CAREERFOUNDRY

Machine Learning with Python: Weather Conditions and Climate Change with ClimateWins

Objective

This project is a continuation of the project you worked on in Achievement 1. In this project, you'll build on the skills you've learned by practicing them in various real-life situations. You'll continue with the same example situation used in the first Achievement: where you're working as a data analyst at a nonprofit organization called ClimateWins.

Introduction and Company Profile

In the first part of this course, you completed your training program with ClimateWins. Now, however, ClimateWins is interested in more-practical applications of machine learning—for instance, if machine learning can be used to help predict the consequences of climate change around Europe and, potentially, the world. ClimateWins still has limited resources, though, so rather than creating a new product immediately, it's tasking its newest data analyst, data scientist, and researcher (you) to come up with some potential best options for moving forward in this area. This will require knowledge of advanced machine learning models!

While you learned a lot about supervised learning in Achievement 1, these more-complex models take advantage of *unsupervised* machine learning. You'll also take your first foray into practical applications of machine learning. Next, you'll look at composite models and newer machine learning algorithms to determine what might hold the answers for ClimateWins' goals. With your newfound knowledge, you'll recommend what resources, algorithms, and data ClimateWins will need to predict weather changes.

Project Background

To reiterate what you know about your employer, ClimateWins wants you to assess what tools are available to categorize and predict the weather in mainland Europe. It's concerned with the increase in extreme weather events, especially in the past 10 to 20 years.

However, it thinks that even weather extremes could be predicted and planned for using advanced tools such as machine learning. With data from the past century, it hopes to create a model for what the future will hold.

In this next iteration, ClimateWins has a few areas it wants to cover:

- Finding new patterns in weather changes over the last 60 years.
- Identifying weather patterns outside the regional norm in Europe.

- Determining whether unusual weather patterns are increasing.
- Generating possibilities for future weather conditions over the next 25 to 50 years based on current trends.
- Determining the safest places for people to live in Europe within the next 25 to 50 years.

Climate refugees are a growing reason for emigration globally, including in Europe.
Source: Steve Evans (CC BY-NC 2.0)

Context

Throughout this Achievement, you'll learn more about advanced machine learning algorithms and how they can be used to predict and generate new data. By the time you finish, you should be able to run and balance complex machine learning algorithms and understand where and how these models are used.

Data Sets

You'll be using a data set based on weather observations from 18 different weather stations across Europe containing data from the late 1800s to 2022. Recordings exist for almost every day with values such as temperature, wind speed, snow, global radiation, and more. This data is collected by the European Climate Assessment & Data Set project.

- Create a "Data Sets" directory where you can download raw data.
- Download the temperature data set (.csv, 16.6mb) (clicking this link will download the data set to your computer). Move it to your "Data Sets" directory so that it's easy to access.

- Create “Supervised” and “Unsupervised” folders inside your “Data Sets” directory.

Project Deliverables

Throughout this Achievement, you’ll learn about some of the most powerful types of machine learning and how to optimize their results. You’ll use these tools to develop a proposal for ClimateWins.

For the task in each Exercise, you’ll submit a deliverable that directly contributes to the final product—in this case, a proposal of a final algorithm to achieve ClimateWins’ goals. Your mentor will review the deliverables for Exercises 2.1 through 2.5, as well as your final proposal in Exercise 2.6, after which they’ll give you feedback on how your project relates to the work of a data analyst.

Learning Objectives

Below is a breakdown of what you will do in each Exercise.

Exercise 2.1 Unsupervised Machine Learning

- Differentiate between unsupervised and supervised learning and identify the key differences in their approaches and applications
- Implement unsupervised learning algorithms to identify patterns and structures in data without relying on predefined labels or target variables
- Compare and contrast between clustering, association, and dimensional reduction algorithms by putting them in practice
- Use unsupervised learning to find insights about European weather conditions over the last 80 years

Exercise 2.2 Complex Machine Learning Models and Keras Part 1

- Set up complex machine learning problems using Keras and Tensorflow
- Set up a machine learning model to predict participant behavior
- Set up CNN and RNN networks

Exercise 2.3 Complex Machine Learning Models and Keras Part 2

- Apply random forests and support vector machines algorithms to real-world machine learning problems
- Use deep learning models with the Keras library to help direct and refine your results.
- Use deep learning models with advanced techniques to accurately detect changes in human works.

Exercise 2.4 Evaluating Hyperparameters

- Analyze the impact of hyperparameters and tuning on the performance of deep learning models
- Compare and contrast the advantages and disadvantages of various hyperparameter settings
- Use the correct hyperparameters for your specific machine learning problem to iterate on your results

Exercise 2.5 Visual Applications of Machine Learning

- Define how machine learning and image recognition can work together
- Create a handwriting discriminator
- Frame a problem as a supervised learning problem using generative adversarial networks (GANs)
- Determine how to decode images to inform your machine learning results for your ClimateWins project

Exercise 2.6 Presenting Your Final Results

- Create a presentation that compiles all your findings for ClimateWins
- Create a thought experiment
- Practice applying the correct mindset when approaching machine learning problems