1
00:00:05,510 --> 00:00:03,350
good morning so there's a lot of words

2
00:00:06,950 --> 00:00:05,520
in that title so I'm just gonna dive in

3
00:00:11,030 --> 00:00:06,960
and we're gonna start talking about

4
00:00:13,370 --> 00:00:11,040
machine learning so machine learning is

5
00:00:16,099 --> 00:00:13,380
a area of computer science and data

6
00:00:18,650 --> 00:00:16,109
science that uses specific algorithms

7
00:00:21,019 --> 00:00:18,660
that learn from data and allow computers

8
00:00:22,490 --> 00:00:21,029
to find connections and data without the

9
00:00:24,679 --> 00:00:22,500
need for them to be explosively

10
00:00:27,019 --> 00:00:24,689
programmed this is a very fundamental

11
00:00:28,970 --> 00:00:27,029
shift in how the computers were

12
00:00:32,900 --> 00:00:28,980
programmed previously and how we think

13
00:00:35,479 --> 00:00:32,910

about data clearly I don't have enough

14
00:00:37,670 --> 00:00:35,489
time to go thing to a clash crash course

15
00:00:39,350 --> 00:00:37,680
on machine learning but basically a

16
00:00:41,810 --> 00:00:39,360
typical machine learning neural network

17
00:00:44,360 --> 00:00:41,820
would look like a series of inputs some

18
00:00:46,010 --> 00:00:44,370
computational nodes called the hidden

19
00:00:48,080 --> 00:00:46,020
layer and then those the results of

20
00:00:50,600 --> 00:00:48,090
those nodes are mapped some to some

21
00:00:53,119 --> 00:00:50,610
outputs when we talk about deep machine

22
00:00:56,180 --> 00:00:53,129
learning it's the same sort of inputs

23
00:00:58,580 --> 00:00:56,190
and outputs but the computational nodes

24
00:01:00,560 --> 00:00:58,590
are there's a lot more of them and

25
00:01:02,810 --> 00:01:00,570
they're highly they're there

26
00:01:07,340 --> 00:01:02,820
interconnection is highly complex and

27
00:01:09,740 --> 00:01:07,350
layered machine learning isn't anything

28
00:01:11,270 --> 00:01:09,750
new and there's lots of applications for

29
00:01:12,980 --> 00:01:11,280
it that are being used right now if

30
00:01:14,570 --> 00:01:12,990
you've used your cell phone this morning

31
00:01:17,020 --> 00:01:14,580
you've probably interacted with a

32
00:01:20,030 --> 00:01:17,030
machine learning system at some point

33
00:01:21,740 --> 00:01:20,040
however there's been a big advance in

34
00:01:24,469 --> 00:01:21,750
the technology in the last few years

35
00:01:27,440 --> 00:01:24,479
where major companies like Google IBM

36
00:01:30,440 --> 00:01:27,450
Microsoft amazon.com and even Facebook

37
00:01:34,010 --> 00:01:30,450
have been investing large amounts of

38
00:01:35,630 --> 00:01:34,020

resources into the technology and what

39
00:01:38,210 --> 00:01:35,640
that's resulted in is a number of new

40
00:01:40,760 --> 00:01:38,220
systems that are either low cost or open

41
00:01:43,609 --> 00:01:40,770
source that make the technology very

42
00:01:46,340 --> 00:01:43,619
very available where once it used to be

43
00:01:48,770 --> 00:01:46,350
sort of a very esoteric and specialized

44
00:01:51,020 --> 00:01:48,780
skill we now have a lot of really

45
00:01:54,460 --> 00:01:51,030
powerful tools that are at our disposal

46
00:01:56,980 --> 00:01:54,470
in addition we now have really robust

47
00:02:00,170 --> 00:01:56,990
libraries that can be accessed through

48
00:02:04,730 --> 00:02:00,180
generally accessible and well-known

49
00:02:06,230 --> 00:02:04,740
programming tools so let's look at an

50
00:02:08,690 --> 00:02:06,240
example of how we might use machine

51
00:02:10,160 --> 00:02:08,700
learning inside research so I've been

52
00:02:12,260 --> 00:02:10,170
working on a project that visualizes

53
00:02:13,110 --> 00:02:12,270
field reg data and when we talk about

54
00:02:14,820 --> 00:02:13,120
field right

55
00:02:17,130 --> 00:02:14,830
typically we start with a group of

56
00:02:20,760 --> 00:02:17,140
people who are engaged in some sort of

57
00:02:22,410 --> 00:02:20,770
common event or task we set up a random

58
00:02:24,030 --> 00:02:22,420
event generator in the background we

59
00:02:26,400 --> 00:02:24,040
collect some data and then we analyze

60
00:02:29,280 --> 00:02:26,410
that data look for deviations of

61
00:02:31,589 --> 00:02:29,290
randomness in the project on doing we

62
00:02:33,140 --> 00:02:31,599
start with a group of people but we ask

63
00:02:37,740 --> 00:02:33,150

them to focus on a very specific

64
00:02:39,750 --> 00:02:37,750
intention or idea within run a read we

65
00:02:42,630 --> 00:02:39,760
collect the data but instead of looking

66
00:02:46,559 --> 00:02:42,640
for statistical differences we model the

67
00:02:49,229 --> 00:02:46,569
data in 3d now and visualize it there

68
00:02:51,630 --> 00:02:49,239
have been other projects that have

69
00:02:53,160 --> 00:02:51,640
visualized 3d are reg data in the past

70
00:02:55,259 --> 00:02:53,170
so this is just sort of my take on it

71
00:02:56,759 --> 00:02:55,269
and when we do this process what we end

72
00:03:00,509 --> 00:02:56,769
up is image it with images that look

73
00:03:03,930 --> 00:03:00,519
like this this is image was created from

74
00:03:06,390 --> 00:03:03,940
data that was collected during Julie by

75
00:03:08,910 --> 00:03:06,400
shells talk at the Vale symposium at in

76
00:03:10,770 --> 00:03:08,920
Vail Colorado where participants were

77
00:03:14,280 --> 00:03:10,780
asked to focus on the idea of their

78
00:03:16,199 --> 00:03:14,290
community this was one where a series of

79
00:03:19,710 --> 00:03:16,209
meditators were focused on a breath

80
00:03:22,170 --> 00:03:19,720
meditation this data was Jim this image

81
00:03:24,030 --> 00:03:22,180
was dinner generated from data excuse me

82
00:03:26,280 --> 00:03:24,040
when an individual was asked to think

83
00:03:30,840 --> 00:03:26,290
about his depression and social anxiety

84
00:03:35,370 --> 00:03:30,850
issues and the intention here was of

85
00:03:37,229 --> 00:03:35,380
abundance so if you look at these images

86
00:03:38,880 --> 00:03:37,239
they're all very colorful and they're

87
00:03:41,280 --> 00:03:38,890
all very pretty and there's a lot of

88
00:03:42,720 --> 00:03:41,290

overlap between the content but there

89
00:03:46,140 --> 00:03:42,730
are some interesting and striking

90
00:03:47,940 --> 00:03:46,150
features in them and when you when you

91
00:03:50,610 --> 00:03:47,950
ask the people that participated in

92
00:03:52,979 --> 00:03:50,620
their creation they the people the

93
00:03:55,259 --> 00:03:52,989
participants will often report that they

94
00:03:57,900 --> 00:03:55,269
feel as if their intention is somehow

95
00:04:00,840 --> 00:03:57,910
reflected in the image now clearly that

96
00:04:02,309 --> 00:04:00,850
is a very subjective process and I may

97
00:04:05,819 --> 00:04:02,319
have done nothing more than create very

98
00:04:07,379 --> 00:04:05,829
very pretty Rorschach pictures but it's

99
00:04:10,199 --> 00:04:07,389
an interesting idea to sort of consider

100
00:04:12,500 --> 00:04:10,209
so to play with that a little bit I ran

101
00:04:15,210 --> 00:04:12,510
two sessions in which we did five

102
00:04:17,490 --> 00:04:15,220
sessions in which the intention was love

103
00:04:19,680 --> 00:04:17,500
and ten five sessions in which the

104
00:04:22,260 --> 00:04:19,690
intention was hate and these are the

105
00:04:23,850 --> 00:04:22,270
resulting images now again a lot of

106
00:04:25,800 --> 00:04:23,860
overlap a lot of similarity between

107
00:04:26,240 --> 00:04:25,810
these images but there's also something

108
00:04:27,850 --> 00:04:26,250
kind

109
00:04:31,940 --> 00:04:27,860
different between these images as well

110
00:04:33,610 --> 00:04:31,950
so how do we quantifiably define the

111
00:04:35,660 --> 00:04:33,620
differences between these two datasets

112
00:04:38,090 --> 00:04:35,670
well one way would be to use a

113
00:04:41,120 --> 00:04:38,100

traditional research method we recruit

114
00:04:44,600 --> 00:04:41,130
participants we collect all we analyze

115
00:04:46,280 --> 00:04:44,610
all the data we do all this stuff but we

116
00:04:48,290 --> 00:04:46,290
and you know we have to develop some

117
00:04:50,000 --> 00:04:48,300
sort of test or sorting or scoring tasks

118
00:04:51,980 --> 00:04:50,010
that may lead us to that difference

119
00:04:54,830 --> 00:04:51,990
eventually however for this particular

120
00:04:59,380 --> 00:04:54,840
project it's very speculative and this

121
00:05:03,920 --> 00:04:59,390
is a kind of a pretty resource-intensive

122
00:05:06,440 --> 00:05:03,930
process so instead I wanted to sort of

123
00:05:09,200 --> 00:05:06,450
remove the human element bias from it

124
00:05:11,240 --> 00:05:09,210
and started looking at deep machine

125
00:05:13,220 --> 00:05:11,250
image learning systems or classifiers

126
00:05:16,010 --> 00:05:13,230
and the one I came up with is a

127
00:05:17,690 --> 00:05:16,020
commercial system called classify which

128
00:05:21,140 --> 00:05:17,700
is available to anyone you can use it

129
00:05:23,270 --> 00:05:21,150
right now if you want it has an open API

130
00:05:24,380 --> 00:05:23,280
that you can use it's free or depending

131
00:05:26,210 --> 00:05:24,390
on how much data you want to use there's

132
00:05:28,400 --> 00:05:26,220
a small cost to using it and it's

133
00:05:31,280 --> 00:05:28,410
specifically designed for image concept

134
00:05:32,990 --> 00:05:31,290
and feature tagging so what does that

135
00:05:34,790 --> 00:05:33,000
mean that means if you get this picture

136
00:05:36,710 --> 00:05:34,800
of a bird that I took you run it through

137
00:05:39,200 --> 00:05:36,720
the classifier it returns a series of

138
00:05:41,300 --> 00:05:39,210

tags that identify it including their

139
00:05:45,260 --> 00:05:41,310
like bird and know person and nature and

140
00:05:49,150 --> 00:05:45,270
tree etc so what happens when we run our

141
00:05:52,310 --> 00:05:49,160
ten images through the system well

142
00:05:53,960 --> 00:05:52,320
excuse me not surprisingly you see a lot

143
00:05:56,090 --> 00:05:53,970
of overlap between the tags that are

144
00:05:57,380 --> 00:05:56,100
returned for the two datasets just like

145
00:05:59,540 --> 00:05:57,390
we would expect because there's a lot of

146
00:06:01,730 --> 00:05:59,550
overlap in these images but the really

147
00:06:05,060 --> 00:06:01,740
weird and interesting bit is that the

148
00:06:08,300 --> 00:06:05,070
classifier tagged four of the five hate

149
00:06:11,450 --> 00:06:08,310
images with a unique identifier and that

150
00:06:14,150 --> 00:06:11,460
identifier is triangular that's a very

151
00:06:16,909 --> 00:06:14,160
very specific feature and I can actually

152
00:06:19,490 --> 00:06:16,919
use that now to track the triangulation

153
00:06:21,740 --> 00:06:19,500
triangular feature back into the

154
00:06:24,350 --> 00:06:21,750
visualization software and potentially

155
00:06:26,060 --> 00:06:24,360
back into the data set so what I've been

156
00:06:27,770 --> 00:06:26,070
able to do here what the classifiers

157
00:06:29,990 --> 00:06:27,780
actually able to do here is actually

158
00:06:33,500 --> 00:06:30,000
start giving me the groundwork for a

159
00:06:35,330 --> 00:06:33,510
testable hypothesis so what is what is

160
00:06:37,159 --> 00:06:35,340
all what have we learned from this the

161
00:06:39,800 --> 00:06:37,169
software made this test possible in just

162
00:06:40,360 --> 00:06:39,810
a couple of hours there was considerable

163
00:06:42,520 --> 00:06:40,370

time

164
00:06:44,230 --> 00:06:42,530
and savings as compared to different

165
00:06:45,850 --> 00:06:44,240
traditional approaches and the

166
00:06:47,560 --> 00:06:45,860
exploratory tests demonstrated the

167
00:06:49,659 --> 00:06:47,570
potential value of machine learning and

168
00:06:53,500 --> 00:06:49,669
efficiency and potential hypothesis

169
00:06:54,879 --> 00:06:53,510
generation so machine learning has a lot

170
00:06:56,860 --> 00:06:54,889
of potential but it's not without its

171
00:06:58,960 --> 00:06:56,870
problems and one of the sort of epic

172
00:07:01,930 --> 00:06:58,970
fails in machine learning has been the

173
00:07:03,670 --> 00:07:01,940
Google Flu Trends project where Google

174
00:07:08,409 --> 00:07:03,680
decided that it could take a bunch of

175
00:07:10,480 --> 00:07:08,419
its search data and convert it and use

176
00:07:12,400 --> 00:07:10,490
it to predict flu outbreaks around the

177
00:07:14,350 --> 00:07:12,410
world and as it turns out it was very

178
00:07:16,510 --> 00:07:14,360
very bad at it

179
00:07:19,450 --> 00:07:16,520
and the Google silently sort of killed

180
00:07:21,310 --> 00:07:19,460
the project after a while another sort

181
00:07:24,460 --> 00:07:21,320
of recent more recent fail was

182
00:07:26,830 --> 00:07:24,470
Microsoft's Twitter bot which used

183
00:07:29,500 --> 00:07:26,840
natural language processing to teach it

184
00:07:31,689 --> 00:07:29,510
to teach to teach the software to talk

185
00:07:35,469 --> 00:07:31,699
like a teenager there's a lot of tease

186
00:07:37,930 --> 00:07:35,479
mid-sentence on Twitter and after being

187
00:07:41,770 --> 00:07:37,940
online for about a day it turned into a

188
00:07:50,170 --> 00:07:41,780

racist jerk and they had to shut it down

189
00:07:51,790 --> 00:07:50,180
and apologized and fortunately they the

190
00:07:56,589 --> 00:07:51,800
system did come back online for another

191
00:07:58,060 --> 00:07:56,599
day or so and before it finally they boy

192
00:07:59,920 --> 00:07:58,070
they probably took it offline completely

193
00:08:03,909 --> 00:07:59,930
it sent out one last very prophetic

194
00:08:05,980 --> 00:08:03,919
tweet which was you are too fast please

195
00:08:08,140 --> 00:08:05,990
take a rest which i think is a really

196
00:08:12,430 --> 00:08:08,150
important message for from our future

197
00:08:14,200 --> 00:08:12,440
robot overlords all that being said

198
00:08:17,170 --> 00:08:14,210
there's a there's been some really

199
00:08:20,740 --> 00:08:17,180
interesting applications for machine

200
00:08:22,510 --> 00:08:20,750
learning in the sciences and this one

201
00:08:24,939 --> 00:08:22,520
has made the rounds in the news it's a

202
00:08:27,850 --> 00:08:24,949
controversial fine but I think it's

203
00:08:29,920 --> 00:08:27,860
worth noting that in just three days a

204
00:08:31,540 --> 00:08:29,930
machine learning system was able to chug

205
00:08:34,060 --> 00:08:31,550
through all the published data on a

206
00:08:36,240 --> 00:08:34,070
particular problem and find the solution

207
00:08:38,500 --> 00:08:36,250
to a hundred year old biology problem

208
00:08:41,199 --> 00:08:38,510
and I think this quote is pretty

209
00:08:43,659 --> 00:08:41,209
interesting but this problem in our

210
00:08:45,370 --> 00:08:43,669
approach is nearly universal it can be

211
00:08:48,220 --> 00:08:45,380
used with anything where functional data

212
00:08:50,500 --> 00:08:48,230
exists and the underlying mechanisms are

213
00:08:52,360 --> 00:08:50,510

hard to guess so when I hear statements

214
00:08:53,690 --> 00:08:52,370
like that I really begin to think about

215
00:08:55,940 --> 00:08:53,700
how we might be a

216
00:08:58,660 --> 00:08:55,950
to apply these kinds of technologies and

217
00:09:01,190 --> 00:08:58,670
these this new set of toolkits to

218
00:09:03,290 --> 00:09:01,200
parapsychology and of course we have all

219
00:09:05,870 --> 00:09:03,300
these interesting sets of data available

220
00:09:09,530 --> 00:09:05,880
to us that may be right for exploration

221
00:09:14,330 --> 00:09:11,930
so in summary researchers now have

222
00:09:17,660 --> 00:09:14,340
access to new powerful low-cost or free

223
00:09:19,910 --> 00:09:17,670
machine learning tools exploratory tests

224
00:09:21,530 --> 00:09:19,920
using visual field req data showed the

225
00:09:23,380 --> 00:09:21,540
potential benefits offered by machine

226
00:09:26,120 --> 00:09:23,390
learning in terms of efficiency and

227
00:09:28,010 --> 00:09:26,130
hypothesis development and all that we

228
00:09:29,330 --> 00:09:28,020
need to be cautious parapsychology may

229
00:09:33,530 --> 00:09:29,340
benefit from machine learning it

230
00:09:35,720 --> 00:09:33,540
analyses of existing side datasets I

231
00:09:37,820 --> 00:09:35,730
want to thank my research partner dr.

232
00:09:39,410 --> 00:09:37,830
Julie by shell all the people to win

233
00:09:41,690 --> 00:09:39,420
bridge Institute that support this work

234
00:09:43,730 --> 00:09:41,700
the volunteers that took place in the

235
00:09:45,560 --> 00:09:43,740
sessions of course you for attending a

236
00:09:48,440 --> 00:09:45,570
quick shout-out to Julia Moss bridge who

237
00:09:50,360 --> 00:09:48,450
dragged me Creek kicking and screaming

238
00:09:52,340 --> 00:09:50,370

into trying to look at this more

239
00:09:54,770 --> 00:09:52,350
empirically and to Michael Dugan for his

240
00:09:57,650 --> 00:09:54,780
never-ending criticism of this project

241
00:10:00,500 --> 00:09:57,660
on Facebook Facebook is where all the

242
00:10:09,380 --> 00:10:00,510
real science happens these days oh thank

243
00:10:11,210 --> 00:10:09,390
you very much thank you Mark oh one

244
00:10:15,170 --> 00:10:11,220
quick question for you was that last

245
00:10:21,790 --> 00:10:15,180
slide by Michael Levin at Tufts last

246
00:10:24,460 --> 00:10:21,800
slide this no farther back yes yes yes

247
00:10:27,260 --> 00:10:24,470
okay thank you

248
00:10:29,810 --> 00:10:27,270
why would you need the intermediate step

249
00:10:31,880 --> 00:10:29,820
of making images can you not just feed

250
00:10:33,800 --> 00:10:31,890
the the random data stream into the

251
00:10:35,660 --> 00:10:33,810
network and give it some condition like

252
00:10:38,180 --> 00:10:35,670
this is condition a this is condition B

253
00:10:44,360 --> 00:10:38,190
to to learn on what would that condition

254
00:10:47,180 --> 00:10:44,370
be sorry I didn't for example you say

255
00:10:49,400 --> 00:10:47,190
this generator was in Intendant made

256
00:10:50,930 --> 00:10:49,410
under love intention and the under the

257
00:10:53,210 --> 00:10:50,940
other under hate and the network is

258
00:10:56,300 --> 00:10:53,220
supposed to learn to to distinguish

259
00:10:58,190 --> 00:10:56,310
between these two um that is completely

260
00:10:59,990 --> 00:10:58,200
a possible that's yes you could actually

261
00:11:01,820 --> 00:11:00,000
you could absolutely do that but that's

262
00:11:03,460 --> 00:11:01,830
not what I was doing with this project I

263
00:11:05,570 --> 00:11:03,470

ended up with these visualizations

264
00:11:07,049 --> 00:11:05,580
because I was interested to see if there

265
00:11:09,089 --> 00:11:07,059
was something else in the data

266
00:11:11,849 --> 00:11:09,099
that wasn't showing up in the statistics

267
00:11:14,159 --> 00:11:11,859
that the visualizations might she might

268
00:11:16,589 --> 00:11:14,169
show so now I have this visualization

269
00:11:21,629 --> 00:11:16,599
data and I'm interested to see what's in

270
00:11:23,189 --> 00:11:21,639
there I am so excited to hear you do

271
00:11:24,809 --> 00:11:23,199
this talk because I've had like maybe no

272
00:11:27,149 --> 00:11:24,819
less than six conversations with people

273
00:11:28,589 --> 00:11:27,159
on the side about like why I mean why is

274
00:11:29,999 --> 00:11:28,599
it anyone using some kind of machine

275
00:11:32,099 --> 00:11:30,009
learning for this stuff and I think it's

276
00:11:34,859 --> 00:11:32,109
death in the future and secondly the

277
00:11:36,959 --> 00:11:34,869
just the step of constructing the images

278
00:11:39,569 --> 00:11:36,969
by itself is I was pretty impressed with

279
00:11:41,459 --> 00:11:39,579
just that but one question I had I know

280
00:11:43,289 --> 00:11:41,469
this is probably you're maybe just kind

281
00:11:44,939 --> 00:11:43,299
of getting your feet wet in this but so

282
00:11:46,439 --> 00:11:44,949
if I understand correctly you're using

283
00:11:47,369 --> 00:11:46,449
these already constructed images and

284
00:11:48,809 --> 00:11:47,379
then trying to look at the relevant

285
00:11:50,099 --> 00:11:48,819
feature and say okay there's this common

286
00:11:53,129 --> 00:11:50,109
feature within the hate images for

287
00:11:55,919 --> 00:11:53,139
example where I've seen this applied is

288
00:11:58,199 --> 00:11:55,929

in like a predictive context so if you

289
00:12:01,049 --> 00:11:58,209
have the images constructed on several

290
00:12:03,809 --> 00:12:01,059
individuals and maybe across multiple

291
00:12:04,979 --> 00:12:03,819
sessions and then trying to extract okay

292
00:12:06,959 --> 00:12:04,989
what are the relevant features that

293
00:12:09,029 --> 00:12:06,969
could kind of classify love versus hate

294
00:12:11,909 --> 00:12:09,039
images and then use the second half of

295
00:12:14,459 --> 00:12:11,919
your data set to try to predict what the

296
00:12:16,409 --> 00:12:14,469
intention was based on the image that

297
00:12:17,489 --> 00:12:16,419
was constructed for that individual then

298
00:12:19,829 --> 00:12:17,499
you could start to see some really

299
00:12:21,989 --> 00:12:19,839
really useful stuff including like you

300
00:12:23,969 --> 00:12:21,999
know hit versus miss insight paradigms

301
00:12:25,589 --> 00:12:23,979
you know rather than did they have a hit

302
00:12:27,359 --> 00:12:25,599
rate that's reached a significant

303
00:12:29,219 --> 00:12:27,369
threshold or what are the relevant

304
00:12:30,749 --> 00:12:29,229
pattern features in like an fMRI

305
00:12:32,909 --> 00:12:30,759
experiment for example that's associated

306
00:12:34,289 --> 00:12:32,919
with the hit versus a Miss and some

307
00:12:39,649 --> 00:12:34,299
remote viewing or something like that so

308
00:12:43,289 --> 00:12:39,659
make sense yes hi excellent presentation

309
00:12:46,049 --> 00:12:43,299
just a clarification question did I

310
00:12:48,749 --> 00:12:46,059
understand correctly that you gave the

311
00:12:51,929 --> 00:12:48,759
clarify network which was pre trained

312
00:12:53,549 --> 00:12:51,939
with some other other images your images

313
00:12:55,109 --> 00:12:53,559

but you didn't drain the network with

314
00:12:56,579 --> 00:12:55,119
your images that is correct okay because

315
00:12:58,679 --> 00:12:56,589
it would be impractical

316
00:13:01,229 --> 00:12:58,689
she was right because the clarify system

317
00:13:03,829 --> 00:13:01,239
has been trained with every image on the

318
00:13:06,380 --> 00:13:03,839
Internet I was just checking to me yeah