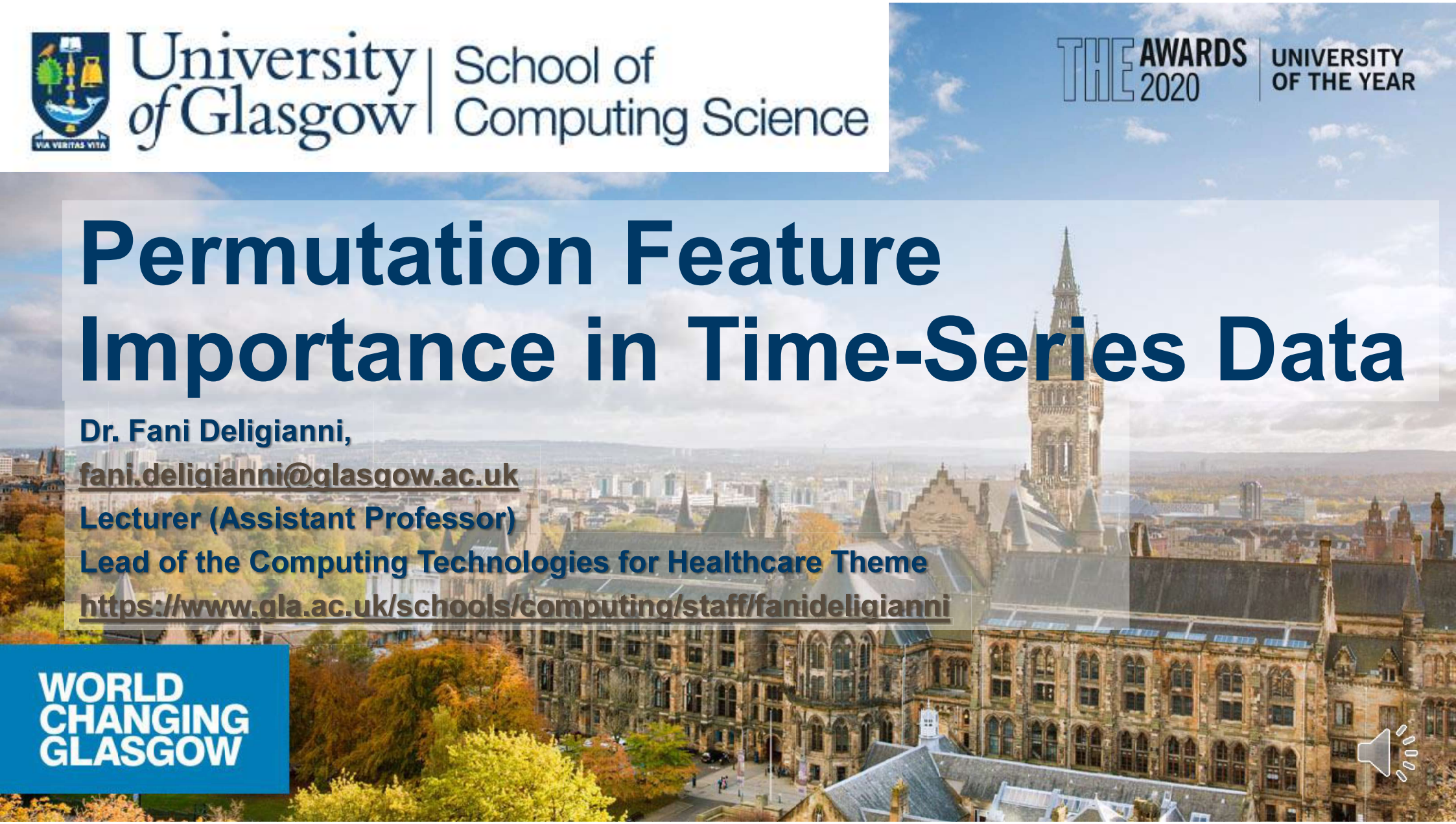
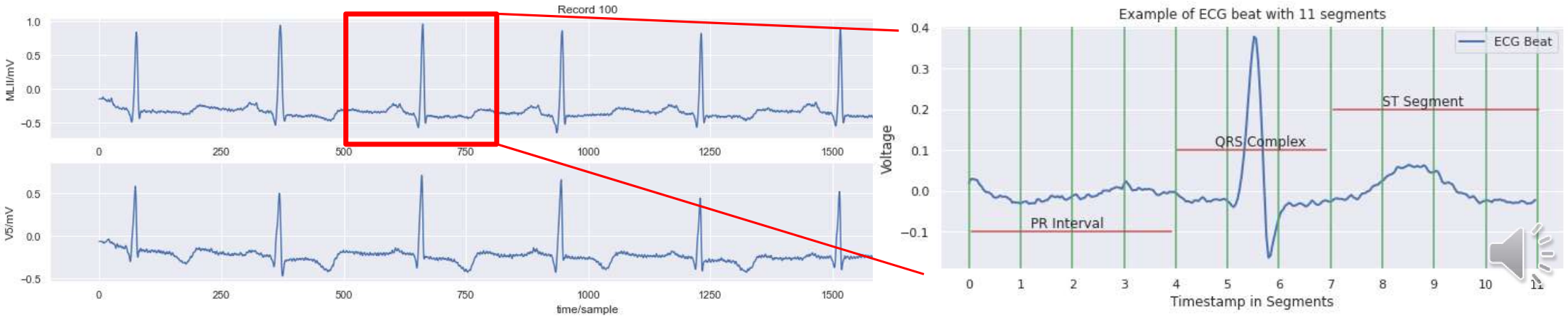


Permutation Feature Importance in Time-Series Data

Dr. Fani Deligianni,
fani.deligianni@glasgow.ac.uk
Lecturer (Assistant Professor)
Lead of the Computing Technologies for Healthcare Theme
https://www.gla.ac.uk/schools/computing/staff/fanideligianni

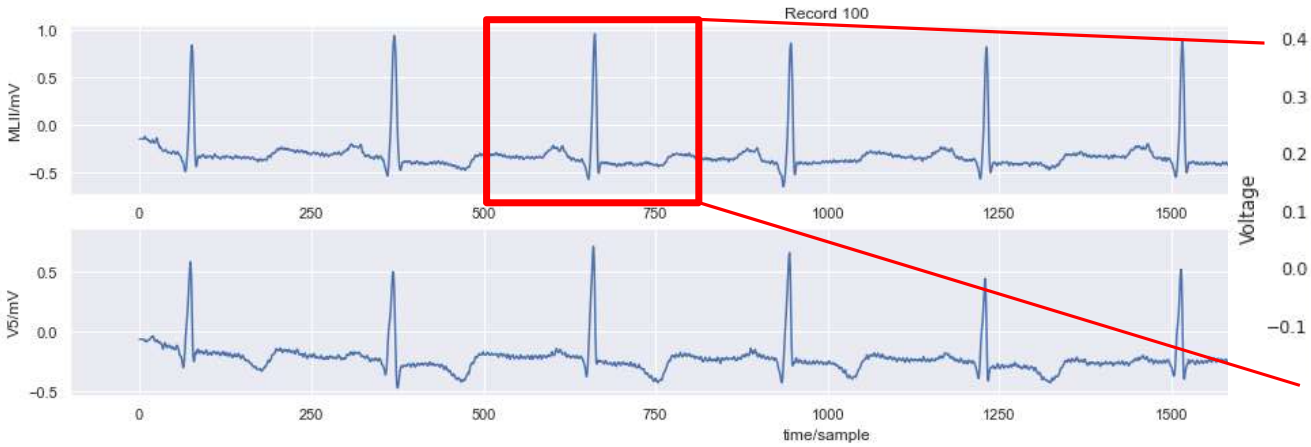
Application of PFI in ECG Classification

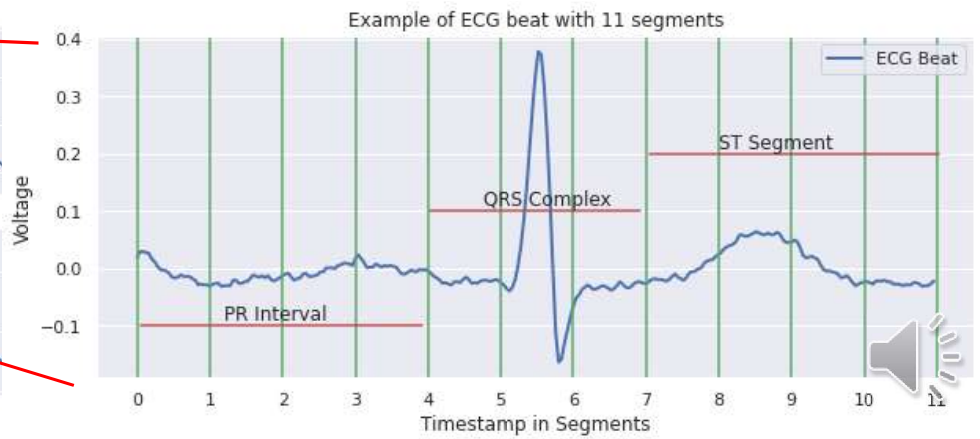
- The ECG beats were divided into slices of 11 segments.
- This helped interpret which segment is being given more importance by the classifier.
- The slices were made by replacing the data points with the average point for each slice.



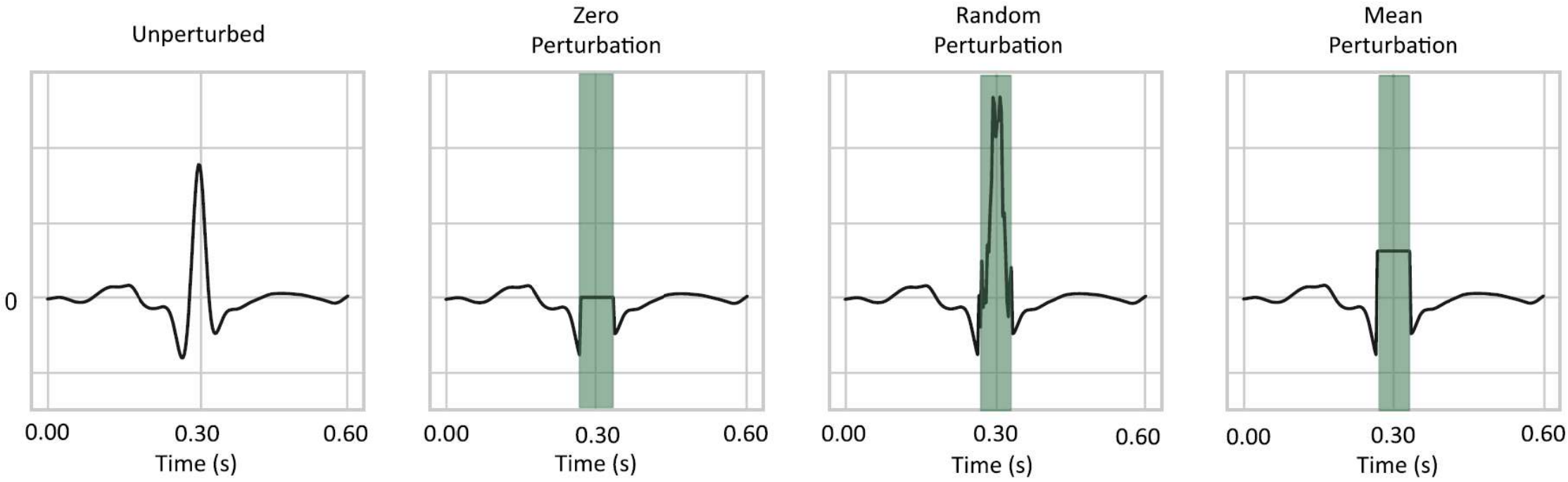
ECG Segmentation

- Segments 1-4 cover the PR interval.
- Segments 5-7 cover the QRS complex
- Segments 8-11 cover the ST segment.
- We expected to see the model focusing on important morphological features of the ECG beat, such as the PR interval, the QRS complex, and the ST segment.

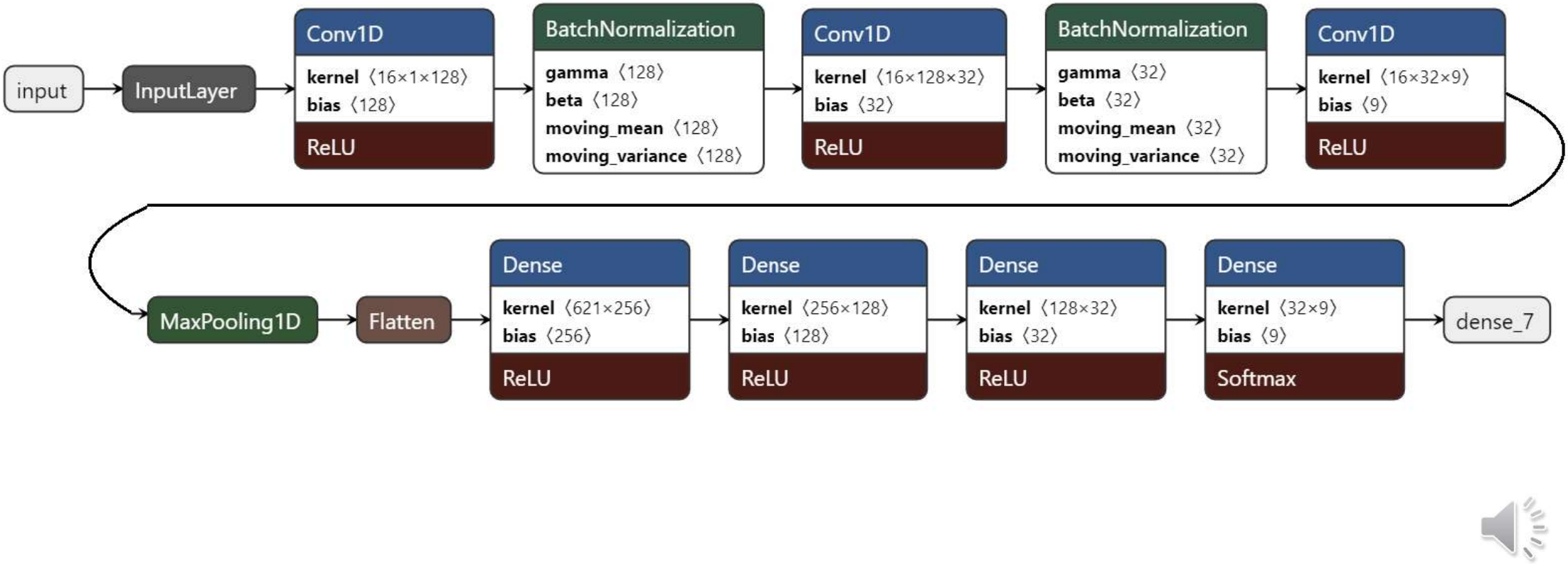




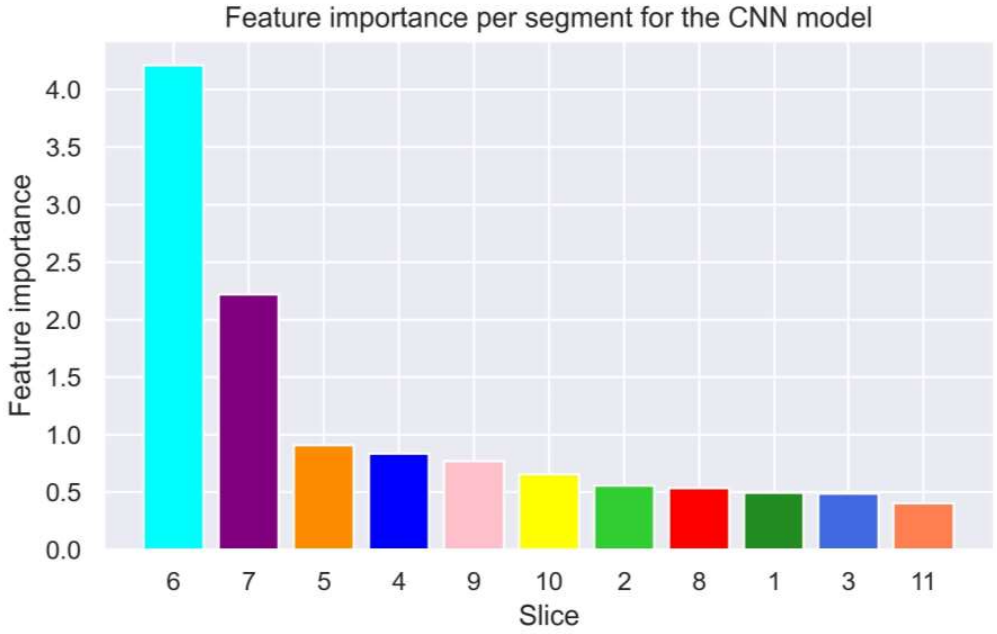
PFI for ECG Classification

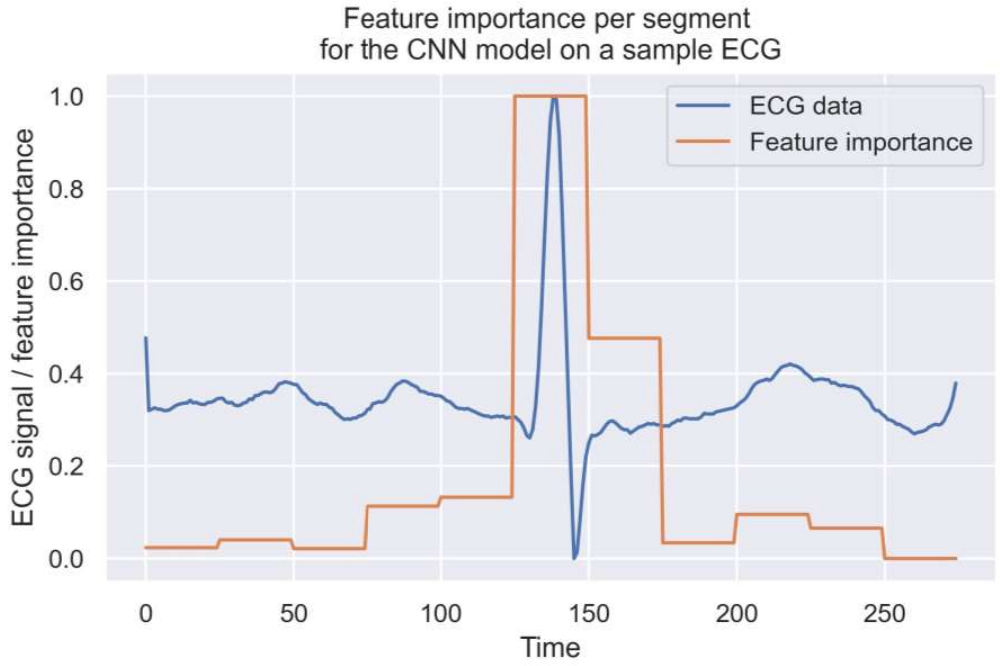


PFI on CNN

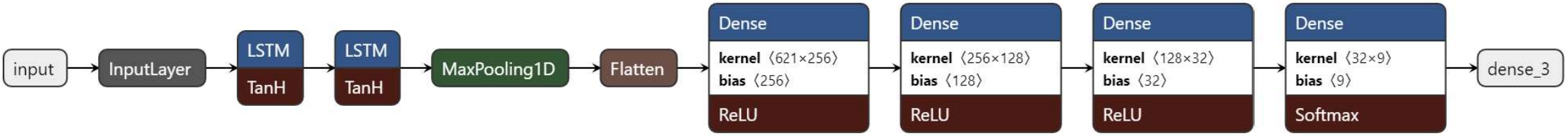

input
InputLayer
Conv1D
kernel 〈16×1×128〉
bias 〈128〉
ReLU
BatchNormalization
gamma 〈128〉
beta 〈128〉
moving_mean 〈128〉
moving_variance 〈128〉
Conv1D
kernel 〈16×128×32〉
bias 〈32〉
ReLU
BatchNormalization
gamma 〈32〉
beta 〈32〉
moving_mean 〈32〉
moving_variance 〈32〉
Conv1D
kernel 〈16×32×9〉
bias 〈9〉
ReLU
MaxPooling1D
Flatten
Dense
kernel 〈621×256〉
bias 〈256〉
ReLU
Dense
kernel 〈256×128〉
bias 〈128〉
ReLU
Dense
kernel 〈128×32〉
bias 〈32〉
ReLU
Dense
kernel 〈32×9〉
bias 〈9〉
Softmax
dense_7

PFI on CNN

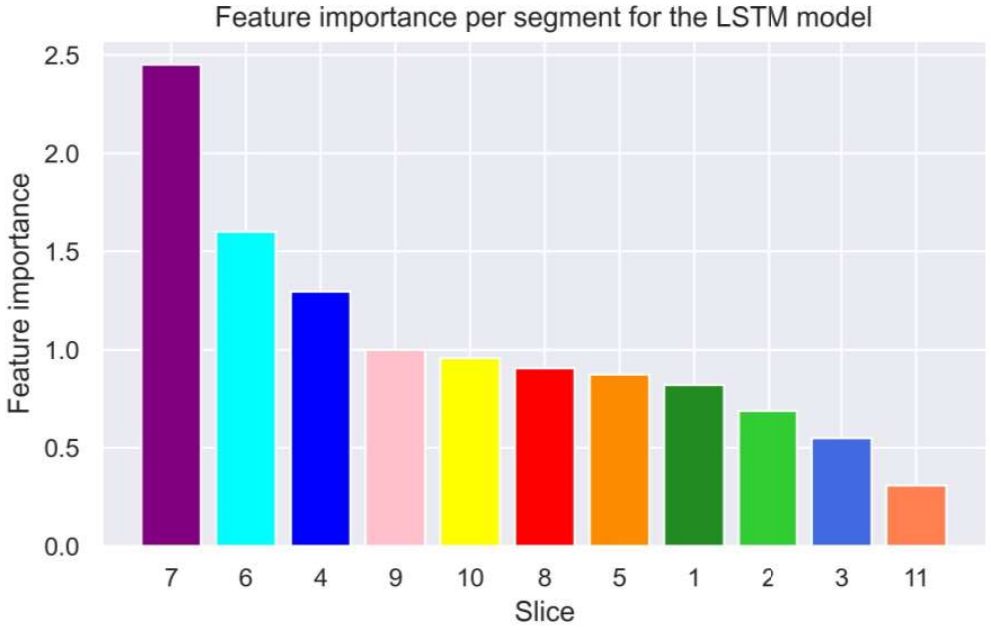

Feature importance per segment for the CNN model
Feature importance
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
6
7
5
4
9
10
2
8
1
3
11
Slice


Feature importance per segment
for the CNN model on a sample ECG
ECG data
Feature importance
ECG signal / feature importance
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
200
250
Time

PFI on LSTM

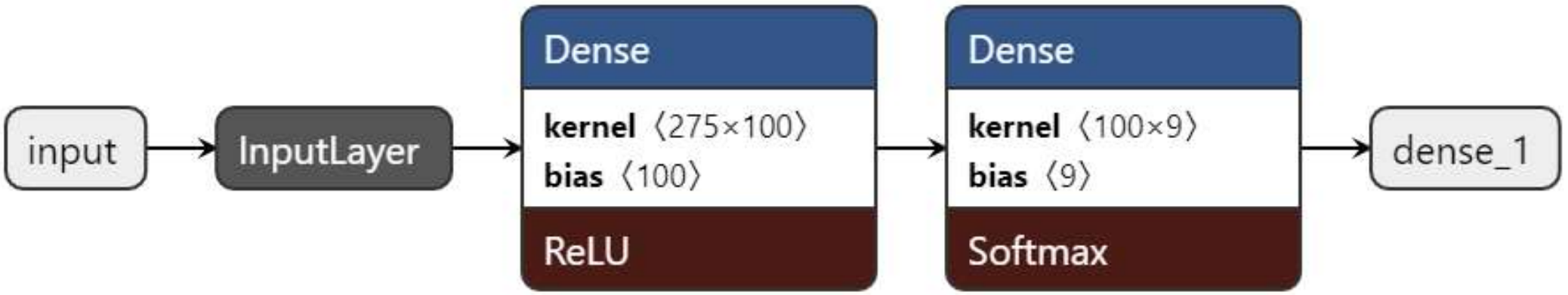

input
InputLayer
LSTM
TanH
LSTM
TanH
MaxPooling1D
Flatten
Dense
kernel 〈621×256〉
bias 〈256〉
ReLU
Dense
kernel 〈256×128〉
bias 〈128〉
ReLU
Dense
kernel 〈128×32〉
bias 〈32〉
ReLU
Dense
kernel 〈32×9〉
bias 〈9〉
Softmax
dense_3

PFI on LSTM

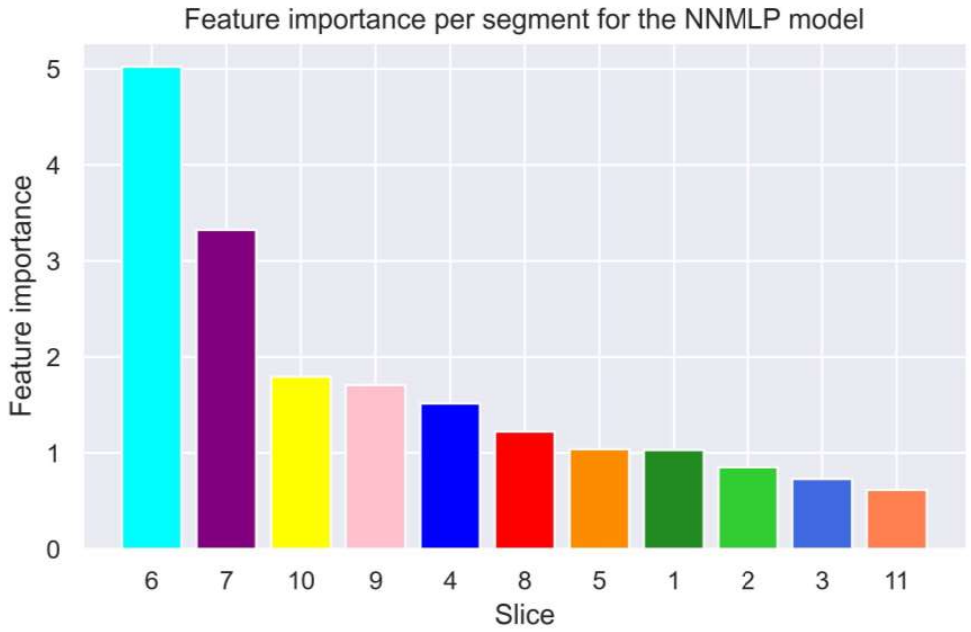

Feature importance per segment for the LSTM model
Feature importance
2.5
2.0
1.5
1.0
0.5
0.0
7
6
4
9
10
8
5
1
2
3
11
Slice

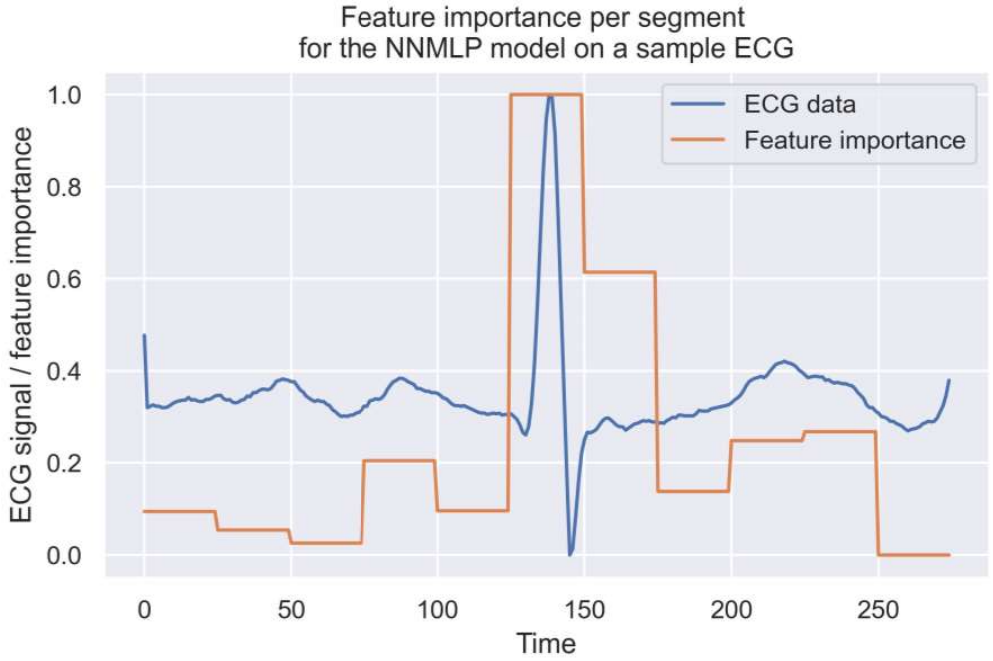

Feature importance per segment
for the LSTM model on a sample ECG
ECG signal / feature importance
1.0
0.8
0.6
0.4
0.2
0.0
ECG data
Feature importance
0
50
100
150
200
250
Time

PFI on MLP


input
InputLayer
Dense
kernel 〈275×100〉
bias 〈100〉
ReLU
Dense
kernel 〈100×9〉
bias 〈9〉
Softmax
dense_1

PFI on MLP


Feature importance per segment for the NNMLP model
Feature importance
Slice
0
1
2
3
4
5
6
7
10
9
4
8
5
1
2
3
11


Feature importance per segment
for the NNMLP model on a sample ECG
ECG signal / feature importance
Time
ECG data
Feature importance
0.0
0.2
0.4
0.6
0.8
1.0
0
50
100
150
200
250

Summary

- Permutation Feature Importance has been originally designed for tabular data
- PFI has applied in time-series data by segmenting the data into segments and shuffle values in each segment
- PFI ignores any spatio-temporal relationships

References

- Ribeiro et al. ‘Model-Agnostic Interpretability of Machine Learning’, ICML Workshop on Human Interpretability in Machine Learning, 2016.
- Mi et al. ‘Permutation-based identification of important biomarkers for complex diseases via machine learning models’, Nature Communications, 2021.
- Y. Jones, F. Deligianni and J Dalton, Improving ECG Classification Interpretability Using Saliency Maps, IEEE BIBE, 2020.
- Neves et al. ‘Interpretable heartbeat classification using local model-agnostic explanations on ECGs’, Computers in Biology and Medicine, 2021.