WEBVTT

1
00:00:03.990 --> 00:00:14.130
Sarah Jeffreson: it's my pleasure today to help him Alex Sandra Cipriano bitch.

2
00:00:18.420 --> 00:00:18.779
Sarah Jeffreson: me.

3
00:00:19.980 --> 00:00:22.710
Sarah Jeffreson: And alexandra's.

4
00:00:24.240 --> 00:00:25.350
Sarah Jeffreson: per search.

5
00:00:27.480 --> 00:00:31.710
Sarah Jeffreson: focuses on the application of arts of.

6
00:00:33.600 --> 00:00:41.580
Sarah Jeffreson: visual intelligence and mercy machine learning to astro.

7
00:00:43.230 --> 00:00:45.630
Sarah Jeffreson: physics and also to high energy.

8
00:00:46.980 --> 00:00:58.350
Sarah Jeffreson: physics and today we're going to be hearing from her on bridging the gap between.

9
00:00:59.490 --> 00:01:05.100
Sarah Jeffreson: simulations and survey data using tables.

10
00:01:09.150 --> 00:01:09.690
Sarah Jeffreson: Learning.

11
00:01:10.950 --> 00:01:14.190
Sarah Jeffreson: So if you have.

12
00:01:17.490 --> 00:01:18.570
Sarah Jeffreson: questions.

13
00:01:24.450 --> 00:01:28.860

Sarah Jeffreson: cheering like killing slowly.

14
00:01:30.360 --> 00:01:34.680
Sarah Jeffreson: Talk please send them in.

15
00:01:35.910 --> 00:01:36.840
Sarah Jeffreson: Private.

16
00:01:38.640 --> 00:01:46.260
Sarah Jeffreson: chat so i'm slowly take it away Alexandra.

17
00:01:47.670 --> 00:01:54.810
Aleksandra Ciprijanovic (she/her): Thank you so much, thank you for the great introduction and thank you so much for inviting me I really am super happy to be here.

18
00:01:56.310 --> 00:01:56.850
Aleksandra Ciprijanovic (she/her): So.

19
00:01:58.410 --> 00:02:01.890
Aleksandra Ciprijanovic (she/her): As you've heard i'm going to tell you a little bit about.

20
00:02:03.360 --> 00:02:15.900
Aleksandra Ciprijanovic (she/her): Using machine learning and deep learning in astronomy and a problem that probably most of us that want to use the birding in astronomy will have already had so.

21
00:02:16.560 --> 00:02:25.170
Aleksandra Ciprijanovic (she/her): This is a problem of how to use deep learning when you have multiple data sets that you want to work on and you want your model to work on both.

22
00:02:25.890 --> 00:02:31.260
Aleksandra Ciprijanovic (she/her): And so i'm from formula, by working the scientific computing division and.

23
00:02:31.890 --> 00:02:45.330
Aleksandra Ciprijanovic (she/her): I put this image here on purpose, so this is an example of to domain situation, so this is the main building of formula, called the Wilson Hall, and obviously you have one detector image, so the photograph and you have a simulated.

24

00:02:46.080 --> 00:02:57.150
Aleksandra Ciprijanovic (she/her): Very roughly simulated image or a drawing of the same building and when you want to use deep learning to you know recognize this image, for example, unfortunately these.

25
00:02:58.110 --> 00:03:10.410
Aleksandra Ciprijanovic (she/her): tiny differences are, in this case, now that tiny will not allow you your model to to work with both data sets and in astronomy will very often have a case where we want to be able to do that.

26
00:03:10.830 --> 00:03:19.260
Aleksandra Ciprijanovic (she/her): So how to do that that's basically the Doc the topic of my talk and i'm Sorry, I do not know what happened.

27
00:03:20.670 --> 00:03:22.950
Aleksandra Ciprijanovic (she/her): One second let's go back.

28
00:03:24.750 --> 00:03:26.130
Aleksandra Ciprijanovic (she/her): it's always happen somehow.

29
00:03:27.390 --> 00:03:46.710
Aleksandra Ciprijanovic (she/her): let's go back to the slideshow okay so we're back, so this is basically my talk outline, so I will try to introduce a little bit more, where do these discrepancies come from in astronomy I will give you an astro example.

30
00:03:46.890 --> 00:03:47.790
Aleksandra Ciprijanovic (she/her): Of the.

31
00:03:48.330 --> 00:04:02.370
Aleksandra Ciprijanovic (she/her): topic that I work on but feel free to extrapolate this to absolutely any other problem astro non astra where you have a similar situation, I will tell you what domain adaptation is what are these methods for deep learning how they.

32
00:04:02.370 --> 00:04:05.550
Aleksandra Ciprijanovic (she/her): Work and hopefully convince you that they are helpful.

33
00:04:05.940 --> 00:04:06.990
Aleksandra Ciprijanovic (she/her): In this kind of problem.

34
00:04:08.280 --> 00:04:08.880

Aleksandra Ciprijanovic (she/her): So.

35
00:04:10.500 --> 00:04:13.620
Aleksandra Ciprijanovic (she/her): Yes, let's see in astronomy.

36
00:04:14.910 --> 00:04:23.190
Aleksandra Ciprijanovic (she/her): You often either have two data sets from two different telescopes, for example, or you have a situation where you have a simulated data and real data.

37
00:04:24.450 --> 00:04:34.890
Aleksandra Ciprijanovic (she/her): They will be different time differences but still, they will be different if you have to telescopes of course they're going to have different specs hands, you know the data is going to look a little bit different.

38
00:04:36.480 --> 00:04:43.980
Aleksandra Ciprijanovic (she/her): If you have simulation and you have real data first these will be different, because the simulated data will not include.

39
00:04:44.340 --> 00:04:53.580
Aleksandra Ciprijanovic (she/her): All of the details that can be either because we don't know all of the physics, so we are not including something or we maybe don't understand it fully or because.

40
00:04:54.450 --> 00:05:02.790
Aleksandra Ciprijanovic (she/her): Just our computational resources are not limitless so we have to approximate something so that we can run our complex simulation in our lifetimes.

41
00:05:03.930 --> 00:05:20.520
Aleksandra Ciprijanovic (she/her): So all of these things, of course, make the data set look slightly different than real data, then, on top of that, we usually add observational effects like batting observational noise are blurring better yourself, so that we can mimic a particular telescope and make a mock data.

42
00:05:21.630 --> 00:05:27.360
Aleksandra Ciprijanovic (she/her): But again, you will not perfectly mimic the real deal with that kind of.

43
00:05:28.500 --> 00:05:48.060
Aleksandra Ciprijanovic (she/her): observational effects and these might not even be obvious when you look at them, but for a deep learning

algorithm, very often, this is enough to make a problem so on the left, you have an example of simulated and real beta, and this is an example of of merging galaxies.

44
00:05:49.110 --> 00:05:58.110
Aleksandra Ciprijanovic (she/her): On top you have a couple of examples from elastic simulation and when you have a simulated data set this is usually the data that you know very well.

45
00:05:58.410 --> 00:06:05.430
Aleksandra Ciprijanovic (she/her): You have labels, if you want to you know distinguish different types of galaxies, for example, or mergers versus non managers, or something.

46
00:06:05.640 --> 00:06:12.150
Aleksandra Ciprijanovic (she/her): You know what's going on in the simulation so you know what the object is representing in each image, you can train them this dataset.

47
00:06:12.600 --> 00:06:20.910
Aleksandra Ciprijanovic (she/her): I will probably very often call this data set the source data set and then the bottom, you have your target data set this can be observation.

48
00:06:21.390 --> 00:06:33.720
Aleksandra Ciprijanovic (she/her): For example, and in particular, for example, if you have new data new observations this data set is very often unlabeled so you can directly use it for a supervised learning algorithm.

49
00:06:34.440 --> 00:06:48.510
Aleksandra Ciprijanovic (she/her): So that's that's the problem basically so Okay, the previous life was a hint we're gonna be talking about merging galaxies and I definitely don't have to convince you why merging galaxies are important for astrophysics.

50
00:06:49.620 --> 00:06:53.490
Aleksandra Ciprijanovic (she/her): They are crucial part of the evolution of matter in the universe.

51
00:06:53.610 --> 00:07:03.180
Aleksandra Ciprijanovic (she/her): And then you know formation of different types of galaxies changes and star formation read chemistry, you know types of galaxies in the end.

52

00:07:04.350 --> 00:07:12.420
Aleksandra Ciprijanovic (she/her): So we really need to understand this process very well because it leads to all kinds of different interesting things that we want to understand.

53
00:07:13.800 --> 00:07:29.340
Aleksandra Ciprijanovic (she/her): How do we do that, well, we have to leverage big data sets of merging galaxies and because yeah, this is a very long process is less millions of years, so we can't really study a pair of galaxies merging We just have to find a lot of them in different stages of imagine.

54
00:07:31.620 --> 00:07:40.260
Aleksandra Ciprijanovic (she/her): The problem with that is that with observations to be sure that something is a merger is is not really easy, especially if you want to do it in large scales.

55
00:07:40.590 --> 00:07:49.080
Aleksandra Ciprijanovic (she/her): Very often in if you want to make a big data set of merging galaxies from observation, you have to visually inspect them that's prone to all kinds of biases and errors.

56
00:07:49.500 --> 00:07:59.280
Aleksandra Ciprijanovic (she/her): From the the people who are looking at the galaxies and sometimes they can, for example, just visually overlap, you need further observations, for example.

57
00:07:59.580 --> 00:08:11.010
Aleksandra Ciprijanovic (she/her): spectroscopy to make sure that they really are near each other, and not just visually overlapping all those things that make all sorts of problems so basically, the solution is.

58
00:08:11.550 --> 00:08:22.800
Aleksandra Ciprijanovic (she/her): For us to find a way to combine observed and and simulated data so that we can learn from from both sources and also learn from the similarities and differences between this data sets.

59
00:08:25.140 --> 00:08:36.390
Aleksandra Ciprijanovic (she/her): In in when we look at through like the the history of how mergers were recognized first we had our standard classification methods, where you have your image you.

60
00:08:37.470 --> 00:08:44.520

Aleksandra Ciprijanovic (she/her): extract you can summarize summarize the data and the shape of the galaxy by calculating some kind of parameter like.

61
00:08:44.820 --> 00:08:53.610
Aleksandra Ciprijanovic (she/her): How concentrated, is how a symmetric or clumpy the the galaxy looks like and then you decide on some rules based on these morphology parameters.

62
00:08:54.270 --> 00:09:09.060
Aleksandra Ciprijanovic (she/her): shape parameters and well you just classify whether something is a merger or non modular based on those things, so the next step and the very logical one was to use a simple machine learning model so, for example in 2019.

63
00:09:10.230 --> 00:09:14.250
Aleksandra Ciprijanovic (she/her): Random forests were used to do the same classification.

64
00:09:15.630 --> 00:09:24.660
Aleksandra Ciprijanovic (she/her): On the same shape parameters, but you don't decide on any of the rules now, so you don't decide how Columbia, or how a symmetric.

65
00:09:25.410 --> 00:09:42.720
Aleksandra Ciprijanovic (she/her): galaxy has to be you just give you all the parameters to the random forest and the random forest decides how to classify these two classes basically mergers versus non mergers and it was shown that actually this outperforms all of the separate methods that were used before.

66
00:09:44.130 --> 00:09:59.220
Aleksandra Ciprijanovic (she/her): And when you think about it you're still kind of doing this classification on some kind of summary which is definitely not a perfect summary of the complex shape of the galaxy so a natural next step, which was done in one of my old papers.

67
00:09:59.250 --> 00:10:00.300
Aleksandra Ciprijanovic (she/her): from last year.

68
00:10:00.780 --> 00:10:15.390
Aleksandra Ciprijanovic (she/her): As to just not some recent data, why would we do that now that we have the computing power, you know to work with with full data, so we decided to use deep learning use delusional neural networks to.

69
00:10:16.500 --> 00:10:17.790
Aleksandra Ciprijanovic (she/her): Use the full.

70
00:10:19.140 --> 00:10:34.710
Aleksandra Ciprijanovic (she/her): image and all the information contained inside to classify things into managers and managers and for that we we use the very simple neural network, so it had three layers and layers and what we.

71
00:10:35.850 --> 00:10:42.030
Aleksandra Ciprijanovic (she/her): Seen there is that we actually outperformed all of the methods before, so it was very promising.

72
00:10:43.080 --> 00:10:59.070
Aleksandra Ciprijanovic (she/her): And this is actually how our data look like, so we were using us for simulation that you saw on the previous slide and we were using distant merging galaxies so redshift to because we wanted to do a tougher task not nearby analysis but also just because this this area of.

73
00:10:59.160 --> 00:11:03.870
Aleksandra Ciprijanovic (she/her): Around redshift to is very important for like large star formation rates and probably mergers very.

74
00:11:03.960 --> 00:11:20.310
Aleksandra Ciprijanovic (she/her): Important in this period so let's see if we can do that, we added we created two levels of like observational realism, when we were making marks first we just added the widespread function of hubble telescope so we just blow it up a little bit.

75
00:11:20.820 --> 00:11:25.020
Aleksandra Ciprijanovic (she/her): And then the second data set, we also added the observational Nice and.

76
00:11:25.260 --> 00:11:39.270
Aleksandra Ciprijanovic (she/her): Now we're still just working with the two data sets separately and not trying to combine them, we saw that we can get accuracy is almost up to 80% which for this data set is really good it's not baptism unfortunately so it's a little bit harder.

77
00:11:41.070 --> 00:11:50.130
Aleksandra Ciprijanovic (she/her): And yes, so we could do it, but that is when I tried to use the model trained on the top row and classify the bottom row, and vice versa.

78
00:11:50.640 --> 00:12:03.180
Aleksandra Ciprijanovic (she/her): Oh, and by the way this is called the galaxy is actually look like, so this is the same thing as the top row, but a logarithmic color map So you see that there's a lot of stuff in there just verifying Okay, so what happens if we try to classify the.

79
00:12:03.180 --> 00:12:03.660
Aleksandra Ciprijanovic (she/her): Other data.

80
00:12:04.440 --> 00:12:14.670
Aleksandra Ciprijanovic (she/her): Well, this is what happens accuracies were 50% and keep in mind that these two data sets are exactly the same the galaxies are exactly the same and we just add the noise.

81
00:12:14.820 --> 00:12:19.920
Aleksandra Ciprijanovic (she/her): So it's not even different looking galaxies which will happen if one of these was actual.

82
00:12:19.980 --> 00:12:20.700
Aleksandra Ciprijanovic (she/her): You know observed.

83
00:12:20.730 --> 00:12:29.190
Aleksandra Ciprijanovic (she/her): Data So how can we combine different data sets this is when I really started thinking about domain adaptation so.

84
00:12:29.910 --> 00:12:35.790
Aleksandra Ciprijanovic (she/her): What happens when you train a neural network for this such a simple like to class classification task.

85
00:12:36.240 --> 00:12:45.000
Aleksandra Ciprijanovic (she/her): The neural network translates your images into Latin space, which you know the internal it can be, you know very multi dimensional space of the neural network.

86
00:12:45.390 --> 00:12:52.620
Aleksandra Ciprijanovic (she/her): It tries to distinguish between two groups of objects and finds a decision boundary between them, when you.

87
00:12:53.430 --> 00:13:00.150
Aleksandra Ciprijanovic (she/her): load a new data set into a model train that way it unfortunately is translated into Latin space.

88

00:13:01.080 --> 00:13:09.480
Aleksandra Ciprijanovic (she/her): On some like it is either rotated a little bit move the side or something like that it just doesn't align perfectly with the old data because.

89
00:13:09.810 --> 00:13:15.360
Aleksandra Ciprijanovic (she/her): The data set just looks a little bit different so the decision boundary from before doesn't work on your layer said.

90
00:13:15.750 --> 00:13:25.890
Aleksandra Ciprijanovic (she/her): Well domain adaptation is added during training and it allows the new data set to align on top of the other one and so that you can find the common decision boundary.

91
00:13:26.610 --> 00:13:40.920
Aleksandra Ciprijanovic (she/her): So this is this is, in essence, what domain adaptation does so on the technical side of how the training works so let's first see how he's doing regular supervised learning algorithm you have your images.

92
00:13:41.700 --> 00:13:54.750
Aleksandra Ciprijanovic (she/her): If they have labels in this case managers are nominated the network trains finds useful features and actually unplugging real useful features for this dataset in the middle, this blue thing.

93
00:13:56.250 --> 00:14:05.730
Aleksandra Ciprijanovic (she/her): This is called the Grad cam It shows you which pixels were the most important for classification into particular class and for mergers, we found that peripheries of the galaxies.

94
00:14:06.090 --> 00:14:14.940
Aleksandra Ciprijanovic (she/her): are the most important it does make sense, you have a lot of famous symmetric interesting features over there, that you need to look at to recognize a merger.

95
00:14:15.600 --> 00:14:25.050
Aleksandra Ciprijanovic (she/her): Okay, so, but when you train that way, and you show it in the testing phase and noisy version of it, it will not work on the nice thing with domain adaptation, the training.

96
00:14:25.680 --> 00:14:36.450

Aleksandra Ciprijanovic (she/her): uses label source data, and very importantly unlabeled target data it just uses it to find somehow the meaning various features.

97
00:14:36.990 --> 00:14:43.350
Aleksandra Ciprijanovic (she/her): So it moves from things that are useful only for one domain, the things that are useful and common in both domains.

98
00:14:44.340 --> 00:14:51.000
Aleksandra Ciprijanovic (she/her): So that in the testing phase it, you can work on both data sets, and this is actually what really happened i'll show you one more example later on.

99
00:14:51.660 --> 00:15:07.440
Aleksandra Ciprijanovic (she/her): The important regions move from the peripherals to the centers because obviously the noise destroys the faint information, so the useful things for the noisy data said only you know remains in the Center where the signal is about this.

100
00:15:08.670 --> 00:15:17.340
Aleksandra Ciprijanovic (she/her): So this is how you technically do it, you have your loss function that the neural network is minimizing from apple apple during training.

101
00:15:17.730 --> 00:15:22.140
Aleksandra Ciprijanovic (she/her): And these are the four losses that we were adding I don't have time to go through all of them.

102
00:15:22.680 --> 00:15:32.730
Aleksandra Ciprijanovic (she/her): But i'll focus on the last one, but transfer loss so first you have your standard task was that's the last that is doing the classification part it's learning how to classify things.

103
00:15:33.360 --> 00:15:43.590
Aleksandra Ciprijanovic (she/her): Then you have transfer the last that's adding domain adaptation and then you have this to middle losses official loss and entropy minimization these are help or losses you don't have to ask them.

104
00:15:44.070 --> 00:15:50.520
Aleksandra Ciprijanovic (she/her): They are both of them do the same thing, but one works visualizing the source domain and entropy in the target domain.

105
00:15:50.850 --> 00:16:02.850
Aleksandra Ciprijanovic (she/her): Both of them are trying to make the classes in the Latin space as compact as possible and as separate as possible, so that when you find a decision boundary it has a lot of clean space in between and.

106
00:16:03.180 --> 00:16:11.070
Aleksandra Ciprijanovic (she/her): Fewer number of examples are going to you know be found on the wrong side basically so sometimes it can help to have these two things.

107
00:16:12.900 --> 00:16:21.240
Aleksandra Ciprijanovic (she/her): So when it comes to transfer loss or domain adaptation, you have two different approaches and two different groups of methods.

108
00:16:21.630 --> 00:16:35.670
Aleksandra Ciprijanovic (she/her): These are methods are very, very often used for self driving cars, for example, because you have to you know, have a car that is trained on images of streets from one town, but it has to know how to drive in all other towns to, so this is kind of tough.

109
00:16:36.720 --> 00:16:49.290
Aleksandra Ciprijanovic (she/her): So the two approaches are on the left, you have distance based methods, and for that we use an example of using maximum in discrepancy i'll tell a little bit more what it is on the next slide.

110
00:16:49.710 --> 00:17:08.220
Aleksandra Ciprijanovic (she/her): it's a way to calculate the distance between two probability distributions and the other methods are our adversary based so you have a method that trains in some kind of adversarial way again show you what we use we use domain adversarial networks.

111
00:17:09.360 --> 00:17:20.100
Aleksandra Ciprijanovic (she/her): So this is the most important part and yeah on top, we use a very standard classification loss, which is the Cross entropy last where why our label white CAP or predicted labels and we're just minimizing.

112
00:17:21.150 --> 00:17:26.640
Aleksandra Ciprijanovic (she/her): This last penalizing all of the wrong classification okay so let's start with them empty.

113

00:17:28.200 --> 00:17:41.130
Aleksandra Ciprijanovic (she/her): So this is a super simplified way of explaining what's going on, so let's see how can we see how similar b&q these two programs distributions how similar.

114
00:17:41.760 --> 00:17:51.720
Aleksandra Ciprijanovic (she/her): We don't know, then we can only observe some data points blue and and read from these two lists evictions, so what Emily does it.

115
00:17:52.530 --> 00:18:04.020
Aleksandra Ciprijanovic (she/her): Transfer transfers the whole problem into the space of kernels interpreters and Colonel cannabis space, so you you substitute each data point with a kernel and so that we can.

116
00:18:04.650 --> 00:18:17.370
Aleksandra Ciprijanovic (she/her): Keep it super simplified on this slide let's say that this is a kernel, but you can use any kind of even kernel and even sums of kernels of an infinite sounds good also still kernels so you can use whatever you want.

117
00:18:18.810 --> 00:18:25.110
Aleksandra Ciprijanovic (she/her): You sum up all of the blue kernel of the red kernels and you get these two curves in the bottom right so.

118
00:18:26.670 --> 00:18:27.270
Aleksandra Ciprijanovic (she/her): If you.

119
00:18:29.130 --> 00:18:38.580
Aleksandra Ciprijanovic (she/her): You can also describe the whole curve by just it's mean embedding which is just the sum of all current thousand divided by the number of.

120
00:18:40.050 --> 00:18:47.880
Aleksandra Ciprijanovic (she/her): This is just how things work in this space, and if you say one mean minus the other you get this blue.

121
00:18:48.540 --> 00:18:56.700
Aleksandra Ciprijanovic (she/her): Black curve or the witness function, obviously the witness function will be zero the two curves are exactly the same, so what we want.

122
00:18:57.120 --> 00:19:10.530
Aleksandra Ciprijanovic (she/her): is to add this as our transfer loss try to minimize it try to force the network to somehow find the features

that will lead to these two distributions looking as similar as possible to the neural network.

123
00:19:11.100 --> 00:19:16.830
Aleksandra Ciprijanovic (she/her): Definitely, this is how the equation looks like so you have three sons of criminals.

124
00:19:17.190 --> 00:19:24.600
Aleksandra Ciprijanovic (she/her): Criminals that are self similarity criminals for the blue and red and one with the minus sign in front with cross similarities.

125
00:19:24.960 --> 00:19:33.210
Aleksandra Ciprijanovic (she/her): In case super simple case of dogs and fish, you will have a matrix like this with all of the examples on the diagonal you have.

126
00:19:33.900 --> 00:19:43.740
Aleksandra Ciprijanovic (she/her): South similarities, these are going to be large and then all the additional elements are crossing loaders we're going to force the Cross similarities to be become as big as we can.

127
00:19:44.340 --> 00:19:55.380
Aleksandra Ciprijanovic (she/her): So that this third term is biggest and handsome md distances is the smallest so basically we're forcing the network to see dog and fish something like this, or in our case managers and managers.

128
00:19:56.730 --> 00:20:03.930
Aleksandra Ciprijanovic (she/her): So that's one method, the other method is dance or domain adversarial neural networks so with these you have.

129
00:20:04.650 --> 00:20:11.160
Aleksandra Ciprijanovic (she/her): The network that slightly different than the regular one you have the convolution layer the feature extract the green part.

130
00:20:11.850 --> 00:20:24.450
Aleksandra Ciprijanovic (she/her): You have the blue part that's doing the classification so it's looking at labels and then you add this second pink branch or domain classifier it doesn't care what is.

131
00:20:28.680 --> 00:20:30.330
Aleksandra Ciprijanovic (she/her): This always happens.

132
00:20:33.270 --> 00:20:40.230
Aleksandra Ciprijanovic (she/her): Okay, so it doesn't care about the labels it doesn't care if something is a merger or non merger it just class knows the domain labels.

133
00:20:40.680 --> 00:20:47.880
Aleksandra Ciprijanovic (she/her): And the most important part here is the gradient reversal part for this branch, so the loss in this part is actually increasing.

134
00:20:48.180 --> 00:20:55.980
Aleksandra Ciprijanovic (she/her): So what happens during training we are forcing the Green Party extract features that are going to be very useful for the blue part so for the classifier.

135
00:20:56.340 --> 00:21:02.670
Aleksandra Ciprijanovic (she/her): But very useless for the domain classifier, this means that we are searching for the meaning various.

136
00:21:03.210 --> 00:21:11.010
Aleksandra Ciprijanovic (she/her): features, so that we want to fall, the pink pink branch and, basically, you can see that the the approach is different, but the result is the same like with them empty.

137
00:21:11.850 --> 00:21:18.930
Aleksandra Ciprijanovic (she/her): So here's how things actually look so you have the two data sets here the two colors are the two classes.

138
00:21:19.260 --> 00:21:27.720
Aleksandra Ciprijanovic (she/her): Before training they're completely separate, this is by the way, a peacenik plot, which shows like a 2d representation of the Multi dimensional Latin space.

139
00:21:28.590 --> 00:21:36.390
Aleksandra Ciprijanovic (she/her): So the sort of the main is are these see through images, if you train a regular classifier on the source dummy you get something like this.

140
00:21:36.960 --> 00:21:49.440
Aleksandra Ciprijanovic (she/her): Now the source of menus on the bottom, you see that the Blue and rather now separate the accuracy is quite large over 80% but the target domain is completely separate and mixed it doesn't work on on the target of me.

141

00:21:49.950 --> 00:22:07.410
Aleksandra Ciprijanovic (she/her): If you now, instead of doing this train with domain adaptation let's move better side you get something like this, so the two domains overlap completely and you can see, the target accuracy increased almost to 80% so roughly 20% increase in accuracy, which is great.

142
00:22:09.450 --> 00:22:18.180
Aleksandra Ciprijanovic (she/her): And if we look at where the neural network is looking so something like the in the some of the starting slides before.

143
00:22:18.660 --> 00:22:24.600
Aleksandra Ciprijanovic (she/her): For the merger class you see in the source domain if there's no domain adaptation, it looks at the periphery.

144
00:22:25.230 --> 00:22:39.090
Aleksandra Ciprijanovic (she/her): But in the target domain, it focuses on the noise and obviously it doesn't work, but if you add domain adaptation it starts for both domains, looking at the Center and i'm also going to show you what happens if you also add.

145
00:22:39.210 --> 00:22:41.040
Aleksandra Ciprijanovic (she/her): Fisher and entropy loss.

146
00:22:41.130 --> 00:22:51.060
Aleksandra Ciprijanovic (she/her): This happens so it's still looking at the Center but also the classes are a lot more separated than before this isn't necessary but sometimes it is useful.

147
00:22:52.710 --> 00:22:59.790
Aleksandra Ciprijanovic (she/her): And finally, can we actually use this for a real problem, so what happens when we don't have a simple.

148
00:23:00.330 --> 00:23:07.020
Aleksandra Ciprijanovic (she/her): The simple difference between data sets like now, with just inclusion of noise but we actually have simulated data and real data.

149
00:23:07.560 --> 00:23:21.300
Aleksandra Ciprijanovic (she/her): So this is the problem that we're trying to do so, we wanted to work with real data, and for that, on the bottom, you have examples of spss galaxies and we use that because we have galaxies available, so we can test.

150

00:23:22.110 --> 00:23:30.750
Aleksandra Ciprijanovic (she/her): If something is okay or not, and, since these are nearby galaxies we also use the last race, but now not distant galaxies we use nearby like final snapshots.

151
00:23:31.830 --> 00:23:44.640
Aleksandra Ciprijanovic (she/her): So this means that in both datasets unfortunately have a very small number of individual objects, we do augmented data, but still it's the whole data set is like three times smaller than the previous example which is not ideal.

152
00:23:45.210 --> 00:23:54.150
Aleksandra Ciprijanovic (she/her): Plus, you can also see that the difference, for example, if you look at the Left row is quite substantial so mergers in as the SS are often like.

153
00:23:54.540 --> 00:24:02.460
Aleksandra Ciprijanovic (she/her): Mostly containing the easiest examples, because human Labor is where the most certain about you know, giving this a positive.

154
00:24:02.940 --> 00:24:09.150
Aleksandra Ciprijanovic (she/her): You know label off a manager, on the other hand, and simulation you can have whatever you, you know the simulation produces.

155
00:24:10.080 --> 00:24:19.470
Aleksandra Ciprijanovic (she/her): On the right, you can see the results, so this is a rock pop if you don't know it What this means is you just look at the area under the curve.

156
00:24:19.800 --> 00:24:32.820
Aleksandra Ciprijanovic (she/her): The bigger area, the better the classifier so the classifier that is perfect, it has the curve is moving towards the top left corner if it's random guessing it's close to the line.

157
00:24:33.660 --> 00:24:40.920
Aleksandra Ciprijanovic (she/her): In this three solid lines are showing us the performance in the target domain, the one that has no problem.

158
00:24:41.370 --> 00:24:51.330
Aleksandra Ciprijanovic (she/her): So you can see that the Blue and violet color are not very good they're kind of very close to the diagonal, so this is training without domain adaptation and river domain adaptation.

159
00:24:51.900 --> 00:24:58.260
Aleksandra Ciprijanovic (she/her): So we were very unhappy about this, but we thought Okay, we have two problems data Center very descriptive.

160
00:24:58.740 --> 00:25:02.820
Aleksandra Ciprijanovic (she/her): and very small let's try to solve it leaves the very small part of the problem.

161
00:25:03.120 --> 00:25:11.790
Aleksandra Ciprijanovic (she/her): So let's use the model training the previous example on guests and merging galaxies because merging galaxies are still merging galaxies there must be something useful.

162
00:25:12.180 --> 00:25:19.890
Aleksandra Ciprijanovic (she/her): So let's start from that model and then not from randomly initialize bates and start we basically do the transfer learning from that.

163
00:25:20.250 --> 00:25:32.160
Aleksandra Ciprijanovic (she/her): train a little bit more with domain adaptation on the new data and see if we can make it work and with it accuracy increased again by 20% so the yellow curve so even in this very.

164
00:25:33.570 --> 00:25:43.170
Aleksandra Ciprijanovic (she/her): tough and and not very good example, we were able to do it, so if you have a bigger data it, this is going to be even easier.

165
00:25:44.220 --> 00:25:46.290
Aleksandra Ciprijanovic (she/her): So finally.

166
00:25:47.520 --> 00:25:59.550
Aleksandra Ciprijanovic (she/her): I guess, I want to say that I hope that convince you the domain of the vision is something that we should care about it's something that is going to be crucial for for real life applications of.

167
00:26:00.360 --> 00:26:12.660
Aleksandra Ciprijanovic (she/her): Deep learning for astronomy is going to allow us to work with new observations, especially if, for example, LSD when it starts running if we do this correctly, we are going to have.

168
00:26:13.710 --> 00:26:26.160

Aleksandra Ciprijanovic (she/her): A model that is going to be able to do something in front in real time, for example, if you want to search for interesting transients during the observations as soon as they start, for example.

169
00:26:27.030 --> 00:26:35.130
Aleksandra Ciprijanovic (she/her): Without them in at the other patient, probably the performance is going to be a lot for us or it's not going to work at all if we just start on the simulations that we have now.

170
00:26:36.750 --> 00:26:45.090
Aleksandra Ciprijanovic (she/her): The minute efficient, can you really increase the robustness to all other kinds of problems, just like telescope noise or pixel level problems.

171
00:26:45.360 --> 00:26:53.040
Aleksandra Ciprijanovic (she/her): I don't have time to go into that that's something that i'm currently working on, if you want, we can talk a little bit about it afterwards I have an example in the next slide.

172
00:26:53.970 --> 00:27:08.910
Aleksandra Ciprijanovic (she/her): So the what I notice is the domain adaptation really just overall helps the model increases the robustness and forces the model to focus on on real things and not a noise or some other kind of problem like that.

173
00:27:10.050 --> 00:27:12.810
Aleksandra Ciprijanovic (she/her): One thing that we have to keep in mind is that.

174
00:27:14.070 --> 00:27:27.420
Aleksandra Ciprijanovic (she/her): It is this is not very easy to do so, the smaller the gap between the data sets the better, so we need to when we were making marks, we need to try to make them as close as possible to the real data set.

175
00:27:28.140 --> 00:27:37.650
Aleksandra Ciprijanovic (she/her): And also, if the two data sets are very weird and different doing domain adaptation with methods that align entire distributions.

176
00:27:38.040 --> 00:27:46.170
Aleksandra Ciprijanovic (she/her): is sometimes not enough, so these that I showed, you are doing that, but there are more sophisticated methods that are aligned class and top of class.

177
00:27:46.500 --> 00:27:57.000
Aleksandra Ciprijanovic (she/her): So, in cases where you have weird looking and different looking data sets and maybe in one you have additional class that you don't have in the other methods that a new class by class thing.

178
00:27:58.050 --> 00:27:59.040
Aleksandra Ciprijanovic (she/her): are going to be better.

179
00:28:00.870 --> 00:28:10.860
Aleksandra Ciprijanovic (she/her): So yeah I am Finally, of course, I want to thank all of the people that work with me on this and that still are working, you can see just part of them here and a part of institutions that are.

180
00:28:11.670 --> 00:28:21.330
Aleksandra Ciprijanovic (she/her): involved in all of this, so you know if you want to talk about any of these things or connect to any of these people, please let me know.

181
00:28:21.900 --> 00:28:33.930
Aleksandra Ciprijanovic (she/her): So thank you so much, and yeah if you're interested in Ai for astronomy check out the deep skies it's an amazing community with a lot of astral people and a lot of non astral people and even from industry, all like.

182
00:28:34.260 --> 00:28:45.540
Aleksandra Ciprijanovic (she/her): interested in Ai for astronomy I learned a lot from them a lot and also it's very good for networking so yeah feel free to ask me about them too, if you're interested Thank you.

183
00:28:49.980 --> 00:28:55.680
Sarah Jeffreson: hi sorry um yeah like so.

184
00:28:58.230 --> 00:28:58.980
Sarah Jeffreson: Thank you.

185
00:29:00.390 --> 00:29:05.940
Sarah Jeffreson: Very much Alexandra for a.

186
00:29:08.190 --> 00:29:14.340
Sarah Jeffreson: feeling like cool cool talk i'm just.

187

00:29:15.570 --> 00:29:19.710
Sarah Jeffreson: mind to every.

188
00:29:20.970 --> 00:29:28.770
Sarah Jeffreson: bond to send your questions in chat.

189
00:29:30.120 --> 00:29:31.320
Sarah Jeffreson: US silly.

190
00:29:32.340 --> 00:29:37.020
Sarah Jeffreson: To me, and to begin.

191
00:29:41.250 --> 00:29:41.640
Sarah Jeffreson: With.

192
00:29:42.720 --> 00:29:45.870
Sarah Jeffreson: We have a question from.

193
00:29:52.440 --> 00:29:53.250
Sarah Jeffreson: Morgan.

194
00:29:56.670 --> 00:29:57.180
Sarah Jeffreson: About.

195
00:29:58.830 --> 00:30:01.800
Sarah Jeffreson: visualizing he.

196
00:30:05.790 --> 00:30:06.630
teaches.

197
00:30:07.650 --> 00:30:08.310
Morgan Elowe MacLeod: Thank you.

198
00:30:09.450 --> 00:30:18.210
Morgan Elowe MacLeod: yeah so thanks so much for this and i'm thinking a little bit about sort of back to where you started.

199
00:30:18.690 --> 00:30:35.730
Morgan Elowe MacLeod: which was saying, we have a simulated data set and the you know real cosmos and we want to compare them and you made this point about the simulated data set never having the kind of like full physics.

200
00:30:37.440 --> 00:30:38.070
Morgan Elowe MacLeod: and

201
00:30:40.530 --> 00:30:44.190
Morgan Elowe MacLeod: So I was wondering if you could talk more about.

202
00:30:45.300 --> 00:30:49.590
Morgan Elowe MacLeod: That aspect of it, in particular with kind of the idea that.

203
00:30:51.240 --> 00:30:54.030
Morgan Elowe MacLeod: What can we learn how do we.

204
00:30:55.500 --> 00:31:04.230
Morgan Elowe MacLeod: How do we use that to basically learn how to constrain the physics in the simulation so say we are doing this comparison with the purpose of.

205
00:31:05.790 --> 00:31:17.640
Morgan Elowe MacLeod: Learning something about this sort of simulated physics, how does the sort of adaptation between the simulation domain and real domain inform.

206
00:31:18.150 --> 00:31:39.570
Morgan Elowe MacLeod: What we do and is like you showed us these kind of amazing visualizations of the key features, is that part of that sort of story of how you know we learned like in terms of say I was somebody who's trying to improve a simulation based on that comparison, how would that work.

207
00:31:39.810 --> 00:31:53.700
Aleksandra Ciprijanovic (she/her): I love where you're going with this, so I think that we can actually use something like this to to try to learn a little bit about our simulations so whether we want to learn.

208
00:31:55.020 --> 00:32:09.360
Aleksandra Ciprijanovic (she/her): about the simulation and the physics, or whether we want to learn about which things are important to add on top of the simulation to have things become closer and closer to.

209
00:32:10.320 --> 00:32:22.050

Aleksandra Ciprijanovic (she/her): Two real data, I feel that in both cases, we can use the main adaptation and, and this is how I kind of see, so I would, in this case, make a test where.

210
00:32:22.590 --> 00:32:28.980
Aleksandra Ciprijanovic (she/her): You include if you are lucky, not to have a super big simulation but something that you can actually work with.

211
00:32:29.520 --> 00:32:47.040
Aleksandra Ciprijanovic (she/her): You simulate something that is kind of a crude estimate of the real beta you train the model Look how distant they are look where the model focuses and then just start adding more and more things on top of it.

212
00:32:47.520 --> 00:33:00.450
Aleksandra Ciprijanovic (she/her): Really simulate but do a better simulation basically and basically observe will this lead the new model to see these two data sets as something closer and closer together.

213
00:33:00.840 --> 00:33:11.010
Aleksandra Ciprijanovic (she/her): Maybe you don't even have to train a new one, you you you train something on a simple simulation you show it the real data is going to be somewhere very far away.

214
00:33:11.460 --> 00:33:24.000
Aleksandra Ciprijanovic (she/her): Then you also show it some other simulations that are somewhere in between and look if they're going to be, you know closer and closer to your data and maybe by looking at these distances, you can kind of.

215
00:33:24.330 --> 00:33:41.670
Aleksandra Ciprijanovic (she/her): See which things are more important, which are less important, and the same thing you can do for adding observational realism so Look how how much the noise level in fact effects the data will move a lot or a little bit.

216
00:33:43.200 --> 00:33:51.360
Aleksandra Ciprijanovic (she/her): and which things or even when you look which things were the most important for classification, like those rings that I showed.

217
00:33:51.720 --> 00:34:06.930
Aleksandra Ciprijanovic (she/her): yeah if you add noise, for example, you can look when which level of noise will make the model like stop

being able to see this as an important feature, for example, and move to something else.

218
00:34:08.130 --> 00:34:09.270
Aleksandra Ciprijanovic (she/her): Something like this.

219
00:34:10.350 --> 00:34:11.670
Aleksandra Ciprijanovic (she/her): Of course, this is not.

220
00:34:12.690 --> 00:34:15.840
Aleksandra Ciprijanovic (she/her): You have to always do this with a grain of salt, because you can't.

221
00:34:17.520 --> 00:34:22.560
Aleksandra Ciprijanovic (she/her): When you're doing things in big a lot of dimensions, things are not easy.

222
00:34:23.820 --> 00:34:36.600
Aleksandra Ciprijanovic (she/her): So I think that we can develop even maybe better metrics or things are think just be creative and think how to use the information we can extract from the data it's kind of.

223
00:34:37.530 --> 00:34:47.400
Aleksandra Ciprijanovic (she/her): not something that is standard now people just you know, try to understand and improvise, but I feel like there is something there, we can we can.

224
00:34:47.940 --> 00:35:00.540
Aleksandra Ciprijanovic (she/her): try to understand stuff from from tests like these, and these are things that I really you know currently i'm thinking about So hopefully i'll you know, in the near future, maybe even will have things to show you.

225
00:35:01.320 --> 00:35:03.090
Aleksandra Ciprijanovic (she/her): Some examples yeah.

226
00:35:03.330 --> 00:35:12.360
Morgan Elowe MacLeod: yeah so I mean I heard a couple of missing things but one is like maybe that this distance measure can become a metric in and of itself.

227
00:35:13.050 --> 00:35:31.020
Morgan Elowe MacLeod: yeah and the other is that you know if you're thinking about like even if you're Designing an observational strategy

and you want to use certain features, you could ask like what noise level obscures the network's ability to recognize those.

228
00:35:31.260 --> 00:35:32.370
Morgan Elowe MacLeod: which I think it's kind of.

229
00:35:33.420 --> 00:35:35.310
Morgan Elowe MacLeod: amazing from a planning strategy.

230
00:35:38.220 --> 00:35:40.410
Morgan Elowe MacLeod: that's super interesting well Thank you so much.

231
00:35:43.380 --> 00:35:49.560
Sarah Jeffreson: Okay cool, so now we have a question from off me about house of.

232
00:35:50.820 --> 00:35:55.560
Sarah Jeffreson: Find onyx protected objects so abby.

233
00:35:56.520 --> 00:35:57.810
Abraham Loeb: Thank you, Sir.

234
00:35:58.920 --> 00:36:05.100
Abraham Loeb: So my question is based on the experience that they had when when I was a postdoc.

235
00:36:06.330 --> 00:36:17.640
Abraham Loeb: many decades ago I remember seeing back then that the astrophysical jonah was in actual volumes in the library not online, and I remember seeing.

236
00:36:18.600 --> 00:36:24.780
Abraham Loeb: That was in the 80s images in the astrophysical Journal of giant arcs around.

237
00:36:25.350 --> 00:36:34.350
Abraham Loeb: A galaxy clusters and until the mid 1980s and nobody spoke about gravitational lensing, is it possible explanation for those and.

238
00:36:34.680 --> 00:36:47.790
Abraham Loeb: People just ignore these they were publishing the astrophysical journal, you could see these arcs but nobody paid attention

to them, and my question is perhaps deep learning we do a better job than.

239
00:36:48.870 --> 00:37:06.180
Abraham Loeb: You know, scientists, in the sense of finding things that you know the simulations do not expect that scientists are not expecting but you know we will see things that, otherwise we will ignore So the question is how to put that into the the software.

240
00:37:07.680 --> 00:37:24.390
Aleksandra Ciprijanovic (she/her): So a lot of people currently some people currently doing deep learning for astro are kind of asking the same question and i've seen a couple of different approaches, but I think that the main approach, where people.

241
00:37:25.470 --> 00:37:37.440
Aleksandra Ciprijanovic (she/her): are trying to use a slightly different methods so not a supervised learning methods like here not train on a specific task and have examples with labels, but just have.

242
00:37:38.280 --> 00:37:59.850
Aleksandra Ciprijanovic (she/her): Really images and and try to leave the model to train on images and try to somehow see what's similar and what's different and and somehow translate that into into its Latin space, and you can look in inside of the network and see where things are located in this in this space and.

243
00:38:00.900 --> 00:38:16.020
Aleksandra Ciprijanovic (she/her): If you have something that is very weird and doesn't look like some things that are common it's going to be somewhere far away, and I remember recently looking at a paper that did this for for galaxy morphology.

244
00:38:17.490 --> 00:38:28.020
Aleksandra Ciprijanovic (she/her): They wanted to see if the network is going to on its own see that they're like, for example, three groups of objects like spires ellipticals and irregulars and mergers.

245
00:38:28.500 --> 00:38:36.990
Aleksandra Ciprijanovic (she/her): Or is it going to decide that maybe, something else is there and that can that the networking clump together as a separate class.

246
00:38:38.040 --> 00:38:45.030

Aleksandra Ciprijanovic (she/her): And yeah well, it turned out, it does have the same classes if they didn't discover anything super interesting and new.

247
00:38:45.390 --> 00:38:59.580
Aleksandra Ciprijanovic (she/her): But those sort of methods will also show you weird things and they will be either very far away or somewhere on the edges of the whole distribution, so those things are useful, I mean something like this, you can also, of course, use.

248
00:39:00.720 --> 00:39:12.570
Aleksandra Ciprijanovic (she/her): But if you're training it on one task everything that's, not that this is going to be strange so it's it's a little bit harder to see if something is really strange or just not included in your training data set.

249
00:39:14.490 --> 00:39:27.240
Aleksandra Ciprijanovic (she/her): So yeah it I think it's definitely going to be useful, because finding weird things with a terabyte per day with our society is not going to be feasible for human, I think.

250
00:39:28.620 --> 00:39:37.980
Aleksandra Ciprijanovic (she/her): And even even for four lenses I know that the four DS I think that people found with domain adaptation.

251
00:39:38.370 --> 00:39:55.890
Aleksandra Ciprijanovic (she/her): Maybe like 102 hundred new lenses just been using that after humans went through the entire data set and found everything that they could so yeah i'm hoping that this will be very helpful and who knows, maybe people device other ways of.

252
00:39:56.040 --> 00:39:57.390
Aleksandra Ciprijanovic (she/her): finding new things to.

253
00:39:58.980 --> 00:39:59.160
Abraham Loeb: Do.

254
00:40:01.890 --> 00:40:05.400
Sarah Jeffreson: yeah like I just like to.

255
00:40:06.930 --> 00:40:08.610
Sarah Jeffreson: follow up on I got.

256
00:40:09.750 --> 00:40:15.660

Sarah Jeffreson: You said, this is the new room.

257
00:40:17.490 --> 00:40:18.120
Sarah Jeffreson: Let.

258
00:40:21.390 --> 00:40:21.630
me.

259
00:40:23.670 --> 00:40:25.860
Sarah Jeffreson: pause able to.

260
00:40:27.270 --> 00:40:27.990
Sarah Jeffreson: Find.

261
00:40:29.910 --> 00:40:36.840
Sarah Jeffreson: More objects and humans are able to I tend to fly.

262
00:40:38.460 --> 00:40:38.940
Sarah Jeffreson: Like.

263
00:40:41.250 --> 00:41:04.920
Sarah Jeffreson: cause they're also callers call is hated to France birth tween the types of objects is I tend to find and the objects feel like slowly by slowly.

264
00:41:07.590 --> 00:41:09.450
Sarah Jeffreson: By that humans are able.

265
00:41:13.590 --> 00:41:14.400
Sarah Jeffreson: To see.

266
00:41:17.490 --> 00:41:26.700
Aleksandra Ciprijanovic (she/her): Very well, yes and no depends on the situation, so I don't remember the exact paper for the lenses the example that I gave you.

267
00:41:29.100 --> 00:41:40.860
Aleksandra Ciprijanovic (she/her): We if you use a supervised learning method and you give it some example of lenses it's usually going to mostly find similar things because that's what it was trained on.

268
00:41:42.480 --> 00:41:53.550

Aleksandra Ciprijanovic (she/her): But it doesn't have to be only that so, for example with with one of them, one of the projects that i'm involved in is using.

269
00:41:55.290 --> 00:41:59.850
Aleksandra Ciprijanovic (she/her): These types of methods so like regular convolution neural networks and also.

270
00:42:01.200 --> 00:42:04.200
Aleksandra Ciprijanovic (she/her): object detection algorithm so algorithms that.

271
00:42:05.310 --> 00:42:12.390
Aleksandra Ciprijanovic (she/her): tell you where the object is not only Oh, we have a galaxy but also galaxies right here, these are the pixels of the galaxy.

272
00:42:12.780 --> 00:42:33.690
Aleksandra Ciprijanovic (she/her): So we are doing applying these two methods to find low suffers brightness galaxies in the US, for example, and the the current biggest like the S catalog I know people who had to look through images they look thousands and thousands of images and created this very painful catalog.

273
00:42:35.220 --> 00:42:47.400
Aleksandra Ciprijanovic (she/her): But after we started using deep learning, first of all, we of course found all of the things that the human could, but we could also find things that have an even lower brightness so.

274
00:42:47.820 --> 00:43:00.570
Aleksandra Ciprijanovic (she/her): It still depends on how you pre process, the images, so I mean it's not like super fully automated but the pre processing is basically the same like was done before, but we could find fainter object, so I think that.

275
00:43:01.950 --> 00:43:14.970
Aleksandra Ciprijanovic (she/her): It can help us go further, maybe not win this way find completely different looking things, but things that are similar to what it knows it will be able to find better than we can.

276
00:43:16.260 --> 00:43:24.750
Aleksandra Ciprijanovic (she/her): But yeah maybe some other methods, like the one that don't require labels will find help us like I mentioned before, find different looking things.

277
00:43:25.800 --> 00:43:35.910
Aleksandra Ciprijanovic (she/her): Just or at least say tell tell to us well this group of objects, is something that I see as something similar, and you inspect it and you see that that's something new, maybe something like that.

278
00:43:37.410 --> 00:43:51.180
Sarah Jeffreson: Okay cool yes, like, I also have a crisis general question like on slides hand you were talking.

279
00:43:53.010 --> 00:43:56.460
Sarah Jeffreson: Yes, I think slide nine or 10.

280
00:43:56.910 --> 00:44:00.480
Sarah Jeffreson: Like you were talking about.

281
00:44:01.740 --> 00:44:03.030
Sarah Jeffreson: hiding the.

282
00:44:06.180 --> 00:44:10.440
Sarah Jeffreson: Data a quarter, according to this vision.

283
00:44:13.860 --> 00:44:15.390
Sarah Jeffreson: pounder ease.

284
00:44:16.200 --> 00:44:16.590
and

285
00:44:17.700 --> 00:44:20.160
Sarah Jeffreson: You drawn them as.

286
00:44:23.160 --> 00:44:24.450
Sarah Jeffreson: Heinz here.

287
00:44:26.310 --> 00:44:32.400
Sarah Jeffreson: Other day he usually assumes who like be.

288
00:44:33.900 --> 00:44:39.990
Sarah Jeffreson: flats surfaces in your your.

289
00:44:42.480 --> 00:44:42.960

Sarah Jeffreson: slowly.

290
00:44:44.220 --> 00:44:44.610
Sarah Jeffreson: Your.

291
00:44:46.230 --> 00:44:47.340
Sarah Jeffreson: parameters.

292
00:44:48.810 --> 00:44:56.820
Sarah Jeffreson: base or how do you control their shapes.

293
00:44:58.560 --> 00:45:04.290
Aleksandra Ciprijanovic (she/her): yeah that's a very good question so they are never.

294
00:45:05.490 --> 00:45:13.590
Aleksandra Ciprijanovic (she/her): Simple looking or flat or anything that we like, but they will be, whatever the depending on the problem so.

295
00:45:14.610 --> 00:45:25.950
Aleksandra Ciprijanovic (she/her): This the blog that I was showing those are called Disney plots and they're not very ideal way of plotting the Multi dimensional space, because they.

296
00:45:27.090 --> 00:45:34.890
Aleksandra Ciprijanovic (she/her): They do translate it into a 2d they're just looking at the similarities between the two distributions one into the one in whatever and the that you have.

297
00:45:36.150 --> 00:45:47.850
Aleksandra Ciprijanovic (she/her): But they're not keeping the distances between the points very realistic, so that, like, for example, if you have a player, a region that has more dense.

298
00:45:49.680 --> 00:46:01.770
Aleksandra Ciprijanovic (she/her): Population you just have that as a big blob some on a tasty but it's kind of weird and these new parts also depend, they have some hyper parameters that you can find you, and so the data will look a little bit different so.

299
00:46:02.280 --> 00:46:12.630
Aleksandra Ciprijanovic (she/her): yeah you can't you, you need to use them just for kind of understanding what's going on not thinking about them very precisely, there are better methods.

300
00:46:14.550 --> 00:46:26.460
Aleksandra Ciprijanovic (she/her): The again, nothing is ideal, because how can you really fathom the distribution in and dimensions, everything is super close the boundaries always near it's very weird things are very oh is near the edges.

301
00:46:27.570 --> 00:46:36.450
Aleksandra Ciprijanovic (she/her): But, for example here it's a different example, I can tell you what it is here, it looks kind of similar You see, but this is a human.

302
00:46:38.430 --> 00:46:39.390
Aleksandra Ciprijanovic (she/her): Oh sorry is a map.

303
00:46:40.410 --> 00:46:52.470
Aleksandra Ciprijanovic (she/her): In ice and maps you do a similar thing, but better you calculate geodesic distances in the end dimensions, you know decide how many examples you want to track.

304
00:46:52.800 --> 00:46:57.510
Aleksandra Ciprijanovic (she/her): And you see how this them, they are in the space and then you translate to to the.

305
00:46:57.930 --> 00:47:15.510
Aleksandra Ciprijanovic (she/her): Trying to keep the dude as existence is the same, so these distances that you calculate somehow here so in Iceland maps the actual positions of the data points is a little bit better it describes the space, the original space, a little bit better.

306
00:47:16.800 --> 00:47:21.450
Aleksandra Ciprijanovic (she/her): So yeah, this is a newer plot, so I moved away from PCs I don't like them.

307
00:47:22.980 --> 00:47:25.380
Aleksandra Ciprijanovic (she/her): they're not very good for quantitative things.

308
00:47:28.020 --> 00:47:38.550
Aleksandra Ciprijanovic (she/her): yeah so, by the way this example is a slightly different is still domain adaptation, but i'm trying to do galaxy morphology and i'm actually studying.

309
00:47:39.690 --> 00:47:45.480

Aleksandra Ciprijanovic (she/her): What happens if you perturb one pixel of the image so like looking at.

310
00:47:48.030 --> 00:48:01.320
Aleksandra Ciprijanovic (she/her): If, like a CCD readout has a problem you have a cosmic Ray whatever something also like compression decompression of images our data is going to be very high resolution we're going to compress compress it.

311
00:48:01.530 --> 00:48:07.050
Aleksandra Ciprijanovic (she/her): And unfortunately this sometimes leads to 10 years it just the way things are, and.

312
00:48:07.740 --> 00:48:25.680
Aleksandra Ciprijanovic (she/her): we're doing one pixel attacks as a proxy for a possible problem like this and and looking if this makes the network, the knows what spiral elliptical and murderer can this make the the thing the classifier wrong and, yes, it can, of course, unfortunately.

313
00:48:26.730 --> 00:48:40.680
Aleksandra Ciprijanovic (she/her): But doing domain adaptation so helping you to understand, so, for example, different noise levels also increases robustness to these kind of problems, so you basically have two Games were just like doing one thing.

314
00:48:41.790 --> 00:48:45.240
Aleksandra Ciprijanovic (she/her): So yeah This is something you were that I started testing.

315
00:48:46.800 --> 00:48:48.600
Aleksandra Ciprijanovic (she/her): But yeah.

316
00:48:50.610 --> 00:48:53.910
Sarah Jeffreson: cool yeah that's us interesting.

317
00:48:55.980 --> 00:48:58.620
Sarah Jeffreson: I think we have time.

318
00:49:00.630 --> 00:49:04.470
Sarah Jeffreson: For one last question.

319
00:49:07.230 --> 00:49:08.850
Sarah Jeffreson: From.

320
00:49:11.850 --> 00:49:14.550
Sarah Jeffreson: From Morgan walk in here.

321
00:49:15.120 --> 00:49:15.870
Thank you.

322
00:49:17.220 --> 00:49:22.080
Morgan Elowe MacLeod: yeah so I mean, I have some questions, but I was hoping, you could talk.

323
00:49:23.400 --> 00:49:27.090
Morgan Elowe MacLeod: I guess a tiny bit more about this last.

324
00:49:28.200 --> 00:49:31.530
Morgan Elowe MacLeod: Point of you know, there are certain.

325
00:49:32.880 --> 00:49:43.740
Morgan Elowe MacLeod: Physical things that we know or observational or physical whatever aspects that we know we need to adapt across and days those might be things like.

326
00:49:44.790 --> 00:49:53.370
Morgan Elowe MacLeod: And I mean you can tell me if i'm interpreting this wrong but it sounds kind of like you know never trust just one pixel to make your classification, all of us.

327
00:49:54.630 --> 00:49:55.380
Morgan Elowe MacLeod: and

328
00:49:57.570 --> 00:49:58.800
Morgan Elowe MacLeod: So i'm wondering.

329
00:49:59.880 --> 00:50:05.040
Morgan Elowe MacLeod: If you can talk more about that point that you made about like.

330
00:50:07.170 --> 00:50:10.890
Morgan Elowe MacLeod: Improving like training the.

331
00:50:12.090 --> 00:50:16.560

Morgan Elowe MacLeod: domain up that station on like one kind of noise or something.

332
00:50:17.580 --> 00:50:32.640
Morgan Elowe MacLeod: Improving the performance, with respect to say cosmic rays, or something like that so like what what is it about the domain adaptation that like encodes that transfer ability or I guess.

333
00:50:33.750 --> 00:50:42.720
Morgan Elowe MacLeod: Like qualitatively I understand this idea that it makes the classification more robust, but how does that actually happen, I don't think I understand.

334
00:50:43.500 --> 00:51:00.120
Aleksandra Ciprijanovic (she/her): So the whole point of domain adaptation methods is that it helps the network try to find the domain invariant features to use So if you have a different noise.

335
00:51:00.540 --> 00:51:10.500
Aleksandra Ciprijanovic (she/her): model or noise level in some of the images that definitely needs to be a feature that you are not going to use because that that's not.

336
00:51:11.070 --> 00:51:23.010
Aleksandra Ciprijanovic (she/her): True yeah so it's forcing it to see that there is this is different, it, I cannot find the same thing into data sets i'm just going to try to focus on something that I do see.

337
00:51:23.610 --> 00:51:30.870
Aleksandra Ciprijanovic (she/her): And for that reason, of course, when the noise is the problem you're just start to disregard the noise it yeah.

338
00:51:31.200 --> 00:51:43.980
Aleksandra Ciprijanovic (she/her): And because in this case, where, for example, when the noise is the problem, or like this is exactly the example that we can use So here we are making marks of analysis team one and 10 years of observations.

339
00:51:45.360 --> 00:51:54.990
Aleksandra Ciprijanovic (she/her): And we're training or the less noisy so around 10 years and we're domain adapting to the one year, but you can do widespread so it doesn't matter.

340
00:51:56.370 --> 00:52:09.780

Aleksandra Ciprijanovic (she/her): And it because it learns not to care about the noise not caring about the noise also helps it not care about that one so that is wrong, for example, if I attack it so that's also kind of part of noise.

341
00:52:10.800 --> 00:52:21.000
Aleksandra Ciprijanovic (she/her): So what we've seen here is that not only, of course, the domain adaptation help it classify things in both domains, we knew that that will happen.

342
00:52:21.480 --> 00:52:26.790
Aleksandra Ciprijanovic (she/her): But what we've seen for like additional problems like those one pixel attack the images.

343
00:52:27.660 --> 00:52:39.120
Aleksandra Ciprijanovic (she/her): The this sense, you have to travel from the original image to the one pixel attack the image that is actually successful in fooling so the cross the decision boundary.

344
00:52:39.930 --> 00:52:50.610
Aleksandra Ciprijanovic (she/her): is I think 2.3 times larger in this example, so you can still theoretically fool the model, of course, because we are attacking it here with a efficient method.

345
00:52:51.240 --> 00:53:04.680
Aleksandra Ciprijanovic (she/her): But in reality, because the distance to cross the border increased so much probably most of the actual problematic because those are not going to be, you know unlucky enough to be able to cross.

346
00:53:05.100 --> 00:53:21.510
Aleksandra Ciprijanovic (she/her): So you basically not only get robustness to the thing that you trained on you get robustness on other kind of similar things, but not the same so yeah when you mentioned distances between stuff we actually try to calculate this here.

347
00:53:21.990 --> 00:53:22.410
Aleksandra Ciprijanovic (she/her): In this.

348
00:53:22.680 --> 00:53:23.610
Aleksandra Ciprijanovic (she/her): sample so I have.

349
00:53:24.420 --> 00:53:33.150

Aleksandra Ciprijanovic (she/her): metrics maybe not that much in this the shorter a paper here but soon ish there will be a bigger paper with a distance isn't that.

350
00:53:33.990 --> 00:53:35.250
amazing yeah.

351
00:53:36.270 --> 00:53:36.840
Morgan Elowe MacLeod: Thank you.

352
00:53:38.640 --> 00:53:40.560
Sarah Jeffreson: Okay, well, we.

353
00:53:41.580 --> 00:53:41.910
Sarah Jeffreson: got.

354
00:53:43.020 --> 00:53:45.330
Sarah Jeffreson: Time and so.

355
00:53:46.980 --> 00:53:49.290
Sarah Jeffreson: we'll end like it will like.

356
00:53:50.370 --> 00:53:52.320
Sarah Jeffreson: Will like in like their.

357
00:53:55.680 --> 00:54:01.050
Sarah Jeffreson: Part Alexandra if you would like to say some.

358
00:54:02.340 --> 00:54:08.190
Sarah Jeffreson: kind of our words to summarize and can.

359
00:54:09.240 --> 00:54:11.670
Sarah Jeffreson: can conclude here.

360
00:54:12.780 --> 00:54:33.060
Aleksandra Ciprijanovic (she/her): Sure um well yeah it was very fun very good Thank you so much for inviting me, I hope I try to to convince you and explain that, with huge data that we already have and where we will have in the future, hopefully next year, this is the will see.

361
00:54:34.260 --> 00:54:42.570

Aleksandra Ciprijanovic (she/her): We are going to have to make models that are better than what we're using now we're going to have to make models that are going to allow us to work.

362
00:54:42.900 --> 00:54:51.780
Aleksandra Ciprijanovic (she/her): With multiple data sets and the main adaptation is definitely going to be one way of doing that because it's going to help us.

363
00:54:52.350 --> 00:54:59.610
Aleksandra Ciprijanovic (she/her): Help us really find common things between data sets and, hopefully, in the future also I really think it will.

364
00:55:00.390 --> 00:55:08.310
Aleksandra Ciprijanovic (she/her): I will the meter the patient will also help us understand the physics models, as we were talking about today and also understand.

365
00:55:08.610 --> 00:55:22.590
Aleksandra Ciprijanovic (she/her): The neural network methods, because we can look inside see what's happening and and use this information to go step further, to build even better models so i'm hoping that we can do all of that, by understanding domain adaptation, a little bit better.

366
00:55:24.630 --> 00:55:25.350
Morgan Elowe MacLeod: Thank you.

367
00:55:27.540 --> 00:55:40.320
Morgan Elowe MacLeod: And if you have your thing and full screen and don't get to see the chat a bunch of people are coming into the site, thank you, so thank you so much we're really grateful and with that Thank you all for coming and.

368
00:55:42.390 --> 00:55:43.950
Aleksandra Ciprijanovic (she/her): Thank you so much for inviting me.