Dwelling Energy Insights – Week 11

Team members: Merel Kreszner, Olivier van Luijk, Rajeev Kalloe,
Santiago Puertas Puchol, Emilio Caba Batuecas & Teo Čurčić

Product Owner: Dr. ir. T.B. Salcedo Rahola

Sprint 5

- Clarification of the project goal:
  - Can we gain energy insights of the characteristics of the dwellings and the people using the data from smart meters?
    - Prediction of the heating system type
    - Prediction of the number of solar panels
    - Prediction of the number of people

- We planned working on these algorithms:
  - Logistic Regression
  - K-nearest neighbor
  - Support Vector Machine
  - Multilayer Perceptron
  - K-means
  - RNN

Done in sprint 5

- We worked on these algorithms:
  - Logistic Regression
  - K-nearest neighbor
  - Support Vector Machine
  - Multilayer Perceptron
  - K-means
  - RNN

- Using the KNMI temperature data
  - T (Temperature)
  - SQ (Duration of the sunshine)
  - Q (Global radiation (J/cm2))
  - N (Cloud cover index (1 - 9))

- Start writing the research paper introduction

SVM with KNMI-data per day for all houses

Before

	precision	recall	f1-score	support
E	0.77	0.39	0.52	6088
WP	0.47	0.37	0.41	5920
Zon	0.46	0.79	0.59	5992
accuracy			0.52	18000
macro avg	0.57	0.52	0.51	18000
weighted avg	0.57	0.52	0.51	18000

After

Score of the model: 0.6074910210364289

	precision	recall	f1-score	support
E	0.60	0.38	0.47	713
WP	0.61	0.85	0.71	876
Zon	0.61	0.47	0.53	360
accuracy			0.61	1949
macro avg	0.61	0.57	0.57	1949
weighted avg	0.61	0.61	0.59	1949

K-nearest Neighbor

Before

	precision	recall	f1-score	support
E	0.49	0.51	0.50	360
WP	0.62	0.66	0.64	465
Zon	0.55	0.39	0.46	175
accuracy			0.56	1000
macro avg	0.55	0.52	0.53	1000
weighted avg	0.56	0.56	0.56	1000

After

	precision	recall	f1-score	support
1	0.89	0.85	0.87	383232
2	0.94	0.94	0.94	497946
3	0.89	0.95	0.92	191824
accuracy			0.91	1073002
macro avg	0.90	0.92	0.91	1073002
weighted avg	0.91	0.91	0.91	1073002

Logistic regression

Before

	precision	recall	f1-score	support
E	0.56	0.56	0.56	351
WP	0.59	0.81	0.68	475
Zon	0.00	0.00	0.00	174
accuracy			0.58	1000
macro avg	0.38	0.46	0.41	1000
weighted avg	0.48	0.58	0.52	1000

After

```
Accuracy on training: 0.715
Accuracy on test: 0.716
True:  [2 2 1 1 2 2 1 2 1 1]
False: [2 2 1 1 2 2 1 1 1 2]
```

	precision	recall	f1-score	support
1	0.72	0.80	0.76	109138
2	0.71	0.92	0.80	142853
3	1.00	0.01	0.02	54581
accuracy			0.72	306572
macro avg	0.81	0.58	0.53	306572
weighted avg	0.77	0.72	0.65	306572

RNN with KNMI-data

delivery	consumption	T	SQ	Q	week	month	season	day	hour
0.0	0.031	14.8	0.0	0.0	37	9	2	12	0
0.0	0.037	14.8	0.0	0.0	37	9	2	12	0
0.0	0.034	14.8	0.0	0.0	37	9	2	12	0
0.0	0.296	14.6	0.0	0.0	37	9	2	12	1
0.0	0.311	14.6	0.0	0.0	37	9	2	12	1

Predictions:

- Heating system type
- No. Of people
- No. Of Solar Panels

LSTM

1 day of data

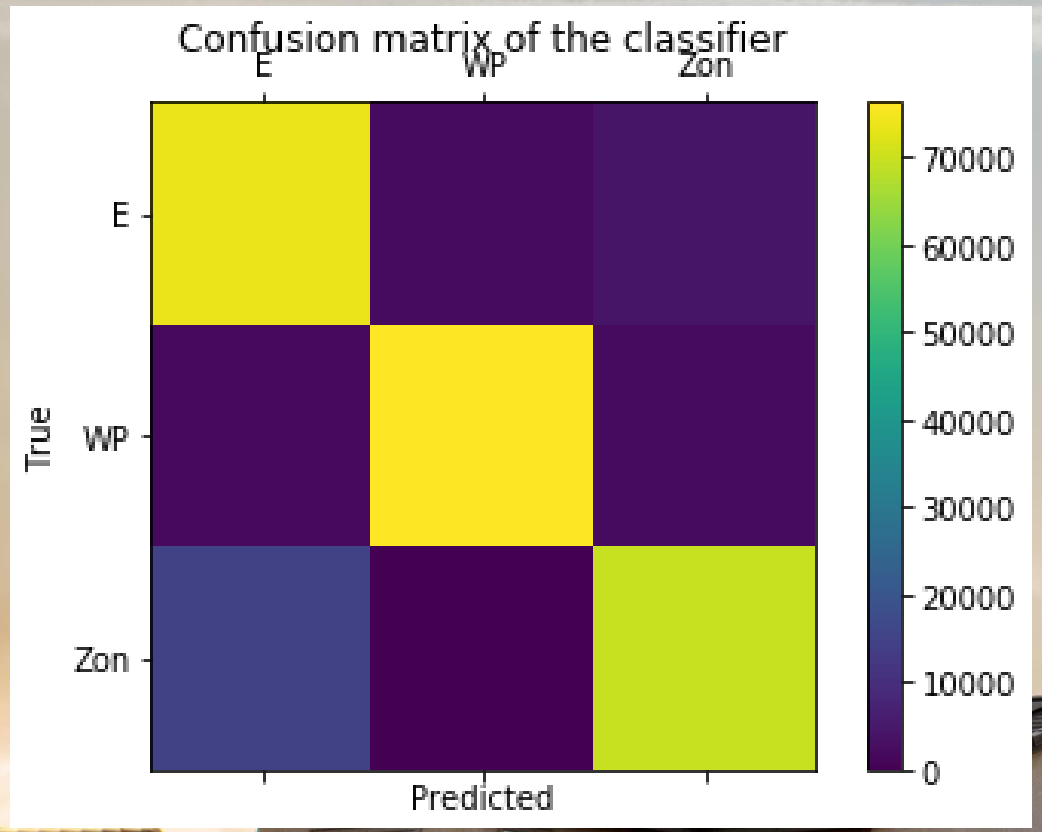
```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm_1 (LSTM)                (None, 92, 30)            4920
_________________________________________________________________
dropout_1 (Dropout)          (None, 92, 30)            0
_________________________________________________________________
dense_1 (Dense)              (None, 92, 4)             124
=================================================================
Total params: 5,044
Trainable params: 5,044
Non-trainable params: 0
_________________________________________________________________
```

possible outputs

```
(samples,rows,df_norm.shape[1])
```

LSTM

Long short-term memory predicting **heating system**

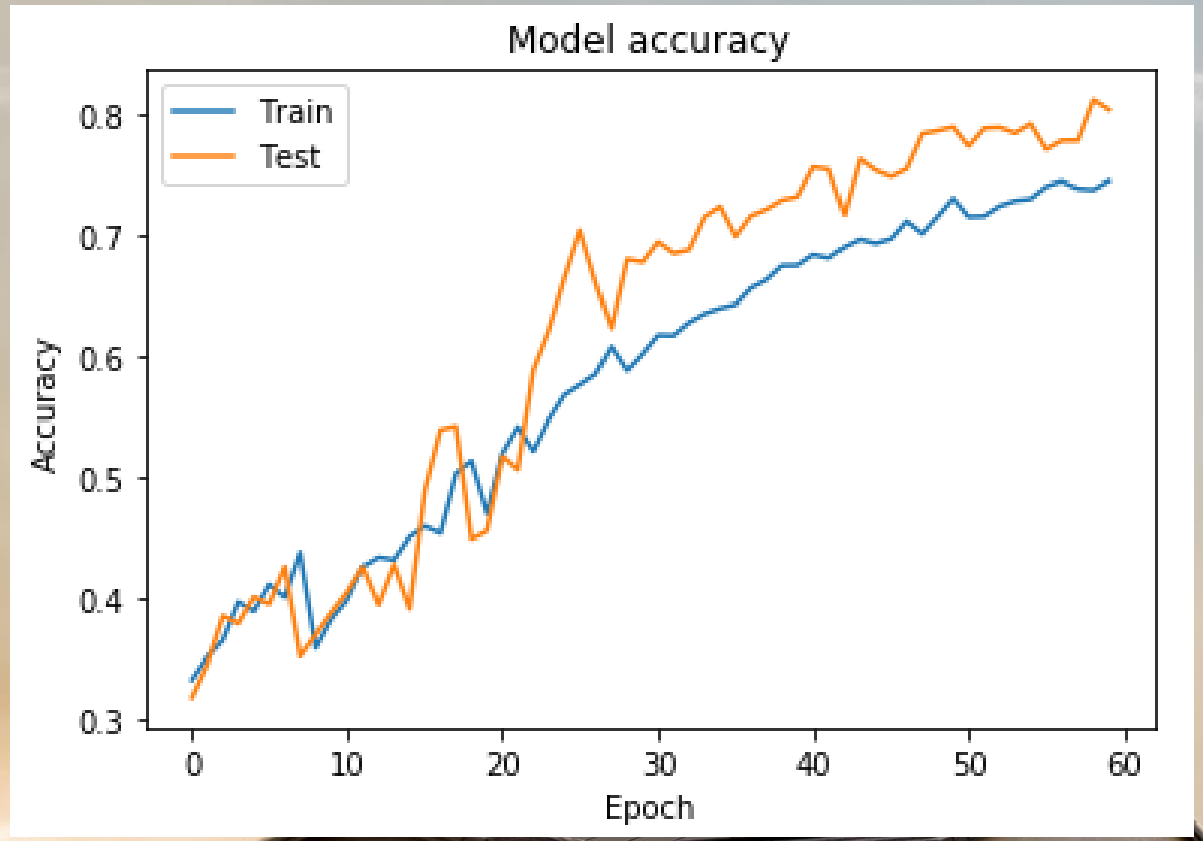


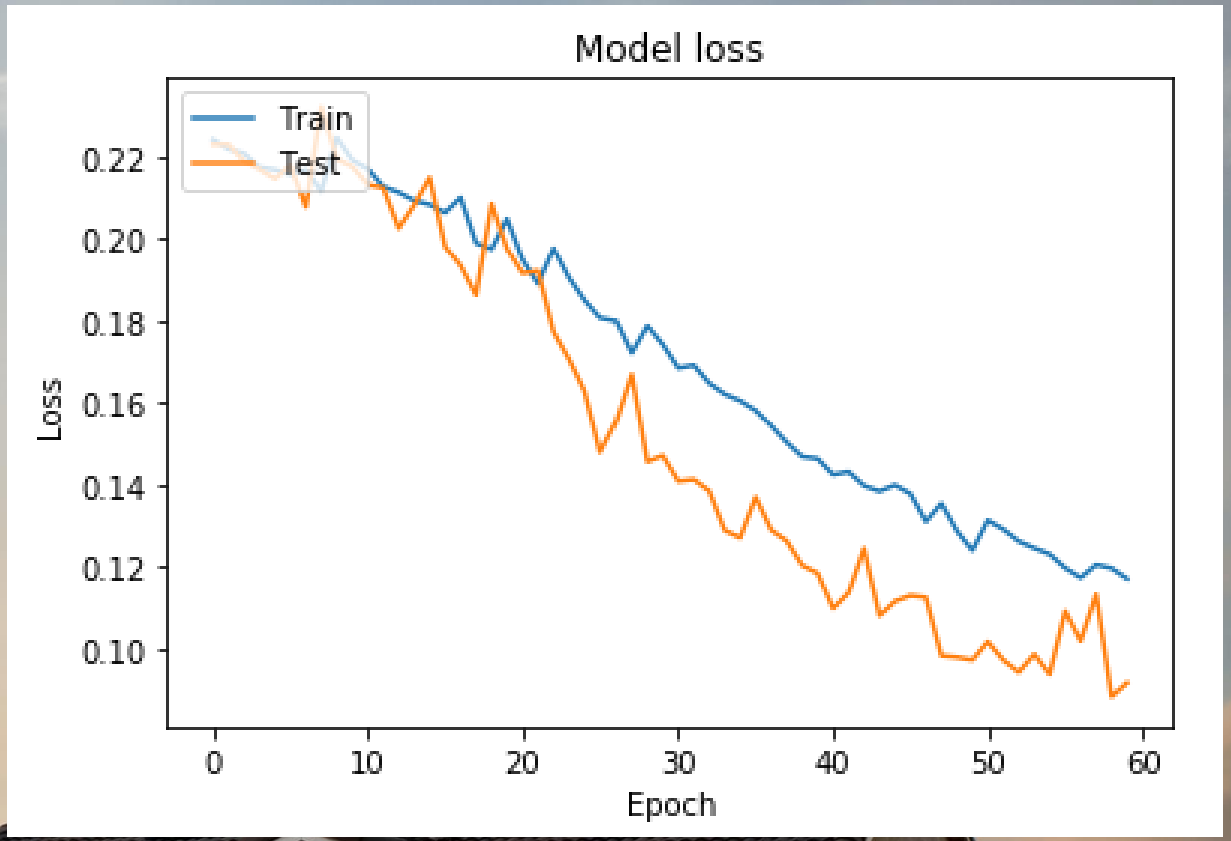
		precision	recall	f1-score	support
E	1.0	0.81	0.92	0.86	80750
WP	2.0	0.97	0.95	0.96	80410
Zon	3.0	0.91	0.82	0.87	85170
accuracy				0.89	246330
macro avg		0.90	0.90	0.90	246330
weighted avg		0.90	0.89	0.89	246330

LSTM

Long short-term memory predicting **heating system**

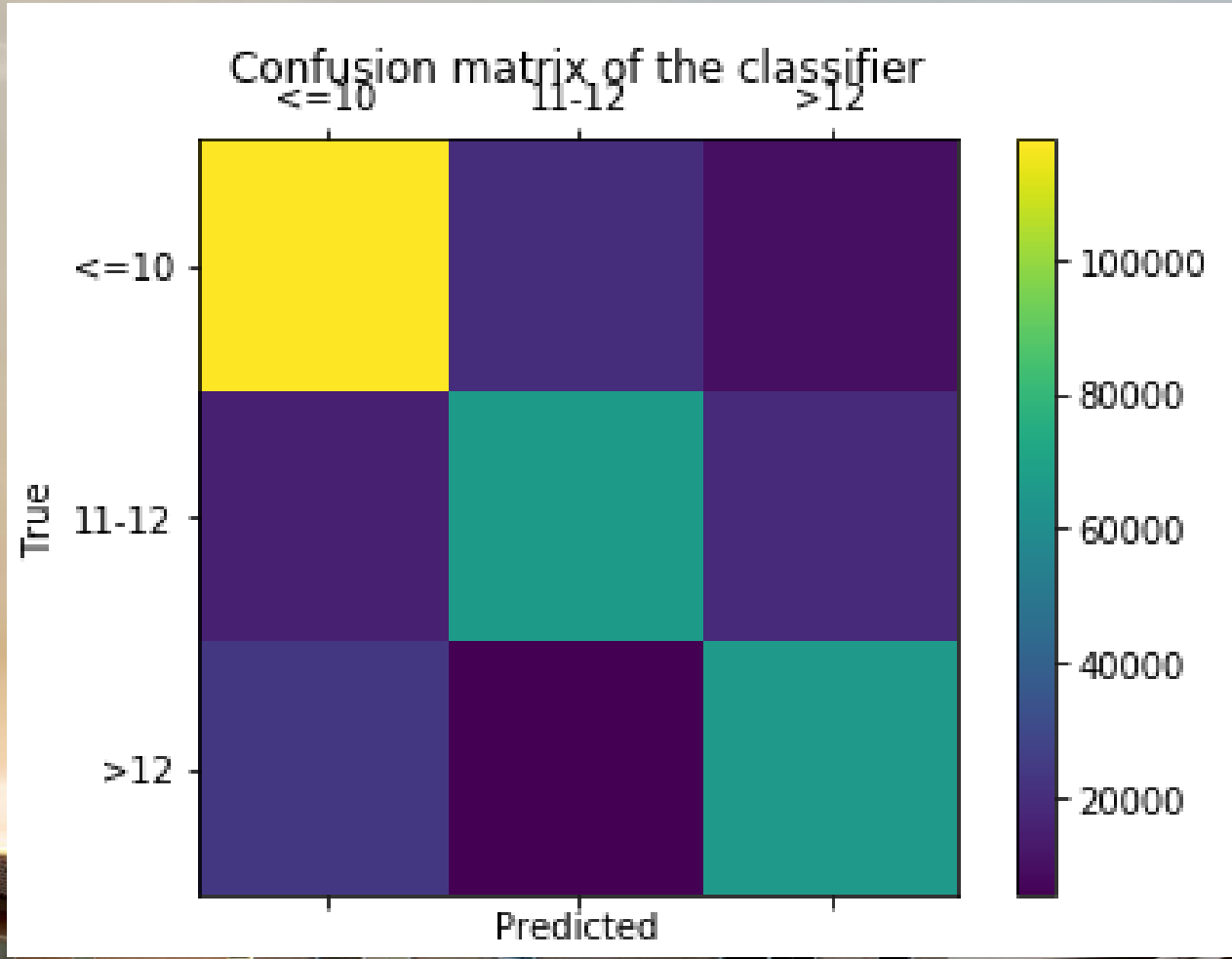
Lr = 0.01
Epoch = 60





LSTM

Long short-term memory predicting **number of solar panels**

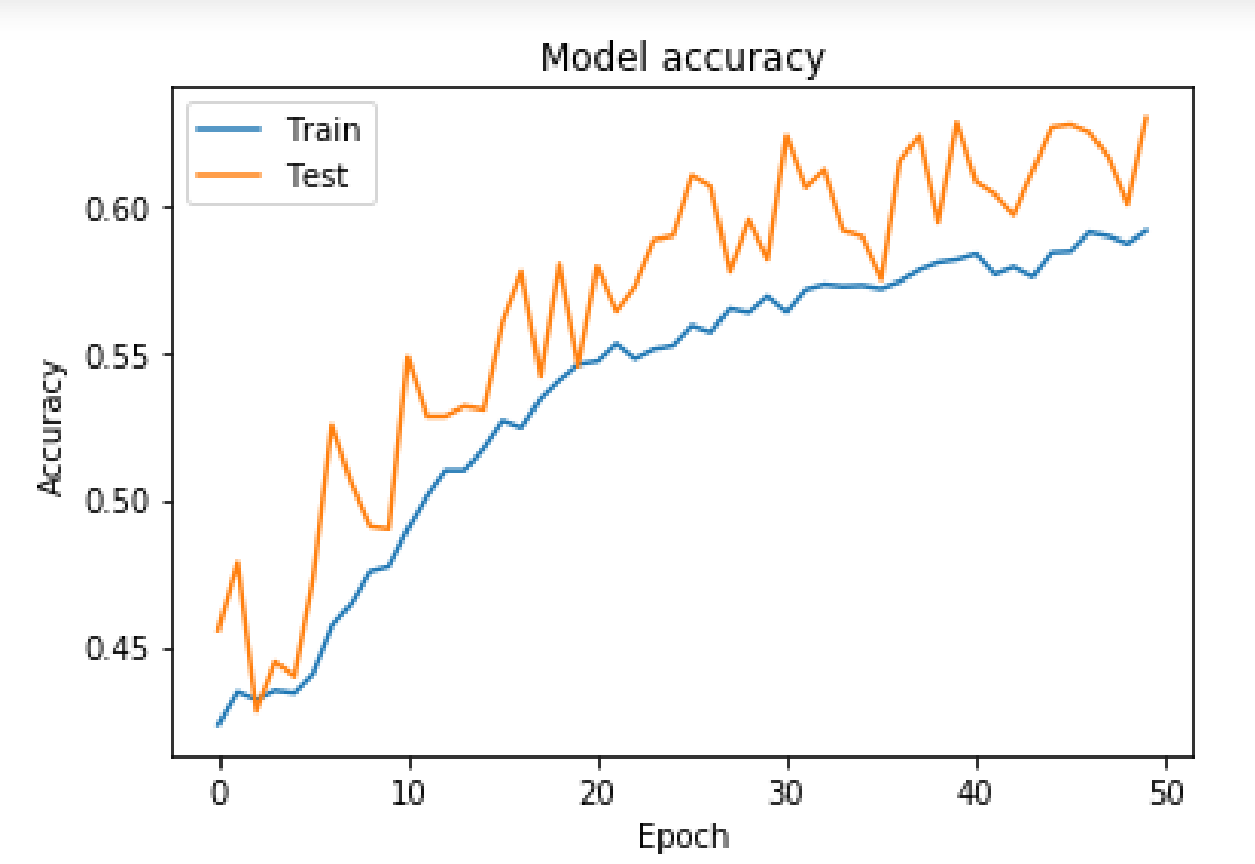


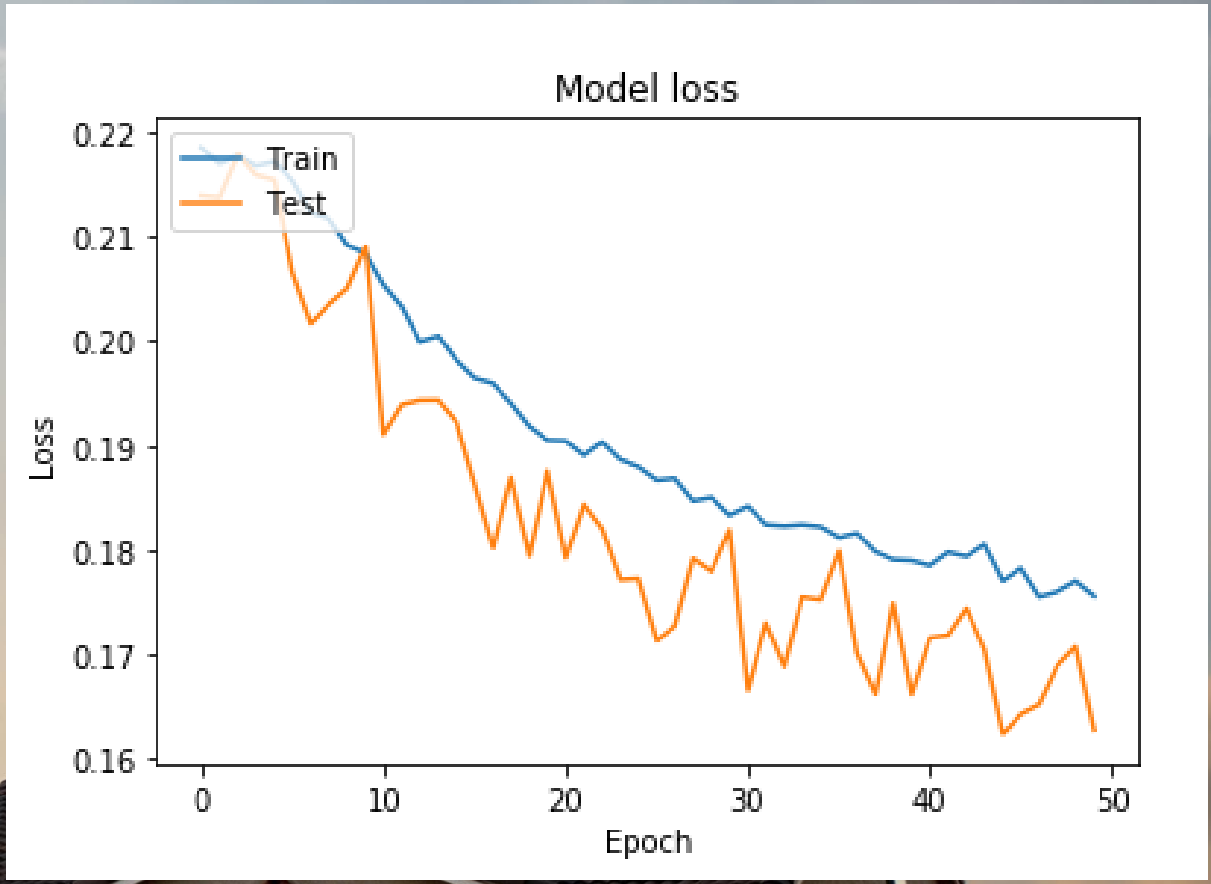
		precision	recall	f1-score	support
<=10	1.0	0.75	0.79	0.77	148488
11-12	2.0	0.72	0.66	0.69	101384
>12	3.0	0.69	0.69	0.69	95036
accuracy				0.73	344908
macro avg		0.72	0.72	0.72	344908
weighted avg		0.73	0.73	0.73	344908

LSTM

Long short-term memory predicting **number of solar panels**

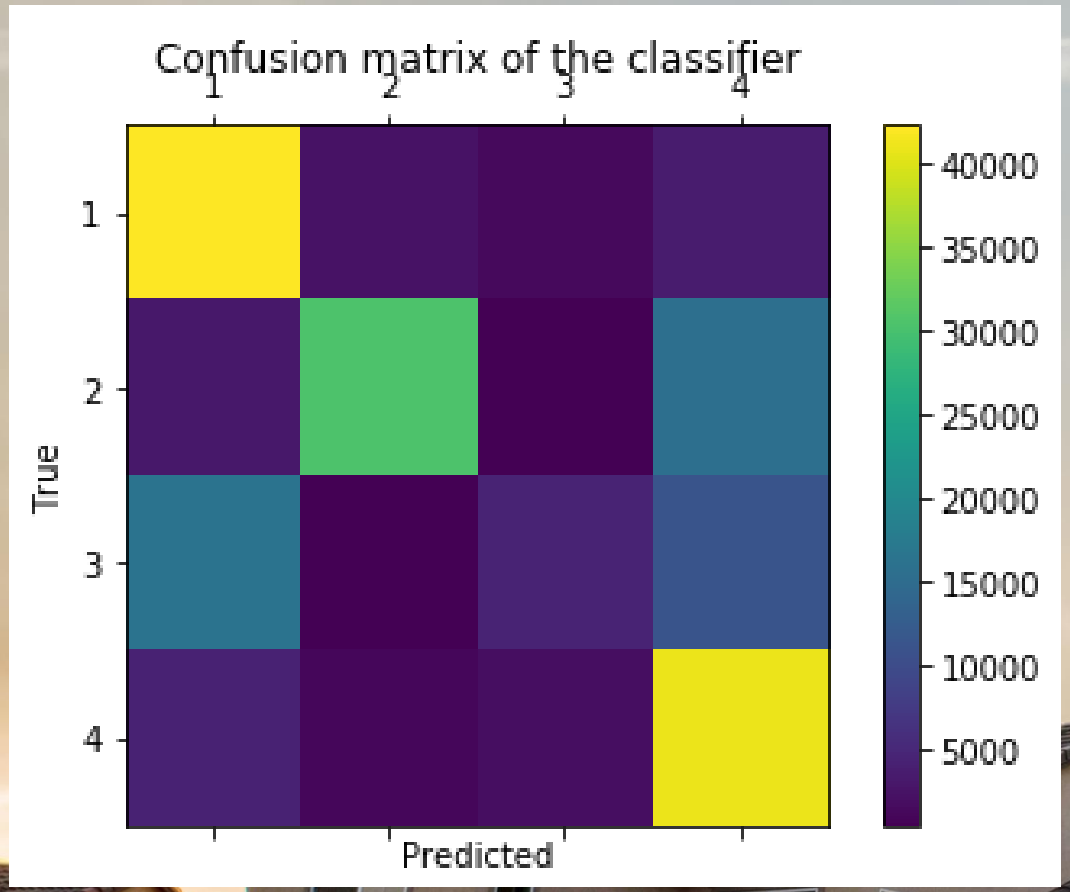
Lr = 0.01
Epochs = 50





LSTM

Long short-term memory predicting **number of people**

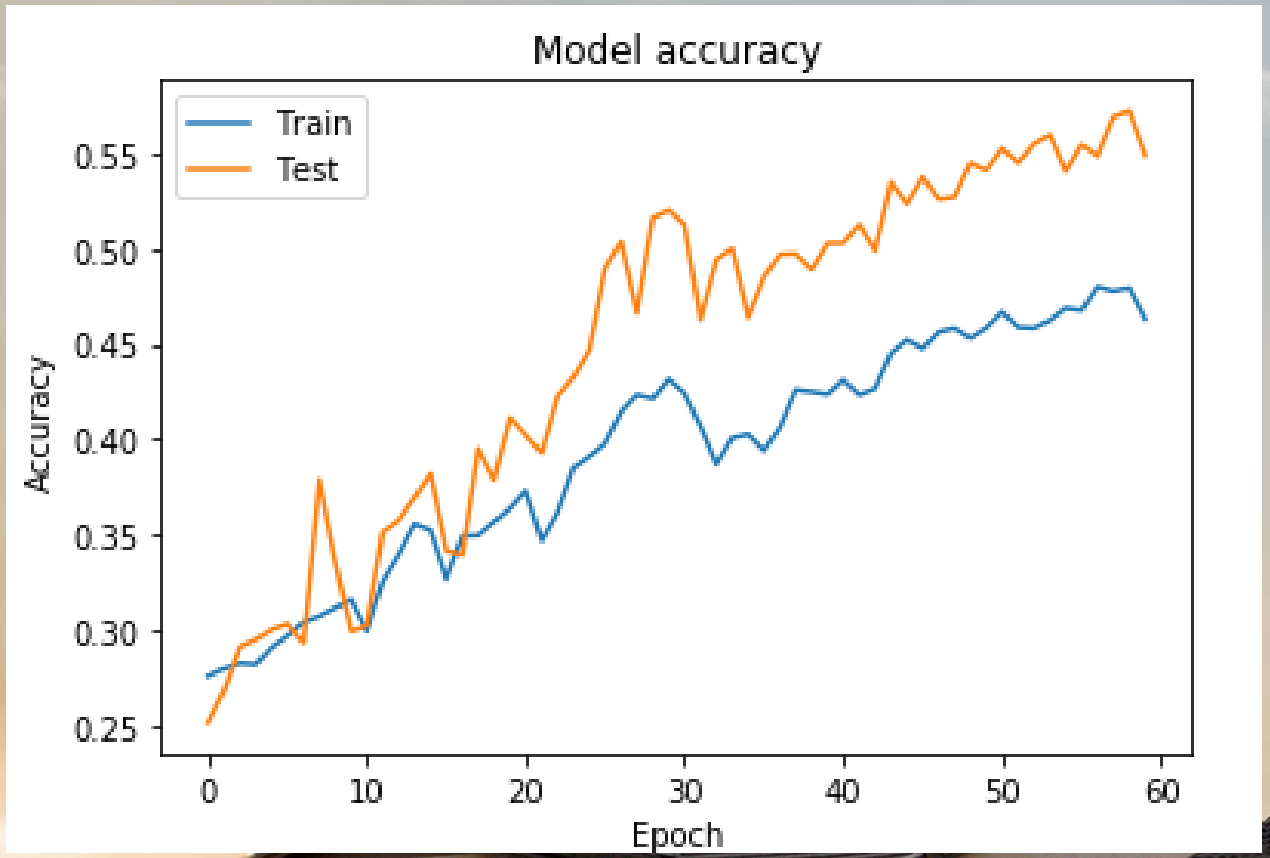


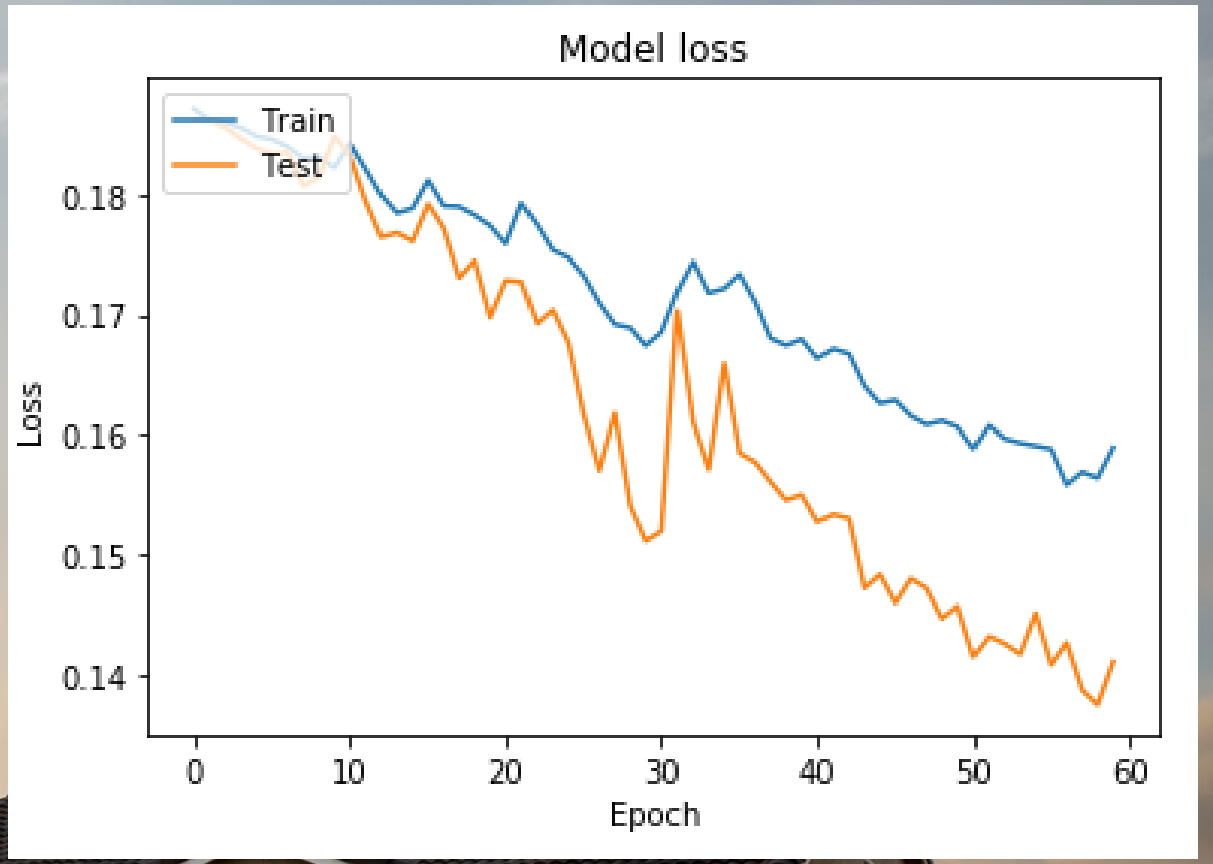
	precision	recall	f1-score	support
1.0	0.64	0.85	0.73	49772
2.0	0.89	0.61	0.73	49956
3.0	0.54	0.14	0.22	32476
4.0	0.57	0.85	0.68	48484
accuracy			0.66	180688
macro avg	0.66	0.61	0.59	180688
weighted avg	0.67	0.66	0.62	180688

LSTM

Long short-term memory predicting **number of people**

Lr = 0.01
Epochs = 60





Remaining tasks

- Improving our models with the one-against-all approach
- Clean up our environment
- Try our models on the new dataset with 120 houses
- Resume writing the research paper

Questions/Feedback

- Are there any questions or feedback based on this presentation?