# EE40098

# Coursework C: Automated Feedback

Submission downloaded 22/01/2026 at 17:25.

Feedback generated on Thursday 22$^{\text{nd}}$ January, 2026 at 17:26.

Saad Ahmed

# University of Bath

## Department of Electronic and Electrical Engineering

**Formative Feedback Only**

All feedback contained within this document is purely indicative of your current solutions performance. Final grades for this assignment will consider a range of additional factors when producing a mark, meaning there may be some slight change between the performance seen here and your final grades. This feedback is intended to provide you with some insight into the performance of your approach prior to submission, allowing you to further improve your solution as the semester progresses.

# 1 Dataset D2

## 1.1 General Data Properties

**Table 1:** Properties of the provided dataset, D2, provided for reference. These may also be found in the coursework specification.

| Property | Value |
|---|---|
| SNR (Approx.) | 60 dB |
| Num. of Classes | 5 |
| Classification Status | Unlabelled |
| Sampling Rate | 25 kHz |

**Table 2:** The D2.mat submission file content metadata. As outlined in the coursework specification, this should include an Index and Class vector. If you have labelled every identified pulse, you should expect the number of class entries to match the number of index entries in your submission data.

| Name | Data Type | Data Shape |
|---|---|---|
| Index | <class 'numpy.ndarray'> | (1, 3758) |
| Class | <class 'numpy.ndarray'> | (1, 3758) |

## 1.2 Submission Performance

The index identification accuracy (i.e. how well you identified where pulses were in the data) was first assessed before class identification was considered. This yielded the following metrics:

**Table 3:** Index identification performance for dataset D2, reported as precision and recall. In both cases, a value closer to 1.0 is better.

| Metric | Value |
|---|---|
| Precision | 0.936402 |
| Recall | 0.883283 |

The class identification was then considered, resulting in the following confusion matrix. Please note, any missed pulses will not be present in this result - so while it may look good, you will also need to consider the results found above in the index performance table.

Predicted Class

|  | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 | 695 | 6 | 3 | 0 | 9 |
| 2 | 4 | 692 | 2 | 1 | 6 |
| 3 | 17 | 25 | 553 | 26 | 15 |
| 4 | 9 | 14 | 21 | 627 | 30 |
| 5 | 5 | 26 | 3 | 175 | 555 |

**Figure 1:** Confusion matrix for the dataset D2 submission file. A good result should have most values sitting along the lead diagonal. The raw count has been used here instead of percentage to provide you with additional insight.

# 2   Dataset D3

## 2.1   General Data Properties

**Table 4:** Properties of the provided dataset, D3, provided for reference. These may also be found in the coursework specification.

| Property | Value |
|---|---|
| SNR (Approx.) | 40 dB |
| Num. of Classes | 5 |
| Classification Status | Unlabelled |
| Sampling Rate | 25 kHz |

**Table 5:** The D3.mat submission file content metadata. As outlined in the coursework specification, this should include an Index and Class vector. If you have labelled every identified pulse, you should expect the number of class entries to match the number of index entries in your submission data.

| Name | Data Type | Data Shape |
|---|---|---|
| Index | <class 'numpy.ndarray'> | (1, 2846) |
| Class | <class 'numpy.ndarray'> | (1, 2846) |

## 2.2   Submission Performance

The index identification accuracy (i.e. how well you identified where pulses were in the data) was first assessed before class identification was considered. This yielded the following metrics:

**Table 6:** Index identification performance for dataset D3, reported as precision and recall. In both cases, a value closer to 1.0 is better.

| Metric | Value |
|---|---|
| Precision | 0.938862 |
| Recall | 0.802643 |

The class identification was then considered, resulting in the following confusion matrix. Please note, any missed pulses will not be present in this result - so while it may look good, you will also need to consider the results found above in the index performance table.

**Figure 2:** Confusion matrix for the dataset D3 submission file. A good result should have most values sitting along the lead diagonal. The raw count has been used here instead of percentage to provide you with additional insight.

# 3   Dataset D4

## 3.1   General Data Properties

**Table 7:** Properties of the provided dataset, D4, provided for reference. These may also be found in the coursework specification.

| Property | Value |
| --- | --- |
| SNR (Approx.) | 20 dB |
| Num. of Classes | 5 |
| Classification Status | Unlabelled |
| Sampling Rate | 25 kHz |

**Table 8:** The D4.mat submission file content metadata. As outlined in the coursework specification, this should include an Index and Class vector. If you have labelled every identified pulse, you should expect the number of class entries to match the number of index entries in your submission data.

| Name | Data Type | Data Shape |
| --- | --- | --- |
| Index | <class 'numpy.ndarray'> | (1, 2423) |
| Class | <class 'numpy.ndarray'> | (1, 2423) |

## 3.2   Submission Performance

The index identification accuracy (i.e. how well you identified where pulses were in the data) was first assessed before class identification was considered. This yielded the following metrics:

**Table 9:** Index identification performance for dataset D4, reported as precision and recall. In both cases, a value closer to 1.0 is better.

| Metric | Value |
| --- | --- |
| Precision | 0.875361 |
| Recall | 0.699769 |

The class identification was then considered, resulting in the following confusion matrix. Please note, any missed pulses will not be present in this result - so while it may look good, you will also need to consider the results found above in the index performance table.

Predicted Class

|  |  | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| | 1 | 549 | 19 | 9 | 2 | 4 |
| | 2 | 70 | 459 | 3 | 4 | 29 |
| Actual Class | 3 | 6 | 9 | 162 | 2 | 7 |
| | 4 | 12 | 11 | 78 | 155 | 29 |
| | 5 | 13 | 42 | 30 | 151 | 266 |

**Figure 3:** Confusion matrix for the dataset D4 submission file. A good result should have most values sitting along the lead diagonal. The raw count has been used here instead of percentage to provide you with additional insight.

# 4   Dataset D5

## 4.1   General Data Properties

**Table 10:** Properties of the provided dataset, D5, provided for reference. These may also be found in the coursework specification.

| Property | Value |
| --- | --- |
| SNR (Approx.) | 0 dB |
| Num. of Classes | 5 |
| Classification Status | Unlabelled |
| Sampling Rate | 25 kHz |

**Table 11:** The D5.mat submission file content metadata. As outlined in the coursework specification, this should include an Index and Class vector. If you have labelled every identified pulse, you should expect the number of class entries to match the number of index entries in your submission data.

| Name | Data Type | Data Shape |
| --- | --- | --- |
| Index | <class 'numpy.ndarray'> | (1, 2087) |
| Class | <class 'numpy.ndarray'> | (1, 2087) |

## 4.2   Submission Performance

The index identification accuracy (i.e. how well you identified where pulses were in the data) was first assessed before class identification was considered. This yielded the following metrics:

**Table 12:** Index identification performance for dataset D5, reported as precision and recall. In both cases, a value closer to 1.0 is better.

| Metric | Value |
| --- | --- |
| Precision | 0.618112 |
| Recall | 0.499613 |

The class identification was then considered, resulting in the following confusion matrix. Please note, any missed pulses will not be present in this result - so while it may look good, you will also need to consider the results found above in the index performance table.

**Figure 4:** Confusion matrix for the dataset D5 submission file. A good result should have most values sitting along the lead diagonal. The raw count has been used here instead of percentage to provide you with additional insight.

# 5 Dataset D6

## 5.1 General Data Properties

**Table 13:** Properties of the provided dataset, D6, provided for reference. These may also be found in the coursework specification.

| Property | Value |
| --- | --- |
| SNR (Approx.) | <0 dB |
| Num. of Classes | 5 |
| Classification Status | Unlabelled |
| Sampling Rate | 25 kHz |

**Table 14:** The D6.mat submission file content metadata. As outlined in the coursework specification, this should include an Index and Class vector. If you have labelled every identified pulse, you should expect the number of class entries to match the number of index entries in your submission data.

| Name | Data Type | Data Shape |
| --- | --- | --- |
| Index | <class 'numpy.ndarray'> | (1, 2937) |
| Class | <class 'numpy.ndarray'> | (1, 2937) |

## 5.2 Submission Performance

The index identification accuracy (i.e. how well you identified where pulses were in the data) was first assessed before class identification was considered. This yielded the following metrics:

**Table 15:** Index identification performance for dataset D6, reported as precision and recall. In both cases, a value closer to 1.0 is better.

| Metric | Value |
| --- | --- |
| Precision | 0.541369 |
| Recall | 0.406546 |

The class identification was then considered, resulting in the following confusion matrix. Please note, any missed pulses will not be present in this result - so while it may look good, you will also need to consider the results found above in the index performance table.

Predicted Class

|  | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 | 297 | 30 | 72 | 20 | 83 |
| 2 | 297 | 87 | 39 | 14 | 101 |
| 3 | 39 | 9 | 39 | 1 | 34 |
| 4 | 47 | 16 | 50 | 3 | 62 |
| 5 | 68 | 34 | 50 | 18 | 80 |

Actual Class

**Figure 5:** Confusion matrix for the dataset D6 submission file. A good result should have most values sitting along the lead diagonal. The raw count has been used here instead of percentage to provide you with additional insight.