VRIJE
UNIVERSITEIT
BRUSSEL

Predicting Zero Point energies

3D CNN v0.1.3

Michiel Jacobs
05/02/2021

Run information

Model:

Author: Michiel Jacobs
Version: 0.1.3
Type: 3D CNN
Feature: Coulomb Matrix
Label: Zero point energy

System information

Platform: Linux-3.10.0-1160.2.2.el7.x86_64-x86_64-with-glibc2.2.5
OS: Linux
Version: #1 SMP Tue Oct 20 16:53:08 UTC 2020
Processorx86_64

Data:

Maximum heavy atoms: 20
Maximum molecule size: 62
Total molecules: 10695

Tensorisation:

Positive dimensions: 0
Negative dimensions: 5

Test and train sets:

Split ratio: 0.8
Molecules for training: 8556
Molecules for testing: 2139

Neural Network

Network compile parameters:

Learningrate: 0.0001
Loss: mean_squared_error
Optimizer: Adam
Metrics: mean_absolute_error, mean_squared_error

Network fit parameters:

Batch size: 128
Epochs: 500
Validation split: 0.2
Shuffle data each epoch: True

Early stopping parameters:

Minimum change required: 0.0001
Epochs no change is allowed before stopping: 5
Restore best weights: True

Neural network Layer settings:

input shape: (6, 62, 62, 1)
kernel size: (1, 3, 3)
activation: relu
pool size: (2, 2, 2)
filters: 64
dropout: 0.2
dense units: 32
output shape: 1

NN summary

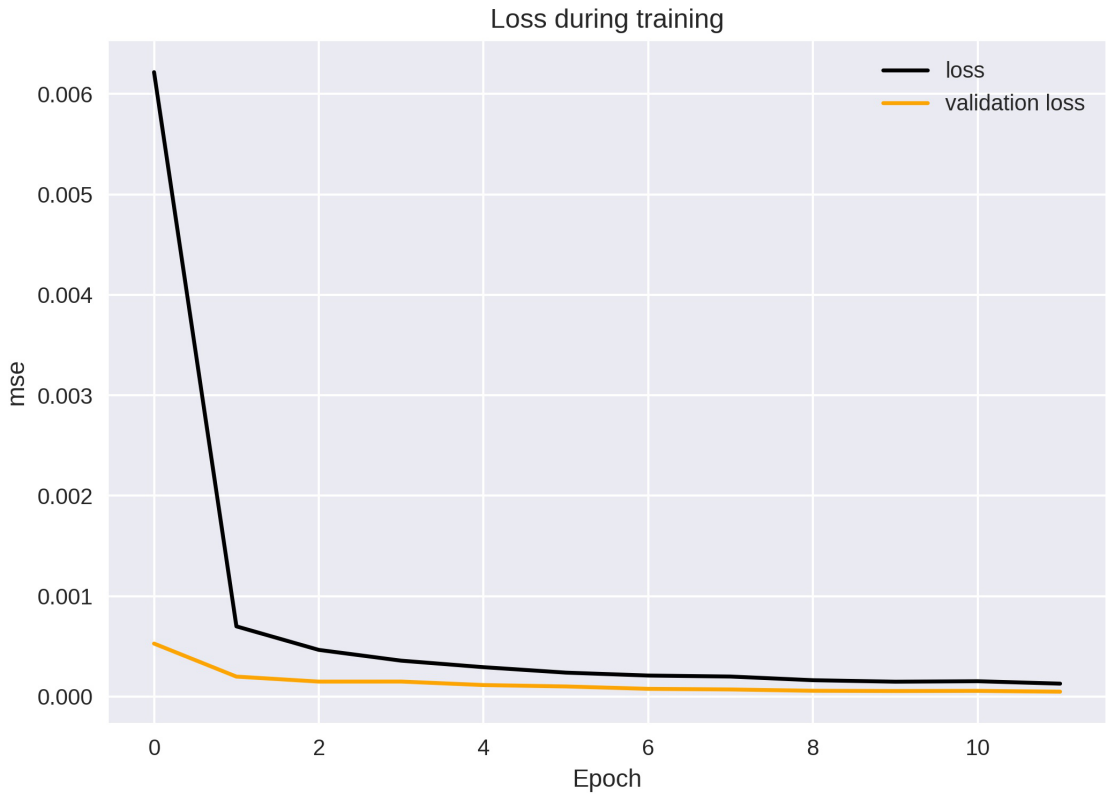
Model: "sequential"

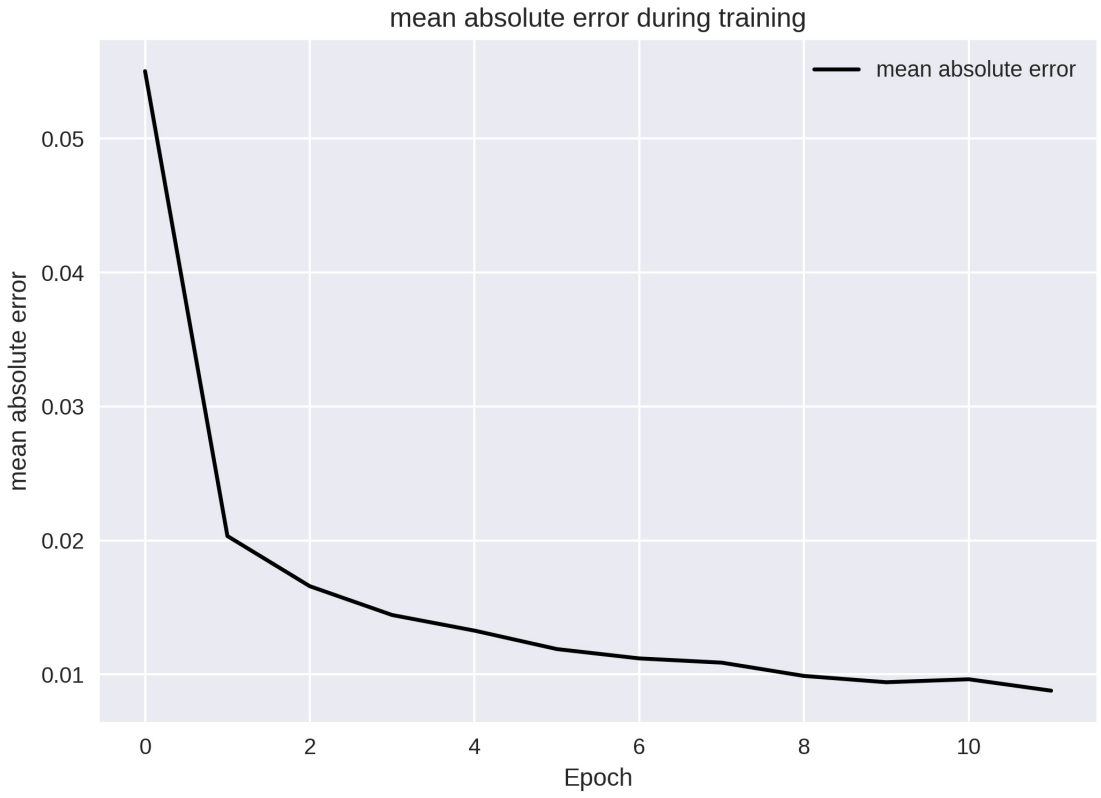
Layer (type)	Output Shape	Param #
conv3d (Conv3D)	(None, 6, 60, 60, 64)	640
max_pooling3d (MaxPooling3D)	(None, 3, 30, 30, 64)	0
conv3d_1 (Conv3D)	(None, 3, 28, 28, 64)	36928
max_pooling3d_1 (MaxPooling3	(None, 1, 14, 14, 64)	0
conv3d_2 (Conv3D)	(None, 1, 12, 12, 64)	36928
flatten (Flatten)	(None, 9216)	0
dropout (Dropout)	(None, 9216)	0
dense (Dense)	(None, 32)	294944
dense_1 (Dense)	(None, 32)	1056
dense_2 (Dense)	(None, 1)	33

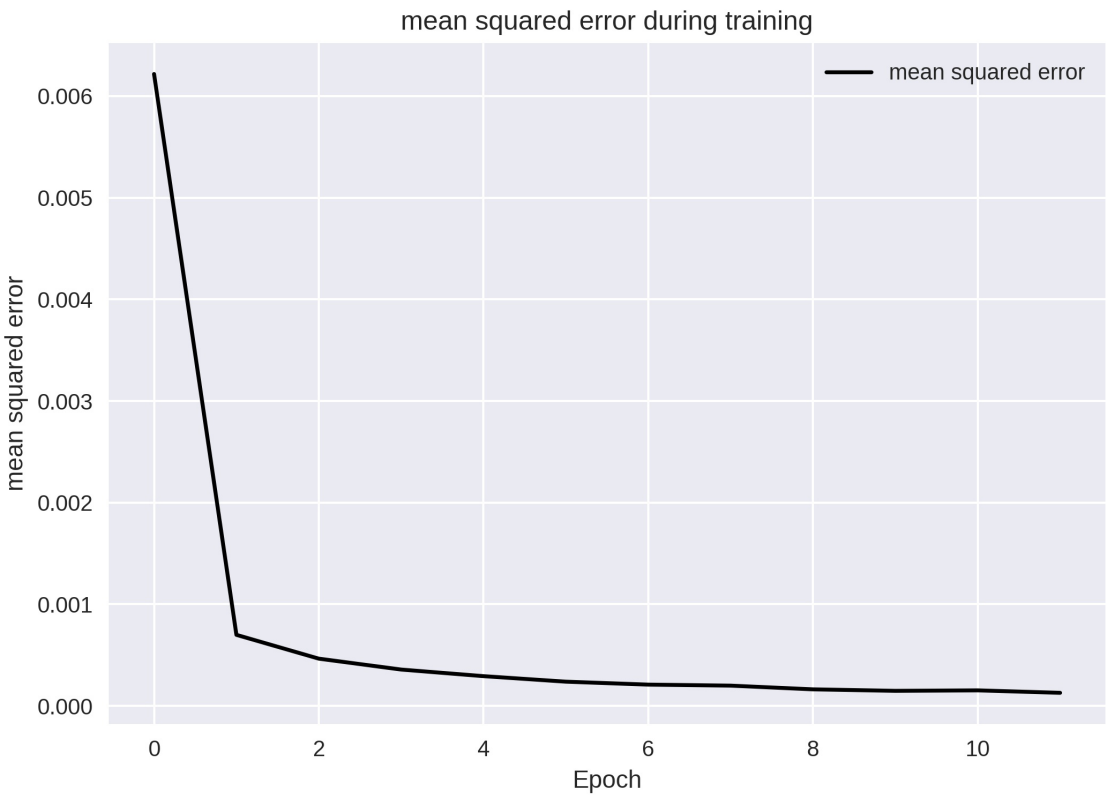
Total params: 370,529
Trainable params: 370,529
Non-trainable params: 0

Results

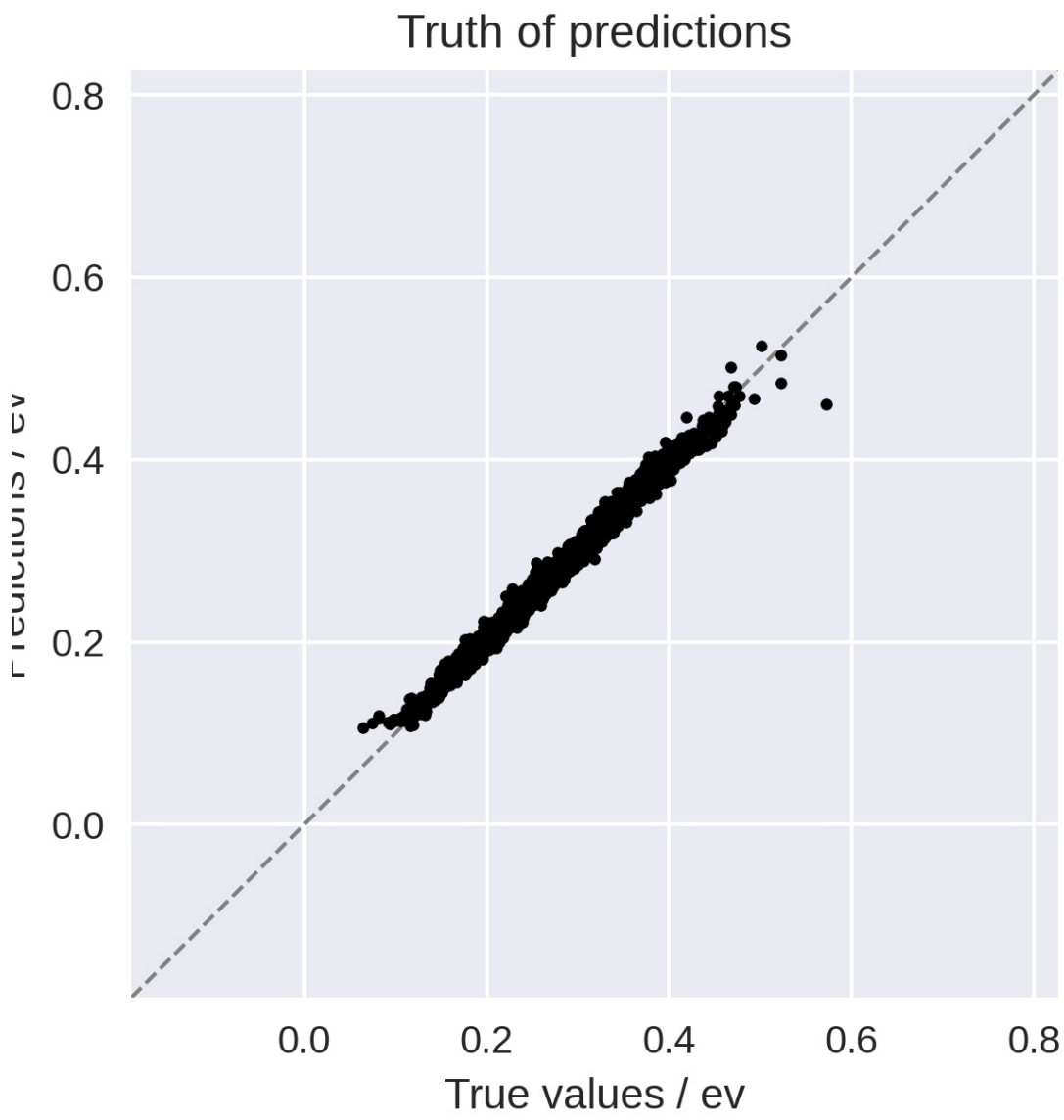
Training evaluation





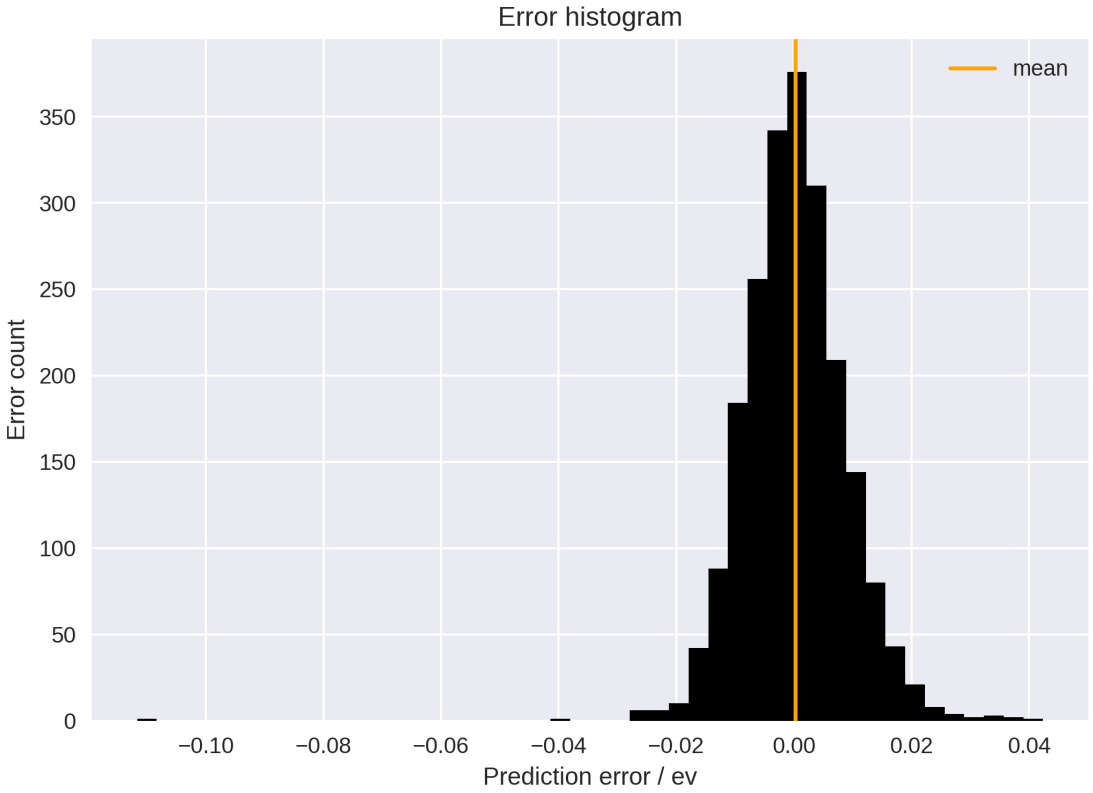


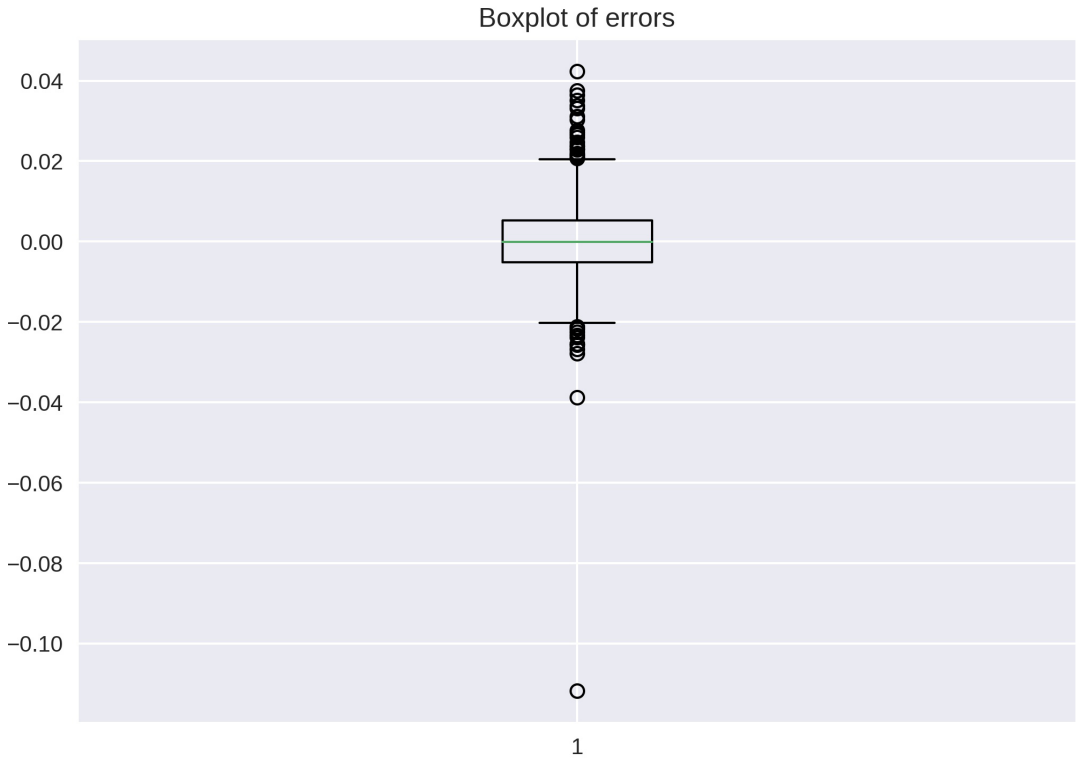
Model evaluation



mean absolute error: 0.0066
mean squared error: 0.0001

Error evaluation





Mean: 0.0002
Median: -0.0
Minimum error: -0.1117
Maximum error: 0.0424
Skewness: -0.6907
Kurtosis: 12.9673
Standard deviation: 0.0088
90% Confidence interval: [-0.0001;0.0005]

Log

```
2021-02-04 17:03:51,539:INFO:Starting model 3D_CNN_0.1.3_on_02-04-2021_17.03.51
2021-02-04 17:03:51,540:INFO:==================== Model info ====================
2021-02-04 17:03:51,540:INFO:Author: Michiel Jacobs
2021-02-04 17:03:51,540:INFO:Version: 0.1.3
2021-02-04 17:03:51,540:INFO:Modeltype: 3D CNN
2021-02-04 17:03:51,540:INFO:Maximum heavy atoms: 20
2021-02-04 17:03:51,540:INFO:Feature: Coulomb Matrix
2021-02-04 17:03:51,540:INFO:Labels: Zero point energy
2021-02-04 17:03:51,540:INFO:DEVELOPMENT: False
2021-02-04 17:03:51,561:INFO:============== Step 1: loading data ==============
2021-02-04 17:03:58,441:INFO:Data loaded
2021-02-04 17:03:58,441:INFO:============== Step 2: data preprocessing ==============
2021-02-04 17:03:58,441:INFO:Trimming dataset...
2021-02-04 17:03:58,572:INFO:Loading arrays...
2021-02-04 17:03:59,344:INFO:Shuffeling data...
2021-02-04 17:03:59,346:INFO:Calculating maximum size of molecules...
2021-02-04 17:03:59,346:INFO:The maximumsize of molecules is 62
2021-02-04 17:03:59,347:INFO:Normalizing data...
2021-02-04 17:04:00,021:INFO:Tensorisation of the coulomb matrices...
2021-02-04 17:09:18,605:INFO:Building channels...
2021-02-04 17:09:18,652:INFO:Calculating train test split...
2021-02-04 17:09:18,652:INFO:There are 10695 entries in this dataset.
2021-02-04 17:09:18,652:INFO:Split ratio set to 0.8.
2021-02-04 17:09:18,652:INFO:Trainingset contains 8556 molecules.
2021-02-04 17:09:18,652:INFO:Building train and test sets...
2021-02-04 17:09:18,655:INFO:Converting train features to tf.tensors...
2021-02-05 02:27:31,469:INFO:Converting test features to tf.tensors...
2021-02-05 04:44:30,121:INFO:Converting train labels to array...
2021-02-05 04:44:30,125:INFO:Converting test labels to array...
2021-02-05 04:44:30,126:INFO:============== Step 3: Model compilation ==============
2021-02-05 04:44:30,126:INFO:Building model...
2021-02-05 04:44:30,210:INFO:Model: "sequential"
2021-02-05
04:44:30,210:INFO:_________________________________________________________________
2021-02-05 04:44:30,210:INFO:Layer (type)                 Output Shape              Param #
2021-02-05
04:44:30,210:INFO:=================================================================
==========
2021-02-05 04:44:30,211:INFO:conv3d (Conv3D)              (None, 6, 60, 60, 64)     640
2021-02-05
04:44:30,211:INFO:_________________________________________________________________
2021-02-05 04:44:30,211:INFO:max_pooling3d (MaxPooling3D) (None, 3, 30, 30, 64)     0
2021-02-05
04:44:30,211:INFO:_________________________________________________________________
2021-02-05 04:44:30,211:INFO:conv3d_1 (Conv3D)            (None, 3, 28, 28, 64)     36928
2021-02-05
04:44:30,211:INFO:_________________________________________________________________
2021-02-05 04:44:30,211:INFO:max_pooling3d_1 (MaxPooling3 (None, 1, 14, 14, 64)     0
2021-02-05
```

```
04:44:30,211:INFO:_________________________________________________________________
2021-02-05 04:44:30,211:INFO:conv3d_2 (Conv3D)           (None, 1, 12, 12, 64)     36928
2021-02-05
04:44:30,211:INFO:_________________________________________________________________
2021-02-05 04:44:30,211:INFO:flatten (Flatten)           (None, 9216)              0
2021-02-05
04:44:30,212:INFO:_________________________________________________________________
2021-02-05 04:44:30,212:INFO:dropout (Dropout)           (None, 9216)              0
2021-02-05
04:44:30,212:INFO:_________________________________________________________________
2021-02-05 04:44:30,212:INFO:dense (Dense)               (None, 32)                294944
2021-02-05
04:44:30,212:INFO:_________________________________________________________________
2021-02-05 04:44:30,212:INFO:dense_1 (Dense)             (None, 32)                1056
2021-02-05
04:44:30,212:INFO:_________________________________________________________________
2021-02-05 04:44:30,212:INFO:dense_2 (Dense)             (None, 1)                 33
2021-02-05
04:44:30,214:INFO:=================================================================
==========
2021-02-05 04:44:30,214:INFO:Total params: 370,529
2021-02-05 04:44:30,214:INFO:Trainable params: 370,529
2021-02-05 04:44:30,214:INFO:Non-trainable params: 0
2021-02-05
04:44:30,215:INFO:_________________________________________________________________
2021-02-05 04:44:30,215:INFO:Compiling the model...
2021-02-05 04:44:30,221:INFO:============== Step 4: Model training ==============
2021-02-05 04:44:30,221:INFO:Enabeling early stopping...
2021-02-05 04:44:30,221:INFO:Start training...
2021-02-05 04:53:15,439:INFO:Plotting loss...
2021-02-05 04:53:15,672:INFO:Plotting metric mean_absolute_error
2021-02-05 04:53:15,850:INFO:Plotting metric mean_squared_error
2021-02-05 04:53:16,034:INFO:============== Step 5: Model evaluation ==============
2021-02-05 04:53:16,034:INFO:Evaluating model...
2021-02-05 04:53:21,347:INFO:Test scores:
2021-02-05 04:53:21,347:INFO:{'loss': 7.694690430071205e-05,
 'mean_absolute_error': 0.006556622218340635,
 'mean_squared_error': 7.694690430071205e-05}
2021-02-05 04:53:21,347:INFO:Making test predictions...
2021-02-05 04:53:27,639:INFO:Plotting ToP plot...
2021-02-05 04:53:27,878:INFO:Plotting Error histogram plot...
2021-02-05 04:53:28,129:INFO:Plotting boxplot...
2021-02-05 04:53:28,291:INFO:============== Step 6: Saving, reporting and cleanup
==============
2021-02-05 04:53:28,291:INFO:Saving model...
2021-02-05 04:53:28,629:WARNING:From
/apps/brussel/CO7/broadwell/software/TensorFlow/2.3.1-foss-2020a-Python-3.8.2/lib/python3.8/site-package
s/tensorflow/python/training/tracking/tracking.py:111: Model.state_updates □□□□□□
tensorflow.python.keras.engine.training) is deprecated and will be removed in a future version.
Instructions for updating:
This property should not be used in TensorFlow 2.0, as updates are applied automatically.
```

2021-02-05 04:53:28,635:WARNING:From /apps/brussel/CO7/broadwell/software/TensorFlow/2.3.1-foss-2020a-Python-3.8.2/lib/python3.8/site-packages/tensorflow/python/training/tracking/tracking.py:111: Layer.updates (from tensorflow.python.keras.engine.base_layer) is deprecated and will be removed in a future version.
Instructions for updating:
This property should not be used in TensorFlow 2.0, as updates are applied automatically.
2021-02-05 04:53:29,970:INFO:Assets written to: /scratch/brussel/102/vsc10255/Experimental-Reactivity-Prediction/code/models/../../models/CM_ZPE_3DCNN/3D_CNN_0.1.3_on_02-04-2021_17.03.51.tf/assets
2021-02-05 04:53:30,224:INFO:Model saved.
2021-02-05 04:53:30,225:INFO:Generating report...
2021-02-05 04:53:32,655:INFO:Page break on page 5 at y=271 for element of height 116