One vs. All

[0:00](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all)

[In this video we'll talk about how to get logistic regression to work for](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [multiclass classification problems.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [And in particular I want to tell you about an algorithm called one-versus-all](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [classification.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all)

[0:12](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all)

[What's a multiclass classification problem?](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [Here are some examples.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [Lets say you want a learning algorithm to automatically put your email into](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [different folders or to automatically tag your emails so](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [you might have different folders or different tags for work email,](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [email from your friends, email from your family, and emails about your hobby.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [And so here we have a classification problem with four classes which we might](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [assign to the classes y = 1, y =2, y =3, and y = 4 too.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [And another example, for medical diagnosis,](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [if a patient comes into your office with maybe a stuffy nose,](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [the possible diagnosis could be that they're not ill.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [Maybe that's y = 1.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [Or they have a cold, 2.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [Or they have a flu.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all)

[0:59](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all)

[And a third and final example if you are using machine learning to classify](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [the weather, you know maybe you want to decide that the weather is sunny, cloudy,](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [rainy, or snow, or if it's gonna be snow, and so in all of these examples,](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [y can take on a small number of values, maybe one to three, one to four and](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [so on, and these are multiclass classification problems.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [And by the way, it doesn't really matter whether we index is at 0, 1,](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [2, 3, or as 1, 2, 3, 4.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [I tend to index my classes starting from 1 rather than starting from 0,](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [but either way we're off and it really doesn't matter.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [Whereas previously for](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [a binary classification problem, our data sets look like this.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [For a multi-class classification problem our data sets may look like](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [this where here I'm using three different symbols to represent our three classes.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [So the question is given the data set with three classes where](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [this is an example of one class, that's an example of a different class, and](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [that's an example of yet a third class.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [How do we get a learning algorithm to work for the setting?](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [We already know how to do binary classification using a regression.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [We know how to you know maybe fit a straight line to set for the positive and](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [negative classes.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [You see an idea called one-vs-all classification.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [We can then take this and make it work for multi-class classification as well.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [Here's how a one-vs-all classification works.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [And this is also sometimes called one-vs-rest.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [Let's say we have a training set like that shown on the left,](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [where we have three classes of y equals 1, we denote that with a triangle,](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [if y equals 2, the square, and if y equals three, then the cross.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [What we're going to do is take our training set and](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [turn this into three separate binary classification problems.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [I'll turn this into three separate two class classification problems.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [So let's start with class one which is the triangle.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [We're gonna essentially create a new sort of fake training set where classes two and](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [three get assigned to the negative class.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [And class one gets assigned to the positive class.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [You want to create a new training set like that shown on the right, and](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [we're going to fit a classifier which I'm going to call h subscript theta](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [superscript one of x where here](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [the triangles are the positive examples and the circles are the negative examples.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [So think of the triangles being assigned the value of one and](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [the circles assigned the value of zero.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [And we're just going to train a standard logistic regression classifier and](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [maybe that will give us a position boundary that looks like that.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [Okay?](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all)

[3:34](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all)

[This superscript one here stands for class one, so we're doing this for](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [the triangles of class one.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [Next we do the same thing for class two.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [Gonna take the squares and assign the squares as the positive class, and](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [assign everything else, the triangles and the crosses, as a negative class.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [And then we fit a second logistic regression classifier and call this h of](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [x superscript two, where the superscript two denotes that we're now doing this,](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [treating the square class as the positive class.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [And maybe we get classified like that.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [And finally, we do the same thing for the third class and](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [fit a third classifier h super script three of x, and](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [maybe this will give us a decision bounty of the visible cross fire.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [This separates the positive and negative examples like that.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all)

[4:22](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all)

[So to summarize, what we've done is, we've fit three classifiers.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [So, for i = 1, 2,](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [3, we'll fit a classifier x super script i subscript theta of x.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [Thus trying to estimate what is the probability that y is equal to class i,](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [given x and parametrized by theta.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [Right? So in the first instance for](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [this first one up here, this classifier was learning to recognize the triangles.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [So it's thinking of the triangles as a positive clause, so](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [x superscript one is essentially trying to estimate what is the probability](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [that the y is equal to one, given that x is parametrized by theta.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [And similarly, this is treating the square class as a positive class and](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [so it's trying to estimate the probability that y = 2 and so on.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [So we now have three classifiers,](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [each of which was trained to recognize one of the three classes.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [Just to summarize, what we've done is we want to](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [train a logistic regression classifier h superscript i of x for](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [each class i to predict the probability that y is equal to i.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [Finally to make a prediction, when we're given a new input x, and](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [we want to make a prediction.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [What we do is we just run all three of our](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [classifiers on the input x and we then pick the class i that maximizes the three.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [So we just basically pick the classifier,](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [I think whichever one of the three classifiers is most confident and so](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [the most enthusiastically says that it thinks it has the right clause.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [So whichever value of i gives us the highest probability](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [we then predict y to be that value.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all)

[6:02](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all)

[So that's it for multi-class classification and one-vs-all method.](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [And with this little method you can now take the logistic regression](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all) [classifier and make it work on multi-class classification problems as well](https://www.coursera.org/learn/machine-learning/lecture/68Pol/multiclass-classification-one-vs-all)

Downloads

* [**Lecture Video**mp4](https://d3c33hcgiwev3.cloudfront.net/06.7-LogisticRegression-MultiClassClassificationOneVsAll.a9eb1050b22b11e4a416e948628da1fd/full/360p/index.mp4?Expires=1470268800&Signature=fmlB0v-nuQ1BaM8ndiX6brB~FtD1vBZJmU6n2AN9XidjDnRxzIafwrLeyYi5UBLn67IXjuBbl30~1eCc~8DPndkLLN-RlIfhgvTS4qCNETU3kqtfhRTnWGhidcJHUdJ20Yzq63ruJCzybQMKCc89szYCYXV-IKdZNac0AKNLg~s_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A)

**[Subtitles (English)](https://www.coursera.org/api/subtitleAssetProxy.v1/b0pgK5zdQROKYCuc3fETZg?expiry=1470268800000&hmac=7AUtg0me0Eo6fFMVzAMl0cSNBRpdS6p0wmJYNryDnhk&fileExtension=vtt)**[WebVTT](https://www.coursera.org/api/subtitleAssetProxy.v1/b0pgK5zdQROKYCuc3fETZg?expiry=1470268800000&hmac=7AUtg0me0Eo6fFMVzAMl0cSNBRpdS6p0wmJYNryDnhk&fileExtension=vtt)

**[Transcript (English)](https://www.coursera.org/api/subtitleAssetProxy.v1/b0pgK5zdQROKYCuc3fETZg?expiry=1470268800000&hmac=X9obFbVDYRlDQm0JgjnkCKlkqYNyzqH0Ro8R8lwBtPQ&fileExtension=txt)**

Neural networks

[0:00](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[In this and in the](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [next set of videos, I'd like](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [to tell you about a learning](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [algorithm called a Neural Network.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[0:07](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[We're going to first talk about](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [the representation and then](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [in the next set of videos](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [talk about learning algorithms for it.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [Neutral networks is actually](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [a pretty old idea, but had](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [fallen out of favor for a while.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [But today, it is the](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [state of the art technique for](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [many different machine learning problems.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[0:23](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[So why do we need yet another learning algorithm?](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [We already have linear regression and](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [we have logistic regression, so why do we need, you know, neural networks?](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[0:32](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[In order to motivate the discussion](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [of neural networks, let me](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [start by showing you a few](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [examples of machine learning](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [problems where we need](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [to learn complex non-linear hypotheses.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[0:43](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[Consider a supervised learning classification](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [problem where you have a training set like this.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [If you want to apply logistic](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [regression to this problem, one](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [thing you could do is apply](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [logistic regression with a](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [lot of nonlinear features like that.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [So here, g as usual](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [is the sigmoid function, and we](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [can include lots of polynomial terms like these.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [And, if you include enough polynomial](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [terms then, you know, maybe](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [you can get a hypotheses](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[1:11](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[that separates the positive and negative examples.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [This particular method works well](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [when you have only, say, two](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [features - x1 and x2](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [- because you can then include](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [all those polynomial terms of](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [x1 and x2.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [But for many interesting machine learning](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [problems would have a](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [lot more features than just two.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[1:30](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[We've been talking for a while](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [about housing prediction, and suppose](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [you have a housing classification](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[1:38](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[problem rather than a](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [regression problem, like maybe](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [if you have different features of](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [a house, and you want](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [to predict what are the](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [odds that your house will](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [be sold within the next](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [six months, so that will be a classification problem.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[1:52](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[And as we saw we can](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [come up with quite a](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [lot of features, maybe a hundred](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [different features of different houses.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[2:00](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[For a problem like this, if](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [you were to include all the](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [quadratic terms, all of](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [these, even all of the](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [quadratic that is the second](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [or the polynomial terms, there would be a lot of them.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [There would be terms like x1 squared,](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[2:12](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[x1x2, x1x3, you know, x1x4](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[2:18](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[up to x1x100 and then](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [you have x2 squared, x2x3](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[2:25](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[and so on.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [And if you include just](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [the second order terms, that](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [is, the terms that are](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [a product of, you know,](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [two of these terms, x1](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [times x1 and so on, then,](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [for the case of n equals](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[2:38](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[100, you end up with about five thousand features.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[2:41](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[And, asymptotically, the](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [number of quadratic features grows](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [roughly as order n](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [squared, where n is the](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [number of the original features,](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [like x1 through x100 that we had.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [And its actually closer to n squared over two.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[2:59](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[So including all the](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [quadratic features doesn't seem](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [like it's maybe a good](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [idea, because that is a](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [lot of features and you](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [might up overfitting the training](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [set, and it can](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [also be computationally expensive, you know, to](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[3:14](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[be working with that many features.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[3:16](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[One thing you could do is](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [include only a subset of](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [these, so if you include only the](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [features x1 squared, x2 squared,](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [x3 squared, up to](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [maybe x100 squared, then](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [the number of features is much smaller.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [Here you have only 100 such](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [quadratic features, but this](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [is not enough features and](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [certainly won't let you fit](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [the data set like that on the upper left.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [In fact, if you include](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [only these quadratic features together](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [with the original x1, and](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [so on, up to x100 features,](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [then you can actually fit very](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [interesting hypotheses. So, you](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [can fit things like, you know, access a](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [line of the ellipses like these, but](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[3:55](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[you certainly cannot fit a more](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [complex data set like that shown here.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[3:59](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[So 5000 features seems like](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [a lot, if you were](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [to include the cubic, or](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [third order known of each others,](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [the x1, x2, x3.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [You know, x1 squared,](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [x2, x10 and](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [x11, x17 and so on.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [You can imagine there are gonna be a lot of these features.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [In fact, they are going to be](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [order and cube such features](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [and if any is 100](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [you can compute that, you](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [end up with on the order](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [of about 170,000 such cubic](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [features and so including](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [these higher auto-polynomial features when](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [your original feature set end](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [is large this really dramatically](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [blows up your feature space and](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [this doesn't seem like a](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [good way to come up with](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [additional features with which](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [to build none many classifiers when n is large.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[4:49](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[For many machine learning problems, n will be pretty large.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [Here's an example.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[4:55](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[Let's consider the problem of computer vision.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[4:59](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[And suppose you want to](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [use machine learning to train](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [a classifier to examine an](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [image and tell us whether](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [or not the image is a car.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[5:09](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[Many people wonder why computer vision could be difficult.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [I mean when you and I](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [look at this picture it is so obvious what this is.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [You wonder how is it](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [that a learning algorithm could possibly](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [fail to know what this picture is.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[5:22](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[To understand why computer vision](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [is hard let's zoom](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [into a small part of the](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [image like that area where the](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [little red rectangle is.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [It turns out that where you](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [and I see a car, the computer sees that.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [What it sees is this matrix,](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [or this grid, of pixel](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [intensity values that tells](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [us the brightness of each pixel in the image.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [So the computer vision problem is](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [to look at this matrix of](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [pixel intensity values, and tell](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [us that these numbers represent the door handle of a car.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[5:54](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[Concretely, when we use](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [machine learning to build a](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [car detector, what we do](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [is we come up with a](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [label training set, with, let's](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [say, a few label examples](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [of cars and a few](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [label examples of things that](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [are not cars, then we](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [give our training set to](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [the learning algorithm trained a](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [classifier and then, you](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [know, we may test it and show](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [the new image and ask, "What is this new thing?".](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[6:17](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[And hopefully it will recognize that that is a car.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[6:21](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[To understand why we](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [need nonlinear hypotheses, let's take](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [a look at some of the](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [images of cars and maybe](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [non-cars that we might feed to our learning algorithm.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[6:32](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[Let's pick a couple of pixel](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [locations in our images, so](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [that's pixel one location and](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [pixel two location, and let's](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [plot this car, you know, at the](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [location, at a certain](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [point, depending on the intensities](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [of pixel one and pixel two.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[6:49](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[And let's do this with a few other images.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [So let's take a different example](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [of the car and you know,](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [look at the same two pixel locations](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[6:56](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[and that image has a](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [different intensity for pixel one](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [and a different intensity for pixel two.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [So, it ends up at a different location on the figure.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [And then let's plot some negative examples as well.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [That's a non-car, that's a](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [non-car .](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [And if we do this for](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [more and more examples using](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [the pluses to denote cars](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [and minuses to denote non-cars,](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [what we'll find is that](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [the cars and non-cars end up](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [lying in different regions of](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [the space, and what we](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [need therefore is some sort](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [of non-linear hypotheses to try](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [to separate out the two classes.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[7:32](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[What is the dimension of the feature space?](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [Suppose we were to use just 50 by 50 pixel images.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [Now that suppose our images were](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [pretty small ones, just 50 pixels on the side.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [Then we would have 2500 pixels,](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[7:46](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[and so the dimension of](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [our feature size will be N](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [equals 2500 where our feature](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [vector x is a list](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [of all the pixel testings, you](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [know, the pixel brightness of pixel](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [one, the brightness of pixel](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [two, and so on down](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [to the pixel brightness of the](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [last pixel where, you know, in a](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [typical computer representation, each of](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [these may be values between say](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [0 to 255 if it gives](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [us the grayscale value.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [So we have n equals 2500,](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [and that's if we](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [were using grayscale images.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [If we were using RGB](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [images with separate red, green](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [and blue values, we would have n equals 7500.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[8:27](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[So, if we were to](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [try to learn a nonlinear](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [hypothesis by including all](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [the quadratic features, that is](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [all the terms of the form, you know,](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [Xi times Xj, while with the](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [2500 pixels we would end](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [up with a total of three million features.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [And that's just too large to](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [be reasonable; the computation would](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [be very expensive to find and](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [to represent all of these](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [three million features per training example.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[8:55](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[So, simple logistic regression together](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [with adding in maybe the](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [quadratic or the cubic features](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [- that's just not a](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [good way to learn complex](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [nonlinear hypotheses when n](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [is large because you just end up with too many features.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [In the next few videos, I](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [would like to tell you about Neural](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [Networks, which turns out](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [to be a much better way to](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [learn complex hypotheses, complex nonlinear](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [hypotheses even when your](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [input feature space, even when n is large.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [And along the way I'll](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [also get to show](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [you a couple of fun videos](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [of historically important applications](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[9:30](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[of Neural networks as well that I](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [hope those videos that](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses) [we'll see later will be fun for you to watch as well.](https://www.coursera.org/learn/machine-learning/lecture/OAOhO/non-linear-hypotheses)

[0:00](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[In this video, I want to start telling you about how we represent neural networks.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [In other words, how we represent our hypothesis or](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [how we represent our model when using neural networks.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [Neural networks were developed as simulating neurons or](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [networks of neurons in the brain.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [So, to explain the hypothesis representation](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [let's start by looking at what a single neuron in the brain looks like.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [Your brain and](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [mine is jam packed full of neurons like these and neurons are cells in the brain.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [And two things to draw attention to are that first.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [The neuron has a cell body, like so, and moreover,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [the neuron has a number of input wires, and these are called the dendrites.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [You think of them as input wires, and these receive inputs from other locations.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [And a neuron also has an output wire called an Axon, and](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [this output wire is what it uses to send signals to other neurons,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [so to send messages to other neurons.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [So, at a simplistic level what a neuron is, is a computational unit that](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [gets a number of inputs through it input wires and does some computation and](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [then it says outputs via its axon to other nodes or to other neurons in the brain.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [Here's a illustration of a group of neurons.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [The way that neurons communicate with each other is with little pulses of](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [electricity, they are also called spikes but that just means pulses of electricity.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [So here is one neuron and what it does is if it wants a send a message what it](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [does is sends a little pulse of electricity.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [Varis axon to some different neuron and here, this axon that is this open wire,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [connects to the dendrites of this second neuron over here,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [which then accepts this incoming message that some computation.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [And they, in turn, decide to send out this message on this axon to other neurons,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [and this is the process by which all human thought happens.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [It's these Neurons doing computations and](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [passing messages to other neurons as a result of what other inputs they've got.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [And, by the way, this is how our senses and our muscles work as well.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [If you want to move one of your muscles the way that where else in your neuron may](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [send this electricity to your muscle and that causes your muscles to contract and](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [your eyes, some senses like your eye must send a message to your brain while it](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [does it senses hosts electricity entity to a neuron in your brain like so.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [In a neuro network, or rather, in an artificial neuron network that we've](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [implemented on the computer, we're going to use a very simple model of](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [what a neuron does we're going to model a neuron as just a logistic unit.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [So, when I draw a yellow circle like that, you should think of that as a playing](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [a role analysis, who's maybe the body of a neuron, and](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [we then feed the neuron a few inputs who's various dendrites or input wiles.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[3:14](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[And the neuron does some computation.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [And output some value on this output wire, or](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [in the biological neuron, this is an axon.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [And whenever I draw a diagram like this,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [what this means is that this represents a computation of h of](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [x equals one over one plus e to the negative theta transpose x,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [where as usual, x and theta are our parameter vectors, like so.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[3:42](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[So this is a very simple, maybe a vastly oversimplified model,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [of the computations that the neuron does, where it gets a number of inputs, x1,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [x2, x3 and it outputs some value computed like so.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[3:59](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[When I draw a neural network, usually I draw only the input nodes x1, x2, x3.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [Sometimes when it's useful to do so, I'll draw an extra node for x0.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[4:11](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[This x0 now that's sometimes called the bias unit or the bias neuron, but](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [because x0 is already equal to 1, sometimes, I draw this, sometimes](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [I won't just depending on whatever is more notationally convenient for that example.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[4:31](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[Finally, one last bit of terminology when we talk about neural networks,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [sometimes we'll say that this is a neuron or](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [an artificial neuron with a Sigmoid or logistic activation function.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [So this activation function in the neural network terminology.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [This is just another term for that function for](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [that non-linearity g(z) = 1 over 1+e to the -z.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [And whereas so far I've been calling theta the parameters of the model,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [I'll mostly continue to use that terminology.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [Here, it's a copy to the parameters, but in neural networks, in the neural network](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [literature sometimes you might hear people talk about weights of a model and](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [weights just means exactly the same thing as parameters of a model.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [But I'll mostly continue to use the terminology parameters in these videos,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [but sometimes, you might hear others use the weights terminology.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[5:27](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[So, this little diagram represents a single neuron.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[5:34](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[What a neural network is, is just a group of this different neurons strong together.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [Completely, here we have input units x1, x2, x3 and once again,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [sometimes you can draw this extra note x0 and Sometimes not, just flow that in here.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [And here we have three neurons which have written 81, 82, 83.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [I'll talk about those indices later.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [And once again we can if we want add in just a0 and](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [add the mixture bias unit there.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [There's always a value of 1.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [And then finally we have this third node and the final layer, and](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [there's this third node that outputs the value that the hypothesis h(x) computes.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [To introduce a bit more terminology, in a neural network,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [the first layer, this is also called the input layer because this is where we](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [Input our features, x1, x2, x3.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [The final layer is also called the output layer because that layer has a neuron,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [this one over here, that outputs the final value computed by a hypothesis.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [And then, layer 2 in between, this is called the hidden layer.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [The term hidden layer isn't a great terminology, but this ideation is that,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [you know, you supervised early,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [where you get to see the inputs and get to see the correct outputs, where](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [there's a hidden layer of values you don't get to observe in the training setup.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [It's not x, and it's not y, and so we call those hidden.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [And they try to see neural nets with more than one hidden layer but](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [in this example, we have one input layer, Layer 1, one hidden layer, Layer 2,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [and one output layer, Layer 3.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [But basically, anything that isn't an input layer and](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [isn't an output layer is called a hidden layer.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[7:29](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[So I want to be really clear about what this neural network is doing.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [Let's step through the computational steps that are and](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [body represented by this diagram.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [To explain these specific computations represented by a neural network,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [here's a little bit more notation.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [I'm going to use a superscript j subscript i to denote the activation](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [of neuron i or of unit i in layer j.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [So completely this gave superscript to sub group one,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [that's the activation of the first unit in layer two, in our hidden layer.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [And by activation I just mean the value that's computed by and](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [as output by a specific.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [In addition, new network is parametrize by these matrixes, theta](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [super script j Where theta j is going to be a matrix of weights controlling](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [the function mapping form one layer, maybe the first layer to the second layer,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [or from the second layer to the third layer.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[8:30](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[So here are the computations that are represented by this diagram.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[8:34](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[This first hidden unit here has it's value computed as follows,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [there's a is a21 is equal to the sigma function of the sigma activation function,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [also called the logistics activation function,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [apply to this sort of linear combination of these inputs.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [And then this second hidden unit has this activation](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [value computer as sigmoid of this.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [And similarly for this third hidden unit is computed by that formula.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [So here we have 3 theta 1 which is matrix](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [of parameters governing our mapping](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [from our three different units, our hidden units.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [Theta 1 is going to be a 3.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[9:35](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[Theta 1 is going to be a 3x4-dimensional matrix.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [And more generally, if a network has SJU units in there j and](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [sj + 1 units and sj + 1 then the matrix theta j](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [which governs the function mapping from there sj + 1.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [That will have to mention sj +1 by sj + 1 I'll just be clear](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [about this notation right.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [This is Subscript j + 1 and that's s subscript j, and](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [then this whole thing, plus 1, this whole thing (sj + 1), okay?](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [So that's s subscript j + 1 by,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[10:21](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[So that's s subscript j + 1 by sj + 1 where](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [this plus one is not part of the subscript.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [Okay, so we talked about what the three hidden units do to compute their values.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [Finally, there's a loss of this final and](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [after that we have one more unit which computer h of x and](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [that's equal can also be written as a(3)1 and that's equal to this.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [And you notice that I've written this with a superscript two here,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [because theta of superscript two is the matrix of parameters, or](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [the matrix of weights that controls the function that maps from the hidden units,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [that is the layer two units to the one layer three unit, that is the output unit.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [To summarize, what we've done is shown how a picture like this over here defines](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [an artificial neural network which defines a function h](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [that maps with x's input values to hopefully to some space that provisions y.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [And these hypothesis are parameterized by parameters](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [denoting with a capital theta so that, as we vary theta,](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [we get different hypothesis and we get different functions.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [Mapping say from x to y.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[11:42](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

[So this gives us a mathematical definition of how to represent the hypothesis in](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [the neural network.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [In the next few videos what I would like to do is give you more intuition about](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [what these hypothesis representations do, as well as go through a few examples and](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i) [talk about how to compute them efficiently.](https://www.coursera.org/learn/machine-learning/lecture/ka3jK/model-representation-i)

Downloads

* [**Lecture Video**mp4](https://d3c33hcgiwev3.cloudfront.net/08.3-NeuralNetworksRepresentation-ModelRepresentationI.d459a4a0b22b11e4beb61117ba5cda9e/full/360p/index.mp4?Expires=1470268800&Signature=W8FZqxBLNGghdgFOQsk3n8kD763XRWaAp-WWz4XEzc7rdY6NCOG8naUewYCKSheaInPFrRqzzraUGFnnKPoHxjXe~CtuTm-OzU76Yxmor3~w-glhZQ1kP-Rgy8m6ES60LBOZ-sa8uVpCkFSUqfTQpPavz5WFwa3nJfAVjPdV5h0_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A)

**[Subtitles (English)](https://www.coursera.org/api/subtitleAssetProxy.v1/dOF88pENQ5qhfPKRDaOaXg?expiry=1470268800000&hmac=yuVRDC6IAHnVfn9iztTlsvqoU7kj7mKwxfjSx9hM4sA&fileExtension=vtt)**[WebVTT](https://www.coursera.org/api/subtitleAssetProxy.v1/dOF88pENQ5qhfPKRDaOaXg?expiry=1470268800000&hmac=yuVRDC6IAHnVfn9iztTlsvqoU7kj7mKwxfjSx9hM4sA&fileExtension=vtt)

**[Transcript (English)](https://www.coursera.org/api/subtitleAssetProxy.v1/dOF88pENQ5qhfPKRDaOaXg?expiry=1470268800000&hmac=gPbzPlTqmQJGNF0CClt3gSTxtzTzs3bIp9o3MxXRJ5Q&fileExtension=txt)**

[0:00](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[In the last video, we gave](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [a mathematical definition of how](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [to represent or how to](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [compute the hypotheses used by Neural Network.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[0:08](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[In this video, I like](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [show you how to actually](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [carry out that computation efficiently, and](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [that is show you a vector rise implementation.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[0:17](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[And second, and more importantly, I want](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [to start giving you intuition about](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [why these neural network representations](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [might be a good idea and how](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [they can help us to learn complex nonlinear hypotheses.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[0:28](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[Consider this neural network.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [Previously we said that](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [the sequence of steps that we](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [need in order to compute](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [the output of a hypotheses](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [is these equations given on](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [the left where we compute](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [the activation values of the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [three hidden uses and then](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [we use those to compute the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [final output of our hypotheses](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [h of x. Now, I'm going](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [to define a few extra terms.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [So, this term that I'm](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [underlining here, I'm going to](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [define that to be](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [z superscript 2 subscript 1.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [So that we have that](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [a(2)1, which is this](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [term is equal to](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [g of z to 1.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [And by the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [way, these superscript 2, you](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [know, what that means is that](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [the z2 and this a2](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [as well, the superscript](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [2 in parentheses means that these](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [are values associated with layer](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [2, that is with the hidden](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [layer in the neural network.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[1:22](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[Now this term here](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [I'm going to similarly define as](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[1:29](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[z(2)2.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [And finally, this last](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [term here that I'm underlining,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[1:34](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[let me define that as z(2)3.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [So that similarly we have a(2)3](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [equals g of](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[1:44](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[z(2)3.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [So these z values are just](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [a linear combination, a weighted](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [linear combination, of the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [input values x0, x1,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [x2, x3 that go into a particular neuron.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[1:57](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[Now if you look at](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [this block of numbers,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[2:01](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[you may notice that that block](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [of numbers corresponds suspiciously similar](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[2:06](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[to the matrix vector](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [operation, matrix vector multiplication](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [of x1 times the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [vector x. Using this observation](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [we're going to be able to vectorize this computation](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [of the neural network.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[2:21](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[Concretely, let's define the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [feature vector x as usual](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [to be the vector of x0, x1,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [x2, x3 where x0](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [as usual is always equal](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [1 and that defines](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [z2 to be the vector](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [of these z-values, you know, of z(2)1 z(2)2, z(2)3.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[2:38](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[And notice that, there, z2 this](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [is a three dimensional vector.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[2:43](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[We can now vectorize the computation](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[2:48](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[of a(2)1, a(2)2, a(2)3 as follows.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [We can just write this in two steps.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [We can compute z2 as theta](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [1 times x and that](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [would give us this vector z2;](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [and then a2 is](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [g of z2 and just](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [to be clear z2 here, This](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [is a three-dimensional vector and](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [a2 is also a three-dimensional](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [vector and thus this](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [activation g. This applies the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [sigmoid function element-wise to each](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [of the z2's elements. And](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [by the way, to make our notation](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [a little more consistent with](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [what we'll do later, in this](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [input layer we have the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [inputs x, but we](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [can also thing it is](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [as in activations of the first layers.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [So, if I defined a1 to](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [be equal to x. So,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [the a1 is vector, I can](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [now take this x here](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [and replace this with z2 equals theta1](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [times a1 just by defining](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [a1 to be activations in my input layer.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[3:44](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[Now, with what I've written](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [so far I've now gotten](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [myself the values for a1,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [a2, a3, and really](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [I should put the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [superscripts there as well.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [But I need one more](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [value, which is I also want this a(0)2](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [and that corresponds to](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [a bias unit in the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [hidden layer that goes to the output there.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [Of course, there was a](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [bias unit here too that,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [you know, it just didn't draw](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [under here but to](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [take care of this extra bias unit,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [what we're going to do is add](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [an extra a0 superscript 2,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [that's equal to one, and after](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [taking this step we now have](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [that a2 is going to](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [be a four dimensional feature](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [vector because we just added](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [this extra, you know,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [a0 which is equal to](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [1 corresponding to the bias unit](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [in the hidden layer. And finally,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[4:35](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[to compute the actual](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [value output of our hypotheses, we](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [then simply need to compute](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[4:42](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[z3. So z3 is](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [equal to this term here that I'm just underlining.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [This inner term there is z3.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[4:53](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[And z3 is stated](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [2 times a2 and finally](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [my hypotheses output h of x which](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [is a3 that is](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [the activation of my](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [one and only unit in](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [the output layer. So, that's just the real number. You can write it as a3](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [or as a(3)1 and that's g of z3.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [This process of computing h of x](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [is also called forward propagation](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[5:19](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[and is called that because we](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [start of with the activations](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [of the input-units and then](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [we sort of forward-propagate that to the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [hidden layer and compute the activations of the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [hidden layer and then we](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [sort of forward propagate that](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [and compute the activations of](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[5:37](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[the output layer, but this process of computing the activations from the input then](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [the hidden then the output layer,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [and that's also called forward propagation](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[5:43](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[and what we just did is](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [we just worked out a vector](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [wise implementation of this](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [procedure. So, if you](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [implement it using these equations](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [that we have on the right, these](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [would give you an efficient way](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [or both of the efficient way of](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [computing h of x.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[5:58](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[This forward propagation view also](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[6:00](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[helps us to understand what](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [Neural Networks might be doing](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [and why they might help us to](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [learn interesting nonlinear hypotheses.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[6:08](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[Consider the following neural network](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [and let's say I cover up](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [the left path of this picture for now.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [If you look at what's left in this picture.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [This looks a lot like](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [logistic regression where what](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [we're doing is we're using](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [that note, that's just the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [logistic regression unit and we're](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [using that to make a](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [prediction h of x. And concretely,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [what the hypotheses is outputting](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [is h of x is going](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [to be equal to g which](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [is my sigmoid activation function times theta 0](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [times a0 is equal](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [to 1 plus theta 1](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[6:45](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[plus theta 2](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [times a2 plus theta](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [3 times a3 whether](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [values a1, a2, a3](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [are those given by these three given units.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[7:01](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[Now, to be actually consistent](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [to my early notation. Actually, we](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [need to, you know, fill in](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [these superscript 2's here everywhere](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[7:12](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[and I also have these](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [indices 1 there because I](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [have only one output unit, but if you focus on the blue parts of the notation.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [This is, you know, this looks](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [awfully like the standard logistic](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [regression model, except that](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [I now have a capital theta instead of lower case theta.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[7:29](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[And what this is](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [doing is just logistic regression.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[7:33](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[But where the features fed into](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [logistic regression are these](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [values computed by the hidden layer.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[7:41](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[Just to say that again, what](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [this neural network is doing is](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [just like logistic regression, except](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [that rather than using the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [original features x1, x2, x3,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[7:52](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[is using these new features a1, a2, a3.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [Again, we'll put the superscripts](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[7:58](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[there, you know, to be consistent with the notation.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[8:02](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[And the cool thing about this,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [is that the features a1, a2,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [a3, they themselves are learned](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [as functions of the input.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[8:10](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[Concretely, the function mapping from](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [layer 1 to layer 2,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [that is determined by some](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [other set of parameters, theta 1.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [So it's as if the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [neural network, instead of being](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [constrained to feed the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [features x1, x2, x3 to logistic regression.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [It gets to](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [learn its own features, a1,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [a2, a3, to feed into the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [logistic regression and as](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [you can imagine depending on](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [what parameters it chooses for](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [theta 1. You can learn some pretty interesting](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [and complex features and therefore](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[8:43](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[you can end up with a](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [better hypotheses than if](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [you were constrained to use](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [the raw features x1, x2 or x3 or if](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [you will constrain to say choose the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [polynomial terms, you know,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [x1, x2, x3, and so on.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [But instead, this algorithm has](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [the flexibility to try](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [to learn whatever features at once, using](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [these a1, a2, a3 in](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [order to feed into this](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [last unit that's essentially](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[9:09](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[a logistic regression here. I realized](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [this example is described as](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [a somewhat high level and so](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [I'm not sure if this intuition](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [of the neural network, you know, having](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [more complex features will quite](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [make sense yet, but if](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [it doesn't yet in the next](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [two videos I'm going to](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [go through a specific example](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [of how a neural network can](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [use this hidden there to compute](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [more complex features to feed](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [into this final output layer](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [and how that can learn more complex hypotheses.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [So, in case what I'm](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [saying here doesn't quite make](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [sense, stick with me](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [for the next two videos and](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [hopefully out there working through](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [those examples this explanation will](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [make a little bit more sense.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [But just the point O. You](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [can have neural networks with](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [other types of diagrams as](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [well, and the way that](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [neural networks are connected, that's called the architecture.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [So the term architecture refers to](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [how the different neurons are connected to each other.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [This is an example](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [of a different neural network architecture](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[10:07](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[and once again you may](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [be able to get this intuition of](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [how the second layer,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [here we have three heading units](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [that are computing some complex](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [function maybe of the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [input layer, and then the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [third layer can take the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [second layer's features and compute](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [even more complex features in layer three](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [so that by the time you get](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [to the output layer, layer four,](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [you can have even more](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [complex features of what](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [you are able to compute in](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [layer three and so get very interesting nonlinear hypotheses.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[10:36](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[By the way, in a network](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [like this, layer one, this is](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [called an input layer. Layer four](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [is still our output layer, and](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [this network has two hidden layers.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [So anything that's not an](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [input layer or an output](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [layer is called a hidden layer.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[10:53](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[So, hopefully from this video](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [you've gotten a sense of](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [how the feed forward propagation step](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [in a neural network works where you](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [start from the activations of](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [the input layer and forward](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [propagate that to the](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [first hidden layer, then the second](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [hidden layer, and then finally the output layer.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [And you also saw how](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [we can vectorize that computation.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[11:13](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[In the next, I realized](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [that some of the intuitions in this](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [video of how, you know, other certain](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [layers are computing complex features of the early layers.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [I realized some of that intuition](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [may be still slightly abstract and kind of a high level.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [And so what I would like](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [to do in the two videos](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [is work through a detailed example](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [of how a neural network can](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [be used to compute nonlinear](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [functions of the input and](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [hope that will give you a](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [good sense of the sorts of](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii) [complex nonlinear hypotheses we can get out of Neural Networks.](https://www.coursera.org/learn/machine-learning/lecture/Hw3VK/model-representation-ii)

[0:00](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[In this video, I want to](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [tell you about how to use neural](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [networks to do multiclass](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [classification where we may](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [have more than one category](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [that we're trying to distinguish amongst.](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [In the last part of](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [the last video, where we](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [had the handwritten digit recognition](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [problem, that was actually](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [a multiclass classification problem because](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [there were ten possible categories](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [for recognizing the digits from](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [0 through 9 and so, if](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [you want us to fill you](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [in on the details of how to do that.](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[0:30](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[The way we do multiclass classification](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[0:32](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[in a neural network is essentially](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [an extension of the one versus all method.](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[0:38](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[So, let's say that we](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [have a computer vision example,](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [where instead of just trying](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [to recognize cars as in](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [the original example that I started off](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [with, but let's say that](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [we're trying to recognize, you know, four](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [categories of objects and given](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [an image we want to](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [decide if it is a pedestrian, a car, a motorcycle or a truck.](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [If that's the case, what](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [we would do is we would](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [build a neural network with four](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [output units so that](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [our neural network now outputs a vector of four numbers.](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[1:09](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[So, the output now is actually](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [needing to be a vector of four](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [numbers and what we're](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [going to try to do is](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [get the first output unit](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [to classify: is the](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [image a pedestrian, yes or no.](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [The second unit to classify: is the image a car, yes or no.](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [This unit to classify: is the](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [image a motorcycle, yes or](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [no, and this would classify:](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [is the image a truck, yes or no.](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [And thus, when the image](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [is of a pedestrian, we](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [would ideally want the network](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [to output 1, 0, 0, 0,](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [when it is a](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [car we want it to output](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [0, 1, 0, 0, when this](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [is a motorcycle, we get it to or rather, we want](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [it to output 0, 0,](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [1, 0 and so on.](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[1:50](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[So this is just like](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [the "one versus all" method](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [that we talked about when we](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [were describing logistic regression, and](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [here we have essentially four logistic](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [regression classifiers, each of](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [which is trying to recognize one](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [of the four classes that](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [we want to distinguish amongst.](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [So, rearranging the slide of](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [it, here's our neural network](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [with four output units and those](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [are what we want h](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [of x to be when we](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [have the different images, and](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [the way we're going to represent the](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [training set in these settings](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [is as follows. So, when we have](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [a training set with different images](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[2:27](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[of pedestrians, cars, motorcycles and](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [trucks, what we're going](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [to do in this example is](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [that whereas previously we had](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [written out the labels as](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [y being an integer from](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [1, 2, 3 or 4. Instead of](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [representing y this way,](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [we're going to instead represent y](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [as follows: namely Yi](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[2:54](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[will be either 1, 0, 0, 0](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [or 0, 1, 0, 0 or 0, 0, 1, 0 or 0, 0, 0, 1 depending on what the](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [corresponding image Xi is.](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [And so one training example](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [will be one pair Xi colon Yi](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[3:04](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[where Xi is an image with, you](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [know one of the four objects and](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [Yi will be one of these vectors.](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[3:10](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[And hopefully, we can find](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [a way to get our](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [Neural Networks to output some](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [value. So, the h of x](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [is approximately y and](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [both h of x and Yi,](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [both of these are going](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [to be in our example,](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [four dimensional vectors when we have four classes.](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[3:31](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[So, that's how you](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [get neural network to do multiclass classification.](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[3:36](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[This wraps up our discussion on](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [how to represent Neural Networks](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [that is on our hypotheses representation.](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[3:42](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

[In the next set of videos, let's](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [start to talk about how take](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [a training set and how to](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification) [automatically learn the parameters of the neural network.](https://www.coursera.org/learn/machine-learning/lecture/gFpiW/multiclass-classification)

Downloads

* [**Lecture Video**mp4](https://d3c33hcgiwev3.cloudfront.net/08.7-NeuralNetworksRepresentation-MultiClassClassification.54188b30b22b11e4beb61117ba5cda9e/full/360p/index.mp4?Expires=1470268800&Signature=Jx3xoo9OetpxfAjKRFmN2HPh2~Ysi6RgX0JQkjsTi34eTEIYu4MCDX-5DHI-wukZVUtRyEtpHbfjzn8xByRIT3gDeOSS2LLzpioLzdnpbJttLRfqaoj7nXTTBCWDFriLSaf~q3vco0FxzxHb5t2zs8ytV0NF0uBxos1bZJU4xPw_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A)

**[Subtitles (English)](https://www.coursera.org/api/subtitleAssetProxy.v1/I2zozepvTzWs6M3qb881iQ?expiry=1470268800000&hmac=7-0BZQUBYc5lKiV5e2-U9SJ0gvKOV17MEe7wdJQMw9c&fileExtension=vtt)**[WebVTT](https://www.coursera.org/api/subtitleAssetProxy.v1/I2zozepvTzWs6M3qb881iQ?expiry=1470268800000&hmac=7-0BZQUBYc5lKiV5e2-U9SJ0gvKOV17MEe7wdJQMw9c&fileExtension=vtt)

**[Transcript (English)](https://www.coursera.org/api/subtitleAssetProxy.v1/I2zozepvTzWs6M3qb881iQ?expiry=1470268800000&hmac=hA5t9Yico2csVJemyNSXgYRs38lQ_j9kt87f_pGvXJ0&fileExtension=txt)**