

Audio file

[Your Recording 20.wav](#)

Transcript

00:00:01 Speaker 2

Welcome to our prerequisites lesson. In this short lesson, we're going to look at three things. First, we can lower the prerequisites for your machine learning journey. 2, how we're really targeting beginners, don't necessarily have any machine learning background. And thirdly, how we're going to be focusing on AWS Sagemaker as our platform for helping us build our machine learning models. So what skills are required to start this course? Well, we expect you to have some Python. Now, we talked about in the previous lesson that we were taking a code first approach. In other words, we're going to be creating and building our models using Python code and Sagemaker, the tool in AWS is going to help us do specific activities along the way to training our model, building our model, and hosting our model ultimately.

00:00:31 Speaker 2

So we do need to have a little bit of Python, but don't worry, we don't need to be a complex Python advanced developer. We just need some basics. We would expect you to be able to write a simple function that's using the def keyword so that we could call a function, it can do something and maybe return a value. But what we'd really like you to be able to do is understand lists. Now here's an example, some Python code where I've got a list defined called numbers, 5 elements in that list, 1234 and five. And why this is important is we are going to be working with tabular data. We're going to be working with house price data. And it will be easy to manipulate when we get the house price data into Python data constructs that we can then manipulate such as lists. So we expect you to be able to store values in a list.

00:01:01 Speaker 2

We're doing here. We expect you to store basic values in normal variables. Here I've got a variable called sum of evens and currently I'm saying that to an integer value of 0.

00:01:08 Speaker 2

We'd expect you to understand what we call iteration. Now iteration is just a really a fancy way of saying to loop over something. So imagine in that list thereof 5 numbers. And I wanted to perform an action on each and everyone. As I wanted to double every element in that list, I would disable for each element in that list do this specific action. So here we're using a for loop. So we're saying for number in numbers. So for each element in that list, I want you to do something. We need to understand conditional statements. Now conditional statements are probably done with the if statement. So if a particular condition is true, then do a specific action. So here we're just doing a test to see if the current number as we are iterating over that list. If that number so modulus 2 operation equals 0, we're basically divided by.

00:01:39 Speaker 2

Remainder 0 is a test of to see if that number is odd or even. So we're saying if that number module 2 = 0, then we know the number is even. Therefore we can perform the action that is indented from the if keyword. And in this case here it looks like we're doing is we're doing a sum of even. So we're gonna add up all the even numbers in the list, but we're only incrementing our sum of evens variable if that condition satisfies true.

00:01:55 Speaker 2

Lastly, we expect you to understand opening to the console, and that's as simple as calling the print function. So you're doing print some of the even numbers is and then using the variable sum of events to provide that numeric output. So when we're looking at how much Python do I need to know if you're happy looking at Python code on screen right now, you have enough Python to be able to train and build a simple machine learning model. And we'll help you along the way with what extra Python packages we need to import and use. We'll teach you that along the way so you can achieve the model building outcome. In terms of AWS experience, what are we going to need you to know and what are we going to cover? Well, we really need you to be aware of how to get in and use AWS. So in other words, you're needing to have an AWS.

00:02:25 Speaker 2

And you need to know how to log into that account using what we call the AWS Management Console, the primary web interface from managing all AWS services. We don't expect you to be an expert on the Amazon S3 service, but having an awareness of what S3 is and the fact that it's a data storage service and we create buckets in there, how we might create the folders on our file system. We can create buckets on S3, but if we can use the Management Console web application to upload content like CSV files, for example, into the S3 bucket or download files from a bucket, that kind of level interaction is going to be required. So if you've not done that before, take a little bit of time out to explore the S3 interface in the Amazon Web Console and try not to load a file to a newly created bucket. Then try and download your file from that bucket.

00:02:55 Speaker 2

Actually be able to do that before we carry on with the rest of the course. Now thirdly, and really do take this as optional, it would be good to have an awareness of some of the compute services in AWS like EC2 from managing virtual machines or ECS for managing containers. Now it's not essential if you've not seen those before. It just might help you relate when we discuss some of the compute platform capabilities of Sagemaker to go, ah, right, that's similar to EC2 or that's similar to ECS, but it's not an absolute prerequisite.

00:03:14 Speaker 2

Having awareness of how names get resolved is also helpful for this course. Now DNS is network name service we use to resolve names like www.codecloud.com and turns out ultimately in an IP address. Now within AWS Route 53 is the name of our native DNS service in AWS. And again, it's helpful to understand how that works, but it's not essential. It will not hold you up if you've not used them before. Now in terms of your awareness of the machine learning process, this course is aimed

at people starting the machine learning journey. So your level of awareness of ML, we're setting a nice low bar for you because we're going to enable you to build models. We're going to enable you to understand what is happening during model training and during model inference. But what I would say is that if we've got a high level view of.

00:03:45 Speaker 2

It would help.

00:03:46 Speaker 2

So we're to start out with some source data. We're going to use some tabular data that's based on some house prices out of the City of London in the UK. We're going to pick a ready made machine learning algorithm. Now there are many ready made algorithms that we can use to solve common problems. We're going to use machine learning algorithm called linear learner, but there are many others like Xgboost or LGBM or PCA, numerous algorithms out there. And it's really the job of the data scientists to choose what off-the-shelf algorithm they want to use given the problem that they are faced with. And they may even try out more than one algorithm in their attempts to solve a particular problem. So with machine learning algorithm and our source data, we're going to be able to conduct some training. And when we run some training, we're going to use that algorithm to try and learn.

00:04:16 Speaker 2

Patterns of data. So we're going to try and get that line of best fit if we're getting 2 dimensions or surface of best fit. If we're thinking 3 dimensions or in a multi dimensional space, we're thinking about that hyperdimensional surface that we're trying to make a best fit. And this is all about machine learning algorithm working out which variables, which essentially input columns in our tabular source data are related to one another. Which one has a greater impact on house price? Is it suburban area versus city? Is it number of bedrooms? Is it number of bathrooms? Is it combination of bedrooms, bathrooms, square footage and the area of the city? And it works out those relationships so that when we present the model that we build with data that is not seen before, it can inference, it can infer.

00:04:46 Speaker 2

What a predicted output would be. So this is the wonderful thing about machine learning is that you are not programming all possible outcomes. You are teaching a model to understand the patterns, to be able to generalize the relationship between data so that it can be shown unseen data that it has never seen before and yet still predict a high accuracy value as an output. So the output of the training process is a trained model. Now, a trained model is nothing more than a file at this point. The R ones will be called like model TGZ doesn't do anything on its own. It's not a binary executable. It's a model file. So if we want to actually generate some predictions from that, what we're going to need to do is host that model somewhere with a little bit of code so that we can pass requests to the model. So I could say, hey, model, I've got a house here.

00:05:16 Speaker 2

4 bedrooms 6000 square foot and this post code this number bathrooms in suburban area, how much it would cost and it would use that generalization database it's been trained upon and it would generate a value for you. And under the hood there's just some complex maths going on and it spits out your value.

00:05:28 Speaker 2

So we need to host that model somewhere on some kind of compute platform. It could be a virtual machine, it could be a physical machine, it could be a container. Now on this course, we've got to concentrate on hosting that model on Sagemaker itself. So users using Sagemaker not just for training, our model for hosting it as well. So that's what we call inference hosting model so that we can pass an inference requests and the response we get back from that is the prediction. So in our example, that is the house price based on house characteristics we're going to pass into the model. Now, if you follow that topology for that structure, then you're going to be absolutely fine. On this course, we're going to show you how to pick the algorithm and show you how to create training job in Sagemaker. We're going to show you how to register your model. We're going to show you how to host your model and ultimately generate an inference request.

00:05:58 Speaker 2

And get a valid prediction back. So that's what we're doing. And that's the level of awareness you need about ML at the stage. Of course, everything else will fill in for you. So summarize, if you've got some very basic Python And ML concepts understood, you've got enough to get started on this training course. Remember, just simple things like understanding Python lists, using if statements and using for statements to iterate over a list and understanding that machine learning that we train a model based on source data to produce a model artifact that we must then host so that we can perform inference. In other words, generate predictions. We've also seen having some familiarity with AWS is going to help you. In particular, if we've seen before S3 uploading files or downloading files from S3 buckets. And if you maybe created a virtual machine in EC2, that might help you a little.

00:06:28 Speaker 2

In terms of understanding what Sagemaker is doing behind the scenes, if you can relate that concept, but again, understanding EC2 is entirely optional and just help your learning. It's not a blocker getting started. So with that, that wraps up the short prerequisite lesson. In our next lesson, we'll be jumping into some machine learning fundamental concepts just so we get solid grounding in what's going on in machine learning. And I might get a bit of maths in there too. See you there.

Agenda

- 01 Lowering the prerequisites for ML learning
- 02 Targeting beginners in their ML journey
- 03 Focus on AWS SageMaker as the preferred ML platform

© Copyright KodeCloud



Python Basics

Write a simple function

Store values in a list

Store values in variables

Iterate over a list

Use conditional statements

Output to console

© Copyright KodeCloud

```
sum_of_evens.py

# Define a list of numbers
numbers = [1, 2, 3, 4, 5]

# Define a variable to hold the sum of even numbers
sum_of_evens = 0

# Iterate over the list
for number in numbers:
    # Check if the number is even
    if number % 2 == 0:
        sum_of_evens += number # Add even numbers to the sum

# Output the result to the console
print("The sum of even numbers is:", sum_of_evens)
```



AWS Basics

01

Must have used
the AWS
Management
Console

02

Can upload
and download
files to an S3
bucket

03

Awareness of
EC2, ECS,
Route53 is helpful
but not required

© Copyright KodeCloud



ML Awareness



© Copyright KodeCloud



Summary

- 01** Basic Python and ML concepts are sufficient to get started
- 02** Familiarity with AWS and its services enhances the learning experience

