Project Name Issue URL Relabel kymatio/kymati https://github.c	Issue numbe					Tuan's Note	Ani Note		
			Tuan	Anj	Scott	tuan's Note	Anj Note	Scott note	
netrack/keras-n https://github.c			888 Tensors & Inputs	Third-party Usage	Third-party Usage				
	-		23 Training - Validation / Testing	Tensors & Inputs	Third-party Usage	Value error when using numpy.precision and recall, could also be third party usage	Could potentially also fall under Training - Validation/ Testing		Depend on what Training - Validation / Testing
ludwig-ai/ludwi https://github.c				Model - Model Type & Properties	Model - Model Type & Properties	Wrong configuration to an encoder, it's not layer properties because it's not a layer		Could be N/A based on software issue	
atomistic-mach https://github.c	-		216 N/A	N/A	N/A	Enhancement, request for a metric calculation feature	"tensor" is used in the physics/engineering sense, this is not an ML issue	Feature request, algorithm category?	
kymatio/kymati https://github.c			449 Tensors & Inputs	Tensors & Inputs	N/A	Some batch size can not be processed, enhance by allowing it	Answer seems to indicate that this relates to batch shape. Could also be "Training - Training Process"	Wasn't an issue	
ludwig-ai/ludwi https://github.c				Training - Validation/ Testing	Training - Training Process	Eval batch size when training not scaled up with train batch size	Could also be "Training - Training Process" as even though name is "eval_batch_size", it could be used as part of training	ning process, e.g. for early stopping	
hackingmateria https://github.c			246 Training - Validation / Testing	Training - Validation/ Testing	Training - Validation/ Testing:	Difficult to evaluate and compare models, add more diagnostic pipeline	Not sure if "Training" is an appropriate top level category, but is still applicable for "Validation/Testing"		
sktime/sktime https://github.c	6 30	009 3	3129 N/A	N/A	N/A	can't predict interval in some cases, SE issue	Possibly "Model Type & Properties"	Hard to know what is happening	
ludwig-ai/ludwi https://github.c	4 18	371 1	1872 Model - model type & properties	Third-party Usage	Training - Training Process	model weight not written, wrong model initialisation	Failed to call in right order		Training - Training Process
kymatio/kymati https://github.c	6 9	02	919 N/A	N/A	Model - Layer - Layer Properties	Deprecated API within the libraryy, fix related issue	Deprecation is within kymatio rather than a third party library		
ludwig-ai/ludwi https://github.c	1 8	328	829 GPU Usage	GPU Usage	GPU Usage				
medipixel/rl_ali https://github.c	6 2	142	279 Model - model type & properties	N/A	N/A	Change config in DQN model	Relates to loss function, but is not an inapproviate or inefficient use, so is software refactoring instead	Configuration	
HazyResearch/f https://github.c	6	56	64 N/A	N/A	N/A	refactoring	Software refactoring	General refactoring task?	
sktime/sktime https://eithub.c	6 13	164 2	2458 N/A	Tensors & Inputs	N/A	Consider adding "Inference - Prediction" Category. There should be one for it	Could potentially also be "Model - Model Type & Properties"	Something for inference?	
ludwig-ai/ludwi https://github.c			2175 Training - Training Process	Training - Hyperparameters	Training - Training Process	Wrong calculation for steps per epoch			
sktime/sktime https://github.c			3340 N/A	N/A	N/A	Wrong default parameter, this is SE issues	Documentation bug	Wrong documentation	
sktime/sktime https://github.c			2855 N/A	Model - Model Type & Properties	Third-party Usage	Feature request	New estimator (type of model)	wrong documentation	third party or mox Model - Model Type & Properties
							new eaumator (type or model)		unit party of mot widder - widder type or Properties
brendanhasz/pr <u>https://github.c</u> sktime/sktime_https://github.c	-		53 Training - Hyperparameter	Training - Training Process N/Δ	Training - Training Process	Allow learning rate to be changed midway	6.66	Need to add parameter related shiff	
					Training - Training Process	refactoring: Move parameter to estimator and out of config file	Software refactoring		
ludwig-ai/ludwi https://github.c			367 Training - Preprocessing of Training Data					Need to add normalising	
ludwig-ai/ludwi https://github.c	5 21		2170 Third-party Usage	N/A	Tensors & Input	Deal with deprecated numpy usage	Software quality issue		Third-party Usage
medipixel/rl_al _i https://github.c			124 N/A	N/A	N/A	Refactoring, move code to bitbucket	Software refactoring / removal of models. Possibly "Model - Model Type & Properties"		
brendanhasz/pr <u>https://github.c</u>		21	39 Training - Training Process	Training - Training Process	Training - Training Process		Was also sampled in initial set of 30 issues		
intel/dffml https://github.c			340 N/A	N/A	Model - Model Type & Properties	Name change for consistency	Software refactoring		
brendanhasz/pi https://github.c	2	19	46 Training - Preprocessing of Training Data	Layer - Activation Function	Tensors & Input	Add parameter to center input data	test_CenteredParameter_fit function in fix PR shows that can be used for specifying prior on weights / activation fu	nction	Either one of the Layer - Activation Function
ludwig-ai/ludwi https://github.c	3 19	91 1	1992 Tensors & Inputs	Tensors & Inputs	Tensors & Input				
sktime/sktime https://github.c	5 39	107 3	3915 Third-party usage	N/A	Third-party Usage		Potentially "Third-party Usage" but is proactive rather than due to inapproprate usage		
ProGamerGov/i https://github.c	4	43	84 Training - Training Process	Training - Training Process	Training - Optimiser	Missing hyper parameter	Was also sampled in initial set of 30 issues		
DeepLabCut/De https://github.c	5 19	81 1	1982 Third-party usage	Third-party Usage	Third-party Usage				
sktime/sktime https://github.c	6 26	96 2	2382 N/A	N/A	Training - Hyperparameters	Documentation	Documentation bug		
Model - Layer - Activation Function: Ins Model - Layer - Layer Properties: Incon Model - Layer - Missing/ Redundant/ Wi	nappropriate act rrect or inappro Vrong Layer: Ad	ivation function priate propert ding, removin	age, PR, project description to label. Here are the 14 or on relate to the use of such functions such a ties for a neural network layer, such as numb ag or changing neural network layers inappro- lel properties, such as the number of layers, a	as ReLu, Sigmoi, Tanh for a neural netwo per of neurons, input/output shape. opriately.					
Inference: Inappropriate behavior during									
			s, wrong reference to GPU device, failed para		subprocesses, faulty transfer of data to	a GPU device.			
			libraries, such as TensorFlow, PyTorch, Kera						
			data, such as noisy data, imbalanced data, i ocesses, such as inappropriate batch sizes, l						
9. Training - Training Process: Inappropriate op				learning rates.					
11. Training - Loss Function: Inappropriat									
			esting procedures or not evaluating the mode	el correctly.					
			eprocessing of training data, such as scaling						
14. Training - Hyperparameters: Inapprop	priate or subopt	imal hyperpar	rameters, such as learning rate, dropout rate	, number of epochs, etc.					
1. Tensors & Inputs: Incorrect shapes of	input, wrong di	mensions, siz	ze, inappropriate file type, encoding, selection	on of data format					
Note for labelers:									
If an issue does not belong to any of the categori									
If an issue does not belong to any of the categori Enhancement and non-critical issues (issues tha	at does not crash t	he program) can	n be label using the categories above (e.g. inappropri	riate use of data format)					
If an issue does not belong to any of the categori Enhancement and non-critical issues (Issues tha If an issue does not belong to any of the categori	at does not crash t ries above, use N/	he program) can A			Dar MA				
If an issue does not belong to any of the categori Enhancement and non-critical issues (issues tha If an issue does not belong to any of the categori If It's related to software issues (e.g. function nee	at does not crash to ries above, use N/ red new parameter	he program) can A) which are not r	related to third-party libraries, fit them in one of the 14		n as N/A				
If an issue does not belong to any of the categori Enhancement and non-critical issues (Issues tha If an issue does not belong to any of the categori	at does not crash to ries above, use N/ red new parameter	he program) can A) which are not r	related to third-party libraries, fit them in one of the 14		n as N/A				
If an issue does not belong to any of the categori Enhancement and non-critical issues (issues tha If an issue does not belong to any of the categori If It's related to software issues (e.g. function nee	at does not crash to ries above, use N/ red new parameter	he program) can A) which are not r	related to third-party libraries, fit them in one of the 14		n as N/A				
If an issue does not belong to any of the categori Enhancement and non-critical issues (issues tha If an issue does not belong to any of the categori If It's related to software issues (e.g. function nee	at does not crash to ries above, use N/ red new parameter	he program) can A) which are not r	related to third-party libraries, fit them in one of the 14		n as NA				
If an issue does not belong to any of the categori Enhancement and non-critical issues (issues that "an issue does not belong to any of the categori If it's related to software issues (e.g., function nee Comment: Consider merge layer property and m	at does not crash t iries above, use N/ sed new parameter model type & Prope	he program) can A ') which are not r erty, hard to disti	related to third-party libraries, fit them in one of the 14	4 categories above, if not appropriate, label ther					
If an issue does not belong to any of the categories for the state of the state of the state of the state of the categories of the state of the state of the state of the categories if it is related to software issues (e.g., function ne Comment. Consider merge layer properly and m the state of the state of the state of the state of the The paper from humbatova does not specify how Calculate inter rate reliability, if 9-0.7 good	at does not crash to iries above, use Ni- sed new parameter model type & Prope ow to fit an issue in	the program) can A ') which are not rerty, hard to distill to a category, we	related to third-party libraries, fit them in one of the f- inguish model and layer e miligate it by getting a list of categories, read from	4 categories above, if not appropriate, label ther					
If an issue does not belong to any of the categoric Enhancement and non-critical issues (exhauncement and the If an issue does not belong to any of the categori If it related to software issues (e.g., function ner Comment: Consider menge skyer properly and in The paper from humbatiova does not specify hox Calculate inter rate reliability, if >0.7 good Add a category for error in inference-prediction.	at does not crash to iries above, use Ni- sed new parameter model type & Prope ow to fit an issue in	the program) can A ') which are not rerty, hard to distill to a category, we	related to third-party libraries, fit them in one of the f- inguish model and layer e miligate it by getting a list of categories, read from	4 categories above, if not appropriate, label ther					
If an issue does not belong to any of the categoric Enhancement and non-critical issues (exhauncement and the If an issue does not belong to any of the categori If it related to software issues (e.g., function ner Comment: Consider menge skyer properly and in The paper from humbatiova does not specify hox Calculate inter rate reliability, if >0.7 good Add a category for error in inference-prediction.	at does not crash to iries above, use Ni- sed new parameter model type & Prope ow to fit an issue in	the program) can A ') which are not rerty, hard to distill to a category, we	related to third-party libraries, fit them in one of the f- inguish model and layer e miligate it by getting a list of categories, read from	4 categories above, if not appropriate, label ther					
If an issue does not belong to any of the categoria- fichancement and non-ordical issues (setThancement and non-ordical issues (setThancement and non-ordical issues (set along of it is related to soften to be any of the categorial if it is related to soften issues (e.g., funders issues	at does not crash to iries above, use Ni- sed new parameter model type & Prope ow to fit an issue in	the program) can A ') which are not rerty, hard to distill to a category, we	related to third-party libraries, fit them in one of the f- inguish model and layer e miligate it by getting a list of categories, read from	4 categories above, if not appropriate, label ther					
If an issue does not belong to any of the categorist- febranement and not croisil issues (issues that If an issue does not belong to any of the categorist If a related to software issues (e.g. function notice Comment: Consider merge layer properly and m The paper from humbatova does not specify hor Cataloutes inter rate reliability, if >0.7 good Add a category for error in inference prediction. There are not performed by an inference or prediction.	at does not crash to iries above, use Ni- sed new parameter model type & Prope ow to fit an issue in	the program) can A ') which are not rety, hard to distill to a category, we	related to third-party libraries, fit them in one of the f- inguish model and layer e miligate it by getting a list of categories, read from	4 categories above, if not appropriate, label ther					
If it is said each not belong to any of the categoric filterancement and not critical issues (issues in it if in its said each not belong to any of the categoric if it is related to software saids (i.e., in a consideration of the categoric in its related to software saids (i.e., in a consideration of the categoric in its related to software saids (i.e., in a consideration of the category of the related to software saids), if it is of pool of the categoric in inference prediction. The categoric in inference prediction. There is no quidation to use the taxonomy. There is no quidation to use the taxonomy. TODO.	at does not crash the irides above, use Ni- bed new parameter model type & Prope ow to fit an issue in put it in the botton	the program) can A ') which are not rety, hard to distill to a category, we	related to third-party libraries, fit them in one of the f- inguish model and layer e miligate it by getting a list of categories, read from	4 categories above, if not appropriate, label ther					
If at issue desa not belong to any of the categor Enfancement and not croiced insues (issues that If it is such each of the categor is and If it is related to software sizes (is g. function except the category is and it is a such as the category and in the category is and it is a such as the category is and in the paper from humbations does not specify ho Calculation into make rate of the category for more than the category for more in tertenous predicts, here are part about to put have in the last Three is no guidance to use the Sacronomy TODO:	at does not crash the strikes above, use Ni the above, use Ni bed new parameter model type & Prope ow to fit an Issue in the botton action	the program) can A ') which are not rety, hard to distill to a category, we	related to third-party libraries, fit them in one of the f- inguish model and layer e miligate it by getting a list of categories, read from	4 categories above, if not appropriate, label ther					
If it is said each not belong to any of the categoric filterancement and not critical issues (issues in it if in its said each not belong to any of the categoric if it is related to software saids (i.e., in a consideration of the categoric in its related to software saids (i.e., in a consideration of the categoric in its related to software saids (i.e., in a consideration of the category of the related to software saids), if it is of pool of the categoric in inference prediction. The categoric in inference prediction. There is no quidation to use the taxonomy. There is no quidation to use the taxonomy. TODO.	at does not crash the strikes above, use Ni the above, use Ni bed new parameter model type & Prope ow to fit an Issue in the botton action	the program) can A ') which are not rety, hard to distill to a category, we	related to third-party libraries, fit them in one of the f- inguish model and layer e miligate it by getting a list of categories, read from	4 categories above, if not appropriate, label ther					