WEBVTT

NOTE duration: ”00:03:15.285”

NOTE Confidence: 0.9938965

00:00:04.400 --> 00:00:05.279 Let's say we take a

NOTE Confidence: 0.9938965

00:00:05.279 --> 00:00:06.559 drop of your blood that

NOTE Confidence: 0.9938965

00:00:06.559 --> 00:00:07.759 has, you know, thousands and

NOTE Confidence: 0.9938965

00:00:07.759 --> 00:00:08.260 thousands

NOTE Confidence: 0.98217773

00:00:08.639 --> 00:00:09.460 of cells.

NOTE Confidence: 0.98410374

00:00:10.080 --> 00:00:11.380 If you think about visualizing

NOTE Confidence: 0.98410374

00:00:11.440 --> 00:00:13.119 the cells, they actually exist

NOTE Confidence: 0.98410374

00:00:13.119 --> 00:00:13.840 in some kind of a

NOTE Confidence: 0.98410374

00:00:13.840 --> 00:00:15.764 hundred dimensional space. Right? But

NOTE Confidence: 0.98410374

00:00:15.764 --> 00:00:16.845 they don't take up all

NOTE Confidence: 0.98410374

00:00:16.845 --> 00:00:17.585 this space.

NOTE Confidence: 0.98094594

00:00:18.045 --> 00:00:19.585 So they take up narrow

NOTE Confidence: 0.98094594

00:00:19.645 --> 00:00:21.664 niches within this high dimensional

NOTE Confidence: 0.98094594

00:00:21.724 --> 00:00:23.244 space. It forms a low

NOTE Confidence: 0.98094594

00:00:23.244 --> 00:00:25.085 dimensional manifold. You can think

NOTE Confidence: 0.98094594

00:00:25.085 --> 00:00:26.044 of a two d piece

NOTE Confidence: 0.98094594

00:00:26.044 --> 00:00:27.325 of paper in three d

NOTE Confidence: 0.98094594

00:00:27.325 --> 00:00:29.150 space. It's forming a lower

NOTE Confidence: 0.98094594

00:00:29.150 --> 00:00:30.750 dimensional manifold, and it solves

NOTE Confidence: 0.98094594

00:00:30.750 --> 00:00:31.630 a lot of the challenges

NOTE Confidence: 0.98094594

00:00:31.630 --> 00:00:32.750 in single cell data if

NOTE Confidence: 0.98094594

00:00:32.750 --> 00:00:34.030 you tried to explicitly model

NOTE Confidence: 0.98094594

00:00:34.030 --> 00:00:35.470 that lower dimensional shape or

NOTE Confidence: 0.98094594

00:00:35.470 --> 00:00:37.470 manifold that the cells reside

NOTE Confidence: 0.98094594

00:00:37.470 --> 00:00:37.970 in.

NOTE Confidence: 0.99726564

00:00:41.045 --> 00:00:43.065 My lab focuses on developing

NOTE Confidence: 0.99726564

00:00:43.125 --> 00:00:45.305 novel machine learning, deep learning,

NOTE Confidence: 0.99975586

00:00:45.685 --> 00:00:47.545 and general AI techniques

NOTE Confidence: 0.96212065

00:00:48.085 --> 00:00:49.365 that are a lot of

NOTE Confidence: 0.96212065

00:00:49.365 --> 00:00:51.525 times either motivated by or

NOTE Confidence: 0.96212065

00:00:51.525 --> 00:00:53.385 specifically designed for,

NOTE Confidence: 0.97196823

00:00:54.100 --> 00:00:56.260 analysis and insight from biomedical

NOTE Confidence: 0.97196823

00:00:56.260 --> 00:00:59.219 data, from sequencing machines, from

NOTE Confidence: 0.97196823

00:00:59.219 --> 00:00:59.719 imaging,

NOTE Confidence: 0.99488986

00:01:00.180 --> 00:01:03.000 from other kinds of specialized

NOTE Confidence: 0.99488986

00:01:03.140 --> 00:01:04.740 technology that can measure things

NOTE Confidence: 0.99488986

00:01:04.740 --> 00:01:06.944 like brain activity, cellular activity.

NOTE Confidence: 0.99488986

00:01:07.165 --> 00:01:08.045 So there's been a lot

NOTE Confidence: 0.99488986

00:01:08.045 --> 00:01:09.165 of this kind of data

NOTE Confidence: 0.99488986

00:01:09.165 --> 00:01:11.185 generated in biology that

NOTE Confidence: 0.9638323

00:01:11.485 --> 00:01:12.925 needs to be analyzed that

NOTE Confidence: 0.9638323

00:01:12.925 --> 00:01:14.225 could give us new insights

NOTE Confidence: 0.9638323

00:01:14.365 --> 00:01:15.985 that can accelerate discovery.

NOTE Confidence: 0.9699707

00:01:19.260 --> 00:01:20.000 We develop

NOTE Confidence: 0.98801494

00:01:20.300 --> 00:01:22.400 a class of methods that

NOTE Confidence: 0.98801494

00:01:22.459 --> 00:01:24.240 we call sort of geometric

NOTE Confidence: 0.98801494

00:01:24.300 --> 00:01:24.800 topological

NOTE Confidence: 0.94793504

00:01:25.500 --> 00:01:27.760 deep learning methods. These are

NOTE Confidence: 0.94793504

00:01:27.819 --> 00:01:29.180 kinda like your normal neural

NOTE Confidence: 0.94793504

00:01:29.180 --> 00:01:30.140 networks that are used to

NOTE Confidence: 0.94793504

00:01:30.140 --> 00:01:31.819 build anything else, but we

NOTE Confidence: 0.94793504

00:01:31.819 --> 00:01:33.020 infuse them with a lot

NOTE Confidence: 0.94793504

00:01:33.020 --> 00:01:34.275 of very specific specific kinds

NOTE Confidence: 0.94793504

00:01:34.275 --> 00:01:34.775 of

NOTE Confidence: 0.9862829

00:01:35.314 --> 00:01:36.595 math that's very important in

NOTE Confidence: 0.9862829

00:01:36.595 --> 00:01:38.595 biology. So geometry is something

NOTE Confidence: 0.9862829

00:01:38.595 --> 00:01:39.634 that we use a lot.

NOTE Confidence: 0.9862829

00:01:39.634 --> 00:01:40.994 Geometry happens to be really

NOTE Confidence: 0.9862829

00:01:40.994 --> 00:01:42.115 important in all parts of

NOTE Confidence: 0.9862829

00:01:42.115 --> 00:01:43.095 biology, including

NOTE Confidence: 0.9791667

00:01:43.475 --> 00:01:44.774 shapes of molecules.

NOTE Confidence: 0.9786328

00:01:46.410 --> 00:01:47.450 A molecule has a certain

NOTE Confidence: 0.9786328

00:01:47.450 --> 00:01:48.490 kind of geometry, and if

NOTE Confidence: 0.9786328

00:01:48.490 --> 00:01:49.530 it didn't have that, it

NOTE Confidence: 0.9786328

00:01:49.530 --> 00:01:51.470 couldn't bind to another molecule

NOTE Confidence: 0.9786328

00:01:51.690 --> 00:01:53.630 or perform its proper function.

NOTE Confidence: 0.9880371

00:01:54.250 --> 00:01:55.770 And then topology is a

NOTE Confidence: 0.9880371

00:01:55.770 --> 00:01:57.630 way of numerically characterizing

NOTE Confidence: 0.998291

00:01:58.570 --> 00:01:59.630 different geometries.

NOTE Confidence: 0.99902344

00:02:00.330 --> 00:02:00.990 We also

NOTE Confidence: 0.9503392

00:02:01.335 --> 00:02:02.535 infused a lot of them

NOTE Confidence: 0.9503392

00:02:02.535 --> 00:02:04.535 with dynamic systems capabilities. So

NOTE Confidence: 0.9503392

00:02:04.535 --> 00:02:05.895 these will be neural networks

NOTE Confidence: 0.9503392

00:02:05.895 --> 00:02:07.575 that can implicitly learn a

NOTE Confidence: 0.9503392

00:02:07.575 --> 00:02:09.255 dynamic system that can generate

NOTE Confidence: 0.9503392

00:02:09.255 --> 00:02:10.315 things like trajectories

NOTE Confidence: 0.9723877

00:02:10.775 --> 00:02:12.615 and forecast what happens to

NOTE Confidence: 0.9723877

00:02:12.615 --> 00:02:14.635 cells and molecules, for example.

NOTE Confidence: 0.96484375

00:02:15.255 --> 00:02:16.875 So these kind of,

NOTE Confidence: 0.9860924

00:02:17.599 --> 00:02:19.599 mathematically enriched neural networks are

NOTE Confidence: 0.9860924

00:02:19.599 --> 00:02:21.299 able to learn more scientifically

NOTE Confidence: 0.9860924

00:02:21.519 --> 00:02:23.919 precise information than the ones

NOTE Confidence: 0.9860924

00:02:23.919 --> 00:02:24.959 that are used out in

NOTE Confidence: 0.9860924

00:02:24.959 --> 00:02:26.719 social media or in other

NOTE Confidence: 0.9860924

00:02:26.719 --> 00:02:27.700 parts of the world.

NOTE Confidence: 0.9887114

00:02:31.135 --> 00:02:32.575 Some prominent areas where we

NOTE Confidence: 0.9887114

00:02:32.575 --> 00:02:33.614 have active grants with a

NOTE Confidence: 0.9887114

00:02:33.614 --> 00:02:35.534 lot of collaborators are cancer,

NOTE Confidence: 0.9887114

00:02:35.534 --> 00:02:37.555 including breast cancer and colorectal

NOTE Confidence: 0.9887114

00:02:37.775 --> 00:02:38.275 cancer,

NOTE Confidence: 0.9967448

00:02:39.055 --> 00:02:39.555 immunology,

NOTE Confidence: 0.9882355

00:02:40.575 --> 00:02:42.754 where we're developing large models

NOTE Confidence: 0.9882355

00:02:42.894 --> 00:02:44.270 that predict immunogenicity,

NOTE Confidence: 0.7861328

00:02:45.690 --> 00:02:46.190 neuroscience,

NOTE Confidence: 0.96589357

00:02:46.650 --> 00:02:48.650 where we're looking both at

NOTE Confidence: 0.96589357

00:02:48.650 --> 00:02:51.050 stem cell development into different

NOTE Confidence: 0.96589357

00:02:51.050 --> 00:02:52.650 neuronal lineages as well as

NOTE Confidence: 0.96589357

00:02:52.650 --> 00:02:54.669 the more neural activity work.

NOTE Confidence: 0.98331326

00:02:55.209 --> 00:02:57.050 Yale is actually unusually well

NOTE Confidence: 0.98331326

00:02:57.050 --> 00:02:58.635 suited for this because you

NOTE Confidence: 0.98331326

00:02:58.635 --> 00:02:59.674 can be at the intersection

NOTE Confidence: 0.98331326

00:02:59.674 --> 00:03:00.635 of all these fields very

NOTE Confidence: 0.98331326

00:03:00.635 --> 00:03:02.075 easily at Yale. So I

NOTE Confidence: 0.98331326

00:03:02.075 --> 00:03:03.355 really love, you know, math

NOTE Confidence: 0.98331326

00:03:03.355 --> 00:03:04.715 and computer science and these

NOTE Confidence: 0.98331326

00:03:04.715 --> 00:03:06.155 kinds of things, and I

NOTE Confidence: 0.98331326

00:03:06.155 --> 00:03:07.215 really wanted to,

NOTE Confidence: 1

00:03:07.594 --> 00:03:08.094 impact

NOTE Confidence: 0.91463214

00:03:08.474 --> 00:03:10.715 a scientific discovery, biology, medicine,

NOTE Confidence: 0.91463214

00:03:10.715 --> 00:03:11.775 these kinds of things.