So what I want to do today is to give a quick overview of the class. And then talk about stuff that you might already know about. So we won't really do any advanced discussions today, but basically bring everybody to the same page regarding basics goal of lectures to just. And things started. So let's begin by. Talking about neural networks in general. And you know why we're all taking this class. So normally in such lectures, I spent the first few minutes motivating the topic and motivating the area. But in this particular case, I feel that it requires more introduction, right? Everybody knows why you're here, right? Everybody knows about deep learning. Everybody at least heard about deep learning, right? About, you know, all the amazing things that. One can do with the techniques developed in the area of deep learning over the last several years. Now things like. Image Understanding. X plus the same. The robotics. Thanks a lot. So we we know and have heard of all these. Amazing. Technologies that have leveraged deep learning to do some very interesting things. So instead of me trying to motivate the different reasons why you should do deep learning, maybe I'll turn it over to you. OK, so. So indeed, a recent success story of deep learning and AI has been this outstanding problem and. On the other. Protein physics, it's called protein folding, which we'll talk about, OK. Notification, OK, we can keep going right? So. Long story short. Let's talk about applications, right? So there's a ton of applications and in this in this course. By the end of the call. We will have gained enough knowledge to kind of sort of address all of these applications. So the ones that you pick up, this course can be used to address all these applications. Of course, since we only have a finite amount of time, I will not be able to get into the details of every one of them, but we will indeed touch upon several of these. So we'll talk quite a bit about image understanding text. Little bit of. Synthesis, I'll talk about gaming and also project folder and maybe a little bit of finance. So these will be some of the applications will dive a bit deeper into, but the list is really much larger than what we can cover in a 14 week class, OK. So. You know, by the end of the class, you know, all of us will be able to at least start thinking about problems in each of these areas. And as we can see, the tools are so broadly applicable and so easy to pick up that it's really not a stretch to say that any. Pretty much any application which involves a reasonable amount of data can be. At least to be solved using the techniques of. So that's very brief. Discussion of the kinds of things that we can do.

What's the definition of? OK, So what is machine learning? I think of machine learning as a box. OK. Machine learning as a procedure, an algorithm if you will, or a technique, A set of techniques which. Texting us input. So this is a systems level view of machine, so the input and output. So I think our machine learning is a procedure. Input to this procedure is data and the output is information. So machine learning, very broadly construed, is a technique or setup techniques. Which receives data as input and produces actionable information. OK. Very broadly, let's think of what we will do in the next several weeks as a form of machine learning. OK. And now the next question is how does. So. What? What are the essential ingredients of machine? What goes into machine? What is needed? What are the components of machine learning or of any machine learning system? What goes into it?

Recipe. This is not the only way to think about these things. So this is my way of thinking about these things so. For this class and I in in our worldview, we will form the following recipe for the machine learning. OK, so here's the recipe for any machine learning. So here's the recipe. You will have three ingredients, so one is the representation. Three ingredients for any machine learning system. The 1st is the representation. The second is the measure of goodness. And the third is optimize the measure. So our definition of machine learning will have three ingredients, right? So there will be three ingredients of representation. And a method to optimize that message. OK, so three things. So what? What are each of these concepts? So let's begin a little bit. So what's your representation? Representation will consist of the way that we encode the input to the system as data. So there's gonna be different ways of coding. Writing out the input data so those will be part of representation. So in this context, talk about things like features. Different types of features which we'll be interested in, different ways to figure out what features work, what features don't work, and so on. This will also include. How we represent the output and what form does the information that we extract? What form is that produced that will also be part of the representation? And this will also include. What the system looks like and how do we computationally talk about the system itself? OK, and there, you know, concepts such as models. OK, so when you talk about a model, we are really talking about the representation of the machine learning system. Features, , output features, output features, models, and. Everything to do with deep learning. machine learning with the particular form of model which we'll come to the end of this. OK, so, so this is the first, the first and perhaps the most important ingredient, OK, So if you come up, if you come across a new application. First thing that you have to think about is their representation. How do we represent the input, how do we represent the output, and how do we represent the model? OK, once. In fact, this is, I would say 75% of the work if you could do this well, the rest are sort of. Let's type standard, right? So, so ingredient #1 and perhaps the most important ingredient is the representation. OK, so that's number one ingredient #2 is a measure of goodness. And I already saw a couple of. How much of the chat talking about loss functions so yes so. Why do we need something like this? Well, if a machine learning system produces a particular output, you know some some kind of information. Figure out whether it's useful or not, right? So we need to figure out whether the information that it's producing is actually good or not. So to measure that, we need the metric of goodness. And then machine learning speak. We call this the loss function. Very often the loss function that we use the learned model is going to be different from the metric used to figure out how good the model is. So a loss function is not always the same as the accuracy metric. So both are important. So sometimes you use one loss function, but you measure performance in terms of a different quantity. So be aware of that distinction in machine learning and also in deep learning as we go. OK, so agree #2 is your goodness. How well does a particular representation produce information relative to a particular task that you're interested in? And finally, the learning part, right? So as many of you already mentioned, the essence of machine learning is an algorithm which iteratively improves its performance as more data comes from, right? So in some sense. Model that measure of goodness that you'll be fine, right? So it incrementally improves that measure of goodness and gets better and better at a particular task as more data comes along. So we need a computational routine or a technique to optimize the measure of goodness that you defined. OK, so this is really the. Training. We'll see several algorithms as we proceed through the course. Lot of few Standard methods merged as best practices, but we will discuss several other algorithms which also are equally in the context of deep learning. But really, you know, that's the that's the third ingredient. In a course like this. Because, you know, this is an academic exercise and we are all in the class.

We spend a lot of so we'll talk about different loss functions, different, different, uh, you know. So on and so forth, you know. Typical in these kinds of problems, right? But I would say that as I let me reiterate that. The real. The real impact consideration comes in the first part, the representation. If you figured out a good representation, that's all you have to do, So keep that in. And uh. In. In summary, what is machine learning? The three-step recipe? OK, so the three ingredients. Representation. Goodness at the door.

Concrete example. Simple. Are you scared of this? For let's put this. Like this. OK, so concrete example and let's start really really simple and that is. The classical example of it. Alright, so let's put this recipe into practice in the context of linear regression. So we will apply these three steps and derive. Using these three steps, it's also. It'll also help us set up some mathematical notation which would be useful for the next couple of months. So what is linear regression? So here's a, here's a perhaps. Motivating application. OK. So finance, yeah, so you are? Currency. Currency Trader. So what what you're doing is you're trying to predict? I don't know the value of some currency of Bitcoin which can be trying to predict the value of Bitcoin. What's the point of? What's the point? Etherium. Not good. Not a good example because I don't know what it depends on. Let's pretend popularity. Let's pretend. I'm going to pick up an example, so let's just for whatever reason, pretend that dodge coin is. Depends on, you know, the SP500. It's not true. But let's just pretend, right? So maybe the fire goes up, you know? Value of gold goes down and vice versa. Ohh, certain, you know, certain companies in the S&P 500 impact the price of gold client more than others, right? Like Elon Musk read something, you know, and then load fan goes up. So maybe you know, buy some Tesla has gone down and that's fine. I don't know I don't know what's happening, but whatever, right So maybe. That big price of somehow dependent on. That's the SP 500 is. A basket of 500 stocks. OK, so it's yeah, Benjamin says index fund yeah, it's it's set of 500 stocks and index it's a measure of how the economy is working. So they made certain stocks higher and certain sort of lower. I believe it's based on. I don't know, not of another trader, but essentially it's a bunch of 500 stocks. OK, so let's hypothesize the price of Dogecoin every day. Depends on It depends on some combination of the 500 stocks in the SMP file. OK, so here's. So that's a problem. OK, I want to predict. I want to build a machine learning model that will predict to me the price of dodge coin given the instantaneous state of the SNP file. OK, that's the problem. So input is all the stocks in the SP500, output is the price. So how will I write this out well? Let me let me think of training a machine learning model. For this problem, so the requirement upfront, we have to give me examples of the SMP file I wanted on different days. OK, so let me write it out so. So I'll use X to denote inputs and Y to denote code puts. OK, so the S&P 500 will be. A bunch of. 500 different numbers. So the instantaneous state of the SP 500 is a bunch of 500 different numbers in general here, so let's call it. D but D is fine. OK, so these 500. In this particular example, and the output is. There's a price of Dodge coin on that particular day. OK, so so I'll write out. The input in the form of a D dimensional vector or an array of size D and the output is one number scalar. OK, so that's that's my prediction problem. So I need a bunch of example data points. This is one data point, right? So input, output. So input features and output. Input output. That's one data point. I'll need a bunch of such. Uh, input output pairs. OK, so this is this is let's say. So one input, 1 output and you give me a bunch of such. OK. Where N is the number of data points that we have observed the price of Dodge Coin over the last say 2 years and also the stock prices over the last two years. So that's like 700 and. 572720 data points, yeah, so and. 20 and is fine. Hey. Funny boy. What? That's a lot. What time is it? That's all. Yeah. Why does the new? George. Remove this table Alright, let me see. Let's see what? At the beginning.

So the first step this is this is, this is what was given to you and you need to learn the machine learning model which will predict the price of. So let me start with the 1st. Step in our recipe The first step in our recipe is representation. But we never figure out. The form of the machine learning model which will transform the input to the output. OK, here there is an arbitrary amount of creativity that can go into this, right? So there's many, many ways in which we can take in a bunch of 500 numbers and produce 1 number, right? Like really large number of ways you can do this. So, uh, let me hypothesize a linear representation, a linear model. Which will do the following. It will it will produce to me an output Y which is a linear combination of the input variables. OK, so a linear combination would be sum of. The sum of the input variables weighted by some weights. OK, so it's a linear model. So there's two actually 2 slightly different definitions of linear, right? So the way I've written it is what a linear algebra is like a linear algebra theorist would call linear in practice. We also add the possibility of an offset. OK, so this essentially, so maybe I'm waving my hands, but hopefully this is clear. A linear model is straight line, right? Didn't model with offset to the same straight line, but shifted away from the origin. OK, So that's sort of. This quantity is 0. Then you know the zero input will produce a zero or OK. So this is my hypothesis for the behavior of Deutschland. It's a linear. So I hypothesize client obeys a linear model. OK, so in machine learning language. These quantities have. I said before that let me. Let me make another simplification here. Everybody's familiar with basic linear algebra. Vectors, dot products, matrices and so on. Everybody familiar with that, Yeah. So if you're not, you know you need to refresh your.

But, you know, instead of having summations and indices all over the place, it's it's going to get very confusing. So instead of doing, you know, writing out everything in terms of summations and indices, we typically like to use linear algebraic notation. Write this out as a. We'll write this out as a dot product dot product of. So if I think of. Factor XI then the output is the dot product of the input. With some other vector w plus the outside b. So this is just an easier way of writing out the same equation. These are special names. Will be called. The set, the set of entries and we call the weights and the offset. The bias OK, so weight and bias. Ohh this is you know I got. Nobody is forcing me to come up with the representation. I came up with this representation out of the blue. OK, so I hypothesized that the behavior is linear and therefore I wrote down a linear model. Next lecture we'll see, you know, other ways or in fact in about half an hour we'll see other ways of. Finding out other different models for the same problem, but for simplicity we'll stick to. Out of curiosity, why are you here? What is nice about linear representation for linear linear models? You're easy to understand. Explain. Yes. Operation Symphonies. Musical Predictions. The difference is we all understand straight lines, right? Princess. Excellent. OK, so first there really simple like I we all understand straight lines, you know, that's the first thing we learn to draw and like electric class, right? Which also. Through that linear model enable interpretability. What does that mean? That means the following, right? So if. I look at the SP500 right and the dodge coin example right? So if I hypothesize, if I hypothesize that the behavior is linear, that means that that implies that if I keep everything else the same. Except the Tesla stock, right? If the if everything else is the same but Tesla increases by 20% then. The price of dodge coin increases by 20 times the weight, right? So there's a very clear meaning of what the weights are in linear models, right? So if everything kept constant. What happens if I 11 star? Approach claim. Goes up by 2 \* W W is the corresponding weight of that particular stock. OK, so linearity is not only conceptually simple and as as you say, easy to compute and so on. It's also human interpretable. As you'll see, it's less. Clear how to interpret deep learning models. You know, we we come to that maybe a couple of lectures time, OK, but yeah, the linearity is both potentially simple as well as intuitively simple. OK, that's a good place to start. So first ingredient representation. We have written down the input, we have written down the output and we have written down the form of the machine learning model which will predict output from.