Project Phase Report - 3

INTP22-ML-5: Power Line Fault Detection

Sshubam Verma

PROJECT OBJECTIVES FOR THE PHASE:

- Decide optimal model metrics and loss function
- Test and implement different model hyperparameters
- Research to make the model architecture reduce overfitting
- Begin Model training process
- Fix code issues

PHASE - 3 PROGRESS

After resolving Layer dimensional issus, I was able to build the final model. Implemented a dropout rate of 20% in the LSTM layer to achieve better model generalisation. Implemented the custom metric Matthew's correlation coefficient to measure the performance of the model. Chose rmsprop as the optimizer. Finalised the model with the following architecture and moved on to the training process.

```
model = Sequential()

model.add(layers.Conv1D(32, 8, padding='same',input_shape=(length_of_sequence, 1),activation='rel
model.add(layers.MaxPooling1D(2, padding='same'))
model.add(layers.Conv1D(64, 8, padding='same', activation='relu'))
model.add(layers.MaxPooling1D(2, padding='same'))
model.add(layers.Conv1D(128, 8, padding='same', activation='relu'))
model.add(layers.MaxPooling1D(2, padding='same'))
model.add(layers.Conv1D(256, 8, padding='same', activation='relu'))
model.add(layers.LSTM(64, dropout = drop_out_rate,recurrent_dropout = recurrent_dropout))
model.add(layers.Dense(1, activation='sigmoid'))

model.compile(loss='binary_crossentropy', optimizer='rmsprop', metrics=['accuracy',matthews_corr_

model.summary()
```

$$\mathrm{MCC} = \frac{\mathrm{TP} \times \mathrm{TN} - \mathrm{FP} \times \mathrm{FN}}{\sqrt{(\mathrm{TP}+\mathrm{FP})(\mathrm{TP}+\mathrm{FN})(\mathrm{TN}+\mathrm{FP})(\mathrm{TN}+\mathrm{FN})}}$$

PHASE - 3 PROGRESS

Implemented custom callbacks, monitoring validation loss along with the training loss, EarlyStopping function to stop model before it overtrains and reduce Learning rate on the basis of validation loss till the step it stops improving.
Also create model checkpoints while training and save only best weights, set the model for training for 50 epochs, set dropout to 20% rate.
The model after training achieved a validation accuracy of 85% and a Matthew's Correlation Coefficient of 0.6
I saved the best model weights after 15 epochs to load it later.

```
weight_path="{}_weights.best.hdf5".format('lstm_model')
early = EarlyStopping(monitor="val_loss",
                      mode="min",
                      patience=10)
lr = ReduceLROnPlateau(monitor='val_loss', factor=0.2, patience=2, min_lr=0.001)
checkpoint = ModelCheckpoint(weight_path, monitor='val_loss', verbose=1,
                             save_best_only=True, mode='min', save_weights_only = True)

callbacks_list = [checkpoint, early, lr]
```

[24] Python

```
train_gen = train_data_gen(metadata_train, train_diff, batch_size=BATCH_SIZE)
```

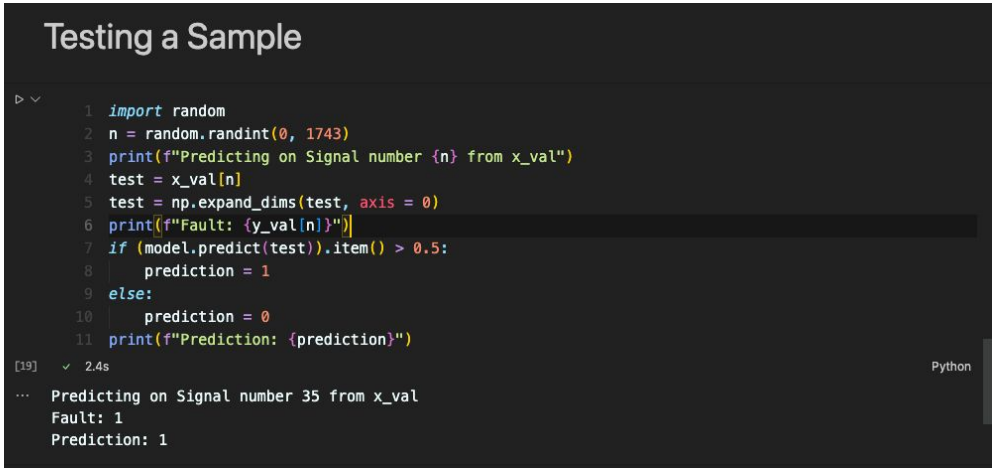
[25] Python

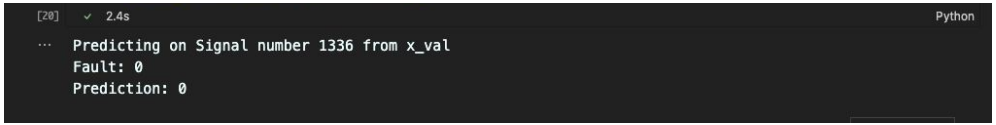
```
history = model.fit_generator(
                train_gen,
                steps_per_epoch=STEPS_PER_EPOCH,
                epochs=EPOCHS,
                validation_data=(x_val,y_val),
                callbacks=callbacks_list)
```

[ ] Python

PHASE - 3 PROGRESS

To test overall generalisation of our model, I implemented a test function to check random samples from unseen data. The Model showed very good results, which concludes the model building, hyperparameter tuning, Fitting the data pipeline ot the model and the model training and evaluation. This forwards me to deploying the model in a Flask App and host it on Cloud.





Gantt Chart

Gantt Chart LINK

PROJECT TRACKING

PROJECT TITLE	Power Line Fault Detection	COMPANY NAME	IAFSM
PROJECT COORDINATOR	DEVESH TARASIA	DATE	01/06/2022

PROJECT DETAILS								DELIVERABLES	
STATUS	PRIORITY	START DATE	END DATE	DURATION	TASK NAME	ASSIGNEE	DESCRIPTION	DELIVERABLE	% DONE
Project Initiation, briefing and planning									**86%**
In Progress	Medium	01/06/2022	06/06/2022	5	Analysis	Sshubam Verma	Problem statement analysis		100%
In Progress	High	07/06/2022	10/06/2022	3	Research	Sshubam Verma	Read and analyse related research papers		100%
In Progress	High	11/06/2022	15/06/2022	4	Data Cleaning	Sshubam Verma	Understanding data and Exploratory Data Analysis		100%
In Progress	High	15/06/2022	24/06/2022	9	Modelling	Sshubam Verma	Model Building and Training		100%
Not Yet Started	Medium	25/06/2022	28/06/2022	3	Tuning	Sshubam Verma	Model Hyperparameter tuning		100%
Not Yet Started	High	29/06/2022	10/07/2022	11	Testing	Sshubam Verma	Model comparison and testing		100%
Not Yet Started	High	11/07/2022	25/07/2022	14	Deployment	Sshubam Verma	Model Deployment		0%
Project Submission and Presentation									**75%**
In Progress	High	05/06/2022	10/06/2022	5	Task	Sshubam Verma	Phase Report - 1		100%
In Progress	High	20/06/2022	30/06/2022	10	Task	Sshubam Verma	Phase Report - 2		100%
Not Yet Started	High	05/07/2022	10/07/2022	5	Task	Sshubam Verma	Phase Report - 3		100%
Not Yet Started	High	15/07/2022	25/07/2022	10	Task	Sshubam Verma	Phase Report - 4		0%