WEBVTT

1
00:00:26.820 --> 00:00:28.680
Gurdip Singh: So good morning everyone.

2
00:00:30.450 --> 00:00:50.430
Gurdip Singh: on behalf of sighs CNS it's a pleasure to welcome you to sys distinguished lecture series I am good eating i'm the division director for the computer network system, starvation and this is one of the units within the computer science and engineering directorate.

3
00:00:51.900 --> 00:01:11.760
Gurdip Singh: So, before I start with the introductions and the distinguished lecture just want to go over some logistics Informations, so we will have the lecture presentation now, and I want to remind everyone that today at 3pm Eastern, we will also have office hours.

4
00:01:12.810 --> 00:01:26.430
Gurdip Singh: That our distinguished lecture will join and you have will have the opportunity to engage in more one on one conversations with him at that point, so please, please, please do join that.

5
00:01:26.880 --> 00:01:46.020
Gurdip Singh: And also want to mention that during the lecture just use the Q amp a feature to type in your questions, we will assemble all those equations and then the Q amp a session begins, we will start with answering those questions so with that it's a great pleasure to.

6
00:01:47.040 --> 00:01:52.110
Gurdip Singh: Have Dr stankovic today to as our distinguished lecture.

7
00:01:54.510 --> 00:02:07.650
Gurdip Singh: And stankovic he's the VP America professor of computer science in the computer science department at the University of Virginia and he's also the director of the link lap there.

8
00:02:08.790 --> 00:02:10.530
Gurdip Singh: He received his PhD.

9
00:02:11.730 --> 00:02:13.170
Gurdip Singh: From brown university.

10
00:02:14.190 --> 00:02:17.400
Gurdip Singh: he's an accomplished researcher and has.

11
00:02:18.570 --> 00:02:23.040
Gurdip Singh: long list of achievement and honors and I will just mention a few he's the.

12
00:02:23.520 --> 00:02:43.680
Gurdip Singh: Father of both it I Tripoli and ECM it was awarded an honorary doctorate from the University of your UK for his work on real time systems, he won he has won the ieee real time systems technical committees Award for outstanding technical contribution leadership's.

13
00:02:45.120 --> 00:02:56.910
Gurdip Singh: is also on the test of time paper award eight best paper award and and and in 2015 he was awarded University of Virginia distinguished research scientist one.

14
00:02:57.390 --> 00:03:02.760
Gurdip Singh: Again, I said the long list is long and you ever refer to you to his bio to.

15
00:03:03.540 --> 00:03:14.790
Gurdip Singh: Look at those his research in past has been in the area of real time systems wireless sensor networks smart and connected health smart city cyber physical system and.

16
00:03:15.390 --> 00:03:32.340
Gurdip Singh: An iot he's been long been working on, you know, looking at all of these technologies applied to the healthcare domain, so today is going to talk about two words ambient intelligence and smart healthcare this, I would like to welcome jack for this presentation.

17
00:03:34.500 --> 00:03:53.850
John A. Stankovic: So it's a pleasure to be here, I thank you for the invitation and just mentioned one thing about the link lab it started in 2018 it's focuses on Cyber physical systems, research and it's grown to 240 faculty and over 200 Grad students so it's a pretty large enterprise.

18
00:03:54.870 --> 00:04:05.910
John A. Stankovic: So I was asked to put one slide in about something about myself before we get into the into the talk so I thought about that and I thought I would say, you know how did I get here i'm.

19
00:04:06.900 --> 00:04:19.800
John A. Stankovic: In a faculty Member for over 40 years I do a lot of research, I do a lot of teaching, so how did I actually get here, and so I want to briefly say something about the teaching side and the research side going teaching.

20
00:04:21.090 --> 00:04:27.570
John A. Stankovic: I was in high school, I was good in math and my friend was terrible at math and his mother begged me to tutor him right.

21
00:04:27.870 --> 00:04:35.940
John A. Stankovic: And so I I did that in it, and it was very satisfying could see how rapidly he improved and he was so happy about it, his mother was even more happy.

22

00:04:36.390 --> 00:04:50.310
John A. Stankovic: And, and so I felt this was this was great and also my sister had her friend, she was in fourth grade and I she had trouble with math so I asked me to to to her, and I did so I got this love of seeing students really.

23
00:04:51.000 --> 00:04:56.910
John A. Stankovic: understand information and learn something, and so I really felt like I should.

24
00:04:57.990 --> 00:05:06.270
John A. Stankovic: Teach In addition I always felt like taking so many math and engineering courses which was all most of my engineering courses are all math.

25
00:05:06.840 --> 00:05:15.420
John A. Stankovic: I went to a small engineering school and it didn't have a lot of hands on and they never told me why practically I said, I have to be able to teach better than that.

26
00:05:15.810 --> 00:05:21.360
John A. Stankovic: You know, when you teach me all this math but I never know why i'm learning it I know all the mechanisms, and so I really want to.

27
00:05:21.780 --> 00:05:27.630
John A. Stankovic: Teach differently, and when somebody is in my class, I want to make sure they understand why we're learning something.

28
00:05:28.290 --> 00:05:36.870
John A. Stankovic: And so that got me into really pushing the notion of I wanted to be a teacher and then a research when I left.

29
00:05:37.410 --> 00:05:49.710
John A. Stankovic: undergrad I had a double a degree, and it was a long time ago and there wasn't many there are many programmers and I ended up going to bell labs and I worked on the Anti ballistic missile system.

30
00:05:50.430 --> 00:06:00.330
John A. Stankovic: And I programs for it in a similar language and I when I got there, the first day they said, you want to learn how to program and I said sure, and he.

31
00:06:00.990 --> 00:06:08.010
John A. Stankovic: taught me how to program and then I worked on this project for four years was such an exciting project with so much research involved with it.

32
00:06:08.400 --> 00:06:18.540
John A. Stankovic: And so I thought you know I really want to do this, but I didn't feel like I had much of a computer science background, so I I said I need to go back to Grad school and learn how to.

33

00:06:19.260 --> 00:06:26.730
John A. Stankovic: Do research and computer science and also have the basis for teaching and so then now 40 years later, you know here I am i'm still doing it.

34
00:06:28.500 --> 00:06:28.860
John A. Stankovic: Okay.

35
00:06:30.270 --> 00:06:39.360
John A. Stankovic: I thought before I get into the core idea brief overview of what our group is doing, and so we do a lot of work with wearables and in situ devices.

36
00:06:39.810 --> 00:06:51.090
John A. Stankovic: We focus on trying to build devices into cognitive assistance, using a lot of machine learning and natural language processing that is going to be the theme for today.

37
00:06:52.260 --> 00:06:57.090
John A. Stankovic: But also there's so many different Apps for medical.

38
00:06:58.950 --> 00:07:07.350
John A. Stankovic: Help for websites with medical information people telling you what to do, for your health and so there's lots of potential.

39
00:07:08.430 --> 00:07:16.680
John A. Stankovic: conflicting information, so we spent a lot of time working on identifying conflicting information and resolving it.

40
00:07:17.700 --> 00:07:24.600
John A. Stankovic: acoustics is a powerful modality we'll talk about that today as well, and we do a lot of work with real deployment time it's.

41
00:07:26.400 --> 00:07:33.030
John A. Stankovic: Just briefly here at some of the projects that emphasize the deployments we did a project was just completed with.

42
00:07:33.900 --> 00:07:43.980
John A. Stankovic: Understanding family dynamics for obese families in Los Angeles, we work with behavioral psychologists there we built the system and then deployed it in 23 real families.

43
00:07:44.760 --> 00:07:55.500
John A. Stankovic: We have this smartwatch reminder systems which we're extending into this notion of cognitive assistance here we work with our va Center for telemedicine.

44
00:07:56.610 --> 00:08:06.810

John A. Stankovic: We currently have a project on the alzheimer's patient care giver interaction, where we focus on the caregiver we want to understand the stress and of the caregiver and improve.

45
00:08:07.950 --> 00:08:17.940
John A. Stankovic: Prove that for them with interventions, basically, and we have three deployments it's this a newer project, and we have three deployments.

46
00:08:18.960 --> 00:08:21.150
John A. Stankovic: undergo undergoing right parody this minute.

47
00:08:22.680 --> 00:08:23.370
John A. Stankovic: We have.

48
00:08:25.290 --> 00:08:36.150
John A. Stankovic: Technology for handwashing which we actually use for a while, in the pediatric intensive care unit and uva and i'll talk about that today as well briefly.

49
00:08:36.930 --> 00:08:46.590
John A. Stankovic: And we work with first responders where we work with North garden fire and rescue is a local rescue outfit here also Richmond fire.

50
00:08:47.280 --> 00:09:01.200
John A. Stankovic: And I did a sabbatical at Oxford and we have an ongoing project with them where there's other first responders in the UK, and this also fits the theme is as cognitive assistant for for that work as well okay oh.

51
00:09:02.340 --> 00:09:12.330
John A. Stankovic: Oh, one more thing is that, in addition to those we have some newer themes that I also worked in smart cities and we try to in that case we applied formal methods.

52
00:09:13.020 --> 00:09:21.240
John A. Stankovic: To improve machine learning models so we want to integrate properties of the system into the models to the models don't give you a.

53
00:09:21.750 --> 00:09:31.980
John A. Stankovic: ludicrous results at times and so he is one more methods to do that and also in complex CPS systems uncertainties are rampant and So how do we.

54
00:09:32.520 --> 00:09:42.240
John A. Stankovic: address the uncertainties in the in the models and in the predictions and now we're trying to think about how do we move this these results into this smart health domain.

55
00:09:43.590 --> 00:09:43.920
John A. Stankovic: Okay.

56
00:09:46.320 --> 00:09:58.230
John A. Stankovic: Today, I mean there's this notion of Internet of healthcare things iot and I want to start with a question, you know, is this height, that this is going to be a revolution, or is it really going to be a revolution.

57
00:09:58.950 --> 00:10:06.810
John A. Stankovic: And so what we see today is, you know they have a lot of smart watches more and more sold like this apple watch here and a lot of.

58
00:10:07.500 --> 00:10:15.780
John A. Stankovic: The sensors are on air, including an ekg, for example, we have disruptive technologies like smart skin, you can kind of build sensor.

59
00:10:16.230 --> 00:10:21.780
John A. Stankovic: circuits and so on, and put them on people and monitor various physiological parameters.

60
00:10:22.170 --> 00:10:28.890
John A. Stankovic: they're smart homes, with a lot of incentive sensors that can be integrated others disruptive technologies like smart textiles.

61
00:10:29.310 --> 00:10:37.290
John A. Stankovic: And I think putting all these together will indeed create a revolution, and let me just give kind of an overall vision of a revolution.

62
00:10:38.220 --> 00:10:42.210
John A. Stankovic: So here's let's say this elderly person is living alone, they have a Walker.

63
00:10:42.990 --> 00:10:54.180
John A. Stankovic: They have maybe in body sensors they might have taken our pills for assessment and they have pacemaker Yvonne body like wearables like a smartwatch or pendants.

64
00:10:54.540 --> 00:11:06.720
John A. Stankovic: That you might have a smartphone which is almost acting like a medical device in many cases, and then there may even be sensors are actually there's on products that help them to be better.

65
00:11:07.380 --> 00:11:14.190
John A. Stankovic: suited for medical situations in addition that's persons living in a smart home the smart home they have.

66
00:11:15.330 --> 00:11:24.630
John A. Stankovic: sensors in a bed to detect sleep quality medication sensors and so on, to try to make the environment, something to improve healthcare.

67
00:11:25.200 --> 00:11:30.120
John A. Stankovic: However, when this person goes out into the city and the Smart Cities smart cities will have.

68
00:11:30.750 --> 00:11:40.290
John A. Stankovic: Various technology also and they may be monitoring things like air quality, and if this person has asthma, then the Air Quality could.

69
00:11:41.220 --> 00:11:54.930
John A. Stankovic: affect that person and morning the person about that would be would be helpful, we said there's also cloud, and all this web md and other data mining data collections and so on collecting information.

70
00:11:56.130 --> 00:12:05.220
John A. Stankovic: personal information so it's personalized say interventions, but also general population information is there.

71
00:12:06.660 --> 00:12:18.420
John A. Stankovic: One concept that this ambient intelligence for for healthcare would be all of this as a holistic viewpoint, and we see more and more actual activations activations on the.

72
00:12:18.900 --> 00:12:39.810
John A. Stankovic: On the end for the patient and, in some cases, the activations are kind of alarms or advice and then pulling all this together, I believe, is is like if all this can work, and all this can work in this synergistic way, you would have this notion of ambient healthcare intelligence.

73
00:12:41.970 --> 00:12:49.800
John A. Stankovic: And so, today I will talk a bit first about using smart watches and wearables as basis for kindness of assistance.

74
00:12:50.460 --> 00:13:05.790
John A. Stankovic: And that drives us towards this ambient health intelligence and also since acoustics is such a powerful modality, especially for healthcare i'm going to briefly talk about some of the results that getting moved from.

75
00:13:07.320 --> 00:13:22.170
John A. Stankovic: A microphone and it is when the microphone is in the environment and not necessarily next to the speaker and also dealing with anxiety and the new ways we can address that mental health issue.

76
00:13:23.190 --> 00:13:35.850
John A. Stankovic: All right, well, the cognitive assistance on a smartwatch is that possible, and you know generally a guy wants to have general intelligence, but general intelligence, I think there's way way far off.

77
00:13:36.420 --> 00:13:43.830
John A. Stankovic: And I think we can make some headway if we talk about more general intelligence for a particular domain like health.

78
00:13:44.520 --> 00:13:55.800
John A. Stankovic: So we want this kind of assistant that say on a smartwatch to interact with healthcare Internet services to support conversations, so that people can just.

79
00:13:56.370 --> 00:14:08.370
John A. Stankovic: You know kind of talk to it and it talks back, especially for elderly they they really can't see or use the smartwatch very easily on such a small kind of interface second.

80
00:14:09.360 --> 00:14:21.900
John A. Stankovic: Typical things like reminders suggestions alarms that's been around a long time, but we want to make those explainable and we want to make them more generic in terms of enabling people to ask questions.

81
00:14:22.950 --> 00:14:37.680
John A. Stankovic: We should support both physical and mental health if the pandemic shows up, we should easily extend it to the technology to add in information about pandemic and, of course, privacy is is key, but we're not going to discuss privacy.

82
00:14:39.150 --> 00:14:39.570
John A. Stankovic: So.

83
00:14:41.040 --> 00:14:49.020
John A. Stankovic: that's what we're doing here So the first thing is, we built a system called i'd here, and it is a variable medication exercise reminder system.

84
00:14:49.500 --> 00:14:57.090
John A. Stankovic: So, working with telemedicine when they're a stroke patients and then they go home they take a smartwatch with them and then.

85
00:14:58.020 --> 00:15:12.480
John A. Stankovic: We want to use that smartwatch for medication reminders and exercise reminders, but we want to extend it towards cognitive assistance, so this screen here shows what the watch shows for a reminder of your medication so.

86
00:15:13.230 --> 00:15:24.750
John A. Stankovic: The medication there's a script that when they leave the script is describing their medication regimen and their exercise regimen and it automatically gets downloads and smartwatch.

87
00:15:26.100 --> 00:15:37.470
John A. Stankovic: When they're home and they get a reminder, for example in the morning, they might get a reminder, it shows up showing the pill, it says take Please take this pill at nine o'clock and the person could say.

88
00:15:38.850 --> 00:15:52.440

John A. Stankovic: Which pill is this again why am I taking a what tray have food with it should I not have food with it and all sorts of additional information about that many of these patients take 678 even 10 different medications.

89
00:15:53.730 --> 00:16:07.530
John A. Stankovic: Then, when they needed to do exercise, you know they can tell them which exercise and they might say how many repetitions am I supposed to do again or do I have to do this every day, so they have a general conversation about exercise and and that could be here.

90
00:16:08.580 --> 00:16:16.260
John A. Stankovic: This has already been all implemented what we're doing with the exercise is we're actually improving it to where we can.

91
00:16:16.740 --> 00:16:33.600
John A. Stankovic: measure, the quality of the exercise as well, and also relate that quality to pain and so that's where we trying to extend the services that exists or dealing with exercise so it's not just simply a reminder, the person can also say remind me later.

92
00:16:34.710 --> 00:16:45.180
John A. Stankovic: Since the Watch has various physiological parameters, they can also say you know kind of here's your ekg or I want to look at my ekg now and it would demonstrate that.

93
00:16:45.810 --> 00:17:00.390
John A. Stankovic: If I said I wouldn't talk about privacy, but we also allow the screen like if a person wants privacy and doesn't want the conversation because other people in the room, or something like that, then it can be just textual.

94
00:17:02.730 --> 00:17:12.480
John A. Stankovic: And so, this is kind of what what I just said it's a summary of the basic thing we see it, a lot of smart watches whatever there, it could be medication exercise reminders.

95
00:17:12.870 --> 00:17:29.040
John A. Stankovic: But we're trying to drive it to general verbal questions about that and then allow rescheduling allow quality and incorporate physiologic parameters in this case, an ekg if you use the empirical watch you could even get additional.

96
00:17:30.210 --> 00:17:42.930
John A. Stankovic: physiological parameters alright, so I think of this as kind of what notion of kind of taking individual Apps and combining them and putting them together to have more comprehensiveness.

97
00:17:44.220 --> 00:17:48.870
John A. Stankovic: Oh, I want to happen is the only video I have and I show a brief DEMO.

98
00:17:49.320 --> 00:18:03.030
John A. Stankovic: The first part, shows the watch and it's only a couple minutes, and the second part shows what's on the web that doctors would see for the second part it's not easy to see on the screen So if you just listen to what's being

said you get an idea of what is.

99
00:18:04.920 --> 00:18:12.480
John A. Stankovic: What we're storing on the on the on the cloud basically let them do the DEMO have made them so i'll receive a reminder.

100
00:18:14.250 --> 00:18:18.150
Is for taking ECC so I tap on the reminder.

101
00:18:19.980 --> 00:18:23.580
it's the time to do an ECG do you want to try it now.

102
00:18:25.860 --> 00:18:26.700
Yes, you.

103
00:18:29.610 --> 00:18:30.780
Please tell me deduce.

104
00:18:38.460 --> 00:18:44.490
Okay press the Crown to go to Apps tap on the ECG APP and follow the instructions.

105
00:18:47.940 --> 00:18:55.860
So no I follow the instructions I go to the ECG Apps and now I can do the ECG.

106
00:19:05.490 --> 00:19:06.960
received a reminder for a bill.

107
00:19:09.450 --> 00:19:14.820
it's time to take your nightly medications take all medications as directed by your physician.

108
00:19:16.800 --> 00:19:17.970
be stimulated is.

109
00:19:27.990 --> 00:19:28.680
That for a minute.

110
00:19:30.930 --> 00:19:34.080
Okay, thank you, can you remind me after one minute.

111
00:19:36.330 --> 00:19:40.800
Okay, remind you after one minutes Okay, thank you.

112
00:19:45.030 --> 00:19:45.870
you're welcome.

113
00:19:47.670 --> 00:20:03.930
So this is the website for the doctors and clinicians to view the patient data it fetches the data from the cloud server with data from the patient gets uploaded so let's find the data for the patient to 000.

114
00:20:05.640 --> 00:20:15.930
So here, you can see that the patient ID and the number of reminders provided the number of reminders the patient responded, the response rate and.

115
00:20:17.910 --> 00:20:21.720
interaction between the patient and the system for each of the treatment.

116
00:20:23.280 --> 00:20:33.510
So by these doctors will be able to see what exercise of what medication, or what repented patient exactly follow Dave it.

117
00:20:41.520 --> 00:20:45.240
John A. Stankovic: That was earlier version of our system and it's been enhanced.

118
00:20:46.530 --> 00:20:52.800
John A. Stankovic: It, for example, that version, you had a tap to watch a couple times to to move ahead, which the goal that's now automated.

119
00:20:54.720 --> 00:20:55.830
John A. Stankovic: Anyway, so.

120
00:20:57.090 --> 00:21:10.500
John A. Stankovic: When the pandemic came, we also wanted to expand the services, and one way is you know kind of integrating additional features into the into the smartwatch and.

121
00:21:11.130 --> 00:21:23.280
John A. Stankovic: We did is we did a solution for handwashing which is useful in for general hygiene, even when it's not in a pandemic, like the work we did with the pediatric intensive care unit hospital.

122
00:21:24.360 --> 00:21:30.360
John A. Stankovic: Also, can say try to detect mood and anxiety which we'll talk about these three things that are in red.

123
00:21:31.830 --> 00:21:44.700
John A. Stankovic: In the rest of the clock, we want to also continue to theme of having voicebase conversation, so the

person can ask for pandemic information, you know how many deaths today, where are the shots available.

124
00:21:45.810 --> 00:21:47.160
John A. Stankovic: and so on.

125
00:21:48.210 --> 00:22:08.070
John A. Stankovic: They could also get reminders and alert about the about in Washington, they can get different physiological parameters and perhaps use those to to detect the whether a person is showing symptoms like coughing and breathing problems those can also be added into this.

126
00:22:09.540 --> 00:22:20.670
John A. Stankovic: OK, so the loving focus on a few things and give a little more technical details and the three things I want to focus on one what is is the hand washing and.

127
00:22:21.480 --> 00:22:27.330
John A. Stankovic: Who has guidelines for how you really should wash and you should be washing hands for at least 20 seconds.

128
00:22:27.720 --> 00:22:33.960
John A. Stankovic: And so the idea is, we want to develop the quality of doing things, not just the two job or not, and our solution.

129
00:22:34.560 --> 00:22:51.570
John A. Stankovic: Is a hybrid CNN rnn solution which i'll show the architecture in a minute, and it also shows supports conversations about yes, you did your hand washing well or you skip this step, where you skip that step, or you didn't do it long enough, and so on.

130
00:22:53.070 --> 00:23:03.060
John A. Stankovic: These are the 10 steps that i'm not going to dwell on this, but there's 10 different steps that you're supposed to follow and we try to detect each and every one of those steps.

131
00:23:04.230 --> 00:23:14.160
John A. Stankovic: That there's a smartwatch APP built here in the middle and it runs on an apple watch and it will do this detection there's a here's a.

132
00:23:14.820 --> 00:23:24.840
John A. Stankovic: Motion sensor that can detect it's actually a bluetooth beacon detective coming in, and when you come in and give you a reminder that you just return home and you need to wash your hands.

133
00:23:25.380 --> 00:23:35.940
John A. Stankovic: And you can have this conversation back and forth with the watch so, for example, one time by log when you enter the home with my say time to watch okay i'll do it.

134

00:23:36.480 --> 00:23:47.490
John A. Stankovic: You did a good job, but you should rub your hands more and wash for a minute and 20 seconds, and so, so the various dialogues that that it supports with respecting to handwashing.

135
00:23:49.080 --> 00:23:53.940
John A. Stankovic: The solution is not trivial and earlier versions.

136
00:23:55.260 --> 00:23:55.800
John A. Stankovic: Had.

137
00:23:57.090 --> 00:23:59.250
John A. Stankovic: different levels of accuracy and.

138
00:24:00.480 --> 00:24:11.550
John A. Stankovic: This architecture seems to work really well and so i'm not going to go into a lot of detail about it but i'll just give you some sense, so the solution really takes this bottom part here.

139
00:24:12.750 --> 00:24:26.130
John A. Stankovic: Is a convolution neural network, and it has three convolution layers 123 and for each of these layers we add this thing called squeeze and excite glock.

140
00:24:26.760 --> 00:24:39.720
John A. Stankovic: Which is another architectural feature and that block is developed by other people, but we use it, and it is it learns the correlation between the filters in the convolution layer

141
00:24:40.140 --> 00:24:49.410
John A. Stankovic: And so we have this represents kind of the relationship of your hands to each other so learns the spatial relationships.

142
00:24:50.130 --> 00:25:09.090
John A. Stankovic: And it in parallel views and lstm and we add to that a self attention architectural unit, so that it focuses on certain parts of the the inputs that are more important, as well as does the dynamics of the handwashing and the temporal relationships.

143
00:25:10.920 --> 00:25:23.070
John A. Stankovic: These numbers down here just say how many filters are in each of these three layers and so anyway, if you want more information about how this is really working, you can you can see our published paper.

144
00:25:24.600 --> 00:25:37.440
John A. Stankovic: We did our own data set where we gave people three practice friends to teach them how to use the WHO approach we had 14 participants and.

145
00:25:37.920 --> 00:25:57.960

John A. Stankovic: They each day 19 sessions over multiple weeks and we have video for ground truth and we compared I wash the name of our system and we compare it to these three state of the art solutions and we basically for the different 10 steps we.

146
00:25:59.190 --> 00:26:05.520
John A. Stankovic: are more accurate roughly by about 10% for most of the steps alone for some of them compared to certain Pearl.

147
00:26:06.840 --> 00:26:08.970
John A. Stankovic: Solutions were significantly better.

148
00:26:11.190 --> 00:26:20.220
John A. Stankovic: Alright, so the second thing is the dealing with mental health and i'm going to discuss our distance emotional recognition and anxiety.

149
00:26:21.240 --> 00:26:33.030
John A. Stankovic: For distance emotion, you know if if a person is living in a home that could be a microphone in the home, that could be if you go to a doctor's office, it could be microphone there so for this.

150
00:26:33.420 --> 00:26:40.260
John A. Stankovic: Ambient intelligence environment, people may not be right next to the microphone and there's been many solutions for detecting.

151
00:26:40.680 --> 00:26:54.900
John A. Stankovic: These kinds of emotions happy sad angry neutral when you're next to the microphone or fix distance, but when you're not needing the microphone it's a much more challenging problem so i'm going to describe our solution for that and also our solution for anxiety.

152
00:26:56.970 --> 00:26:57.210
John A. Stankovic: So.

153
00:26:58.710 --> 00:27:06.840
John A. Stankovic: As I said, if you're near the microphone or you're gonna fix this is from the microphone many solutions already exists or detecting emotion.

154
00:27:07.470 --> 00:27:17.460
John A. Stankovic: But if you're walking around and some ambient intelligent environment, then you might be a different distance you might be further away, there might be multiple people there might be.

155
00:27:18.120 --> 00:27:26.460
John A. Stankovic: Even more people at different distances some right next to each other, and so on, and we end up with like these four problems over here to sell.

156

00:27:27.090 --> 00:27:39.120
John A. Stankovic: there's a reverberation from the Environment there's ambient noise, a lot of it sometimes there's the amplification speech, which just means that has further away, you are from the microphone the less.

157
00:27:41.160 --> 00:27:41.940
John A. Stankovic: amplified.

158
00:27:43.050 --> 00:27:47.250
John A. Stankovic: detection is happening and then people could be speaking over each other and so on.

159
00:27:48.420 --> 00:27:50.550
John A. Stankovic: So our solution has three parts.

160
00:27:51.600 --> 00:27:59.730
John A. Stankovic: i'll go through each one, the first is we create what we call distance agnostic features and then we convert these code words.

161
00:28:00.750 --> 00:28:12.150
John A. Stankovic: And then, secondly we create a new feature modeling technique called mo to vector, which is a vector embedding For those of you know what that is, and then we just input that to a standard lstm.

162
00:28:12.930 --> 00:28:23.280
John A. Stankovic: So those are three parts these two we believe our novel this one is it's just standard okay So the first thing we did was we said if we're not any other microphone.

163
00:28:24.690 --> 00:28:32.790
John A. Stankovic: Can we identify features which are agnostic to distance and so we started with 231 low level features.

164
00:28:33.810 --> 00:28:40.080
John A. Stankovic: And these 231 which shows in from the literature, where they really work well with.

165
00:28:41.130 --> 00:28:51.600
John A. Stankovic: When you're near the microphone, and so this is the kind of features, there are, and when you take these and it's delta and delta delta you end up with 231 low level.

166
00:28:52.020 --> 00:29:06.600
John A. Stankovic: descriptors and we did experiments, where we had people speaking, and we had seven or eight microphones all at different distances and we can tell like that it turned out that these 48 we found 48 will level descriptors features that were.

167
00:29:07.680 --> 00:29:16.080
John A. Stankovic: Somewhat or very agnostic to distance, so we use these features and not these features so that's the

first step.

168
00:29:16.830 --> 00:29:34.950
John A. Stankovic: But when we get these features, we convert them to what's called audio code worth it, let me explain that and so let's say I have this signal here acoustic signal shown above and i'm going to typically look at a portion of the signal like that's what this.

169
00:29:36.450 --> 00:29:37.200
John A. Stankovic: Little.

170
00:29:38.490 --> 00:29:46.710
John A. Stankovic: bars down here mean and usually they have a sliding window on them and then, when you look in here you look at the signal you try to figure out what the features are.

171
00:29:47.160 --> 00:29:54.540
John A. Stankovic: And so, normally in our solution we use 48 features, so we figure out, what are the 48 features for this part of the single.

172
00:29:55.290 --> 00:30:05.670
John A. Stankovic: However, I can't draw a 48 dimensional space so imagine that there's only X and y here and there's two features so, then I can draw it so.

173
00:30:06.300 --> 00:30:21.570
John A. Stankovic: What I might end up with is that I process this segment maybe i'm in training and I process the segment and for every one of these, I get a point two dimensional point so and I do this over and over again, and I end up with.

174
00:30:23.940 --> 00:30:32.790
John A. Stankovic: clusters of points in different places and then I would choose the central eight of those as an o'neill word.

175
00:30:33.810 --> 00:30:45.150
John A. Stankovic: And so, all of these clusters are here and the centroid of each one is is called a audio word and then, when it's being used for apple.

176
00:30:46.500 --> 00:30:54.240
John A. Stankovic: A new point might come in, I have an audio signal and I determine its in this case X, Y features and it's.

177
00:30:54.780 --> 00:31:09.120
John A. Stankovic: In this part of the space we map it to the closest audio work, so we don't have to have infinite numbers of points or infinite numbers of features, we end up with some fixed numbers, which is the word number of words in your vocabulary.

178
00:31:11.940 --> 00:31:13.650
John A. Stankovic: And another another word is.

179
00:31:14.790 --> 00:31:29.670
John A. Stankovic: detected another feature set as expected, we map it and they may be the same so that word and that word are so close to each other, we just seen code them as a single entity and don't treat them as slightly different from each other.

180
00:31:30.810 --> 00:31:34.470
John A. Stankovic: And we tested anything like from 500 to 2500.

181
00:31:35.760 --> 00:31:58.020
John A. Stankovic: Using K means clustering you have to choose how many code words you're going to create and one interesting result was that different emotions, for example, I think, sad required 250 code code words and something like happy Hello you require to thousand code words to get good results.

182
00:31:59.520 --> 00:32:06.750
John A. Stankovic: Now, another thing that's the second step is is creating a vector representation of the code words.

183
00:32:07.260 --> 00:32:12.870
John A. Stankovic: Because the code, whereas only represent a little slice right they don't represent how slices are related to other slices.

184
00:32:13.530 --> 00:32:29.220
John A. Stankovic: And so what we've done is we, in this example let's assume here's a whole bunch of training data and there's a code word we label a one B and C D and all of these are code words that appear in happy.

185
00:32:30.570 --> 00:32:52.440
John A. Stankovic: signals and if I look at a I want to look at what code where it's around day, and here I might have P Q R s T, for example, and I look at B and b has a P Q R s T also surrounding it so A and B are very much related to each other in terms of happiness and so we would basically.

186
00:32:53.940 --> 00:33:01.200
John A. Stankovic: create a vector representation of A and B that are similar to each other, because they were both representing something about having this.

187
00:33:02.340 --> 00:33:02.730
John A. Stankovic: Now.

188
00:33:03.810 --> 00:33:13.590
John A. Stankovic: Alternatively, if we have see a different code word it might also have P Q R s T relationships, but see is actually say not happy.

189
00:33:14.880 --> 00:33:28.200
John A. Stankovic: So, even though it has a similar surrounding context it's not as part of the happiness so vector representation would be not close to NP.

190
00:33:30.930 --> 00:33:38.670
John A. Stankovic: Alright, so that's more or less what I was saying, then we just simply take an audio signal, as shown up here, we would.

191
00:33:39.660 --> 00:33:48.390
John A. Stankovic: You know, look at in our case 48 features here, and then we would, these are the robust low level features and now we.

192
00:33:49.260 --> 00:34:04.260
John A. Stankovic: convert that to their vector embedding and then we put it into a cm and as we move in time through the voice sample we're going to the next stages in the lstm right, and then the output is whether this is a.

193
00:34:05.310 --> 00:34:06.510
John A. Stankovic: Happy or sad or.

194
00:34:08.730 --> 00:34:14.490
John A. Stankovic: neutral basis for anger, we evaluate this with different data sets.

195
00:34:15.660 --> 00:34:22.770
John A. Stankovic: We actually even create our own data set we work with people and at usc were 20 families were brought into.

196
00:34:23.250 --> 00:34:44.100
John A. Stankovic: A lab and they were eating dinner, and there were spontaneous discussions and we try to generate different emotions from them, based on questions that they were given to talk about so on and and turns out the solutions data sets and this data set give us basically similar kinds of results.

197
00:34:46.020 --> 00:34:56.790
John A. Stankovic: We also compare our work to for baselines and we show that the solution is basically like 16 hours said, better than the best baseline.

198
00:34:57.840 --> 00:35:08.100
John A. Stankovic: or details, you know I found here, but let me also mentioned that Just to give you 16% better is good, but you know what's the absolute.

199
00:35:09.060 --> 00:35:20.010
John A. Stankovic: kinds of accuracy So here we have both accuracy and recall and in the accuracy and recall for angry happy sad are in the 90% range, which is.

200
00:35:20.520 --> 00:35:31.110
John A. Stankovic: Quite quite good, especially given people are not near the microphone and they have all these realistic issues about what noise levels are there and Berber ation and so on.

201
00:35:33.930 --> 00:35:43.350
John A. Stankovic: One observation, we made it looking at the data with that in the past, people originally had trouble distinguishing between happy and angry.

202
00:35:43.980 --> 00:36:06.360
John A. Stankovic: And many times that this data and information was desert observed by I could datasets having people act out happy and sad and but both of those use energy based features and so both happy and angry speech tend to have very strong volume and so on, so then.

203
00:36:08.010 --> 00:36:16.110
John A. Stankovic: We got rid of those features, basically, and so we were able to distinguish between happy and a pretty much, much better, it also in the real world.

204
00:36:16.860 --> 00:36:34.410
John A. Stankovic: Laughter is some extra information that you can build your solution on that also helps discriminate not always I mean people are not always laughing when they're happy, but, but if they are then it helps identify a happy state.

205
00:36:36.180 --> 00:36:57.540
John A. Stankovic: We also looked at what if we you know didn't eliminate the distorted features and we just use the original 231 features, well, we end up with about a 10% in accuracy and recall, and so on for for angry happy and sad, all of them, so it is a helpful feature to do helpful approach.

206
00:36:59.460 --> 00:37:09.810
John A. Stankovic: Also, you know, the fact of distance we do lose accuracy, as people move away, but we with person moves from the MIC or to around six meters away.

207
00:37:10.410 --> 00:37:21.180
John A. Stankovic: dropping accuracy is only about 5% and we saw a state of the art that accuracy level is drops about 12%, so I think we've improved proves a situation there.

208
00:37:22.350 --> 00:37:22.530
Right.

209
00:37:24.030 --> 00:37:33.690
John A. Stankovic: The last point I want to bring up is looking at you know kind of mental disorder and one really significant problem is anxiety.

210
00:37:34.590 --> 00:37:47.400

John A. Stankovic: And 11% of Americans suffer from anxiety and, in general, most people lose fully supervised learning, which is really, really difficult for speech because it's hard to.

211
00:37:48.660 --> 00:37:59.550
John A. Stankovic: kind of label all the different parts of speech or what's representing anxiety and so we applied a solution based on multiple what's called multiple distance learning.

212
00:38:00.330 --> 00:38:13.920
John A. Stankovic: And it's a weekly supervised learning technique, we also created a new feature modeling you know kind of vector embedding but this case it's called and into deck it's different from previous example I gave for.

213
00:38:15.240 --> 00:38:34.020
John A. Stankovic: For kind of mood and then we have a new classifier where we combine this is a bi directional lstm with multiple assists learning and combining those two together make for about accuracy is 90% range and.

214
00:38:35.040 --> 00:38:40.260
John A. Stankovic: Overall, these are 17% better than using the state of the art baselines.

215
00:38:41.430 --> 00:38:58.050
John A. Stankovic: So let me explain a little bit about what multiple instance learning is and how we how we use it so let's assume that this top curve is a voice sample from somebody and it is a boy sample from someone suffering from mental disorders such as anxiety, however.

216
00:38:59.220 --> 00:39:05.760
John A. Stankovic: it's i'd say they talking for 10 seconds here, so I might be a little label the whole thing is anxiety right but.

217
00:39:06.540 --> 00:39:16.890
John A. Stankovic: Where is the anxiety and the speech is that here to here, maybe it's just here, and this, the only place in the speech, where it's it's the label, so if I if I think of it.

218
00:39:17.490 --> 00:39:27.150
John A. Stankovic: If I label this all as anxiety, then I fall asleep labeling this part is anxiety and so on, it really hurts the classifier a lot.

219
00:39:27.810 --> 00:39:35.040
John A. Stankovic: However, and other case, maybe that's where the anxiety disorder is or maybe it's here, and maybe it's all three of them.

220
00:39:35.760 --> 00:39:58.350
John A. Stankovic: Okay, but we don't know all we know is this whole episode was treated as a positive example of anxiety and then this bottom line here is that we label, this is not anxiety, so there is no representative example of the speech inside this signal that represents anxiety right.

221
00:39:59.460 --> 00:40:11.850
John A. Stankovic: So what we're going to do is very similar to the previous solution we're going to use code words again so let's say i'm training and I have this speech coming in, and I, and I know that.

222
00:40:13.080 --> 00:40:29.970
John A. Stankovic: This whole thing is a positive audio clip and I I don't know that this is false here yet and i'm basically like say looking at this part of the speech which happens to be true, positive and again we're sliding window, looking at the features and the signal.

223
00:40:31.500 --> 00:40:33.240
John A. Stankovic: Those features are being.

224
00:40:34.290 --> 00:40:41.790
John A. Stankovic: Created into code words just like here, so we segment into these little segments, we use code words for this segments.

225
00:40:42.150 --> 00:40:52.620
John A. Stankovic: And then we're going to take those card words and find relationships of the code words to each other by a vector and Bernie just like we did before, except in this case vector embedding.

226
00:40:53.820 --> 00:41:06.600
John A. Stankovic: The emo to VAC we built in the previous solution doesn't work here, and we had to come up with a different one which we call an end to that where the N n stands for a neural network, so we use a neural network to learn these embedding.

227
00:41:09.900 --> 00:41:10.560
John A. Stankovic: And so.

228
00:41:11.610 --> 00:41:19.020
John A. Stankovic: As an example, if we have of the Senate and to vet feature modeling is let's say we have these three.

229
00:41:20.220 --> 00:41:33.060
John A. Stankovic: positive examples of speech, that is, anxiety and in here we have these code words you know 21231 and here we have similar code words in here we have similar code words again then.

230
00:41:35.100 --> 00:41:49.890
John A. Stankovic: Each of these kind of words would would be trained to be positive examples of of anxiety because they're connected for other ones that also have that kind of characteristic.

231
00:41:54.030 --> 00:41:57.720
John A. Stankovic: And then taking this and putting it into a.

232

00:41:59.790 --> 00:42:07.320
John A. Stankovic: into a classifier people have done work with multiple instance learning and but they've used an SDM.

233
00:42:08.670 --> 00:42:21.030
John A. Stankovic: And a dnn neither of these really aren't ends and so these are more static like descriptions of something and and they use multiple instance learning.

234
00:42:21.480 --> 00:42:29.430
John A. Stankovic: And we feel like these fail to account for some of the temporal dynamics in a speech segment, so we applied it to an rnn.

235
00:42:29.970 --> 00:42:36.870
John A. Stankovic: And this is kind of a busy busy diagram but i'll walk through it a little bit to give an idea of the solution.

236
00:42:37.680 --> 00:42:46.560
John A. Stankovic: So on the left here is the audio signal, instead of I just draw it vertically now let's say it starts from the top, and it goes down in time.

237
00:42:47.160 --> 00:42:57.570
John A. Stankovic: And so we again have a segmentation we segment the signal we create look at all the features in a signal that creates this audio word we use the audio or.

238
00:42:58.290 --> 00:43:07.260
John A. Stankovic: To do the vector embedding and then we collect a bunch of these and put them into what's called a bag for the training and.

239
00:43:08.250 --> 00:43:17.640
John A. Stankovic: Then going through this you have multiple of these bags and then each of these bags is going to train is bi directional, this is a bi directional lstm.

240
00:43:18.270 --> 00:43:29.430
John A. Stankovic: So in training, you can predict what's going forward, and you can also predict what's going backwards and you basically can then train this network to detect anxiety.

241
00:43:30.180 --> 00:43:44.280
John A. Stankovic: And, in the end, each of these layers of the network are producing an instant score and after you look at a bunch of these signals you looking at whether the instance core is saying that.

242
00:43:46.140 --> 00:43:50.940
John A. Stankovic: This is anxiety or not I mean that's basically that the output is done to be anxiety or not.

243

00:43:51.540 --> 00:43:58.680
John A. Stankovic: And so, for the training you kind of do this and then you run through and learn all the weights and so on in the CRM.

244
00:43:59.160 --> 00:44:13.290
John A. Stankovic: And then we at runtime when you're using it, we do the same thing we still collect a set segment of these and then put it through the network and then it creates an instant score it decides whether or not this is anxiety.

245
00:44:15.660 --> 00:44:27.540
John A. Stankovic: Here here are the evaluation we used to have five participants now this anxiety was for college students giving talks okay so.

246
00:44:28.200 --> 00:44:47.130
John A. Stankovic: there's still a lot more to do with anxiety, I believe, because it's different for different populations, and so this was dying and there's another paper here that describes these these solutions we're not going to talk about the depression, but the solution also works for detecting depression.

247
00:44:48.780 --> 00:44:53.310
John A. Stankovic: and give you a little bit of sense something of the ideas and results.

248
00:44:54.690 --> 00:45:04.380
John A. Stankovic: And so, one thing is, you know how much data, do we we collect and listen to before we make a decision about whether this person is anxious or not.

249
00:45:04.920 --> 00:45:27.090
John A. Stankovic: And so, this chart shows the F1 score on the y axis and the segment window that we're looking at so, for example, if we only look at one second intervals of speech, we only get F1 score of 83% Okay, but if we look at about eight to 15 or 13 to 16.

250
00:45:28.260 --> 00:45:32.400
John A. Stankovic: segments of speech, we get close to 90 over 90%.

251
00:45:34.290 --> 00:45:45.840
John A. Stankovic: And this is in contrast to mood in mood you normally see very good performance when you're doing about five seconds of speech.

252
00:45:47.010 --> 00:45:54.780
John A. Stankovic: And it looks like to get attacked accuracy anxiety, more accurately, you need some more data basically all right.

253
00:45:56.280 --> 00:46:09.600
John A. Stankovic: This hair says, you know what it's kind of the best level of numbers of audio states that work best and we found for for anxiety and in this context with these these.

254
00:46:11.040 --> 00:46:16.920
John A. Stankovic: undergraduate students what we needed about 30 3500 audience states that was kind of the maximum.

255
00:46:18.720 --> 00:46:33.750
John A. Stankovic: We also studied whether the dnn to vet that I show here that's the one we developed for this, and where we get F1 scores and 90% we compared to what if we use the the mo to fact that I just presented in the previous.

256
00:46:35.160 --> 00:46:43.530
John A. Stankovic: Part of the talk or I vector, which is very well known and used in the literature, or we have other kinds of.

257
00:46:44.760 --> 00:46:54.510
John A. Stankovic: baselines and we see a basically about a 15% improvement by using the specialized you know embedding.

258
00:46:56.130 --> 00:47:05.310
John A. Stankovic: Similarly, another thing to compare against is that I said that this our solution here, the Boc with multiple distance learning.

259
00:47:06.180 --> 00:47:15.300
John A. Stankovic: We also compared to what if we use the dnn multiple insists learning or am I stands for multiple instance learning actually stm these are from the.

260
00:47:15.810 --> 00:47:28.020
John A. Stankovic: The state of the art or we even if we don't use weekly supervised, but we can somehow torturously label everything and we use supervised learning techniques.

261
00:47:29.670 --> 00:47:51.420
John A. Stankovic: dnn CNN and a combination of CNN and bi bi directional lstm Here we see even the very best solution turn is ours, with 90% but the second best was this the SEM when it's supervised, and we still get higher accuracy than.

262
00:47:52.770 --> 00:47:54.450
John A. Stankovic: A supervised learning technique.

263
00:47:55.920 --> 00:48:06.450
John A. Stankovic: So briefly in summary and i'm trying to say a smart and connected health is moving forward it's based on bear wearables as well as in situ systems.

264
00:48:06.870 --> 00:48:23.220
John A. Stankovic: I think we should be pushing towards you know, have the intelligence, where this cognitive assistant to be on your smartwatch and in the environment, and they would need to exchange information between each other,

and we need to incrementally improve this.

265
00:48:24.570 --> 00:48:41.730
John A. Stankovic: The cognition of these systems, the intelligence of the systems in that I do think still acoustic modalities are going to still be very, very prominent and I want to thank some of the students that did most of this work and.

266
00:48:43.200 --> 00:48:50.910
John A. Stankovic: Unfortunately, for me, most of them are gone now they're they're taking positions at other universities, for the most part and.

267
00:48:51.960 --> 00:48:53.880
John A. Stankovic: I certainly give them a lot of credit.

268
00:48:55.440 --> 00:48:57.120
John A. Stankovic: So i'm happy to take questions.

269
00:48:58.050 --> 00:49:05.190
Gurdip Singh: So thank you jack for the excellent presentation like to give virtual round of applause to you for that.

270
00:49:06.270 --> 00:49:18.810
Gurdip Singh: So we have some questions coming in, so I will go through them one, at times, so the first one is you know, would like to hear about.

271
00:49:19.290 --> 00:49:29.640
Gurdip Singh: applications of Ai for the hard of hearing, such as smarter hearing aids that perform better in the real world environments hearing it today of.

272
00:49:30.210 --> 00:49:41.130
Gurdip Singh: Far from accurately replicating the human a performance and it also says that you're hearing, as you know, daughter health impacts they've started living in more restricted.

273
00:49:42.240 --> 00:49:48.720
Gurdip Singh: spaces so things progress that can be made towards this would really go about putting that.

274
00:49:50.700 --> 00:50:01.410
John A. Stankovic: On the important thing is, I actually need hearing, so I really shouldn't be going to get one myself, and then I could look into it more I really don't know the technology and and hearing aids.

275
00:50:01.650 --> 00:50:12.930
John A. Stankovic: So i'm not sure I could ask answer that particularly certainly in a general sense I would expect that it should be possible that using Ai can.

276
00:50:14.250 --> 00:50:15.120
John A. Stankovic: improve.

277
00:50:16.200 --> 00:50:32.070
John A. Stankovic: The way that the hearing aid works in different contexts, like noise and background noise and multiple speakers and overlap speakers so technically, I believe that that could happen, I don't know what they're doing and I don't know the solutions there.

278
00:50:32.730 --> 00:50:33.360
Gurdip Singh: So, could the.

279
00:50:34.410 --> 00:50:40.350
Gurdip Singh: The techniques that you have helping sort of recognizing speech from others, and then replaying it, you know.

280
00:50:41.220 --> 00:50:42.060
John A. Stankovic: yeah yeah.

281
00:50:42.510 --> 00:50:45.360
John A. Stankovic: And and almost all the pipelines I didn't explain.

282
00:50:45.450 --> 00:50:54.120
John A. Stankovic: You know, when we really use this there's there's more to it than just the core part of saying if it's say happy or sad or neutral so.

283
00:50:54.780 --> 00:51:03.090
John A. Stankovic: The the systems that we have have to first monitoring the environment and ignore silence, first of all, so you have to detect silence.

284
00:51:03.630 --> 00:51:14.190
John A. Stankovic: And then ignore other sounds like the refrigerator and so so detecting those sounds are also necessary and then finding that oh wait, this is speech.

285
00:51:15.030 --> 00:51:26.460
John A. Stankovic: Okay, speech and then, if you have a targeted for a particular people like what we do in home solutions, and we have this a caregiver and patient outside as patient.

286
00:51:27.060 --> 00:51:39.660
John A. Stankovic: We only really are interested in their voice right if somebody else's visiting me say that's not them, so we have to do speaker speaker identification and then, once we detected, this is a.

287
00:51:40.080 --> 00:51:48.090

John A. Stankovic: This is a speech and it's a you know kind of on the right person, then we can apply the the solution for detecting mood.

288
00:51:49.380 --> 00:52:00.060
John A. Stankovic: And these other techniques signal processing techniques, you know I think are important, that would solve some of these issues about filtering out of you know, noise and so on.

289
00:52:00.720 --> 00:52:09.930
John A. Stankovic: And by the way, I don't think every every part of the solution doesn't need machine learning necessarily I mean there's a lot of strong signal processing solutions that could be used.

290
00:52:11.700 --> 00:52:12.120
Gurdip Singh: Thank you.

291
00:52:13.230 --> 00:52:14.460
Gurdip Singh: Next, one is from.

292
00:52:15.540 --> 00:52:21.300
Gurdip Singh: bonnie asks is is the code word for an emotion different from one person to.

293
00:52:24.270 --> 00:52:25.080
John A. Stankovic: not see.

294
00:52:29.880 --> 00:52:49.320
John A. Stankovic: meaning, it can be so if you train just on one person, you would get a certain say distribution and you would maybe come up with certain code words but what we're doing is trying to have enough training data that it's general cross many people.

295
00:52:51.120 --> 00:52:54.330
John A. Stankovic: And so the code words are more generic.

296
00:52:55.350 --> 00:53:08.910
John A. Stankovic: Now we haven't tried say it's a really good good observation actually we actually haven't tried say training just on one person and protecting that one person's mode.

297
00:53:12.870 --> 00:53:23.640
Gurdip Singh: Next is from grants Cleveland so he first to mention the great talk and then he says he's wondering about system level performance issues in general, and in particular about.

298
00:53:24.720 --> 00:53:38.790
Gurdip Singh: The first application hard voice based interaction, which has video involving mobile devices, how do you deal with performance energy consumption, issues of running neural networks on such platforms.

299
00:53:40.410 --> 00:53:54.480
John A. Stankovic: yeah that's a great question also and one one easy way to answer is we're going towards you know, like we're going toward we don't have all the answers, but you know, for example, when I was at Oxford that people there had this really nice result where.

300
00:53:55.500 --> 00:54:00.630
John A. Stankovic: You can build your classifier is your very complex solutions and then you put it into a.

301
00:54:01.140 --> 00:54:12.330
John A. Stankovic: compiler and you give the compiler the original network, as well as the requirements for the platform, so you can give it the requirements of the smartwatch you only have you know.

302
00:54:13.110 --> 00:54:28.620
John A. Stankovic: 64 megabytes of memory, you have this and it produces a version of the model that fits on that device, so you know those kinds of technologies which we have not deployed, yet that they could be helpful, I think, in and mapping it down.

303
00:54:29.040 --> 00:54:30.420
John A. Stankovic: But, for example.

304
00:54:32.340 --> 00:54:33.420
John A. Stankovic: We normally.

305
00:54:34.770 --> 00:54:39.600
John A. Stankovic: We don't usually use voice to extract text.

306
00:54:40.740 --> 00:54:41.340
John A. Stankovic: and

307
00:54:43.320 --> 00:54:53.580
John A. Stankovic: For mood like because of privacy issues like when we deploy it in real people's homes, they don't want us to record their text, and so we we using that just the kind of.

308
00:54:57.060 --> 00:55:05.430
John A. Stankovic: properties and features of this of your speech and not the actual words, but when we do work with first responders we do need the first words.

309
00:55:05.820 --> 00:55:15.000
John A. Stankovic: And so the first step is to send it to to Google translator or to some standard translator that they are really soft better.

310
00:55:15.420 --> 00:55:23.400

John A. Stankovic: because then we couldn't fit those things on a smartwatch, for example, something like that, and so we make use of tools offline when we can.

311
00:55:24.030 --> 00:55:37.170
John A. Stankovic: If the delay is too long, or there is a separation say you lose the Internet right, and then we have like a in the first responder project, we have a backup to use.

312
00:55:37.650 --> 00:55:47.760
John A. Stankovic: The voice translation on the watch but it's a much simpler kind of system and it's not nearly as accurate as as you would get something like from Google.

313
00:55:49.980 --> 00:55:50.280
Gurdip Singh: Thank you.

314
00:55:52.020 --> 00:55:55.260
Gurdip Singh: The next one is conflated you touched upon briefly.

315
00:55:55.500 --> 00:56:06.450
Gurdip Singh: Earlier so it's it's my friend, yes curious how to ensure privacy and how's the acceptance of this among people in general.

316
00:56:09.390 --> 00:56:09.840
John A. Stankovic: yeah.

317
00:56:11.220 --> 00:56:14.880
John A. Stankovic: I mean that's that's a pretty broad and important questions.

318
00:56:17.700 --> 00:56:24.300
John A. Stankovic: we'll take the easier, one is, you know how to get acceptance I think my experience has been.

319
00:56:25.170 --> 00:56:33.150
John A. Stankovic: In a lot of cases, people will not want any technology they just don't want it right and there's not much you can do for that.

320
00:56:34.110 --> 00:56:47.040
John A. Stankovic: However, I also found that if the quid pro quo is strong enough, then they say Oh, this is useful, so, for example, one time, I gave a lecture at an Assisted Living.

321
00:56:47.790 --> 00:56:54.360
John A. Stankovic: And I first asked you know how many would like us to put all our sensors in your home like nobody wanted them.

322

00:56:54.960 --> 00:57:11.430
John A. Stankovic: Right and I said, well, what if we keep you out of a nursing home longer and they all raised their hand, so they have to see some benefit to it and even in the project, we have now with alzheimer's patients, even though we're only have you know, three families so far.

323
00:57:14.250 --> 00:57:20.730
John A. Stankovic: One person seems like they wanted to do the study to help them with their stress and dealing with their alzheimer's patient.

324
00:57:21.360 --> 00:57:35.730
John A. Stankovic: But now that the study is underway, they change their mind something some sense saying oh it's too much work, you know i'm too busy to do the interventions and, and so they don't see I guess that that it's valuable yet enough.

325
00:57:37.590 --> 00:57:38.100
John A. Stankovic: So.

326
00:57:39.720 --> 00:57:57.540
John A. Stankovic: We, we think that, as the younger generation comes along and they are so much more attuned to the technology, it will be less and less of an issue, especially if you can show value to it right, I mean, in the end that's what has to happen, otherwise people won't want to use it.

327
00:57:59.490 --> 00:58:01.140
John A. Stankovic: Now privacy privacy.

328
00:58:04.770 --> 00:58:12.480
John A. Stankovic: You can do all the standard things like you know you can anything you send to the cloud or to a to a server you're going to encrypt and.

329
00:58:12.960 --> 00:58:27.420
John A. Stankovic: you're going to have passwords and you know, try to protect the data, you could you know allow people to opt out for so, for example, we do allow people to turn their systems off so when.

330
00:58:28.440 --> 00:58:37.410
John A. Stankovic: Like, especially the alzheimer's patients if they have visitors, you know they don't want to be recorded anymore right, so they can they can turn the system off, and it can turn it back on so.

331
00:58:37.860 --> 00:58:48.690
John A. Stankovic: So that seems that to help some people saying Okay, then i'll take the system or if we explain we're not we're not keeping your words, so they don't want us to record their cursing and so on, right.

332
00:58:49.860 --> 00:58:58.770
John A. Stankovic: But there okay if we record that they're really agitated but not what exactly they're saying, and so, so you do have to make some accommodations I believe.

333
00:59:00.090 --> 00:59:00.360
Gurdip Singh: yeah.

334
00:59:01.140 --> 00:59:01.650
John A. Stankovic: that's the.

335
00:59:04.980 --> 00:59:07.050
Gurdip Singh: Next, one is by an invalid.

336
00:59:09.330 --> 00:59:17.850
Gurdip Singh: Have you tried models trained with US Data on other countries data forecasting and if, yes, how they performed.

337
00:59:22.140 --> 00:59:40.350
John A. Stankovic: We have done that, to some extent, where some of the ETS are from other countries like there's one famous German data set for depression, and so we actually use that and then some of the depression studies, the results that we have.

338
00:59:41.610 --> 00:59:43.770
John A. Stankovic: seemed to work okay for.

339
00:59:45.300 --> 00:59:54.450
John A. Stankovic: Across the German data set in the US data set and and you might think that German is not very similar to English so.

340
00:59:55.050 --> 01:00:08.820
John A. Stankovic: It might be a problem, but but it turns out that you know if you can get the core features that pendant on the accent, then I then you can do a kind of a pretty good job across languages.

341
01:00:11.550 --> 01:00:12.300
Gurdip Singh: Thank you.

342
01:00:14.250 --> 01:00:24.180
Gurdip Singh: The next one is in your anxiety accuracy experiments what's your definition or the criteria of labeling a sample as anxiety.

343
01:00:25.620 --> 01:00:27.030
John A. Stankovic: So I didn't quite get it.

344
01:00:27.510 --> 01:00:39.180
Gurdip Singh: was just it's basically saying that you know when you are doing the experiments anxiety accuracy

experiments, so it says what's your definition of the criteria of labeling a sample.

345
01:00:40.290 --> 01:00:41.340
Gurdip Singh: As anxiety.

346
01:00:43.170 --> 01:00:44.220
Gurdip Singh: How do you recognize.

347
01:00:44.430 --> 01:00:44.850
yeah.

348
01:00:46.470 --> 01:00:56.850
John A. Stankovic: So the way this these experiments were done by psychologists and they were using students and they were defining anxiety as.

349
01:00:57.540 --> 01:01:09.060
John A. Stankovic: When students were required to give two or three minute talks and that they the students themselves labeled how anxious they felt and so on, and then they were labeled or.

350
01:01:09.810 --> 01:01:21.270
John A. Stankovic: Also, that label whether they thought that the student was being kind of anxious and and so that was more or less the way ground truth was was determined on it.

351
01:01:24.630 --> 01:01:34.200
Gurdip Singh: next question is by early notion third sets what are you able to accommodate low vision in developing Apps for a smartwatch.

352
01:01:37.440 --> 01:01:41.400
John A. Stankovic: Where we developing now absence my my with first part of the question, I know.

353
01:01:42.210 --> 01:01:46.260
Gurdip Singh: So, what are you able to accommodate low vision in developing Apps.

354
01:01:48.090 --> 01:01:51.450
Gurdip Singh: So, then, I kind of cool for the smartwatch.

355
01:01:54.060 --> 01:01:54.450
John A. Stankovic: Still.

356
01:01:54.510 --> 01:01:55.920
John A. Stankovic: Yes, what kind of vision.

357
01:01:56.400 --> 01:01:57.330
Gurdip Singh: Low vision.

358
01:01:58.830 --> 01:02:05.580
Gurdip Singh: So maybe less I think yeah it's it's somewhat maybe maybe the person who posted it Eleanor can.

359
01:02:05.940 --> 01:02:08.610
Gurdip Singh: clarify with the letter, but what they meant by because.

360
01:02:08.700 --> 01:02:10.830
Gurdip Singh: I don't see vision sensors being used yeah.

361
01:02:12.120 --> 01:02:15.000
John A. Stankovic: Like I ve is io n vision.

362
01:02:15.600 --> 01:02:16.950
John A. Stankovic: There is no vision on the smart.

363
01:02:17.190 --> 01:02:19.380
Gurdip Singh: TV on the smartwatch that's what I said.

364
01:02:19.830 --> 01:02:25.050
Gurdip Singh: yeah for the Ek so next question is for the ekg watch APP.

365
01:02:26.250 --> 01:02:38.250
Gurdip Singh: wondering if you got a chance to evaluate how much time a physician needs to spend to actually go through the recorded data in the cloud what's their feedback on the extra workload.

366
01:02:39.360 --> 01:02:48.090
John A. Stankovic: yeah so so that's really a very good question but we haven't gotten that far, so we don't we don't have a good answer for that.

367
01:02:49.800 --> 01:02:50.430
John A. Stankovic: You know the.

368
01:02:51.570 --> 01:02:59.340
John A. Stankovic: Developing the technology is is a difficult process and we're we've we've been we've done that, but we haven't had enough.

369
01:03:00.480 --> 01:03:16.230
John A. Stankovic: kind of cases and data up there, we do work with you know physicians who will look at that and

maybe give us some feedback but, but I think that in the end what's needed and talking to them about this is that they only should be.

370
01:03:18.000 --> 01:03:20.250
John A. Stankovic: kind of notified when something looks wrong.

371
01:03:21.540 --> 01:03:31.410
John A. Stankovic: You know, and so they they won't have time to sit down and look at all these different patients every day going through all the data and trying to find something.

372
01:03:31.860 --> 01:03:40.260
John A. Stankovic: That that's not possible, but all the data and everything would be there and then, if something is detected, even even the apple.

373
01:03:41.340 --> 01:03:55.650
John A. Stankovic: ekg by itself has a alarm system, and so on, right, so if if it points out that you, like you, better see your doctor or something like that, then we can easily just highlight that to the doctor and say you don't need to look at this.

374
01:03:58.560 --> 01:04:04.440
John A. Stankovic: But I don't have any any data on the line, so the question, which is a very important question.

375
01:04:08.700 --> 01:04:09.450
Gurdip Singh: Thank you.

376
01:04:12.870 --> 01:04:23.640
Gurdip Singh: So the next one is if we have an application, where we do not have enough data to train and ei algorithm How would you how would you approach to that problem.

377
01:04:25.560 --> 01:04:27.360
Gurdip Singh: So, how does it impact the accuracy.

378
01:04:29.850 --> 01:04:47.370
John A. Stankovic: yeah so I mean the standard techniques and from the literature is, if you use data augmentation or you use transfer learning, you know Those are two standard techniques and we have been doing a bunch of work on transfer learning and so, for example, if we were going to.

379
01:04:48.480 --> 01:04:50.880
John A. Stankovic: Say look at anxiety and elderly.

380
01:04:52.620 --> 01:04:53.700
John A. Stankovic: say even dementia.

381
01:04:53.700 --> 01:05:05.760
John A. Stankovic: Patients that living at home, then we could we learn from the anxiety data that we've already collected on college students and so there's a transfer learning, we would be able to say let's build upon that.

382
01:05:06.480 --> 01:05:21.630
John A. Stankovic: Information and then have just a little bit of training data and then try to have a new classifier that is effective for for elderly and there's it's a very active research area, to identify and develop new transfer learning techniques.

383
01:05:23.010 --> 01:05:25.380
John A. Stankovic: The data augmentation thing is more like.

384
01:05:28.620 --> 01:05:45.360
John A. Stankovic: You know, you can modify your sample So if you have a data set with 100 samples in it and you need 500 samples, then you randomly adjusting the different samples in some reasonable way so to make it look like you have more data and sometimes that that works.

385
01:05:49.560 --> 01:05:50.220
Gurdip Singh: So.

386
01:05:52.260 --> 01:06:05.190
Gurdip Singh: The next question is so from tomorrow, it said that at the beginning of your presentation, you mentioned about addressing uncertainties in animal models, would you please elaborate on that.

387
01:06:06.780 --> 01:06:07.710
John A. Stankovic: yeah sure, thank you.

388
01:06:09.330 --> 01:06:14.370
John A. Stankovic: So it's a lot of machine learning models at the there.

389
01:06:15.720 --> 01:06:34.050
John A. Stankovic: At the end I use soft Max which basically gives you a probability right so let's say you're looking at a picture and you want to know if it's a dog or a cat or a mouse or a lion and it goes through a CNN and the end, instead of just saying dog.

390
01:06:35.070 --> 01:06:47.160
John A. Stankovic: It says dog with 70% probability and cat with 20% probability things like that, so that that's one level of uncertainty, you might be able to use soft Max, however.

391
01:06:48.360 --> 01:06:56.880
John A. Stankovic: When we don't think that's really sufficient and what we do some of our smart cities where we're looking at a very, very complex situation.

392
01:06:57.270 --> 01:07:06.330
John A. Stankovic: And we're trying to predict the state of the city into the future, so, given the past state of the city and our current sensor readings, what is the city going to look like in the future.

393
01:07:06.960 --> 01:07:14.160
John A. Stankovic: And so, obviously we can't be precise, right now going to be precise, into the future, when there's so many uncertainties and changes.

394
01:07:14.580 --> 01:07:25.050
John A. Stankovic: And so what we end up doing is running the future in multiple multiple times like even hundreds of times, and each time we run the future we change something in the model.

395
01:07:25.530 --> 01:07:33.360
John A. Stankovic: So that we get essentially like it's a simulation right now we're simulating many different possible future paths and those.

396
01:07:34.140 --> 01:07:44.610
John A. Stankovic: The Cone if you can draw a Cone around this that's called a flow pipe and so you can have a flow pipe of possible outcomes into the future, with a certain confidence level.

397
01:07:45.660 --> 01:07:50.520
John A. Stankovic: And and that's that's what we have been doing for smart cities work.

398
01:07:53.790 --> 01:07:54.510
Gurdip Singh: Thank you.

399
01:07:56.160 --> 01:08:07.620
Gurdip Singh: next question is that curious if you have looked at how to detect emotions in the wise when Boston is speaking a language that has got multiple tones.

400
01:08:08.070 --> 01:08:20.220
Gurdip Singh: which indicate the meaning of the sound versus the emotions so, for example, Mandarin it says as for different bones of the things and based on the tone, the emotion like might be different.

401
01:08:23.130 --> 01:08:29.880
John A. Stankovic: Now we just have done that, with me we focus on English, so we haven't really looked at those issues.

402
01:08:31.200 --> 01:08:35.910
Gurdip Singh: So there's also mean like if a person if if i'm you know.

403
01:08:36.960 --> 01:08:40.680

Gurdip Singh: If I have let's say a cough or a cold so dusty.

404
01:08:42.750 --> 01:08:43.380
Gurdip Singh: dusty our.

405
01:08:43.440 --> 01:08:47.700
Gurdip Singh: models are sort of accommodate for any kind of perturbations that may occur because of.

406
01:08:49.140 --> 01:08:50.880
Gurdip Singh: These these aspects of.

407
01:08:53.190 --> 01:08:54.180
John A. Stankovic: Well, you know.

408
01:08:55.350 --> 01:08:59.220
John A. Stankovic: At the moment, go, I would say you don't take that into account, but.

409
01:09:00.900 --> 01:09:13.800
John A. Stankovic: I think this comes back to the notion of you know how robust is your model, you know, and so, if a person has a cold and then he sound very different and and the model was built on a.

410
01:09:14.640 --> 01:09:27.420
John A. Stankovic: kind of very sensitive brittle form, then that slight change is going to be not detected very well and it won't work, however, you few power trying to build a robust models, then.

411
01:09:28.050 --> 01:09:38.100
John A. Stankovic: Hopefully that even though wasn't trained on that it would still do accurate assessment and so that that's what a lot of people are working on is how to build these robots models.

412
01:09:39.690 --> 01:09:40.110
Gurdip Singh: and

413
01:09:42.900 --> 01:09:52.830
Gurdip Singh: So yeah so I guess the question that was there, earlier on novation say it was English related to the fact that someone with the you know.

414
01:09:53.550 --> 01:10:07.170
Gurdip Singh: which may have a may not be able to see as as good might not be able to see the icons on the watch or read the small print you know they may have, because there are so many of them on small space.

415
01:10:09.360 --> 01:10:23.430
John A. Stankovic: Well that's one of the motivations for using voice, and you know, in the DEMO you saw that the

person is still had a top a few things but but that that's gone away completely so so that he can just talk and that that's.

416
01:10:24.870 --> 01:10:25.500
John A. Stankovic: But if.

417
01:10:27.390 --> 01:10:37.710
John A. Stankovic: But there are things like, for example, if if you detect somebody is stressed and you ask that you provide some interventions and an intervention could be to.

418
01:10:38.340 --> 01:10:52.680
John A. Stankovic: You know, do a mindfulness activity like something if you're familiar with calm as as an APP and if if calm is on your watch, then you know it's really small you can't see it so wouldn't be effective for a person with bad mission.

419
01:10:53.790 --> 01:11:03.240
John A. Stankovic: Even if it was on a smart phone probably still too small for them, and so that intervention wouldn't make sense for a person with with poor vision.

420
01:11:07.110 --> 01:11:14.790
Gurdip Singh: Here anymore, I see that you have, at the end of the q&a list anymore last call for any questions.

421
01:11:17.280 --> 01:11:20.160
Gurdip Singh: If not, then I would again like to thank.

422
01:11:21.300 --> 01:11:22.620
Gurdip Singh: jack for the excellent.

423
01:11:22.620 --> 01:11:24.870
Gurdip Singh: presentation and for the.

424
01:11:25.350 --> 01:11:34.950
Gurdip Singh: rate Q amp a at the end, and thank you and also, I want to remind the audience once again that we will have office hours today at.

425
01:11:35.160 --> 01:11:36.570
Gurdip Singh: 3pm and you are.

426
01:11:36.930 --> 01:11:40.200
Gurdip Singh: very welcome to join that so thanks once more jack.

427
01:11:40.890 --> 01:11:42.600
John A. Stankovic: Thank you see you two or three.

428
01:11:43.650 --> 01:11:45.030
John A. Stankovic: All right, bye.