Course Description

At the end of day, the value of Data Scientists rests on their ability to describe the world and to make predictions. Machine Learning is the field of teaching machines and computers to learn from existing data to make predictions on new data - will a given tumor be benign or malignant? Which of your customers will take their business elsewhere? Is a particular email spam or not? In this course, you'll learn how to use Python to perform supervised learning, an essential component of Machine Learning. You'll learn how to build predictive models, how to tune their parameters and how to tell how well they will perform on unseen data, all the while using real world datasets. You'll do so using scikit-learn, one of the most popular and user-friendly machine learning libraries for Python.

---

1   Classification   FREE

0%

In this chapter, you will be introduced to classification problems and learn how to solve them using supervised learning techniques. Classification problems are prevalent in a variety of domains, ranging from finance to healthcare. Here, you will have the chance to apply what you are learning to a political dataset, where you classify the party affiliation of United States Congressmen based on their voting records.

VIEW CHAPTER DETAILS

Continue Chapter

---

2   Regression

0%

In the previous chapter, you made use of image and political datasets to predict binary as well as multiclass outcomes. But what if your problem requires a continuous outcome? Regression, which is the focus of this chapter, is best suited to solving such problems. You will learn about fundamental concepts in regression and apply them to predict the life expectancy in a given country using Gapminder data.

VIEW CHAPTER DETAILS

Continue Chapter

---

3 Fine-tuning your model

0%

Having trained your model, your next task is to evaluate its performance. What metrics can you use to gauge how good your model is? So far, you have used accuracy for classification and R-squared for regression. In this chapter, you will learn about some of the other metrics available in scikit-learn that will allow you to assess your model's performance in a more nuanced manner. You will then learn to optimize both your classification as well as regression models using hyperparameter tuning.

VIEW CHAPTER DETAILS

Continue Chapter

4 Preprocessing and pipelines

0%

This chapter will introduce the notion of pipelines and how scikit-learn allows for transformers and estimators to be chained together and used as a single unit. Pre-processing techniques will be then be introduced as a way to enhance model performance and pipelines will be the glue that ties together concepts in the prior chapters.

VIEW CHAPTER DETAILS