Build and deploy Node.js, Java, and Python apps with Azure. LEARN MORE
HIDE AD • AD VIA BUYSELLADS

Saving a machine learning Model

Difficulty Level : Easy • Last Updated : 24 Jun, 2021

In machine learning, while working with scikit learn library, we need to save the trained models in a file and restore them in order to reuse it to compare the model with other models, to test the model on a new data. The saving of data is called *Serialization*, while restoring the data is called *Deserialization*.
Also, we deal with different types and sizes of data. Some datasets are easily trained i.e- they take less time to train but the datasets whose size is large (more than 1GB) can take very large time to train on a local machine even with GPU. When we need the same trained data in some different project or later sometime, to avoid the wastage of the training time, store trained model so that it can be used anytime in the future.
There are two ways we can save a model in scikit learn:

1. **Pickle string**: The pickle module implements a fundamental, but powerful algorithm for serializing and de-serializing a Python object structure.

Pickle model provides the following functions –
pickle.dump *to serialize an object hierarchy, you simply use dump().*
pickle.load *to deserialize a data stream, you call the loads() function.*

1. **Example:** Let's apply K Nearest Neighbor on iris dataset and then save the model.

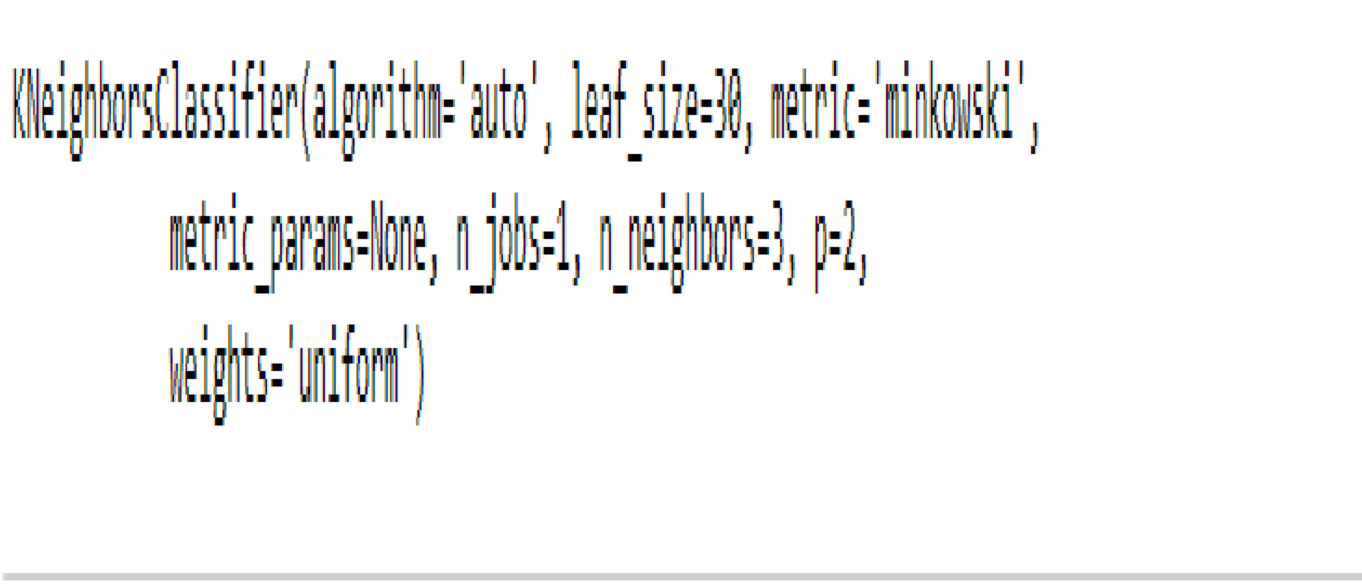
▲

Build and deploy Node.js, Java, and Python apps with Azure. LEARN MORE

HIDE AD • AD VIA BUYSELLADS

```
import numpy as np

# Load dataset
from sklearn.datasets import load_iris
iris = load_iris()

X = iris.data
y = iris.target

# Split dataset into train and test
X_train, X_test, y_train, y_test = \
    train_test_split(X, y, test_size = 0.3,
                        random_state = 2018)

# import KNeighborsClassifier model
from sklearn.neighbors import KNeighborsClassifier as KNN
knn = KNN(n_neighbors = 3)

# train model
knn.fit(X_train, y_train)
```

1.



```
KNeighborsClassifier(algorithm='auto', leaf_size=30, metric='minkowski',
           metric_params=None, n_jobs=1, n_neighbors=3, p=2,
           weights='uniform')
```

1. Save model to string using pickle –

Build and deploy Node.js, Java, and Python apps with Azure.

LEARN MORE

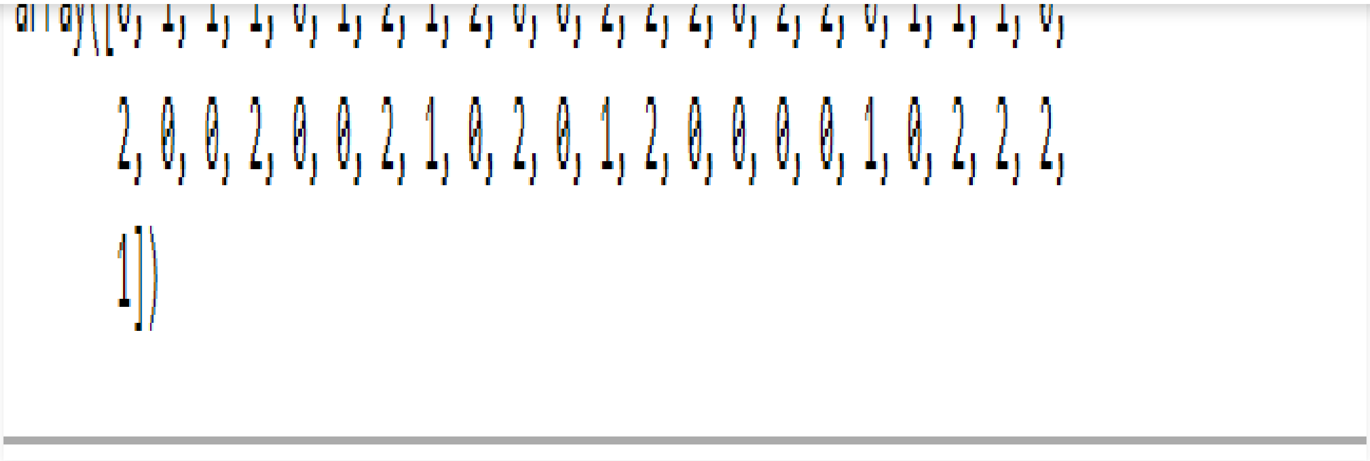
HIDE AD • AD VIA BUYSELLADS

```
import pickle

# Save the trained model as a pickle string.
saved_model = pickle.dumps(knn)

# Load the pickled model
knn_from_pickle = pickle.loads(saved_model)

# Use the loaded pickled model to make predictions
knn_from_pickle.predict(X_test)
```

1. **Output:**

Build and deploy Node.js, Java, and Python apps with Azure. LEARN MORE

HIDE AD • AD VIA BUYSELLADS



```
array([0, 1, 1, 1, 0, 1, 2, 1, 2, 0, 0, 2, 2, 2, 0, 2, 2, 0, 1, 1, 1, 0,
       2, 0, 0, 2, 0, 0, 2, 1, 0, 2, 0, 1, 2, 0, 0, 0, 0, 1, 0, 2, 2, 2,
       1])
```

1.
2. **Pickled model as a file using joblib**: Joblib is the replacement of pickle as it is more

Related Articles

joblib.dump *to serialize an object hierarchy*

joblib.load *to deserialize a data stream*

1. Save to pickled file using joblib –

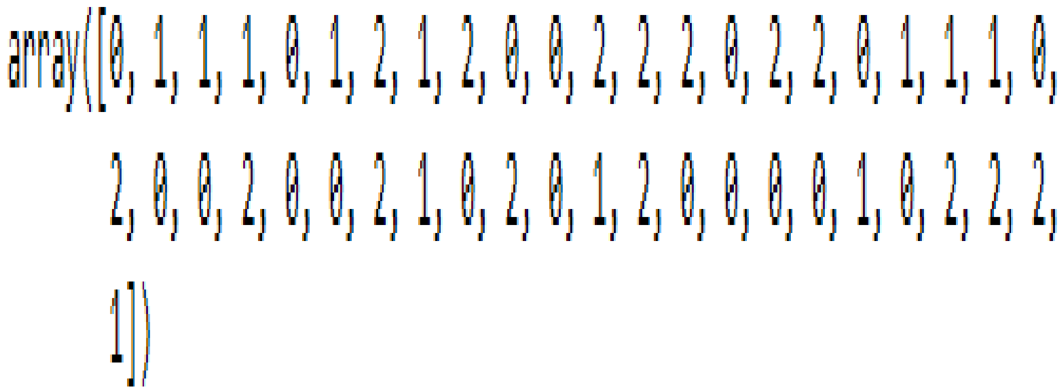
Python3

```
from sklearn.externals import joblib

# Save the model as a pickle in a file
joblib.dump(knn, 'filename.pkl')

# Load the model from the file
knn_from_joblib = joblib.load('filename.pkl')

# Use the loaded model to make predictions
knn_from_joblib.predict(X_test)
```

▲

Build and deploy Node.js, Java, and Python apps with Azure. LEARN MORE

HIDE AD • AD VIA BUYSELLADS



```
array([0, 1, 1, 1, 0, 1, 2, 1, 2, 0, 0, 2, 2, 2, 0, 2, 2, 0, 1, 1, 1, 0,
       2, 0, 0, 2, 0, 0, 2, 1, 0, 2, 0, 1, 2, 0, 0, 0, 0, 1, 0, 2, 2, 2,
       1])
```

Attention reader! Don't stop learning now. Get hold of all the important Machine Learning Concepts with the **Machine Learning Foundation Course** at a student-friendly price and become industry ready.

Build and deploy Node.js, Java, and Python apps with Azure.

LEARN MORE

HIDE AD • AD VIA BUYSELLADS

Previous

Next

ADVERTISEMENT BY ADRECOVER

Rawdacious Pack

This footwear collection is ready to help unleash your potential and achieve your goals.

Nike

RECOMMENDED ARTICLES

Page : 1 2 3

01 ML - Saving a Deep Learning model in Keras
12, May 20

02 Learning Model Building in Scikit-learn : A Python Machine Learning Library
17, Feb 17

03 Creating a simple machine learning model
13, Jun 18

04 seq2seq model in Machine Learnir
05, Dec 18

05 Deploy Machine Learning Model using Flask
30, Sep 19

06 Machine Learning Computing at the edge using model artifacts
26, Apr 20

07 Deploy a Machine Learning Model using Streamlit Library
28, Aug 20

08 Python - Create UIs for prototyping Machine Learning model with Gradio
27, Aug 20

Build and deploy Node.js, Java, and Python apps with Azure. LEARN MORE

HIDE AD • AD VIA BUYSELLADS

Article Contributed By :

aakarsha_chugh

@aakarsha_chugh

Vote for difficulty

Current difficulty : Easy

Easy Normal Medium Hard Expert

Improved By : as5853535

Article Tags : Advanced Computer Subject, Machine Learning

Practice Tags : Machine Learning

Improve Article Report Issue

Writing code in comment? Please use ide.geeksforgeeks.org, generate link and share the link here.

Load Comments

ADVERTISEMENT BY ADRECOVER

Rawdacious Pack

This footwear collection is ready to help unleash your potential and achieve your goals.

Nike

▲

Build and deploy Node.js, Java, and Python apps with Azure.
LEARN MORE
HIDE AD • AD VIA BUYSELLADS

Sector-136, Noida, Uttar Pradesh - 201305

feedback@geeksforgeeks.org

Company

- About Us
- Careers
- Privacy Policy
- Contact Us
- Copyright Policy

Learn

- Algorithms
- Data Structures
- Languages
- CS Subjects
- Video Tutorials

Practice

- Courses
- Company-wise
- Topic-wise
- How to begin?

Contribute

- Write an Article
- Write Interview Experience
- Internships
- Videos

@geeksforgeeks , Some rights reserved

▲