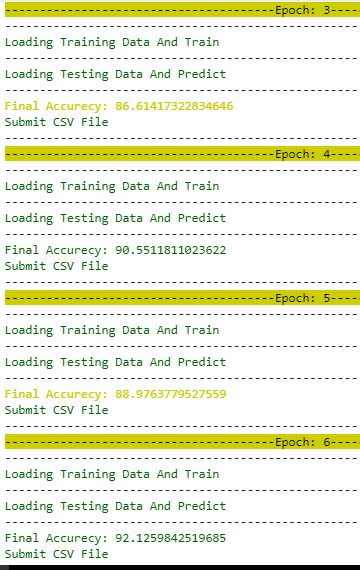
**Video Action Recognition**

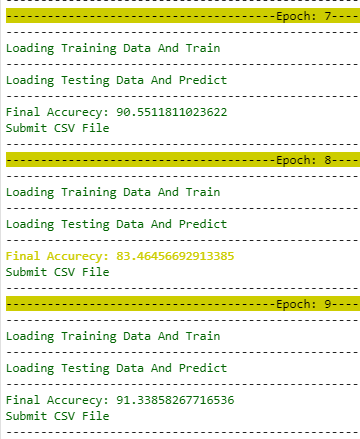


**Team Id : 5**

First Model

- Overview: first we have a bunch of videos before training we have to apply preprocessing we will take from each video bunch of frames each half a second and feed each frame to vgg pre-trained model to extract the features from each frame then calculate the average frames number then make all video frames equal then feed all the feature to the temporal model which contain Bidirectional LSTM

- Hyperparametars : it seems that changing the split percentage, epochsChange the accurcy around 7% , Results 3 ~ 9 Epochs



But split percentage is playing hight role to determine the accuracy Also, we changed the model parameters like the number of units in LSTM 256, 512, 1024’gpu fail'

Second model

kinetics-i3d

preprocessing:

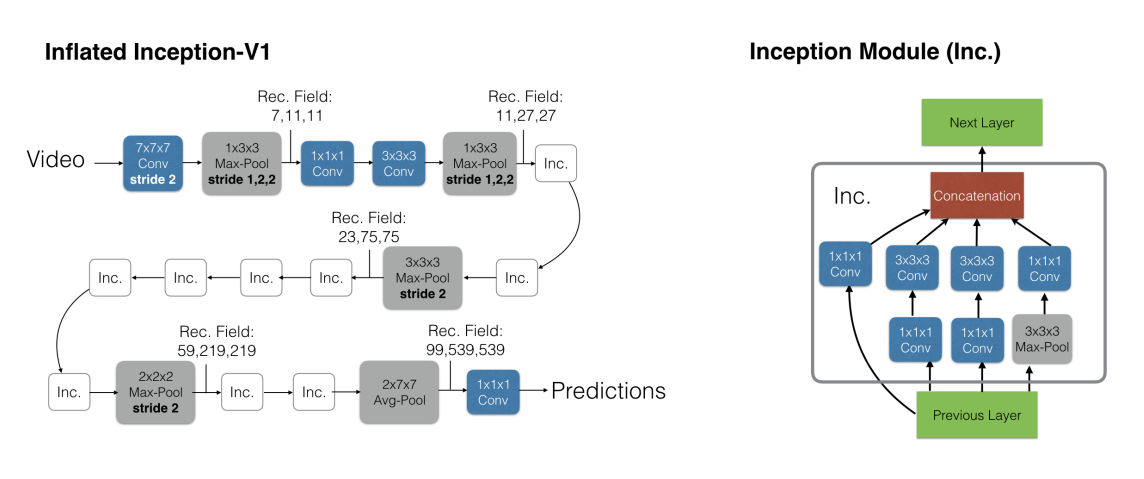
It’s a 3D 2 stream model one stream accepts rgb frames as input and the other stream accepts optical flows as input, any amount of frames or flows can be passed to the model during training or testing, so preprocessing is done by:

1. Extracting a number of rgb frames from a single video with stride that is calculated depending on the number of frames and for cases when the number of frames of the video is less than number of requested frames, video frames are repeated.
2. Extracting a number of optical flows from the video by the same process in 1

Predictions can be done by one of the 2 streams or the summation of both.

Model is pretrained on kinetics datatset.

Architecture:



Results:

Best results are running the rgb stream only

Running both or the flow stream only gets lower results in general

#epochs between 10 and 15

Pretrained weights on kinetics dataset is loaded

Lr = 0.00001

|  |  |  |
| --- | --- | --- |
| Type | Public | Private |
| RGB stream with 64 frames input in training and testing | 0.89333 | 0.98039 |
| RGB stream with 90 frames input in training and testing | 0.94666 | 0.92156 |
| RGB stream with 250 frames input in training and testing | 0.96000 | 0.92156 |
| RGB stream with 64 frames input in training and 250 in testing | 0.94666 | 0.96078 |