

EE32009

# Coursework: Automated Feedback

Submission downloaded 17/11/2025 at 17:32.

Feedback generated on Monday 17<sup>th</sup> November, 2025 at 18:57.

Joshua Poole

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, 3746)
Class	<class ‘numpy.ndarray’>	(1, 3746)

## 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	0.991725
Recall	0.93248

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
Actual Class	1	83	152	188	203	122
	2	93	154	178	192	133
	3	71	131	151	170	124
	4	84	164	186	180	153
	5	91	153	211	209	139

**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, 3314)
Class	<class ‘numpy.ndarray’>	(1, 3314)

### 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	0.909475
Recall	0.905377

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
Actual Class	1	4	139	190	228	57
	2	12	143	212	245	69
	3	8	88	153	192	46
	4	12	134	200	225	55
	5	9	128	184	221	60

**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, 3868)
Class	<class ‘numpy.ndarray’>	(1, 3868)

### 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	0.658221
Recall	0.839987

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
Actual Class	1	0	99	242	253	16
	2	1	97	228	226	24
	3	0	48	112	121	12
	4	0	89	206	197	19
	5	1	89	237	210	19

**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, 5068)
Class	<class ‘numpy.ndarray’>	(1, 5068)

### 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	0.375296
Recall	0.736638

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
Actual Class	1	0	139	188	97	26
	2	0	132	236	99	40
	3	1	49	87	38	17
	4	0	89	146	61	18
	5	0	131	176	93	39

**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, 5708)
Class	<class ‘numpy.ndarray’>	(1, 5708)

### 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	0.412754
Recall	0.602403

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
Actual Class	1	16	64	126	231	171
	2	20	63	145	266	141
	3	5	34	41	101	72
	4	3	42	62	153	102
	5	6	37	92	229	134

**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.