0:03 So good afternoon.
0:04 My name is Renee Landers and I'm a member of the law faculty and the faculty director for the Master of Science in Law Life Sciences Program,
0:13 which the course that Robert's going to talk to you about today is a part, it's one of the statistics courses,
0:20 one of the required courses for the program. But Associate Dean Brigid Sandusky is going to give you a description of the MSLL program after the mini-class.
0:32 So in case you're interested in learning more about that opportunity.
0:37 So we have been working with Robert over the last two years in helping structure a program for the students in the MSLL program,
0:48 which as Dean Sandusky will explain to you, and I should call him Professor Smith.
0:54 So as Dean Sandusky will explain to you, is designed for people who are interested in going into the life sciences industry.
1:05 It's a master of science degree in law.
1:09 However, people do not have to be lawyers or in a law program in order to enroll in the program.
1:15 It's designed to be multidisciplinary and a law degree or legal experience is not required,
1:21 although people who have JDs are enrolled in the program and some of, there is an opportunity to pursue a joint program with JD students.
1:30 Anyway, I want to say a little bit more about Robert and Professor Smith.
1:35 It was really, it's really an honor for me to be here, to be able to introduce him today,
1:39 because he really is one of the outstanding faculty members of the Sawyer Business School of the University
1:44 generally. He has, being from New Orleans,
1:48 he has a degree from LSU. Sorry about that
1:50 loss the other night and a master's degree from Mississippi State University and his Ph.D. from Florida State University.
2:01 His research focuses on the manifestations of consumer psychology,
2:06 including consumer memory, biases, perceptions and engagement within, within the economy.
2:12 And of course, none of those issues are relevant to anything that's in the headlines today.
2:18 One of the, he is very well recognized for his teaching on our faculty.
2:24 In 2021, he won an innovative teaching award.
2:29 In 2017, he won the Student Government Association, outstanding faculty, the Faculty of the Year Award.
2:37 He has received a best paper award from the Eastern Academy of Management and the best paper award from the New England Journal of Entrepreneurship.
2:48 So his work is recognized both in teaching and his scholarship.
2:54 He, in conjunction with receiving the award from the Student Government Association in 2017,
3:03 the University published a profile of him in the online newsletter,
3:06 which I commend to all of you, because it really explains in great detail his philosophy of teaching.
3:12 And the most, I think the most important feature of his teaching philosophy for our purposes today is that he tries to make

3:21 connections between learning statistics and quantitative methods and real world applications of those tools.
3:29 And some, a lot of people come to law school or to this program with sort of, you know, insecurities about their quantitative skills.
3:40 And I think Professor Smith is one of the faculty members who is able to embrace that
3:47 uncertainty and move students along to a really good place in learning how to use those skills.
3:54 Yesterday afternoon, when I was up at the, the Suffolk I.T. Department,
4:04 getting my passwords changed and various other things realigned on these new systems that we have here.
4:12 I was chatting with the undergraduate students who were working at the desk yesterday afternoon,
4:19 and I mentioned that I was meeting Professor Smith in person for the first time this afternoon,
4:27 and both students randomly had taken his courses and had heard and knew fabulous, fabulous things about him.
4:35 So it's really an honor that he's teaching in this MSLL program.
4:41 And we look forward to his conversation today about one application of statistics in the life sciences.
4:50 Awesome. Thank you very much. And I feel like really beefed up after that.
4:57 Yeah, I appreciate that introduction.
5:03 So it's funny that you give me such a thorough introduction because I was going to start off by making a joke that throughout my time here at Suffolk,
5:12 I've been the other Robert Smith. In the law school,
5:19 I've never actually had the opportunity to meet this person,
5:22 but there's been another professor named Robert Smith has been here a lot longer than me and much more established than I am.
5:29 And so I've always had this.
5:33 Let me see. Do I have a laser pointer here? Yeah. I've always had this little "S" of shame in my name because he's Robert Smith,
5:39 and I'm Robert Smith. And then it's also my email.
5:44 And you know, I had the weird experience last year where I was on an email thread to the whole university announcing that I was retiring.
5:54 But it was, yeah, the other Robert Smith was retiring, and I was incredibly excited.
6:01 They were talking about the celebration for his retirement.
6:04 I was going to be the main person celebrating because I am now officially the Robert Smith at Suffolk University.
6:13 And it's a great feeling. So, yeah, as Renee alluded to, I teach primarily the statistics classes.
6:22 I am in the business school mostly, but I had the awesome opportunity to teach in the law school, to teach this biostatistics class.
6:32 And it's kind of a secret. I don't mention this to my business students so much, but it's actually become my favorite class for reasons that I'll explain.
6:42 So today what I'm going to be doing is presenting some work that me and one of my former students collaborated on, and it involves machine learning.
6:53 Okay, so I was thinking that I really want to create like a theme for this presentation.

6:58 So I was like, you know, I'm talking about my findings from machine learning.
7:02 Maybe what I'll do is I'll use ChatGPT to come up with a title for my talk, right?
7:08 So I took a summary of my class and the project that I worked on with Brittania and I put in into ChatGPT, I told it, to ChatGPT that I was creating,
7:19 I was giving a presentation on this and that was the title it came up with, which I actually don't love because it's a little like snazzy.
7:28 “Blending the power of” is kind of corny, but I was like “nah, I'll stick with it”.
7:32 So I created that title. Yeah, Professor Smith, we're not seeing your slides.
7:39 Oh, is it not sharing? Okay, let's see.
7:44 Doo doo doo So.
7:50 Sure. Is that working?
7:56 Yes, it is. Okay, we're good.
8:04 So then I was like, okay, well, now that I've got a title for it, I was like, Well, do you guys know what DALL-E is?
8:10 You guys familiar with DALL-E? All right, so OpenAI is the company that owns ChatGPT.
8:17 I'm guessing you guys have heard of ChatGPT, and they also have AI that generates images, right?
8:25 Using the same textual prompts.
8:27 So I was like, Well, now that I've created a title for my presentation, it'd be cool to create a background for the presentation, right?
8:35 So I put in some information about my class and I created this background.
8:41 It's faded it out so that I could use it.
8:44 But this is unfaded, okay? And it's all gibberish, right?
8:50 Like the, these codes and signs don't mean anything.
8:53 It just took some prompts about artificial intelligence and, and learning and kind of human insights, and it created this.
9:04 But it's every bit as good as what you would see on, like, a textbook cover art for like an undergraduate level class.
9:12 But the reason why I'm talking about this is to get here.
9:18 I actually had to sift through quite a bit of absolutely terrifying images that created, and I'm going to show you guys these images.
9:27 But there is a trigger warning here.
9:30 They are, they're scary because the actual topic of this project deals with predicting fetal health outcomes.
9:40 And the software really fixated on this idea of the fetus.
9:44 And then it created some.
9:50 Demon spawn babies that,
9:53 I don't know why, I think as a parent like these babies, just this is what it felt like taking home one the first time, right?
10:03 Oh, I can't even look at that one. This is what it came up with for my backdrop.
10:09 And this is the one that I'll pause on because it's my favorite.
10:12 I like how that baby looks kind of smug, right?
10:17 There's kind of an arrogance to this baby. And it also seems to be like dreaming of some sort of phantom baby, some spirit baby or something.

10:28 It's also oddly above the baby. So I decided, I almost used that as the backdrop for the presentation, but
10:35 I was like, That's too weird.
10:37 So anyway, yeah, the topic today is machine learning in artificial intelligence, which is permeating every industry in every sector.
10:46 Right. As I'm sure you're all well aware and I'll talk a little bit about,
10:50 but to kind of zoom out a bit about my class, well, my class doesn't focus necessarily on machine learning,
10:57 so it's not quite as structured as this makes it look.
11:01 But my class, what's pretty awesome about it is that the class sizes are really small,
11:08 so I'm able to customize the class to the interests of my students. So as students show up,
11:15 we discuss what they're interested in through a clinical lens, and then we, we kind of, we start pulling up some of the major papers,
11:24 some of the scientific discovery that's happening in those areas.
11:28 And we started reading them and my students get interested in something and then I go and I find them data on that topic.
11:36 And then we try to replicate some of these major studies.
11:39 Right? So this topic is focused on machine learning not because I designed the class to teach machine learning,
11:46 but because my student wanted to use machine learning to improve health outcomes for pregnant women in emerging economies.
11:58 Right.
11:59 And I thought that was a really cool interest, and I was able to find some cool data, and it just necessitated the use of machine learning as our tool.
12:07 And so let's talk about machine learning a little bit.
12:12 Originally, I was going to work through this whole timeline,
12:16 but I am worried about time because I already talked a lot about demon babies and that slowed me down a bit,
12:24 but I kind of hit on the major points of the evolution of machine learning.
12:30 In 1956, there was a conference in Dartmouth where the idea of artificial intelligence really picked up.
12:38 A few years later, we developed our first supercomputer, the IBM 7090.
12:45 This was a big breakthrough, and I am going to focus on this because I want to compare where we are now to where we were in 1960.
12:54 I'll skip a few bullets. Deep Blue. Does anybody know what Deep Blue refers to?
13:00 Any chess players in here? Deep Blue was the computer program that finally beat the best humans in the world.
13:08 At the time it was Garry Kasparov was our best chess player and it was in 1997 that computers got better than humans at chess.
13:18 Now computers are so much better than humans that completely publicly available software,
13:25 not even software,
13:26 open access programs that you can get on to your laptop right now could beat Magnus Carlsen, who's the best chess player in the history of the world.
13:36 So my four year old can open up the computer and he knows the basic rules of chess.

13:42 And he could beat the best chess player in the world using this program that’s available to everybody.
13:46 It's ruined online chess. Self-driving cars.
13:50 You guys are aware of all these things?
13:52 I definitely recommend googling some of the phrases I've put down here if you want to go down a wormhole regarding
13:59 how artificial intelligence is being used today because some of the applications are absolutely mind blowing.
14:06 So for instance, what Amazon is doing with their Rekognition platform.
14:12 Most of us, when we hear about facial recognition and what Google and Amazon are using it for,
14:17 we think, Oh, it lets me access my phone or something like that.
14:21 But actually advertisers are able to use facial recognition,
14:28 artificial intelligence to get a sense of like the emotions you're experiencing in real time as you use these products,
14:37 which is pretty awesome or scary
14:41 depending on how you look at that? My lawyers might be more interested in what JPMorgan is doing,
14:47 where they are using artificial intelligence to process legal documents and generate legal documents.
14:54 So that's also awesome or scary depending on how you look at it.
14:58 And plenty of other applications are on this at the bottom, three’s Scout where they use artificial intelligence to scan
15:07 crop yields at the scale of thousands of miles at a time to identify the crop fields that are distressed.
15:16 Right. The crops, the individual crops are distressed at a rate that obviously humans can't compete with.
15:24 What I want to focus on, though, is this gap. You'll see like a 33 year gap where there weren't any major breakthroughs.
15:32 And if you ever read a book on artificial intelligence, they will refer to this as the AI Winter, right?
15:39 And it's usually framed as this stretch in time where we just weren't making progress towards this goal of artificial intelligence.
15:48 And I actually want to dispel that notion a bit because we made a ton of progress in those 30 years.
15:55 Right. So if you know anything about Moore's Law, I don't know if you've ever heard this before,
16:03 but it's the idea that the number of transistors that we can fit on a microchip doubles every two years.
16:11 Right. So to put that in perspective, because it's exponential growth, to give you an idea of what that means.
16:19 Well, that IBM 7090 that I referenced from 60 years ago, our first supercomputer, it could perform 200,000 operations per second.
16:29 And it kind of sounds like a lot.
16:31 The computer used a room about this size, the entire computer and it took manual punch cards from human hands to operate.
16:40 I bet it could do 200,000 operations per second.
16:44 Well, the first iPhone could do 400 million operations per second right and it fits in your pocket.
16:51 And we get annoyed when the battery dies. The iPhone Ten can do 600 billion operations per second.
17:00 A modern supercomputer, which is a better comparison to the IBM 7090, can do 30 quadrillion operations per second.

17:10 So in one second, a modern supercomputer can do what it would take IBM 7090, 5000 years to do.
17:18 We made that progress in 60 years. Right.
17:21 That's incredible. And what's really incredible is Moore's Law states in two years it'll double, right?
17:30 That's just mind blowing, right? So that's kind of the scale that we're dealing with.
17:34 And obviously it has applications to clinical settings, right?
17:40 So application number one, which is probably most immediate, is diagnostic discovery.
17:48 So this is Google's LYNA technology which identifies cancer in lymph nodes, and it can beat out a panel of the best doctors in the world.
18:00 Right. So it can identify at a 99% accuracy, and the best doctors in the world combined,
18:08 the best they've ever done is 96%. So it's pretty incredible.
18:13 We're applying technology similar to this, to identifying Alzheimer's.
18:18 And the, the AI that identifies Alzheimer's performs that are 82%, humans are usually around 70%.
18:29 Again, when I say humans, I mean experts. We are also, you can't see it because of the header, but we're also using this technology to – perfect -
18:44 we're also using this technology to, to diagnose diabetic retinopathy,
18:51 which I don't know if you guys have any loved ones or know anybody that has diabetes,
18:56 but one of the side effects of diabetes is blindness because of hemorrhaging in the blood vessels in the retina.
19:05 And this is then something that doctors have long struggled with,
19:11 but that computers have been able to tackle by processing at a very large scale these images and images such as this
19:22 that I'm talking about, thousands and thousands of these images.
19:26 And so I don't want this to turn into an artificial intelligence talk because I can talk about this topic forever.
19:32 Right. But some of what we're doing with artificial intelligence is pretty incredible and also pretty scary.
19:40 Another application in a clinical setting is personalizing treatment.
19:45 Right.
19:45 So without getting derailed on this too much, doctors obviously, alright, I’m going to state an opinion, have a lot of biases when they assign the treatment.
19:56 Right. And while the artificial intelligence could also be biased,
20:03 removing human bias in industry, so there's different treatments that can be used to tackle the same condition.
20:12 And obviously this is removing some of that human bias.
20:17 And then finally, drug discovery.
20:19 We're actually discovering new antibiotics simply by dumping a ton of compounds into AI models and letting the AI discover new treatments,
20:31 which is pretty incredible. All right.
20:34 So my student and then intern, my student and I became interested in applying this idea to predicting perinatal health outcomes.
20:45 Right. So I want to take a second and talk about why this is important.

20:51 So this topic is important,
20:54 I think it's obvious to some extent why it's important, but I want to contextualize it a bit because if you think about the evolution of humans,
21:04 right, and you think about how life expectancy has changed, for some people, the rate at which life expectancy has changed is a little surprising.
21:14 Okay, so if you were to plot this on a timeline, it actually wouldn't look like this,
21:20 because for the first half of this graph we're covering 12,000 years, in the second half of the graph, we're only covering 400 years.
21:28 Right. I don't have enough room to actually show the evolution of humans.
21:34 All of our progress has happened in the last hundred years.
21:38 Like if you look at that, we give you guys more precise numbers.
21:43 For 2000 years of human existence or for 12,000 years of human existence, we improve life expectancy by about 12% total.
21:55 And in the last 120 years, we've increased it by like 140%.
22:00 Right. So it's an incredibly, it's obviously due to improvements in health care technology.
22:08 Right. But the problem is access to that isn’t equivalent across the globe, right?
22:15 So if you look at how child mortality has changed,
22:20 see experts attribute much of this gain in this graph to two factors: reductions in child mortality and reductions in maternal mortality.
22:30 And if you look at how child mortality has evolved in the last 200 years, I generated this visualization myself.
22:43 And you can see obviously across the world we're seeing improvements in child mortality, but it does differ for emerging economies.
22:52 And so that was a selection of eight different nations.
22:56 But here is everybody this everybody had data on.
23:00 So I created this visualization. Our Y-axis is child mortality rates.
23:06 Right. Our X-axis is income, so how wealthy the nation is.
23:11 The size of the bubble is the size of the population of that nation.
23:16 And then you can see the year. Right. And we're going to watch it change.
23:20 Now, I sped this up really fast because I was worried about time.
23:23 The visualization when I originally created it takes about 30 seconds.
23:27 So you have to watch fast and. So you watch the bubbles grow.
23:33 Watch the population growing larger as our nations become more wealthy, then what
23:39 you'll see here is I'm zooming in on the y axis that you can see the effect here
23:46 and you can see our African countries are in blue, that it's not, it's basically
23:56 these, these emerging economies are where the United States was in the thirties.
24:01 Right. And the thing about machine learning and artificial intelligence is it's so much more scalable than humans are.

24:12 Right. In these emerging economies, they don't have access to the same doctors that we have access to here. By creating machine learning algorithms,
24:24 we are thereby removing the need for as many of those health care professionals.
24:29 So this is why the topic is important.
24:31 Also looking at maternal mortality you’ll see the exact same effect, I’ll slow down the video and around the sixties because I didn't have data before the sixties.
24:41 So I really sped it up. And you see this in video that same relationship as with child mortality.
24:49 All right. So let me talk about our data. Let me talk about what we actually did with this.
24:53 So we use CTG data. And if you don't know what CTG data is, it's it's how we monitor a fetus's heart rate
25:05 during the pregnancy. We were looking at the third trimester, the final trimester, and we had 21,
25:12 we had access to a dataset of over 2100 fetuses.
25:17 Right. So we were looking at things like what's the baseline, fetal heart rate, how many acceleration were there,
25:24 how many deceleration, were there uterine contractions, did the heart rate zero and how often did zero out and peak, etc., etc.
25:34 A lot of features focusing on the volatility of the heart rate.
25:40 All right. So I'm going to have to explain very, very, very quickly what machine learning actually is and what binary classification is.
25:50 Now, I want you to know that I'm going to blend in maybe three different topics that I would normally spend like a month on.
25:58 Right. So this is going to be a speed run on machine learning and binary classification.
26:04 All right. So first, I'm guessing you guys, I know my students do, but what about the law students?
26:12 If you guys learned linear regression.
26:16 You've heard of it. Yeah. All right. Real fast recap on linear regression.
26:22 With linear regression, we're predicting a continuous outcome.
26:27 In other words, we’re predicting an amount of something, right?
26:30 So in this case, if we wanted to predict someone's blood pressure using their age, we might use linear regression.
26:37 Now, the simplest form of linear regression is conveniently called simple linear regression.
26:43 And it's where we have one independent variable, one x variable on the x axis gauge, predicting one dependent variable.
26:51 We would plot this relationship using a scatterplot.
26:54 This should sound pretty familiar to most of you. We would create a model.
26:58 The model we're going to visually represent is a line of best fit.
27:03 It is a line that minimizes the difference between every real observation in the predicted observation, right.
27:12 What's awesome about linear regression in the simple form is we can calculate the best line.
27:18 There is no, it's not an approximation, it is the best line.
27:23 You guys have probably seen this formula before. Y is what we're predicting,
27:27 blood pressure, beta zero is our intercept, right,

27:30 it's where our model intercept intercepts, so it's basically, it's the model's prediction when your x variable is zero, right?
27:37 So when you're zero years old, this is the prediction for Y.
27:41 And then we don't really need that for today or we do need to understand sort of is what a beta coefficient is.
27:48 A beta coefficient is what the model's predicting will change in y - blood pressure - when we increase x – age - by one.
27:58 Right. So as I increase age, how much does my blood pressure change. As I was making these slides, I realized why I'm using a hypothetical data set?
28:10 I have real data on this as I took my real data and I created this simple linear regression, right?
28:17 So this is literally what a model would predict.
28:20 Now, again, this is a very simple model, so it's not going to be very accurate.
28:25 But here's real data looking at the relationship between age and systolic blood pressure, and what we're seeing here,
28:32 here's our data, our data is conveniently, basically one, so that every year we age, our blood pressure is predicted to go up by one.
28:41 Our intercepts 82. So this model's predicting a zero year old
28:45 to have a blood pressure of 82, which is actually very inaccurate based on my Google searching, maybe the blood pressure like 40-something currently.
28:55 All right. But a simple linear regression model is not going to be very accurate because it's only got one feature.
29:00 It's only got one variable it's using to predict off of - age.
29:04 Okay, enter logistic regression.
29:08 Now, I would normally spend about a month teaching logistic regression, right, if I wanted to teach this to you properly.
29:15 Logistic regression is based on the same mathematical principles as linear regression, but it's more complicated.
29:23 With logistic regression, we're predicting a categorical outcome.
29:28 Right. And believe it or not, that's a more common need is to predict a categorical outcome.
29:34 So to think about meteorology, a meteorologist would use linear regression to predict how much precipitation we're going to get tomorrow,
29:44 but they would use logistic regression to predict will it rain?
29:48 Right. And a lot of times that's what we care about, is will something happen, yes or no?
29:53 And that's when we need logistic regression.
29:57 So the problem with logistic regression is we want to model this linearly like we do with linear regression, but look at our data.
30:07 It's all weird because you only have either heart disease or you don't have heart disease, right?
30:14 It's yes or no with logistic, it's binary in this case.
30:18 Right? So because of that, we have to perform a logistic function on our model. It creates this little
30:25 S-curve and this is hypothetical data, so don't freak out,
30:29 but basically what this is saying is at a young age, this is your probability of heart disease, by the way,
30:37 so at a young age, as you get older, there's almost no impact on your probability of heart disease.
30:43 And then what you get to about my age, that changes pretty rapidly and then it slows down. Again,

30:50 this is made up data. And as I was putting this together, I realize, again, I have data this, let me just use the real data.
30:58 Problem is you don't see the full S curve. The traditional logistic regression model function has an S curve to it.
31:06 Well, but that's good news because, based on the data that I have access to, your probability of heart disease never clears 50%.
31:15 So it's good. The bad news is, is because something else kills you, kind of a dark direction.
31:21 Okay, last thing. Our output in logistic regression is in terms of odds ratios.
31:30 Okay. So what an odds ratio is, is it's the probability of something happening relative to the probability of it not happening.
31:39 Right. Kind of simple. So if there's a 75% chance that you become MSLL students, the odds ratio would be three, right?
31:53 Three times as likely to become a student is not, right?
31:57 So it would be your odds ratio if there was a 90% chance, your odds ratio would be nine.
32:02 Now, the reason why this matters is that we can't model things in terms of probabilities, because probabilities cap at 100%.
32:12 So if you remember what a beta coefficient is, it's showing. well, every time we increase x, what's happening to y?
32:19 Well, if we did this with a probability, alright, let's say our beta is five.
32:24 So every time you get a year older, your probability of heart disease goes up by 5%.
32:31 They won't take long before your probability was over 100%.
32:35 That'd be very bad. So that's why we have to use odds ratios, because odds ratios go up to infinity.
32:42 There is no artificial cap on them. If you think about what an odds ratio is for long enough, you'll realize, Oh yeah, they can go up forever, right?
32:51 As a probability, asymptotically gets closer to one and odds ratio just rises to infinity.
32:58 All right. That's the background that you need to know to kind of understand what our output is.
33:03 I'm going to skip these slides because I'm a little worried that I'm talking too long.
33:08 My student and I performed exploratory data analysis, as you would do before running any major model like this.
33:15 So the first slide, actually, I will pause.
33:19 This is our dependent variable. This is what we're predicting.
33:23 There's three categories. We've got a normal fetus, a suspect fetus,
33:29 (this is not my coding) and a pathological fetus. The way we have those classes
33:35 Those classifications is based on a team of obstetricians.
33:41 I think that's right. A team of obstetricians look at the data and they assign them into those three categories, all 2200 fetuses.
33:52 Right. And what we're doing is we're going to take those classifications and we're going to build a model that can replicate that.
34:02 That we're going to use it to teach the model how to identify a fetus as either normal, suspect or pathological.
34:11 And then we're going to test its ability to be as good as those doctors were going forward.

34:16 It's really powerful stuff. So the first thing we did is look at those classifications.
34:22 You can see 78% of the fetuses were normal, 8% were pathological.
34:31 I'm going to skip these. We did a box plot to get rid of our outliers.
34:35 We did Nested Histograms to just kind of look in on every variable and its distribution, get rid of any messy data.
34:43 We did this correlation heatmap where we can look at some of the really hot relationships and you want to remove those redundancies from your model.
34:54 If you guys have questions about that,
34:56 I can elaborate after this presentation. This is what I want to pause at because this is what we're doing with machine learning.
35:03 We're taking all this data and we're going to separate the data into two groups.
35:09 70% of it is what we call training data.
35:13 Okay, so this is where we take the machine learning algorithm and we give it the answers, right?
35:19 We give it all the CTG data and we say this was a pathological virus.
35:24 And then here this is a this is a normal fetus. And we do that for 70% of it.
35:30 And it creates an algorithm itself. It creates beta weights itself and optimizes these data once it learns.
35:38 Right, this is what is called supervised learning.
35:42 The model gets smarter and smarter because it has the answers and eventually it gets to where it's got basically,
35:49 basically perfect accuracy, at least in our case.
35:54 But then we take that other 30% of the data. We left that out, right?
35:58 And then once we've built the model, what really matters is how it performs on the test data.
36:04 That's our approximation of how this model is going to perform without us there in an unsupervised way.
36:11 How is this model? Can it replicate what that panel of expert obstetricians
36:16 could do, right, without us there. And if it can, then that means it's, it's got utility.
36:22 It's something that we can offer to, to these settings that they need more resources in their health care practices.
36:35 Right. So this is what Brittania and I were dedicated to doing for the semester.
36:41 Here's our syntax. If you don't like programing, this will scare you.
36:44 But I shield the students from it that don't want to get messy with it.
36:49 So you don't have to do this. This is 200 lines of code that I wrote,
36:54 but the program is much more sophisticated than this will even look because for those that don't know anything about coding,
37:03 the first 20 lines are actually importing entire libraries and modules of other programs.
37:09 So it's obviously, like we're doing machine learning, the program is pretty sophisticated.

37:16 All right. So our findings, we're going to gloss over the beta coefficients, mostly due to time.
37:23 But to kind of remind you or maybe to point out the complexity of logistic regression,
37:30 these coefficients are not intuitively interpretable because they're in units of log odds and I'm
37:38 not even going to take the time to log odd is, but these odds ratios are a bit more interpretable,
37:45 right? So what this odds ratio is saying is that when we increase this independent variable by one, this is how much the odds of being a one,
37:56 in this case, a pathological fetus goes up, the odds,
38:00 remember the odds, not the probability. Right. So what you could do is subtract one from this odds ratio and think of it as a percentage of change.
38:09 So when you increase a normal short term variability of the heart rate by a one,
38:14 the odds that this is a pathological fetus goes up by 550%, it's a very big impact.
38:22 So I'll skip right over to our takeaway slide.
38:27 This is how good the me and Britannia's model was. The model was 97% accurate.
38:34 The best way to think about this is this model is 97% as good as a team of doctors.
38:42 Me and the student put this together in one semester. So that's the predictive power,
38:48 the precision. Precision here is how accurate the model is when it identified the pathological fetus.
38:55 Right. So the precision was 95%. That means if it said that we have a pathological fetus here, 95.3% of the time it was right.
39:05 Precision is a, is a, oh yeah that’s why I just said. Sensitivity is a little more difficult,
39:12 though, because if you look at this, most fetuses are healthy, right?
39:18 Most of these fetuses were correct. What's hard is to find the needle in the haystack.
39:23 Right. So the model was right, 70.7% of the time at identifying
39:29 those needles in the haystack, which is actually pretty good. The specificity in this case, that's the model's ability to find a negative case.
39:37 So when it said when we had a healthy baby, I shouldn’t say baby, when we had a healthy fetus, how accurate was it at identifying is healthy, it was was 99.7% accurate.
39:48 So that's pretty powerful stuff, especially for a student with almost no statistical background to pull off in one semester.
39:59 So the implications of this are pretty big, right?
40:02 Constant CTG monitoring reduces birth asphyxia and fetal mortality rates.
40:08 The model approximated expert human performance, not regular human performance, but a team of expert obstetricians.
40:15 And most importantly, you know, CTG is that the ability to perform ongoing monitoring of CTG data is not accessible to all continents, right?
40:31 To be able to monitor, to replicate this with machine learning technologies would be pretty impactful for those situations.

40:41 So that's everything. I have no idea how I did on time. I'm sorry if I'm way over.
40:46 Okay. I think we're all right. I'm sorry if I was talking fast
40:52 there are guys who started to realize I packed a lot in this presentation.
40:58 Thank you so much Professor Smith. I am sure both in Zoom land in here in the classroom today
41:11 we'll have some questions for you. But first, we're going to kind of continue with our presentation.
41:17 My name is Bridgett Sandusky. I am the Assistant Dean of Graduate Law Programs here.
41:22 And the Master of Science in Law: Life Sciences program is one of several degree programs that I oversee here at Suffolk.
41:29 Okay. So I'm going to - bear with me for a moment – as I start to share.
41:37 So I know some of you are Suffolk students either at Suffolk University or at the law school,
41:43 but I also know that we have some visitors online and some visitors here in the classroom that aren't familiar with Suffolk and who we are.
41:50 So I thought it'd be helpful just to tell you a little bit about Suffolk and our vision of what it is that we do.
42:00 And I think Professor Smith here today really demonstrates Suffolk's vision accurately for you, of faculty members
42:07 who are able to take really complicated and complex information and make that so much more readily accessible.
42:16 We are grounded in a practical education from our founding.
42:21 In 1906, we started as a law school and then grew as a university,
42:27 but we always focused on providing our students experiential opportunities and that is becoming ever more important.
42:36 And that's why many of the classes that you that will find within the Masters of
42:40 Science in Law: Life Sciences program are going to have that same grounding,
42:46 that grounding and providing information, so you're going to learn within the classroom,
42:51 but many of the classes are going to have a project based, similar to Professor Smith today,
42:56 but he is certainly not alone, in which students are taking that information and developing projects.
43:02 So information that's going to be helpful and useful for you when you're taking this degree and going out into the real world with it.
43:10 And so we do that both with this brand new degree, which I'm really excited to talk with you about,
43:17 but also within our JD program and at the University.
43:27 [unintelligible]
43:38 So the Master of Law: Life Sciences, it is,
43:44 we launched it during COVID, probably not the best time to be launching a degree program, but it was an opportunity.
43:51 We have been developing it for a number of years. So first and foremost, we did a lot of research.
43:58 We went out into the market to talk with individuals.
44:02 We went and talked to the scientists and said, What is it,
44:06 when you talking with an attorney in the biotech pharma sector, what is it you want them to know?

44:12 And we went and we talked with the attorneys and the same thing.
44:15 So what this degree is supposed to be doing for you is filling in those gaps.
44:21 So you don't need to be a scientist to come into this program.
44:26 You don't need to be a lawyer to come into this program. We are going to fill in whatever educational gaps that you might have.
44:33 So you will have a firm grounding in both science, business and law.
44:40 So it's really unique.
44:42 We believe that it's still one of the only programs structured specifically like this in the United States, potentially in the world as well.
44:52 There are some other programs that might be geared to specific subsets.
44:56 So just for people who are scientists, for example.
45:00 But our program is really wonderful in that it allows individuals from all different types of backgrounds.
45:07 So if you're in a class like Professor Smith's class, you're going to have students that are attorneys,
45:12 students that are coming from a business development background,
45:15 students that are scientists that are looking to learn more about business and law.
45:25 And it really is reflective, this degree is really reflective of what jobs are today.
45:32 We're no longer, jobs are not just siloed. Right.
45:36 We know that in our jobs were expected to wear so many hats nowadays.
45:41 So shouldn't we be engaging in education
45:45 that's going to be reflective of that? And so this and offering an interdisciplinary degree in which you have faculty members from Suffolk Law School,
45:54 from Suffolk, from Sawyer Business School and from our College of Arts and Sciences, really reflects that.
46:02 And we're really excited about it. And of course,
46:07 the life sciences in the state of Massachusetts is really the main reason why
46:13 we developed this program was because there has continued, long before COVID,
46:18 there has been a continuous growth within this sector.
46:22 And we were hearing and you can read in the Boston Globe, for example,
46:27 there is always such a, there was the need for people in Master's and higher level positions within biotech pharma companies,
46:38 and they're actively recruiting and stealing talent away from the various different companies because there weren't enough people.
46:46 So Mass BioEd every year puts out the Life Sciences Workforce Trends report.
46:52 And so that's where we're grabbing this data from. A 67% increase in life sciences jobs over the past decade.
47:01 They're projecting 20,000 new biotech pharma jobs by 2024, and it's highly concentrated.
47:11 And this next slide really reflects that quite well.
47:15 You can see from 2017 to 2019, Massachusetts life sciences jobs grew by 17%.

47:25 Blowing away what it grew over the rest of the United States at 4%, which is really great,
47:30 still terrific. But. So much more needed outcomes.
47:37 It was Lifesciences 9% for those, but more than double,
47:41 almost double the growth in Massachusetts for life sciences positions versus life sciences across the United States.
47:50 And then in reference to 2019 and 2020,
47:54 still Massachusetts continued to grow a reasonable 3% when we saw growth across the United States and in Massachusetts,
48:02 other private sector jobs retracting quite substantially.
48:07 So it is a stable but also continuing to grow.
48:13 And we expect that to not be stopping any time in the foreseeable future.
48:19 Over this five year time period, we saw 55% increase in job demand during that time period.
48:26 10% alone from 2019 to 2020.
48:30 And there's a projected 9.2% by 2027.
48:36 So that means what? Jobs. And we think our program is able to prepare you to enter into a variety of different roles.
48:45 Obviously, employment outcomes are going to be dependent upon what it is you're coming into the program with already.
48:52 So if you're not already an attorney, we're not going to make you an attorney in this program.
48:58 And if you're not already a scientist, we're not going to make you into a scientist, but we are going to teach you the terminology.
49:04 We're going to teach you the language so that you can enter into a variety of different roles.
49:11 So the structure of the program, the program itself is comprised of 30 credits, and that will include ten required courses.
49:22 Those courses are three credits each. Now you'll see in the next slide that the different courses, biostatistics with Professor Smith,
49:30 is one of those ten courses. You can pursue it in on a full time basis or on a part time basis.
49:36 There's a great deal of flexibility built into this program's structure.
49:43 Part time: you can go really as fast or slow as you want.
49:47 So a year and a half to over two and a half years. And then we're completely test optional.
49:53 So no, GRE, GMAT, LSAT is required.
49:59 And for any international students that might be joining us here today, this is a STEM approved degree.
50:05 And so that is really great news for students that are F-1 visa holders because it allows
50:13 them to stay for the 12 months of OPT and then renew for an additional 24 months.
50:18 So three years in total. So they complete the year the degree in one year and then they're able to stay work in a organization,
50:29 company, government as long it is, as it is related to the degree itself for the next three years.
50:37 So that is a huge benefit for any of our international students.
50:42 And you can start full time and you can start in the fall semester.
50:46 So that would be an August start date. Part time students, you have the option to start either in August or in January.

50:52 There's two start dates for that. Our priority application deadline is in May 15th for the fall,
51:00 so there's still plenty of time to be submitting applications for fall of 2023, and most of our classes will be held at 4:30 p.m. or later.
51:12 So that's convenient both for our full time students that might end up securing internships within the life sciences field,
51:22 and that is something that we help and we assist with.
51:25 So particularly if you do not already have any sort of background working within biotech or pharma,
51:33 that is something that we would be working with you to help you pursue an internship and experiential experience during your time in your program.
51:44 And then for part time students, that means you're working during the day and then you're coming to class at night.
51:51 And the nice thing about these courses are they're in blocks.
51:56 So it's potentially you might only be, you could as a part time student be in class once a week.
52:02 You could have class from 4:30 to 7:10 and then 7:15 to 9:50.
52:07 So if you think about that as a person approaching this as a part time student, you know could I do graduate a graduate program one night a week?
52:21 Yeah, I could do that. I think most of us could say we could do that.
52:25 It's not something that you need to be here four days a week and in doing that.
52:30 So I think there is that flexibility for students.
52:36 Every single application that is submitted will be automatically reviewed for merit based scholarship potential.
52:44 So you don't need to be submitting any separate documentation, for example.
52:51 And the vast majority of the admitted students will receive some level of merit based funding, ranging up to about one third of total degree cost.
53:04 Here are our current ten courses.
53:07 It's really structured as half the program is more geared toward law courses, and half the courses are geared more towards science courses.
53:15 This introduction to molecular biology for professionals is really our foundational science course.
53:22 And so that's going to be unlocking your other more sophisticated science courses,
53:27 so applied genetics therapies, clinical research, etc., and then you are going to have them intermixed.
53:36 So if you're a full time student, five courses in one semester, five courses in the next semester; part time students,
53:43 it could be as little as two courses in a semester, upwards of four courses, if that's the intensity that you wish to pursue.
53:53 I think it is worth sharing, although I know Professor Smith, makes tonight's data really approachable and this is so interesting,
54:03 I also know sometimes that might be – you've got to take note,
54:06 he said "I would be teaching this for a month. I'd be teaching this for two months."
54:11 So if things were really kind of going over your head,
54:15 keeping that into perspective that he was trying to give us, kind of focusing on outcomes,

54:22 because we know it's impossible to be teaching a biostatistics course in 30 minutes.
54:28 You can't do that. So we decided it would be best for him to really focusing on what a student actually did within the class, within the classroom.
54:39 But we are not making you into statisticians, unless you already are.
54:44 As I said, we're not making you into an attorney if you aren't already one.
54:48 What our job is to do is to make sure that you understand the language and the terminology.
54:54 And so if you're already an IP attorney and you have applied to this program, you're not going to take the Intellectual Property Survey course.
55:03 So we custom designed the program to your own individual needs and educational background.
55:09 So if you already have a STEM degree and you said I took statistics, you're not going to have to take it again.
55:16 We will evaluate that, we’ll ask for you to secure the syllabus for the course that you took.
55:21 And then we send it to the experts. So I'll send it to Professor Smith and he'll say, Yeah, they cover 90% of what I would be covering.
55:29 There's no need for them to replicate that and take this course.
55:33 And then we find a course that's going to add better value to your degree.
55:39 So maybe that's a health law course, maybe it's a FDA food and drug policy course,
55:47 and maybe it's another business related course if you're interested in that.
55:51 That is really what makes this program so, that we're able to gear it and structure it to a wide variety of individuals.
56:02 And so if you have a degree in molecular biology, you're not taking the intro to molecular biology.
56:07 And that's the flip, too,
56:09 because often students that are coming from a legal background will say there's going to be people that are molecular biologist in the classroom.
56:17 No, we're not having them in that classroom with you.
56:20 It's going to be a subset of students that should be in that classroom that haven't taken that course.
56:27 So it's really important to focus on that, that the courses themselves are geared toward the particular student type within that course.
56:37 And they're small enough so that these courses, particularly with the science based courses, are intentionally small.
56:45 So a faculty member is going to know if no one's getting it, they're going to be able to see it and they're going to be able to, you know, reevaluate.
56:53 Okay, let's go back, review this and make sure kind of everybody's on board moving ahead.
57:01 So if we're talking about ten students, it's much easier to be able to do that.
57:05 We're not talking about 100 students in a science course, and none of the science courses do not have labs.
57:12 So they are just purely course based courses -

57:18 there are no labs attached to those particular courses.
57:26 Here's a listing of potential employment outcomes.
57:29 These are going to be, as I said, really variable dependent upon your prior educational background, your prior professional experience potentially.
57:40 So as I said, we're not going to make if you aren't already an attorney,
57:43 you're not going to have an employment outcome of a biotech transactions attorney, for example.
57:49 And, but there there's going to be a wide variety of different types of employment outcomes that we can also discuss with you on a one on one basis,
58:00 taking into consideration. And I am here as part of a team.
58:05 And so you met today earlier, Professor Landers, who is the academic director.
58:11 She also teaches the data privacy course. And then online tonight and on Zoom,
58:18 we also have director Jennifer Karnakis, and.
58:24 she is the executive director for the Intellectual Property Center here at Suffolk, and she was a life sciences attorney for a number of years,
58:36 so has, in Massachusetts, in Boston, so has a really wonderful an amazing, strong and deep network.
58:44 And because this program is interdisciplinary,
58:48 we're able to draw upon both the resources of the university's career center, as well as the law school's career center.
58:58 So those, the MSLL students are really lucky in that they're able to, because of the interdisciplinary aspect of this degree,
59:05 able to draw upon both the strengths and the alumni base that we have across the entire university.
59:15 And I think it also bears saying that many of the faculty members that are teaching,
59:22 whether or not they're full time faculty members or if we are drawing on a couple of individuals teaching within the program are adjuncts,
59:29 so they’re a compliance, they have a compliance consultancy firm that they're doing every day,
59:37 and then they're coming and they're teaching compliance in the life sciences industry.
59:41 And so really bringing in individuals that are teaching with a lot of real life, hands on knowledge.
59:50 So this isn't just what it says in the book. You know,
59:53 let me tell you about what happened today and really bring in prior work experience that they've had prior to coming and teaching within the program.
1:00:05 Here are several upcoming events. So Director Karnakis will be holding several virtual meet and greets and next month, they are held monthly.
1:00:16 And then we have a graduate open house later this month, on March 25th, from 10 a.m. to 1 p.m. Eastern Standard Time.
1:00:26 So I would encourage you to come and visit us at one of those times as well.
1:00:32 And so this is my contact information as well as for Professor Landers and Director Karnakis and I would be happy to open this up,

1:00:41 I know we're a little bit past our 6:00 time frame,
1:00:45 but if you have any questions for me about the program or some questions for Professor Smith as well,
1:00:53 those that are online, I can see, I can go back and I see that there's at least one chat question over here.
1:01:02 So if anybody here has questions, feel free to questions.
1:01:06 So one of the things I really like when you were talking about how hard it is to see something going on,
1:01:20 was I just kind of like defining [unintelligible]?
1:01:25 Yes. So that. Oh, yeah, that's a great question.
1:01:29 So what I'm referencing there is a mathematical computational method.
1:01:37 So the way that you heard "neural network"?
1:01:41 Yeah. Oh, okay. Yeah, yeah.
1:01:44 Have you heard of it in the context of computer science or artificial intelligence? It's the same thing.
1:01:49 It's like what you learn about, right? So each connection in the human brain is literally just electronic signals from one neuron to another.
1:02:05 Well, in a neural network of a computer, it's the same thing.
1:02:09 And there is a mathematical operation that's going determine whether or not the signal is sent.
1:02:14 That's right. So that they can perform 30 quadrillion mathematical operations a second, determining whether that's sent.
1:02:25 Yeah, that's a great question. Just trying to know that.
1:02:30 Yeah, I'm sorry, but a bigger question, I guess, we were talking very briefly at the beginning of your talk,
1:02:38 I was just wondering when you were constructing your model, I guess I try to avoid some of those biases too.
1:02:54 So I didn't know that was an issue.
1:02:57 There's a common misconception that there's not bias involved in data science, and that's just not true. There's going to be bias and the complexity of data
1:03:06 science is trying to remove all that bias. That's why statistics class is so complicated, we are trying to remove the bias.
1:03:16 So what's actually interesting right now, in artificial intelligence, is that these neural networks are so sophisticated,
1:03:25 so complicated that humans are struggling to even figure it out because the output, without getting, making this
1:03:34 is a whole lecture on artificial intelligence, the output data that you get, we're not even like, well, ChapGPT generates some
1:03:44 text for you, the model that creates that text is so complicated that the humans that created that model
1:03:51 don't even know why that data got generated. Right.
1:03:54 So it's becoming more difficult for humans, if that makes sense.
1:04:08 There's a question about OPT extending to 18 months, you can extend to 18 months.
1:04:16 That is not a problem. You have to be at least a minimum of nine credits or more to be a full time student on an F1 student visa.
1:04:25 So if you wanted to be completing it over a period of nine months, you would have the flexibility to do that.
1:04:33 But there would be very sort of specific courses that we would have you be doing just to make sure that lock step you're taking the right sequence

1:04:43 of courses. Simple question,
1:04:48 so most people on the program, are they full time or part time? It's a mix.
1:04:53 It really is. I think we do, it’s still a what we would say a small program, or it’s still in the building blocks of even doing that,
1:05:02 but there is there's a good mix of both and part time. We also offer this as a dual degree for our J.D. student population as well.
1:05:13 And so we do have some J.D. students that are that are in the program.
1:05:18 Generally speaking, they have been full time as well. Other questions?
1:05:29 Okay. We'll make sure online.
1:05:34 No, I think we are good. Well, thank you so much, everyone, for coming and attending tonight's mini class on biostatistics.
1:05:41 I hope many of you will consider in learning more about the Masters of Science: Law Life Sciences at Suffolk University Law School.
1:05:50 Thank you.