Project Name	Issue URL	Issue number	Fix PR	Tuan	Anj	Scott	Tuan comment	Anj comment	Scott comment
ludwig-ai/ludwig	https://github.com/l	484	1174	3. Tensor and Input	3. Tensor and Input	3. Tensor and Input	Selection of data format		
sktime/sktime	https://github.com/s	3748	3760	4. Training Process	5. Third party usage	5. Third party usage	Inefficient testing procedure		
intel/dffml	https://github.com/i	435	634	6. Other	6. Other	6. Other	Documentation		
medipixel/rl_algorithms	https://github.com/r	129	140	2. Model	6. Other	4. Training Process	Inference, check performance of model. However, the fix is the hyperparam in the training process		
catalyst-team/catalyst	https://github.com/c	130	182	5. Third party usage	5. Third party usage	5. Third party usage	Tensorboard, tensorflow usage		
blue-oil/blueoil	https://github.com/b	484	1072	5. Third party usage	5. Third party usage	5. Third party usage	Remove pandas dependency to reduce disk size		Could be tensor and input
sktime/sktime	https://github.com/s	2968	2970	3. Tensor and Input	3. Tensor and Input	3. Tensor and Input	Data related		
sktime/sktime	https://github.com/s	2386	2418	6. Other	6. Other	6. Other	Warning message		
sktime/sktime	https://github.com/s	3169	3231	3. Tensor and Input	3. Tensor and Input	3. Tensor and Input	Data related		
sktime/sktime	https://github.com/s	1687	1874	6. Other	6. Other	6. Other	code refactor to add new kinds of prediction	New functionality for a model	
medipixel/rl_algorithms	https://github.com/r	243	272	3. Tensor and Input	3. Tensor and Input	3. Tensor and Input	Change input type		
sktime/sktime	https://github.com/s	1079	2609	5. Third party usage	3. Tensor and Input	5. Third party usage	Prophet is a forecasting library from Facebook, wrong setting for it		
kymatio/kymatio	https://github.com/k	568	578	6. Other	6. Other	6. Other	Documentation		
aiqm/torchani	https://github.com/a	511	512	3. Tensor and Input	3. Tensor and Input	3. Tensor and Input	Add check for input type		Could be model but related to data
ludwig-ai/ludwig	https://github.com/l	1181	1326	3. Tensor and Input	3. Tensor and Input	3. Tensor and Input	Division by 0 with a single record, inference error, if they provide multiple audio records it works. Data related		
hackingmaterials/automatminer	https://github.com/h	241	246	4. Training Process	2. Model	4. Training Process			Occurs during training
kymatio/kymatio	https://github.com/k	775	883	5. Third party usage	5. Third party usage	5. Third party usage			Could also be "Tensor and Input". However, as issue mentioned "TF2.4 works flawlessly" seems to be a library issue.
sktime/sktime	https://github.com/s	3474	3481	2. Model	2. Model	2. Model			
ludwig-ai/ludwig	https://github.com/l	1093	1103	4. Training Process	4. Training Process	4. Training Process			
sktime/sktime	https://github.com/s	2444	2609	5. Third party usage	3. Tensor and Input	5. Third party usage	Similar to #13, sktime is DL framework lib which use other library, this could be treated as hyper param, config in trai	Duplicate of 1079	
sktime/sktime	https://github.com/s	3816	3820	4. Training Process	5. Third party usage	5. Third party usage	optimiser		Could also be "Training Process". However, issue discussion seems to suggest is due to an incompatible version of keras.
sktime/sktime	https://github.com/s	4067	4075	2. Model	2. Model	2. Model	model initialisation		
kymatio/kymatio	https://github.com/k	461	459	3. Tensor and Input	3. Tensor and Input	3. Tensor and Input	Data preprocessing preparation		
intel/dffml	https://github.com/i	1181	1226	6. Other	4. Training Process	4. Training Process	Rename accuracy to score		Is a naming issue, but is important for understanding the purpose of the metric, thus could be treated as incorrect validation/testing procedure
kymatio/kymatio	https://github.com/k	542	576	3. Tensor and Input	3. Tensor and Input	3. Tensor and Input			Is within documentation, but still a code issue
blue-oil/blueoil	https://github.com/b	805	807	2. Model	6. Other	2. Model	Parameter in prediction function, inference		Appears to be benchmarking the time to process an image (at inferrence time)
sktime/sktime	https://github.com/s	1290	1232	3. Tensor and Input	3. Tensor and Input	3. Tensor and Input			
qubvel/efficientnet	https://github.com/q	51	52	2. Model	6. Other	2. Model			Appears to be an issue with using different backends (e.g. TensorFlow), but isn't an issue with the third-party library itself
sktime/sktime	https://github.com/s	3084	3085	4. Training Process	3. Tensor and Input	3. Tensor and Input	Pre/post-processing issue		
ludwig-ai/ludwig	https://github.com/l	1389	1402	4. Training Process	4. Training Process	4. Training Process			Possibly GPU related, but this does not seem to be the main issue

Instruction: Either create a copy of this spreadsheet or download it to your local machine, add a column for the label a note for discussion if needed, when done, send Tuan via Slack
Open the Issue URL, read the issue title, issue description, comments, code change, PR, project description to label.
Go through the list of categories bellow from top to bottom and pick the first one you think it's a match, there are 6 categories in total.

Note for labelers:
Enhancement and non-critical issues (issues that does not crash the program) can be label using the categories above (e.g. inappropriate use of data format)
Code refactoring can be in the first 5 categories, if not, label as "Other"

1. GPU Usage	**Incorrect or inefficient usage of GPUs, wrong reference to GPU device, failed parallelism, incorrect state sharing between subprocesses, faulty transfer of data to a GPU device.**
2. Model	**Inappropriate, inefficient or incorrect model initialisation, choice of architecture**
	Inappropriate use of activation function, incorrect properties for a neural network layer, missing, redundant or wrong layer. Error occur during inference.
3. Tensor and Input	**Error or inefficiency in data quality such as low quality data, noisy data, imbalanced data, insufficient data.**
	Inappropriate preprocessing of data such as scaling, normalisation, feature engineering.
	Incorrect shapes of input, wrong dimensions, size, inappropriate file type, encoding, selection of data format
4. Training Process	**Inappropriate or inefficient training processes excluding data related problem, such as inappropriate batch sizes, learning rates.**
	Hyperparameter issues such as learning rate, dropout rate, number of epochs. Inappropriate optimiser
	Inappropriate choice of loss function when using during training.
	Inefficient or incorrect validation/ testing procedure.
5. Third party usage	**Inappropriate usage of third-party programs or libraries, such as TensorFlow, PyTorch, Keras.**
6. Other	**Documentation issues or anything unrelated to the 5 categories above.**

Tuan - Anj

Cohen's Kappa for 2 Raters (Weights: unweighted)

Value of Kappa	Level of Agreement	% of Data that are Reliable
0–.20	None	0–4%
.21–.39	Minimal	4–15%
.40–.59	Weak	15–35%
.60–.79	Moderate	35–63%
.80–.90	Strong	64–81%
Above.90	Almost Perfect	82–100%

Subjects = 30
Raters = 2
Kappa = 0.571 (WEAK)

z = 6.19
p-value = 6.07e-10

Average: 0.671 (MODERATE)

Tuan - Scott

Cohen's Kappa for 2 Raters (Weights: unweighted)

Subjects = 30
Raters = 2
Kappa = 0.787 (MODERATE)

z = 8.43
p-value = 0

Scott- Anj

Cohen's Kappa for 2 Raters (Weights: unweighted)

Subjects = 30
Raters = 2
Kappa = 0.655 (MODERATE)

z = 7.17
p-value = 7.52e-13